FIBRE CHANNEL FUNDAMENTALS

FIBRE CHANNEL INDUSTRY ASSOCIATION

www.fibrechannel.org

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

Cover Design: Kit Orn, Phoenix, Arizona
Interior Design and Production Services: Kelly Johnson, El Cajon, California
Author: Scott Kipp, McDATA Corporation, Broomfield, Colorado
Illustrator: Ginny McDonald, McDATA Corporation, Broomfield, Colorado
Transceiver Illustrator: Cleaver Brinkerhoff, Molex Incorporated, Lisle, Illinois

ISBN 1-885982-03-8

FOREWORD

Fibre Channel Fundamentals describes the basic functions of Fibre Channel and is a publication of the Fibre Channel Industry Association. The companion publication, *Fibre Channel Storage Area Networks*, investigates the benefits and application of Fibre Channel in Storage Area Networks (SANs).

Fibre Channel Fundamentals explores the nuts and bolts of Fibre Channel. Through a fictitious video application, the underlying processes of Fibre Channel are investigated. The application shows how Fibre Channel scales and uses different topologies. The physical layer and interconnect devices that form the Fibre Channel fabric are described and illustrated. End devices, such as storage devices, Host Bus Adapters (HBAs), and translation devices, are also discussed with the Fabric Services that empower Fibre Channel.

Fibre Channel Fundamentals is a handy reference for further explorations into the protocol upon which SANs are founded. The underlying ANSI standards that form the foundation of Fibre Channel are described, as well as other standards used in SANs. Efforts to ensure that products comply with these industry standards are documented in this publication as well.

A number of valuable references are listed in the reference chapter of this publication, including books and websites. These references contain information on a variety of SAN topics. Acronyms are also listed for quick reference.

[This is a blank page.]

CONTENTS

Chapter 1

INTRODUCTION

Fibre Channel is the networking technology that reliably connects storage and servers at gigabit speeds to create Storage Area Networks (SANs). Fibre Channel was designed to be the I/O interface of the future and to overcome the limitations of Direct Attached Storage (DAS). DAS is tied to the host bus of a computer or to a Small Computer Systems Interface (SCSI) bus as seen in Figure 1. DAS has scalability, availability, and distance (SAD) limitations that are overcome by Fibre Channel.

A comparison of DAS and SANs shows the benefits of Fibre Channel. DAS typically allows only one initiator on the SCSI bus. SANs, including ones based on arbitrated loop, enable multiple initiators to access data through the same port. SANs, including ones based on arbitrated loop, enable multiple initiators to access the same data. Accessibility of data is key, since databases and applications have become increasingly powerful. The value and accessibility of terabytes of data should not be held captive by one server or several directly attached servers.

SANs allow hundreds of servers to access large databases that may be stored across multiple storage devices. Since multiple servers can access the data, the availability of the data increases with SANs. Fibre Channel also uses fiber-optics that allow connectivity over hundreds of kilometers for individual links. The improvements in scalability, availability, and distance provided by Fibre Channel overcome the SAD limitations of DAS.

Fibre Channel is evolving as the interface solution for the data center. As data centers grow from terabytes (10^{12} bytes) of data to petabytes (10^{15} bytes) of data, Fibre Channel based storage devices are evolving into massive repositories. Disk drives can store hundreds

Fibre Channel Advantages

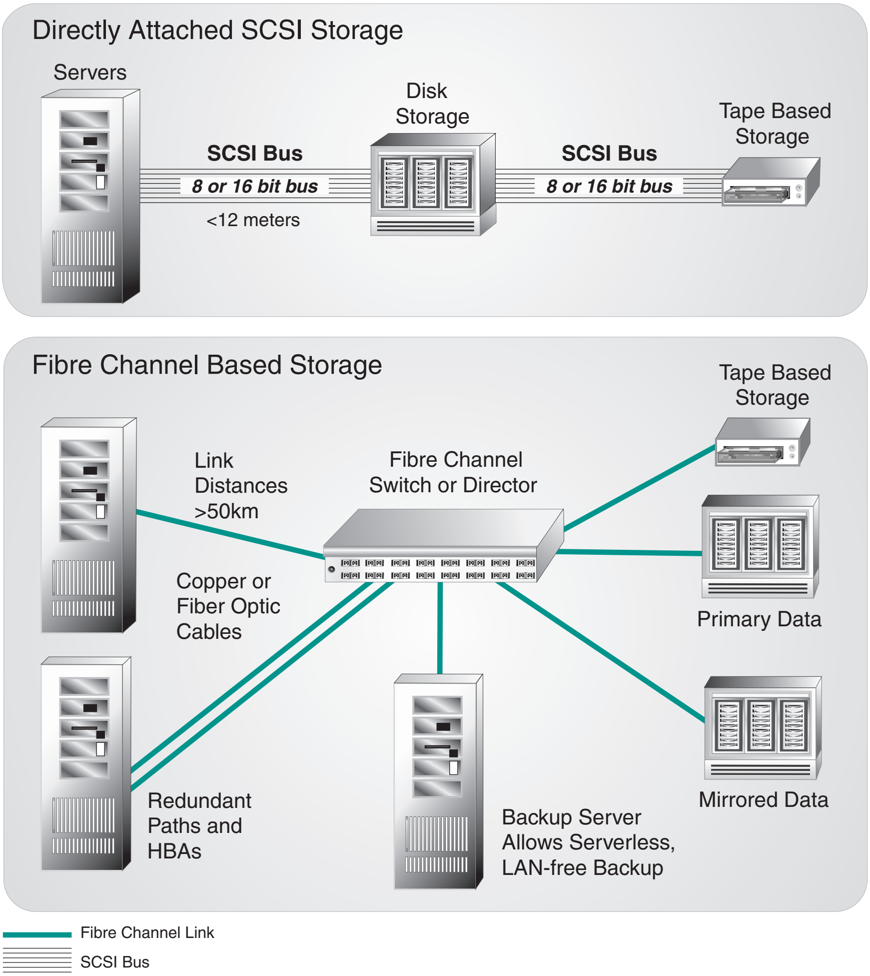


Figure 1: Traditional open storage networks based on the SCSI architecture limit the scalability, availability, and distance (SAD) of the storage network. SANs based on the Fibre Channel switched serial architecture go beyond each of these limitations and also offer advanced management and security.

of gigabytes, while Just a Bunch of Disks (JBODs) can store terabytes. Storage subsystems can store hundreds of terabytes, while tape libraries store petabytes. As data centers grow from hundreds of servers to thousands of servers, Fibre Channel fabrics provide thousands of ports of connectivity. As data centers grow from campus wide to intercontinental, Fibre Channel uses Wavelength Division Multiplexing (WDM) and Fibre Channel over Internet Protocol (FCIP) to go the distance. The Fibre Channel standards will continue to evolve to solve the problems that confront the information age.

THE SERVANT OF UPPER LEVEL PROTOCOLS

Upper Level Protocols (ULPs) have many connectivity options, as seen in Figure 2. The connectivity layer of Fibre Channel allows connectivity to millions of devices, while the SCSI bus limits connectivity to about 10 devices. IEEE-1394 limits connectivity to thousands of devices, but lacks the facilities to properly manage such a large network. Fibre Channel is scalable for data center applications, while other physical interfaces are limited.

What is amazing about Fibre Channel is that it allows multiple protocols to run simultaneously over the same physical connection. Fibre Channel was designed to be protocol independent so that many types of applications can utilize the SAN. Examples of ULPs that use Fibre Channel are shown in Table 1. The wide variety of ULPs shows how adaptable Fibre Channel is to a cornucopia of computer applications.

Table 1. Upper Level Protocol Mappings

Upper Level Protocol	**Application**
Fibre Channel Protocol for SCSI (FCP)	FCP is the most widely used ULP and maps SCSI commands to the Fibre Channel interface.
Fibre Channel Protocol for SCSI – Second Version (FCP-2)	The second version of FCP improves error recovery and uses the SCSI-3 Architecture Model (SAM-2).
Fibre Channel Single Byte Command Code Sets -2 Mapping Protocol (FC-SB-2)	FC-SB-2 maps FICON commands to the Fibre Channel interface so that IBM's zSeries mainframes can utilize Fibre Channel.
Fibre Channel Virtual Interface (FC-VI)	FC-VI defines how VI Information Units are mapped to Fibre Channel and how Fibre Channel Services are used to perform VI services.
IP and ARP over Fibre Channel (IPFC)	This mapping defines how the Internet Protocol can use Fibre Channel fabrics.
Fibre Channel Audio-Visual (FC-AV)	This maps digital audio and video protocols like MPEG-2 and MPEG-3 to Fibre Channel.

ULP Connectivity

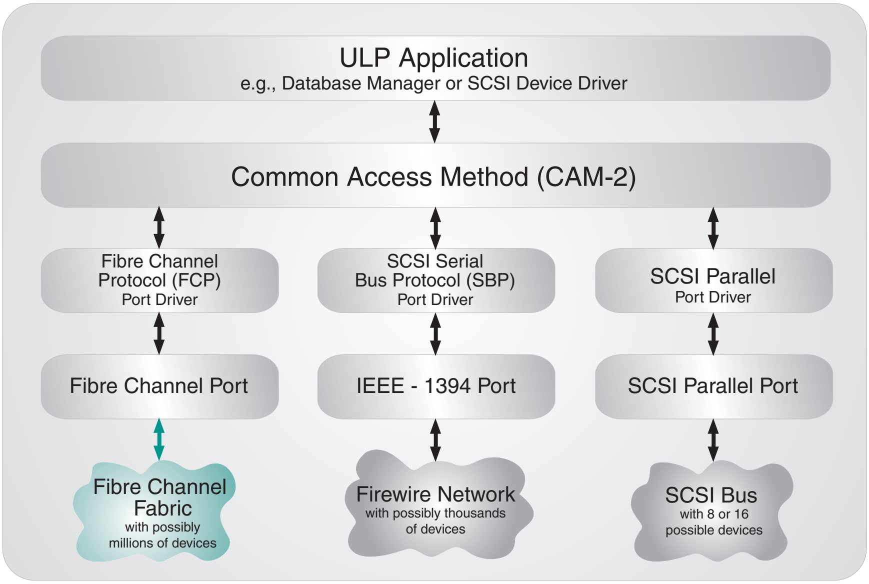


Figure 2: Upper Level Protocols can use a variety of interconnect technologies to connect to other devices. The most scalable solution is Fibre Channel that already supports thousands of ports in single fabrics and can scale to millions of ports. The SCSI bus limits connectivity to only 8 or 16 addresses. IEEE-1394, also known as FireWire, can theoretically support thousands of devices, but is practically limited to tens of devices.

INDEPENDENCE OF SERVERS AND STORAGE

Fibre Channel defines a fabric or storage network between servers and storage. Before Fibre Channel, most computers used DAS to provide storage for a server. As the server/client model grew and succeeded, the amount of data that was stored behind an individual server grew to levels that made it impossible for one server to serve all of the data. Storage devices developed multiple ports but could still only service a small number of servers.

The Fibre Channel SAN was created to allow a single storage port to support multiple servers, as seen in Figure 3. From the perspective of a storage port, fan-out refers to the number of server ports a single storage port services. Fan-out allows multiple servers to access the data behind a storage port. Fan-out ratios are used in SAN design to ensure that storage ports and Inter-Switch Links (ISLs) are not overutilized.

Applications that use large databases that hold terabytes of data may require storage to be held on multiple storage devices, as seen at the bottom of Figure 3. The database application running on a server will need to access multiple storage devices. Fan-in is the ratio of the number of storage ports a single server can access. If a server port has a high fan-in rate, the server may rarely use a single storage port, and fabrics can be designed accordingly. Fan-in demonstrates how Fibre Channel allows a server to be independent from storage in a SAN.

Server and Storage Independence

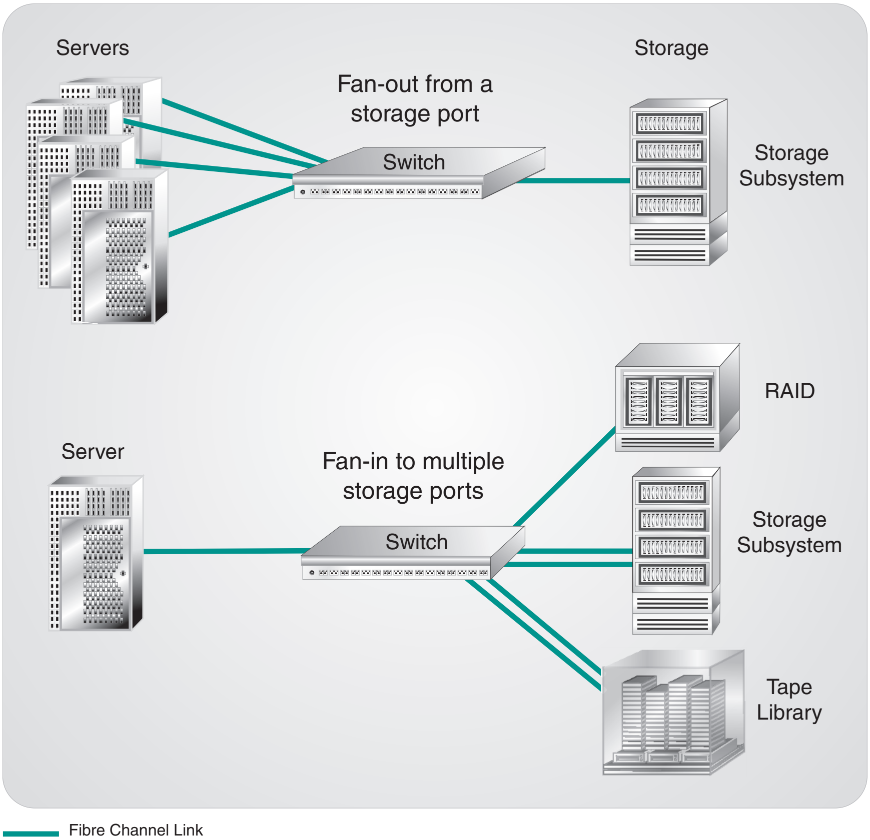


Figure 3: Independence of servers and storage are enabled by Fibre Channel SANs. Fan-out refers to how many server ports a single storage port can service. Fan-in refers to how many storage ports a single server port can access.

HISTORY AND FUTURE OF FIBRE CHANNEL

Fibre Channel began in American National Standards Institute (ANSI) committees in the late 1980s. The first Fibre Channel products were released in the mid 1990s and were largely based on the arbitrated loop topology. In the late 1990s, switches became more prevalent and fabrics found wider deployment. The new millennium has seen thousand-port fabrics being deployed across multiple data centers.

As the size of these data centers has grown, the complexity of managing and securing the fabric has grown. ANSI, the Fibre Channel

Industry Association, the Storage Network Industry Association and the Fibre Channel vendors are addressing these concerns so that Fibre Channel will scale through the 21st century. Management and security of SANs is a paramount concern that is being tackled from many angles.

The basic functionality of Fibre Channel continues to grow as the Fibre Channel standards evolve. The T11 technical committee of the InterNational Committee of Information Technology Standards (INCITS) actively continues to develop Fibre Channel standards for ANSI. The Internet Engineering Task Force (IETF) has developed several protocols that enable SANs to traverse global IP networks. Management applications are also evolving to provide effective management and security for large SANs with thousands of ports. Fibre Channel continues to evolve to meet the escalating demands of corporate data centers.

SUMMARY

The goal of this publication is to demonstrate how Fibre Channel is the dominant interconnect technology for SANs. Chapters 2 and 3 describe how the different Fibre Channel topologies are used in different applications. The remaining chapters explain the latest developments in hardware and software that make Fibre Channel the best implementation for today and tomorrow. The implementation of Fibre Channel continues on a steep growth curve as more and more Information Technology (IT) professionals discover the power of Fibre Channel. Fibre Channel will remain the dominant SAN protocol for many years to come.

- *The Servant of Upper Level Protocols*: Fibre Channel provides a conduit for multiple ULPs.
- *Independence of Servers and Storage*: Fibre Channel SANs liberate storage from an individual server and allow servers to access multiple storage devices.
- *History and Future of Fibre Channel*: The Fibre Channel standards and fabric continue to evolve to meet the demands of corporate data centers.

[This is a blank page.]

Chapter 2

A SIMPLE APPLICATION

A simple application of Fibre Channel technology shows how some of the basic processes of Fibre Channel work. Simple point-to-point connections and arbitrated loops show how Fibre Channel can be used by a small TV station. The Fibre Channel technology discussed in this chapter has been available since the mid 1990s.

KLLR is a television station in a metropolitan area with over 500,000 people. KLLR has decided to make its broadcast available on the Internet. Cable TV companies will pick up the high quality digital signal from the Internet and will rebroadcast KLLR's signal for a fee. KLLR will also archive their signals digitally and allow old programs and video clips to be downloaded from the Internet. KLLR's leap into the digital age makes their broadcasts available around the world.

SAN PLANNING

At the start of any large project, KLLR always develops a fairly thorough plan. One of the key early decisions they made was to store their video in the MPEG-2 digital video format. MPEG-2 is the digital video standard used on DVDs, digital cable broadcasts, and direct broadcast satellite like DirectTV. By using the MPEG-2 format, end users can edit the video on a standard home computer or in a high-end, video-editing studio.

KLLR chose an encoding rate of 587 KBps (4.7 Mbps). At this quality setting, a 30 minute program without commercials consumes less than one gigabyte of storage. Table 2 shows some relative storage requirements for storing digital video. It's easy to see that KLLR had some serious storage requirements. KLLR was no stranger to storing large amounts of video. Most of their material had been stored on Betamax tapes, but since 1999 most of their production was stored digitally on DV (which is discussed later in this chapter). Even with the high compression rates of MPEG-2, KLLR would fill their 2 TB RAID to capacity in about a month and would need to back it up to tape.

Table 2. Digital Video Programs vs. Storage Requirements

Video Time	**MPEG-2 (DVD) Storage Requirements**
1 Second of Video	587 kB
30 Second TV commercial	17.6 MB
3 Minute Music Video	106 MB
5 Minutes of Video	264 MB
22 Minute Program without Commercials	776 MB
30 Minute Program with Commercials	1.1 GB
1 Hour Program	2.1 GB
2 Hour Movie	4.2 GB ~ 1 DVD
3 Hour Football Game	6.3 GB < 2 DVDs
1 Day of Programming	51GB 11 DVDs
1 Week of Programming	355 GB 77 DVDs
1 Month of Programming	1.5 TB 320 DVDs
1 Quarter of Programming	4.6 TB 960 DVDs
1 Year of Programming	18.5 TB 3,840 DVDs

Table 2: Video stored in the MPEG-2 format seems minor at less than one MBps, but its relentless accumulation adds up to terabytes every month. A mid-sized tape library would be needed to back up over 20 terabytes of digital video a year. However, the relatively small size of a music video allows many Internet users to download the files quickly.

For the initial rollout of the service, KLLR installed the SAN shown in Figure 4. The SAN was a combination of an encoding server, Redundant Array of Independent Drives (RAID), video servers, and a tape library. This SAN used a dedicated point-to-point connection between the encoding server and the RAID and between the RAID and the cable TV video server. To reach several servers for a low price, arbitrated loop topology was used between the RAID and the video servers.

KLLR's Initial SAN

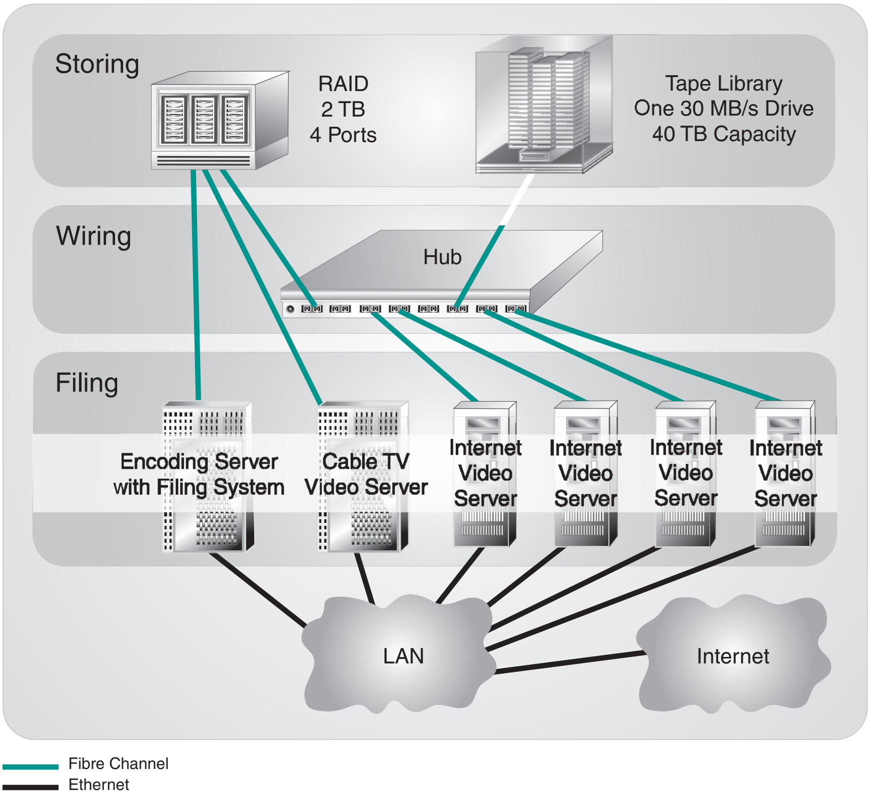


Figure 4: KLLR's initial SAN used a dedicated point-to-point connection from the RAID to the encoding server. The RAID was also connected to a hub for distributing the video to the several Internet video servers and the tape library. Three basic aspects of the SAN are storing, wiring and filing. Storing is done by the storage devices and filing is typically held in the servers. Wiring is the domain of Fibre Channel and no other technology compares to its proven reliability, scalability, and performance.

KLLR's SAN revolved around the 2 TB RAID that consisted of 40 disk drives that held 50 GB of data each. The RAID controller supported four Fibre Channel ports. The RAID had a total capacity for about one month's worth of programming and other data. The

RAID also supported the filing system, the metadata on the video files, and the 200 GB of cached programs that were most recently accessed. The filing system that contained the metadata for the video files was maintained in a database that was constantly updated by the encoding server.

POINT-TO-POINT TOPOLOGY

This chapter continues with a discussion of the basic topologies and data transfer process for this application. The point-to-point topology uses the simplest type of connection between two devices or nodes. When two Node_Ports (N_Ports) are connected and powered on, the link goes through the simple procedure shown in Figure 5. The initialization procedures are defined in Fibre Channel Framing and Signaling (FC-FS), the standard that defines the basic protocols for Fibre Channel.

Speed Negotiation is the second step in the initialization process after the port is powered up. This optional procedure determines the highest speed at which the link operates. Speed Negotiation enables a transceiver to run at either one and two Gbps. The third step in the initialization process is link initialization that follows the Port State Machine (PSM). The four stages of the PSM are Offline Primitive Sequence (OLS), Link Reset Primitive Sequence (LR), Link Reset Response Primitive Sequence (LRR), and Idle. When the link reaches the Idle state, then it is capable of transferring frames.

After the link is initialized, the initiator (server) sends the Port Login (PLOGI) command to the target (storage device) in the fourth step of the initialization process. The PLOGI establishes the service parameters between the N_Ports so that they can communicate efficiently. The service parameters include time-out values, credit between the ports, supported classes of service, and port and node attributes. When the PLOGI is complete, the N_ ports are ready for ULPs to introduce themselves.

The final step that enables an application to use the link is the Process Login (PRLI) command. The PRLI establishes protocol specific information so that ULP specific processes can be established.

The PRLI command establishes the service parameters for the ULP so that data can be sent over the Fibre Channel link.

The benefits of the Fibre Channel point-to-point topology are that it is a dedicated connection and that it is simple to implement and manage. KLLR wanted a dedicated connection with minimal disruptions for the video encoding, so they used the point-to-point topology. The disadvantages of point-to-point links are that the link may be underutilized and single points of failure exist. The switched Fibre Channel architecture is more commonly deployed today because it enables redundancy between the devices and allows access to multiple devices. Switched Fabrics will be explained in the next chapter.

Point-to-Point Connection

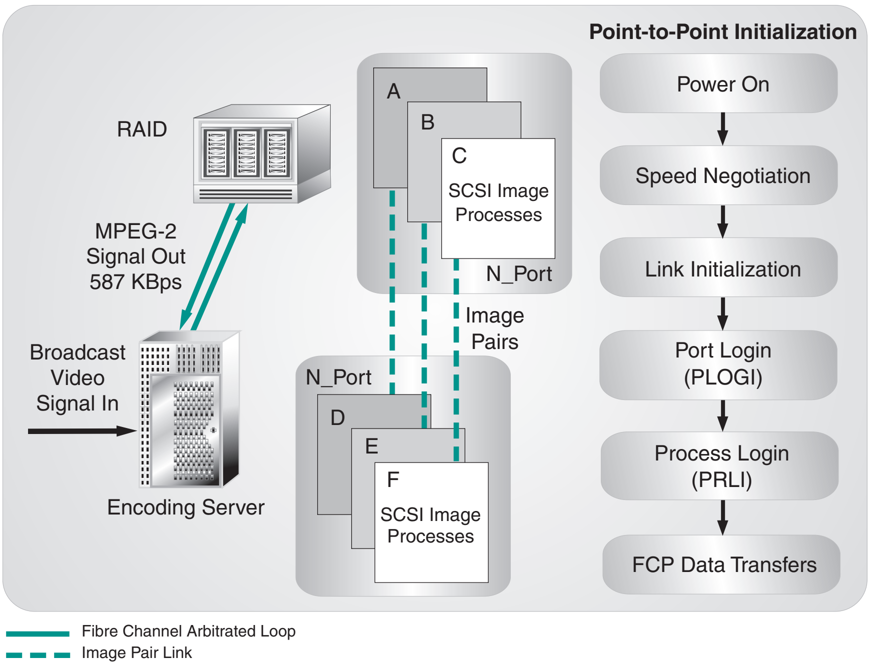


Figure 5: KLLR uses a point-to-point connection between the encoding server and the RAID. The encoding server also uses the link to send metadata about the video files to the RAID. In this example, the N_Ports have three SCSI Image Processes that are sending SCSI block traffic over the Fibre Channel point-to-point link.

DATA TRANSFERS

For this application, the ULP that uses the point-to-point link is the Fibre Channel Protocol for SCSI (FCP). FCP defines the mapping between the SCSI Architecture Model (SAM) and the Fibre Channel interface. The PRLI command establishes image pairs between the two N_Ports as seen in Figure 6. The image pairs send FCP commands so that FCP Information Units (IUs) can be exchanged between the devices.

Multiple image pairs can be established simultaneously. In this application, one image pair writes the digitized video signal onto the RAID. Another image pair is responsible for writing the filing structure and metadata about the video files onto the RAID. Several FCP processes run simultaneously to give Fibre Channel its power.

Data Transfers

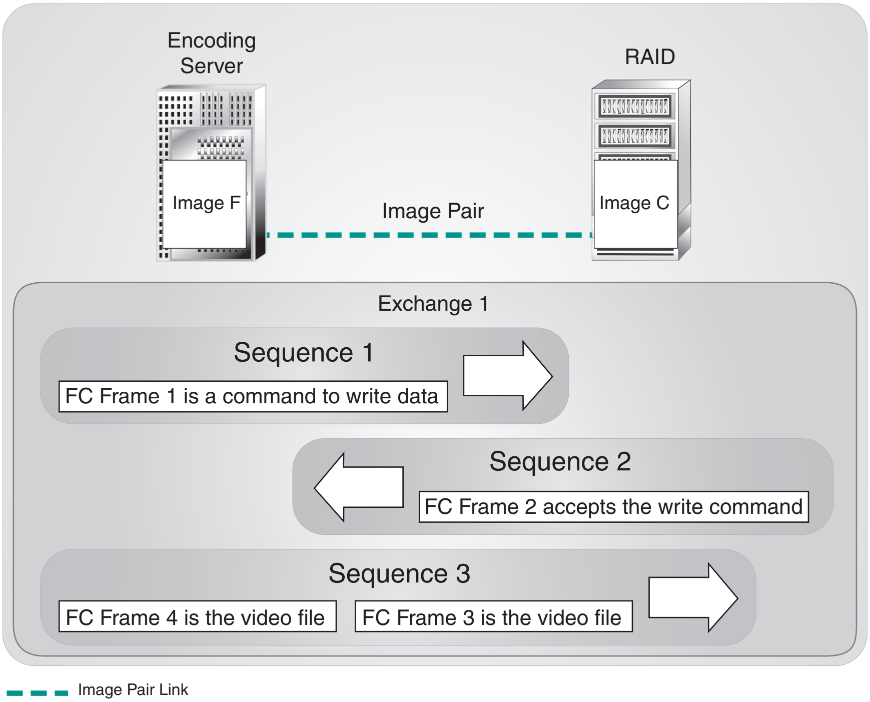


Figure 6: The transfer of data between ULP images is completed within an exchange. Each exchange may consist of several sequences, and each sequence may contain several Fibre Channel frames.

After an FCP image pair is established, data is transferred between the images in an exchange. An exchange may consist of several sequences that contain commands or data. While exchanges travel in both directions, individual sequences travel in only one direction. For example, the first sequence is a command to request a write of video data to the RAID. The second sequence is sent from the RAID to the encoding server, telling the server that the RAID is ready for the data. In the third sequence, the encoding server sends the video file to the RAID in multiple frames. Each frame can hold a up to 2,112 bytes, so a large sequence might consist of thousands of frames. An exchange may be open for an extended period of time and contain hundreds of sequences with thousands of frames. The hierarchy of exchanges, sequences, and frames ensures that error recovery is maintained at multiple levels.

ARBITRATED LOOP TOPOLOGY

KLLR also used a hub in their initial deployment, so that multiple servers would have access to the RAID. The arbitrated loop topology allows up to 126 Node Loop Ports (NL_Ports) to be connected on one continuous loop. KLLR only used six devices on its loop. While the SCSI bus can hold up to eight or sixteen storage devices on the bus, it has difficulty supporting several initiators (servers). Fibre Channel does not have this limitation, so it is often used with large storage devices to support multiple servers.

Each Internet video server can support up to twenty five video streams, so a total of one hundred Internet users can access the site at any given time. Since the arbitrated loop can only support one circuit at a time, each server loads the desired video clip onto its internal drive and then serves the data to the client. KLLR expected most users to have broadband connections at less than 1 Mbps. If 100 broadband users were all downloading files simultaneously, the single 100 MBps loop should easily handle the 100 Mbps streaming demand.

KLLR's arbitrated loop and the start-up process for NL_Ports is shown in Figure 7. After power up, loops do not go through speed negotiation or link initialization, but go directly into loop initialization. The loop initialization process that is defined in Fibre Channel Arbitrated Loop 2 (FC-AL-2) is a logical procedure that establishes the Arbitrated Loop Physical Address (AL_PA) of each port. The

Loop Ports (L_Ports) use the Loop Port State Machine (LPSM) to discover the environment and react appropriately. Extensive SANmark tests (see Chapter 10 for details) have corrected most interoperability issues with loop initialization. At the end of loop initialization, the NL_Ports have their AL_PA.

After the loop initialization process has completed, a single NL_Port needs to gain control of the shared loop through arbitration. Since only one device can communicate on the loop at a given time, the device must arbitrate for the loop before it can communicate with another device. When a device wins arbitration, it sends an Open Primitive Signal (OPN) to the destination port. The destination port responds to the OPN and completes the circuit across the rest of the loop. When the loop circuit is established, the link acts like a point-to-point connection and proceeds to PLOGI and PRLI before sending IUs (data) between the ports. After the data is sent, the device sends the Close Primitive Signal (CLS) to open the loop for arbitration once again.

Arbitrated Loop

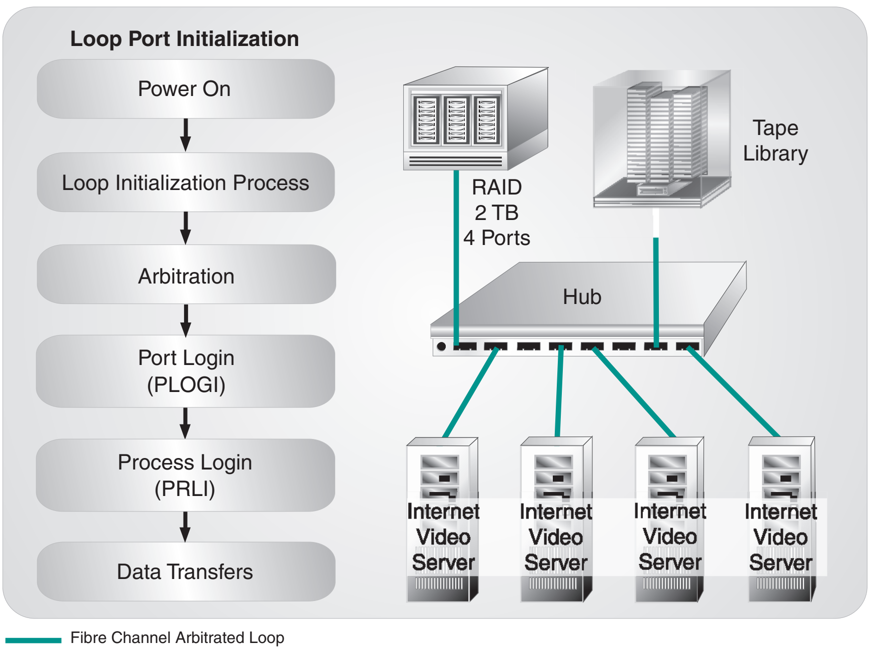


Figure 7: KLLR's arbitrated loop allows multiple servers to be attached to the RAID so that about 100 clients can access the same information. A device attached to the hub goes through the six-step start-up process before data can be transferred between the devices.

INITIAL RESULTS

After a couple of weeks of testing the SAN and digitizing, KLLR began advertising their Internet streaming capabilities on the news broadcasts. Response was moderate, but repeat customers began to grow in numbers. Most of the repeat users were amateur digital video editors who wanted specific video clips. KLLR contacted some of the users and found out that they edited the downloaded video clips and combined them with their own videos. They preferred the downloaded video because the quality was higher than video recorded from on-air broadcasts.

KLLR was making minimal money from the downloads. KLLR charged a small fee of $1 for featured video clips and $2 for 30-minute shows. KLLR is affiliated with ABN, one of the five national broadcasting companies. ABN had agreed to the download rates on a trial basis and received 10% of the revenues. The payments for the downloads were made online with a web-based e-money company. The e-money company had a promotion of offering $5 worth of e-payments for new enrollees. This encouraged many first time users, because they could sign up for free and get up to five free downloads.

KLLR got a larger response when a popular boy band played at a local music hall. KLLR recorded the live performance and made a feature video of one of the band's songs. They advertised the three-minute music video on the news, and the servers were soon at full capacity with 100 users. A problem with KLLR's configuration occurred when many of the teenage users were downloading at 28.8 Kbps on a dial-up modem. At this rate, downloading took over an hour. Best case download times for various length video clips can be found in Table 3. As word got out on the band's website, the servers were soon overwhelmed with thousands of requests for the hot video.

Table 3. Video Downloads

Video Time	Dialup Modem @56kbps	DSL @ 250 kbps	Cable Modem @ 3 Mbps	Fast Ethernet @ 40 Mbps
1 Second of Video	10 seconds	2.4 seconds	0.2 seconds	.01 seconds
30 Second TV commercial	5 minutes	1.2 minutes	5.9 seconds	.44 seconds
3 Minute Music Video	32 minutes	7.1 minutes	35 seconds	3 seconds
5 Minutes of Video	1.3 hours	17.6 minutes	1.5 minutes	7 seconds
22 Minute Program w/o Commercials	3.8 hours	52 minutes	4.3 minutes	19 seconds
30 Minute Program with Commercials	5.2 hours	1.2 hours	5.9 minutes	26 seconds
1 Hour Program	10.5 hours	2.4 hours	12 minutes	53 seconds
2 Hour Movie	21 hours	4.7 hours	24 minutes	1.8 minutes
3 Hour Football Game	32 hours	7.1 hours	35 minutes	2.6 minutes

Table 3 shows the time to download MPEG-2 videos from the KLLR website depending on the connection rate. Times could be considerably longer if congestion causes lower transfer rates. While dialup modem times are almost intolerable for a 30-minute television show, broadband connections at over 1 Mbps make downloads feasible. The broadband connection provides Internet Protocol Video-On-Demand (IP-VOD).

After the digital video file was posted on fan sites, the popularity of the video spread until it was finally aired on MTV. The lawyers for the boy band were soon in contact with KLLR. The lawyers first wanted the clip taken off the website, but soon realized that it was too late. Instead, an agreement was reached wherein the band would receive 30% of the revenue from the music video. It was a win-win situation. KLLR began making similar deals with other bands that played at the local concert hall.

Downloads were averaging hundreds of dollars a day, and the audience continued to grow. Another unexpected phenomenon was that users were paying $0.25 or $1 to download weather reports and television commercials. The initial phase of the project was deemed a success when the company broke even on their investment after one year and continued to bring in thousands of dollars a day.

SUMMARY

KLLR's simple SAN utilizes two of the three Fibre Channel topologies (point-to-point and arbitrated loop). KLLR wanted a dedicated connection for the station's initial recording, so the point-to-point topology works well. KLLR also had multiple servers with low bandwidth requirements, so the arbitrated-loop topology worked better. The SAN's capacity and performance along with the ubiquitous web-based access were key enablers to the success of the program.

The initialization process for the Fibre Channel links shows how the Fibre Channel protocol adapts to different environments. Simple processes negotiate the speed of the link, the service parameters, and access to the arbitrated loop. The applicability of Fibre Channel to many environments ensures its success.

- *SAN Planning*: Recording KLLR's broadcasts and distributing them over the Internet required disk storage with terabytes of data, a tape library, and a few servers.
- *Point-to-Point Topology*: The point-to-point topology has a simple initiation process that involves establishing the link and the Upper Level Protocol.
- *Arbitrated Loop Topology*: The arbitrated loop topology allows several servers to share the RAID and the tape library for a low cost.
- *Initial Results*: KLLR began generating revenue by selling popular music videos, weather reports, and other small video clips over the Internet. As more broadband customers come online, entire TV programs can be downloaded for IP-VOD.

[This is a blank page.]

Chapter 3

SWITCHED APPLICATIONS

The main benefits of Fibre Channel compared to DAS are seen in larger switched applications. When a switched fabric connects many devices, the bandwidth of the network grows with each attached port. Managing a small network or loop is relatively easy, but as the size of the network grows, the difficulty of managing it grows exponentially. Fibre Channel has built-in intelligence that enables the fabric to be managed easily. The intelligence of Fibre Channel is contained in the fabric services. Fabric services enable discovery, reporting, and zoning so that the SAN can be managed effectively.

VIDEO PRODUCTION SAN

KLLR's broadcasting parent, ABN, has an annual meeting of the CIOs from ABN's affiliate stations. At the annual meeting, KLLR's CIO saw a presentation about how another TV station was using Fibre Channel in their production environment. The CIO envisioned a network that connected all of his production facilities in one seamless digital pool. The affiliate station reported that production costs fell after the initial investment, and information could be shared between affiliate stations very easily after it was centrally stored.

Since the late 1990s, KLLR, like much of the news industry, had been using digital video cameras based on the DV format. The station accumulated thousands of the little 60-minute DV tapes, and most of them were less than half filled. The employees were also tired of running the tapes around the production area via "sneaker net." Tapes were often misplaced, and they had to be reloaded each time the tape arrived at a new workstation. Expensive DV tape drives often wore out after two years from repeated use, and it was time consuming to load the tapes multiple times. Figure 8 shows the production facilities of KLLR before the SAN was deployed.

The typical video editing bay in KLLR's production room consisted of a DV tape drive, a non-linear editing workstation, and a directly attached SCSI RAID. The computer was also attached to KLLR's Local Area Network (LAN), but most data transfers to the control room, where the video was broadcast, were still done by sneaker net. The RAIDs varied in capacity from 40 GB to 240 GB and varied in their performance and reliability considerably. Several editors would work at the various edit stations throughout the week. The RAID in an edit bay would often fill to capacity, and no one knew which files could be deleted. Managing the DAS for each edit bay was causing production problems.

The CIO decided to pool the storage resources into one high-performance storage subsystem as seen in Figure 9. The new storage subsystem allowed the video (storage) to be managed from one location. Satellite video feeds, tapes from the field, and all videos produced at KLLR could now be centrally stored and accessed from any workstation. Tapes from the field were now recorded in the tape loading room so the edit bays could access the video immediately. Their most valuable assets—their programming—were now readily accessible.

Each editor was allotted 60 GB of centralized storage, which is equivalent to five hours of uncompressed video. Editors managed their own storage space, which was automatically backed up to a tape library every night. The control room accessed the new videos over the SAN and could broadcast them within seconds of completion. The Fibre Channel SAN turned the laborious tape-based production facility into an efficient networked production facility. The News Department saw a 20% decrease in the amount of storage required, as well as less capital expenditure on new MiniDV tape drives and MiniDV tapes, which were now being reused. The best return on investment was the increased efficiency of the whole department.

With the large number of Fibre Channel connections in the new production environment, KLLR decided to deploy a switch. The switched fabric enabled several editors and the control bay to exchange material simultaneously while news feeds were being recorded continuously. The switched fabric also allowed the SAN to scale by adding another switch when more servers or storage is needed. Redundancy is also used for the control room computer to ensure that the station never delivers dead airtime. Fibre Channel's switched fabric offers performance, scalability, and redundancy that DAS cannot.

KLLR's Production Before

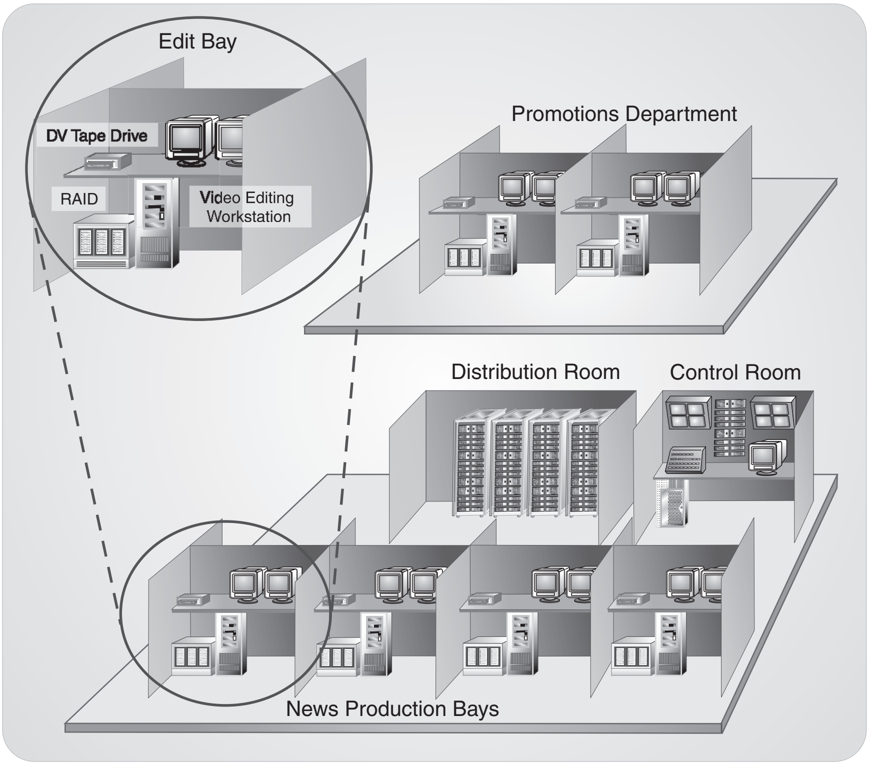


Figure 8: KLLR's production facility used digital tapes from video cameras to transfer data between workstations. Thousands of tapes were stored in the facility, and most of them were less than half filled. RAIDs were used in each edit bay, and significant resources were required to manage the storage.

KLLR's Production After

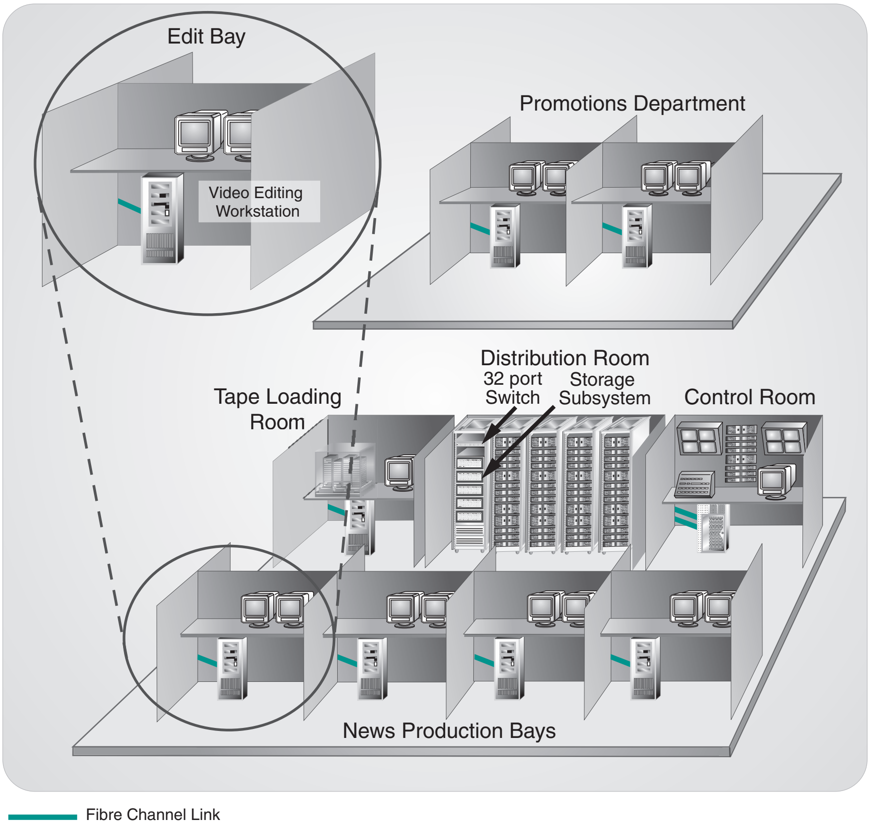


Figure 9: KLLR upgraded their production facilities by replacing the directly attached storage (RAIDs and tape drives) with storage subsystems and a tape library. The Fibre Channel switch connects the storage to all of the edit bays and the control booth. The efficiency of production increased tremendously after the Fibre Channel SAN connected the various departments and work areas.

SWITCHED FABRICS

When a new device (node) is connected to the fabric, each Node_Port (N_Port) follows the procedure shown in Figure 10. After link initialization, the N_Port logs in to the fabric with a Fabric Login (FLOGI) command. The FLOGI exchanges service parameters similar to a PLOGI for effective fabric-to-N_Port communications. The reply to the FLOGI also contains the Fibre Channel address that the fabric assigns the N_Port. The Fibre Channel address and E_Ports are discussed later in this chapter.

Port Initialization

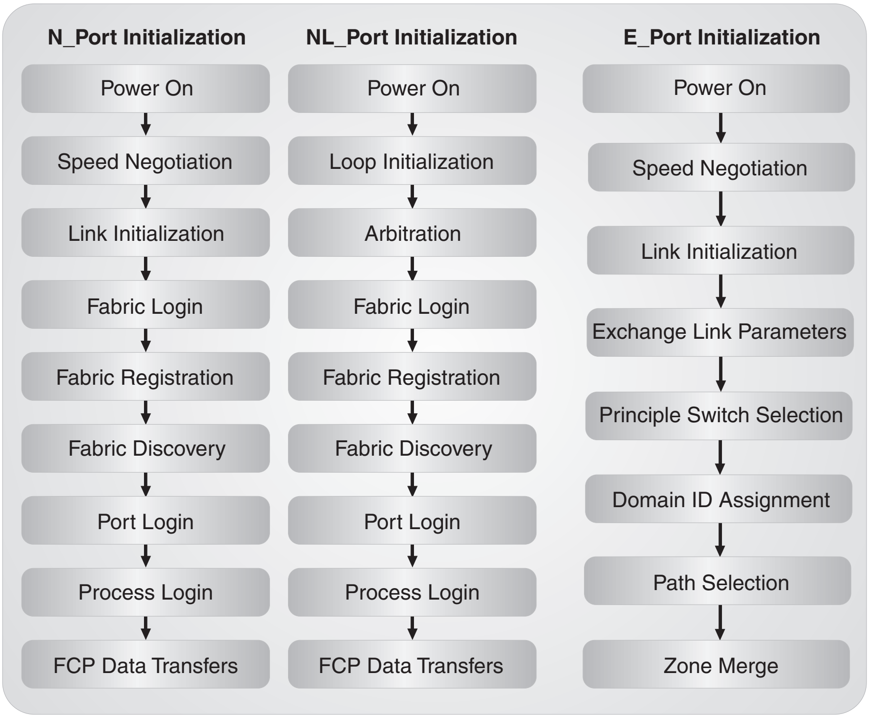


Figure 10: The initialization process for a port depends on the type of port that is connecting to the fabric. N_Ports and NL_Ports must log in to the fabric to receive their Fibre Channel Addresses and discover the fabric's topologies. E_Ports connect switches and follow a sophisticated initialization process to form a seamless fabric.

After fabric login, the port needs to register its information with the Name Server. A number of attributes, shown in Figure 11, may be registered for discovery and management. When the port finishes registration, the fabric announces the arrival of the new port to all ports that have registered for Registered State Change Notifications (RSCNs). The port is also placed in a zone by the SAN administrator. A zone is a list of devices that can communicate. The port may also query the Name Server to discover the devices with which it is allowed to communicate. The port may also register to receive fabric updates and link incident reports. These fabric services defined for Fibre Channel offer the fabric intelligence that is unmatched by any other interconnect technology.

Name Server Attributes

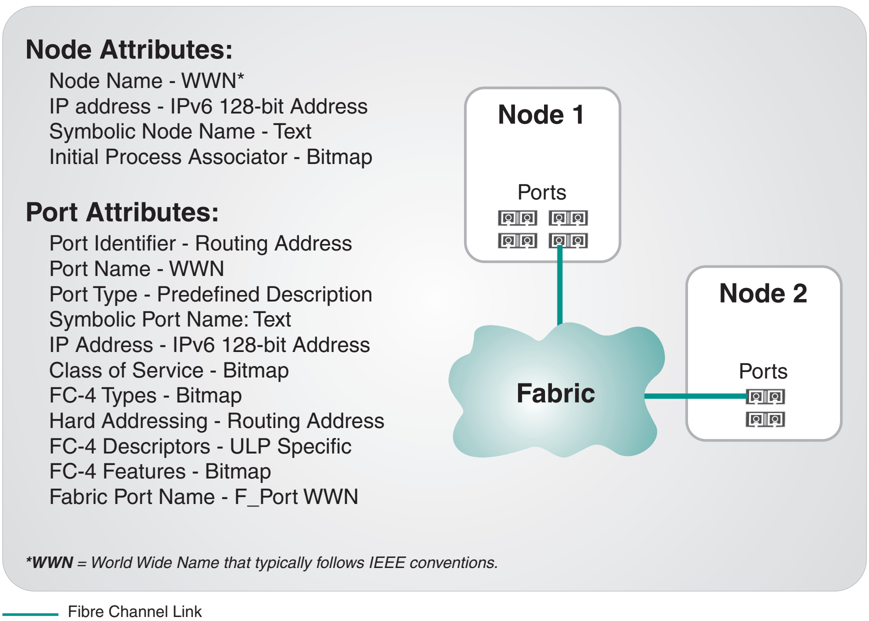


Figure 11: When a port logs in to the fabric, it may register any of these attributes for discovery purposes with the Name Server.

After registration and discovery are complete, the N_Port sends a Port Login (PLOGI) command to a desirable port and then sends a Process Login (PRLI) command for the ULP. NL_Ports follow a different initialization procedure, as seen in Figure 10. The NL_Port initialization is a mixture of the loop start-up procedure and the fabric login process. After the initializations, the ports are ready to transfer data at gigabit speeds.

VIDEO-ON-DEMAND DATA CENTER

At the annual meeting of ABN's affiliate stations, KLLR's presentation on their web-based video distribution system turned some heads. Executives were surprised that people would pay for high-quality material on the Internet. ABN asked KLLR's CIO to give a presentation to the board of directors about their experience. After one year in operation, KLLR had broken even and was now generating positive cash flow. ABN's executive board decided to carry out a nationwide market research project to see if high-quality web-based video was feasible on a national level.

The market study found a number of interesting points. First, people were ready to pay a monthly fee for unlimited program downloads or minimal fees for individual programs, if no commercials were included on the downloaded programs. Users were also interested in ordering programs like rare sitcom episodes, news, and sporting events. Users repeatedly asked for an easy-to-use interface to the large volume of videos that were available. Overall, it appeared that ABN could make money distributing video over the Internet. The executive board was encouraged enough to give IP-VOD a try.

Several synergistic factors helped ABN deploy VOD on a large scale. One practical factor was the lower cost of disk-based storage. Hundreds of gigabytes could be stored on a single disk, which correlated to hundreds of hours of video. Besides being able to store days of video on one storage subsystem in the data center, users could store hours of video at their homes on Digital Video Recorders (DVRs) —often referred to as Personal Video Recorders too. The latest models of the DVRs have at least a 20 GB hard drive that could preload up to 35 hours of video. When DVRs were connected to broadband networks, users could see the benefits on their TV sets. With over ten million broadband subscribers in 2001, broadband reached critical mass. Another crucial technology was the user interface to the vast collection of videos. Similar to a web browser interface, DVR users could order shows with relative ease, and download multiple programs overnight. Finally, digital rights management software was required to protect their and others' intellectual property. Low storage costs, DVRs, broadband, and a friendly user interface enabled the success of IP-VOD.

ABN has 168 affiliate stations and each station produces about three hours of local programming per day. The remaining 21 hours of programming is broadcast from ABN. If ABN stored the national broadcasts and all of the local broadcasts, ABN would need to store the equivalent of 22 TV stations recorded around the clock. Since ABN had a limited budget, they decided to store the local broadcasts for the top 64 affiliate stations in the first phase of the project. The storage requirements for the data center are summarized in Table 4.

ABN decided to store one week's worth of affiliate shows for each station which resulted in over 2 TB of storage. Shows that had high hit ratios would stay on the 4 TB storage subsystem longer. A database tracked the number of hits per show and discarded shows that did not yield revenue.

Other storage requirements involved making past productions available. ABN had about 10 TB of sporting events and movies on tape. They also digitized their hottest sitcoms and dramas resulting in over 3 TB of storage. They literally had decades of material available, so they would have to rotate the shows that were available and promote them. Since these programs would be promoted and in high demand, they placed them on a storage subsystem as well.

Table 4. Storage Requirements for Data Center

Programs or Length of Time	**Number of Shows**	**Storage per Show**	**Storage Requirements**
22-Minute Sitcom Episodes	2000	775 MB	1.55 TB
45-Minute Sitcom Episodes	1000	1.6 GB	1.60 TB
90-Minute Movies or Specials*	500	3.2 GB	1.6 TB
Two-Hour Sporting Events*	1500	4.2 GB	6.3 TB
One Week of Programming for 64 stations	2688	775 MB	2.1 TB
Totals	7688		13.2 TB

* Tape based and free of commercials.

MULTI-SWITCH SANS

Although it was difficult to store and serve all of the local stations' broadcasts from one location, it was easier than each station storing and serving its own material. One information management team could service the data center for all the affiliates and provide a consistent theme for all TV stations and programs. This consolidation of storage and resources saved the company millions of dollars. The first phase of the SAN is shown in Figure 12.

The new SAN required the high availability and connectivity of a high-port-count director to interconnect the core of the SAN. Multiple edge switches on the edge were needed to accommodate the large number of servers. Switches or directors could be added to the fabric in the future to increase bandwidth and port counts. When switches are connected to other switches, the ports become Expansion Ports (E_Ports) and follow the initialization process shown in Figure 10. The complete initialization process is defined in Fibre Channel Switched Fabrics 2 (FC-SW-2). After speed negotiation and link initialization, the E_Ports exchange link parameters. Exchanging link parameters is analogous to the port login process in that basic parameters are exchanged to initiate communication between the ports. If the exchanged parameters are unacceptable, the link isolates so that no data can flow across the link.

Data Center SAN

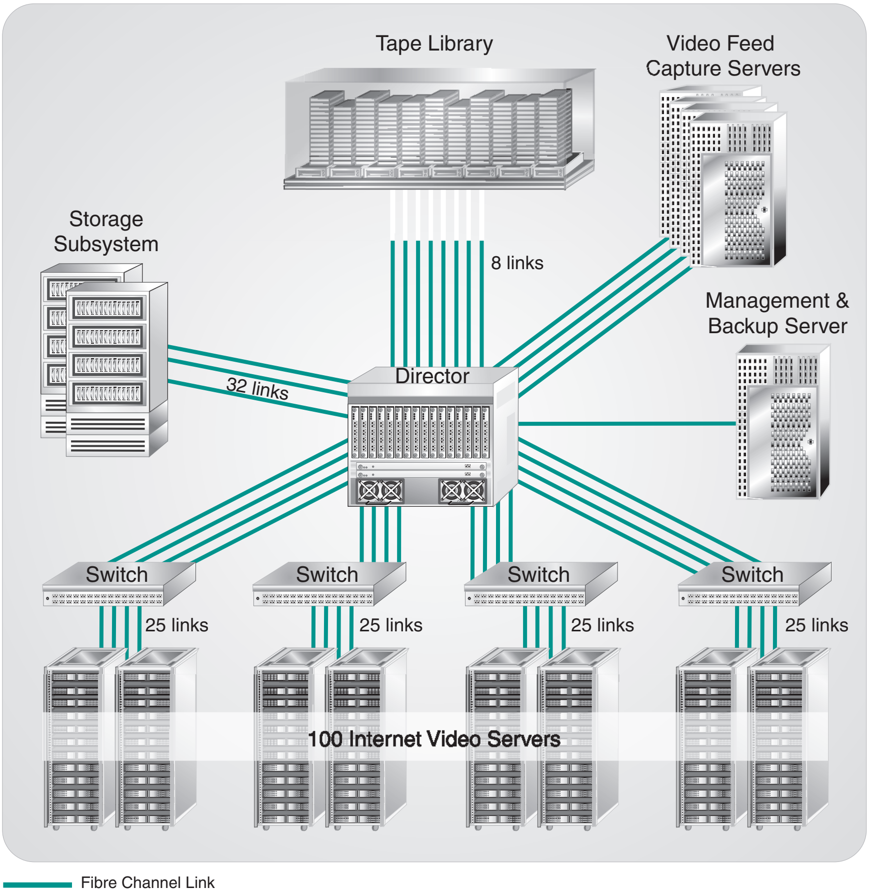


Figure 12: ABN's data center holds over 10 TB of storage and over 100 servers. This two-tier design has storage attached to the director and has servers attached to the switches.

If the ELPs are successfully exchanged, then the links transmit exchange fabric parameters (EFP) commands to ensure that the fabric can join together without conflicts. If the EFPs are successful, then the joined fabrics decide on one principle switch. The principle switch is determined by the switch priority that is assigned by the SAN administrator. Once the principle switch is selected, the principle switch grants Domain IDs to the other switches in the fabric. The Domain ID is the first byte of the Fibre Channel Address, as shown in Figure 13.

After the domains are assigned to the switches, the Fabric Shortest Path First (FSPF) routing protocol selects paths between switches. Consistent routing between E_Ports ensures that frames are delivered in order. The last step in the fabric initialization process is merging zones. The fabric needs to have consistent zoning throughout the fabric. If the zoning is incompatible for any reason, the E_Ports segment to prevent unauthorized access to devices. After the E_Ports have finished initialization, the fabric acts like one big switch. End devices are unaware of the distances that the fabric may span or of devices outside of their zone.

Three Byte Fibre Channel Address

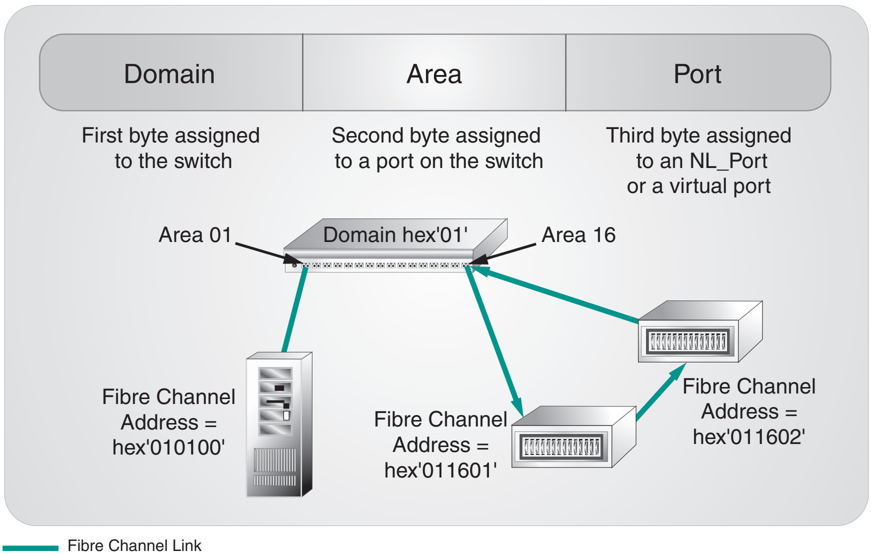


Figure 13. The Fibre Channel address consists of three bytes that are used for hierarchical routing. The domain ID is assigned to a switch during fabric initialization. N_Ports and NL_Ports receive their Fibre Channel addresses when they log in to the fabric.

SUMMARY

The benefits of Fibre Channel become obvious when SANs are applied to large-scale applications like ABN's data center. The SAN allows access to large pools of information from hundreds of servers. To take full advantage of the power of storage subsystems and tape libraries, a switched connection technology like Fibre Channel is necessary.

Besides scalability, a Fibre Channel SAN also offers advanced fabric services like device discovery, topology discovery, and zoning. Fibre Channel also scales when switches are connected to form larger fabrics. Fibre Channel enables consolidation of storage so that clusters of computers can access and serve vast quantities of data.

ABN's IP-VOD system proved to be a big success. The venture turned out to be a highly collaborative effort between Cable TV networks, DVR providers, and movie studios. The data center was replicated many times across the nation, and ABN currently plans to go international.

- *Video Production SAN*: KLLR was able to consolidate its storage with Fibre Channel so that tape drives and storage could be removed from each edit bay. The SAN also enabled electronic transfers to the control room and between the edit bays.
- *Switched Fabrics*: Devices that connect to the switched fabric can have simultaneous communications with multiple devices. The fabric also enables advanced fabric services like device discovery, topology discovery, zoning and reporting.
- *Video-on-Demand Data Center*: ABN's data center serves decades worth of programming over the Internet. Only Fibre Channel SANs can allow so many servers to access such large volumes of data at gigabit speeds.
- *Multi-Switched SANs*: Fibre Channel switches can be connected together to form seamless fabrics that are highly scalable. High port count directors are used at the center of SANs to connect high-end storage devices to hundreds of servers.

PHYSICAL LAYER

It all starts at the physical layer. Every technology is based on the physical layer, and Fibre Channel was the second widely deployed digital technology to use fiber-optics at gigabit speeds; telephony was the first. Fibre Channel optical transceivers can send native signals over fifty kilometers. Fibre Channel signals have also been mapped to Wavelength Division Multiplexing (WDM) equipment to send signals for hundreds of kilometers over Metropolitan Area Networks (MANs). Fibre Channel continues to keep pace with other data communication protocols by sending data at two and ten gigabits per second (Gbps).

GIGABITS AND MORE GIGABITS

Fibre Channel standards committees developed high-speed serial communications similar to the physical layer of IBM's Enterprise Systems Connection (ESCON) architecture. Fibre Channel's serial architecture is based on a SERDES (SERializer/DESerializer) chip to convert parallel signals to a serial signal and back to parallel signals, as seen in Figure 14. The SERDES encodes and decodes the data stream using 8B/10B encoding. This encoding scheme enables control signals to be inserted into the serial data stream as well. This complex encoding algorithm, ordered sets, and primitive sequences are defined in FC-FS to ensure low bit error rates over long distances. The SERDES conversion to serial signals allows devices to communicate over long distances at high speeds while running at lower parallel speeds within the end devices.

SERDES Encoding

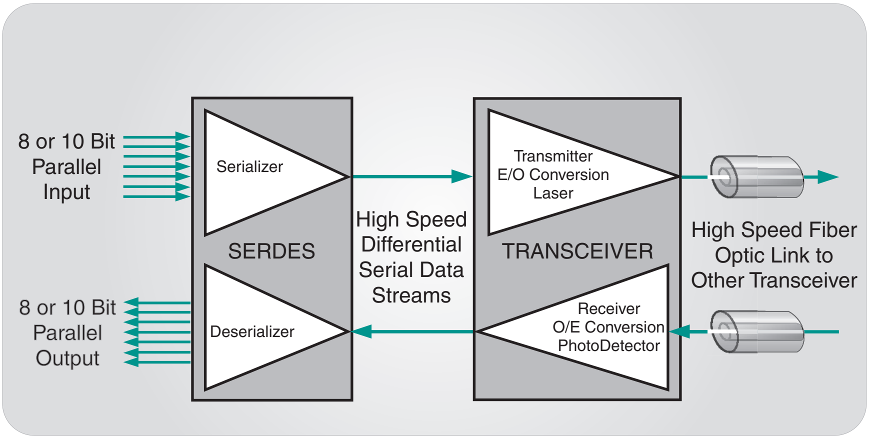


Figure 14: The SERDES chip encodes and decodes low-speed parallel signals into high-speed serial signals. A transceiver amplifies the high-speed serial signal so that the signal can travel over very long distances. This illustration shows the transceiver converting the electrical signal into an optical signal (E/O conversion) so that the signal can travel over optical fibers for long distances. The transceiver may be an electrical transceiver that sends the signal over electrical cables for thirty meters. The receiver in the transceiver will convert the optical signal into an electrical signal (O/E Conversion) and feed the signal to the Deserializer.

The physical layer of Fibre Channel runs at multiple serial speeds. The first deployed Fibre Channel products ran at 265 Mbps to yield a data rate of 25 MBps. The vast majority of Fibre Channel fabrics before 2002 were deployed with serial rates of 1.0625 Gbps to yield data rates of 100 MBps. 2.125 Gbps fabrics are being deployed rapidly and interoperate with legacy 1 Gigabit Fibre Channel (GFC) equipment. Table 5 includes the speeds that Fibre Channel uses, including future speeds. 10 GFC keeps pace with other networking technologies that are converging at 10 Gbps. Fibre Channel is highly adaptable and runs at speeds higher than most LAN technologies.

Table 5. Fibre Channel Link Speeds

Speed	Throughput (MBps)*	Line Rate (Gbaud)	Release Date (Year)
1 GFC	200	1.0625	1998
2 GFC	400	2.125	2000
4 GFC**	800	4.25	2003
10 GFC	2400	10.2 or 4 X 3.1875	2003

* Throughput for duplex connections
** Intrabox applications

FIBRE CHANNEL MEDIA

The high-speed serial communications of Fibre Channel are mainly carried over fiber-optics today. Figure 15 shows the various types of media that are supported by Fibre Channel. The same media are used for Fibre Channel and other networking technologies, such as Ethernet and SONET.

Depending on the distances required for a given implementation, a particular media and transceiver must be chosen for the application. For intra-room applications, electrical cables and transceivers provide a cost-effective solution. For ease of use, future proofing, and implementations that span several rooms or floors, multimode fiber with 850 nanometer (nm) transceivers offers the ideal solution. The 850 nm transceivers use Vertical Cavity Surface Emitting Lasers (VCSELs) that couple to either 50 micrometer (um) or 62.5 um multimode fibers. Single-mode fibers use either 1310 nm transceivers to achieve distances over 10 km or 1550 nm transceivers for distances over 50 kilometers. These two laser types are commonly used in telecommunications as well and provide connectivity over large distances. The relative cost of each transceiver type increases with the distance achievable by that transceiver.

Fibre Channel Media

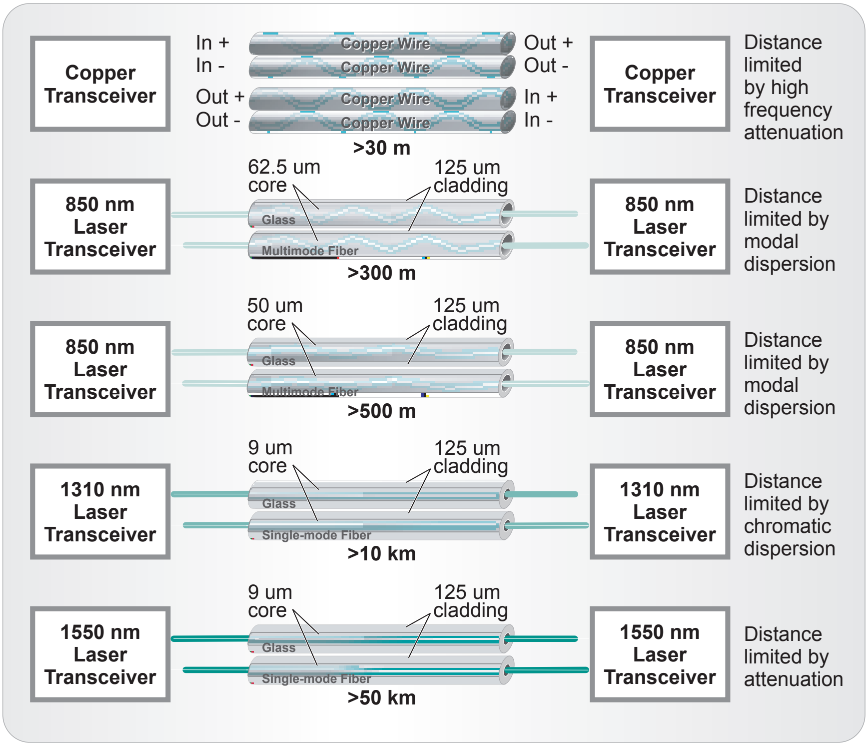


Figure 15: Fibre Channel media, collectively known as fibre, consist of four cable types combined with four transceiver types. Electrical media is limited to very short-range applications, while multimode fiber is limited to a few hundred meters. Single-mode fiber is typically used in campus environments, but 1550-nanometer transceivers can extend operable distances to over 50 km for MAN applications. All distances are noted as "greater than" because the link is defined to operate for at least the specified distance.

WDM EQUIPMENT

One of the most revolutionary technologies to reach the communications world multiplexes several layers of traffic over a single fiber-optic cable. Optical fibers can carry multiple wavelengths of light to increase the bandwidth of the media dramatically. Figure 16 shows the general concept of how various types of signals can be carried concurrently with Wavelength Division Multiplexing (WDM) equipment. Besides increasing the number of signals carried on a single fiber, the high-powered lasers in the WDM equipment can increase the reachable distance to over 100 km. This increases the bandwidth-length product of the network.

WDM Implementation

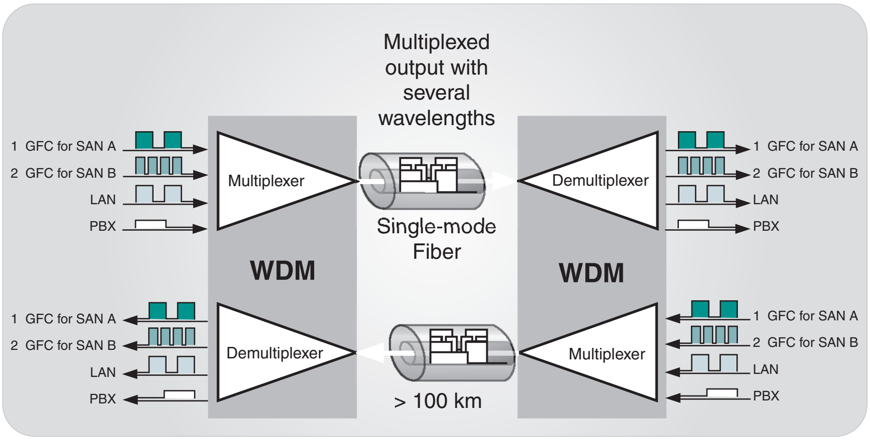


Figure 16: Wavelength Division Multiplexing (WDM) equipment has dramatically increased the bandwidth of a single optical fiber. WDM equipment combines the signals from different sources and transmits them over the same fiber at different wavelengths. Each signal can run at a different speed and use a different protocol.

TRANSCEIVERS

Figure 17 shows the four basic transceiver types used in Fibre Channel. Multi-Sourcing Agreements (MSAs) between transceiver vendors have yielded the backend electrical interface and latching mechanism for these transceivers. The frontend connection of the transceiver depends on the physical connector and whether the output is optical or electrical. The independence of the output connector from the input to the transceiver allows flexibility, so that a user can plug in various types of transceivers to the same port. If a user's implementation requires using Fibre Channel over long distances, a 1550 nm transceiver is used with single-mode fiber. The various form factors of the transceivers improve their applicability to many environments.

Although Figure 17 shows the most common fiber-optic connectors used in Fibre Channel, other connectors have been specified in FC-PI. Transceivers with different front end connectors may be used in a port but may not be qualified with a given vendor.

10 GFC

A new standard and technology, 10 GFC, will revolutionize the way that Fibre Channel converges with other technologies, such as 10 Gigabit Ethernet and OC-192. Because serial communication at 10 GHz is in the range of radio waves, lower cost solutions utilize parallel technologies running at lower speeds.

The physical layer of 10 GFC requires new approaches. The backend electrical interface to transceivers in 10 GFC is a four-channel parallel interface known as XAUI. The XAUI interface has four input and four output channels, each running at 3.1875 Gbps, to yield a data rate of 2,400 MBps. The 10 GFC transceivers convert the parallel XAUI signals into one 10.2 Gbps serial signal or into a parallel optical signal that can reach distances of up to 40 km. The 10.2 Gbps data stream was designed to allow it to be mapped to SONET for long distance communications. 10 GFC is ready to take Fibre Channel to the next level of revolutionary throughput.

Fibre Channel Transceivers

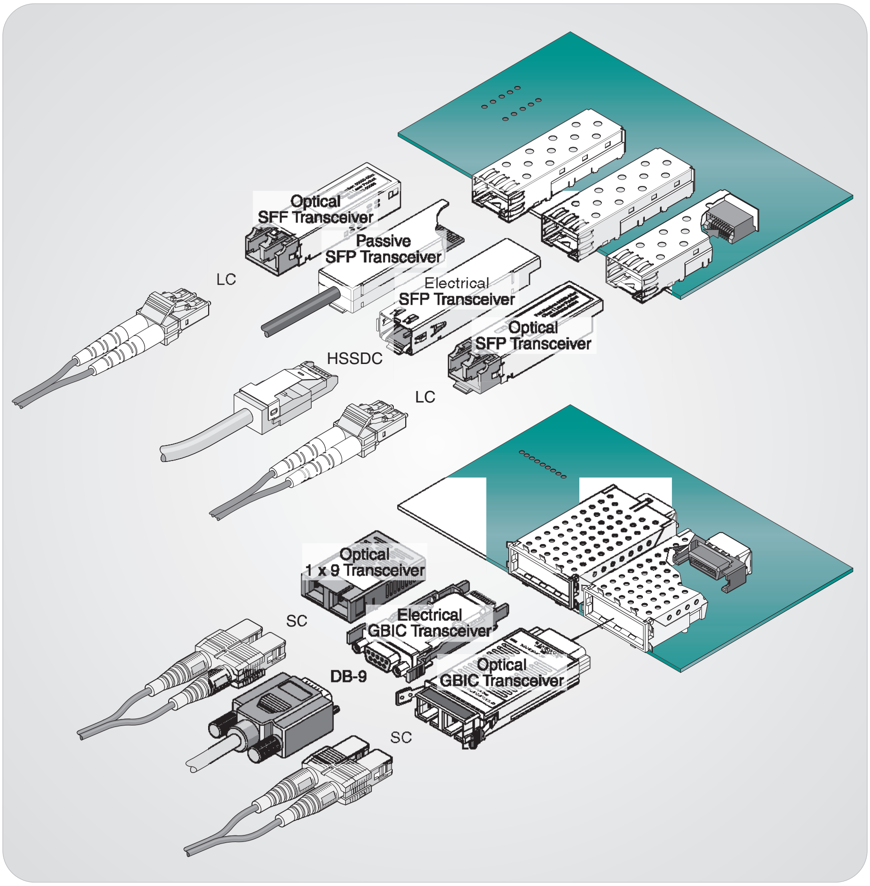


Figure 17: The four basic transceiver types used in Fibre Channel are Small Form Factor Pluggable (SFP), Small Form Factor (SFF), Gigabit Interface Converters (GBICs) and 1X9s (1 row of 9 pins). Small Form Factor connectors, shown in the top half of the drawing, have less than half of the width of traditional SC fiber optic connectors or the DB-9 electrical connectors, shown in the bottom half of the drawing. The same optical or electrical output requirements are placed on transceivers regardless of the form factor of the transceiver. Pluggable transceivers (SFPs and GBICs) require Serial ID support so that the port can query the transceiver to find its capabilities.

SUMMARY

The physical layer of Fibre Channel uses fiber-optics and high-speed serial communications to drastically increase the reach of fabrics. Before 2002, most Fibre Channel SANs were deployed at 1 GFC. The 2 GFC systems are now being deployed to meet the needs of demanding real-time applications like High Definition Television (HDTV). 10 GFC promises to take Fibre Channel to the next level of connectivity. Instead of having twelve 1 GFC ports from a mainframe or storage subsystem to a switch, a single 10 GFC connection transports the same amount of data.

Some of the revolutionary technologies that have enabled the Fibre Channel physical layer are SERDES chips, high speed semiconductors, fiber-optic cables, low-cost VCSELs, long-wavelength lasers, high-powered lasers, and WDM equipment. Fibre Channel is an ever evolving industry that adapts and uses the latest available communication technologies. In fact, Fibre Channel is driving down the cost points of many high-end communication products.

- *Gigabits and More Gigabits*: Fibre Channel was the second technology to widely deploy gigabit technologies over fiber-optics. Fibre Channel has recently doubled its speed to 2 GFC and finished specifying 10 GFC in 2002.
- *Fibre Channel Media*: Fibre is a term coined by the Fibre Channel industry that means copper wires and optical fibers that carry Fibre Channel signals. Fibre Channel has paired four types of fibre with four transceiver types for communications from one meter to over 50 km.
- *WDM Equipment*: Serial optical signals are ideal inputs to WDM equipment. WDM equipment uses high-powered lasers to send Fibre Channel signals over 100 km.
- *Transceivers*: The four basic transceiver types are SFP, SFF, GBIC and 1X9. Each of these transceivers can use various fiber-optic or electrical connectors. The four basic transceiver types are either pluggable or solderable and either small form factor or standard form factor.
- *10 GFC*: 10 GFC is the next progression in the ever increasing speeds of Fibre Channel. 10 GFC uses the same physical layer components as 10 Gigabit Ethernet.

Chapter 5

INTERCONNECT DEVICES

Interconnect devices are the devices that form the topology of the Storage Area Network. The first such device is the hub, which connects arbitrated loop devices. A switching hub is the next level of sophistication in interconnect devices and incorporates some of the features of a switch. The switch is a very powerful device that enables multiple, concurrent, full-duplex communications and provides scalability when multiple switches are connected. The most powerful interconnect element is the director, which combines carrier class reliability with high port counts to form the backbone of corporate SANs. The last type of interconnect devices are the gateway devices and FCIP devices that extend SANs across Wide Area Networks. This variety of interconnect devices is what makes Fibre Channel SANs so adaptable.

HUBS

The first widely deployed Fibre Channel interconnect device is the hub. The hub is a very simple device that allows up to 126 arbitrated loop devices to share connectivity. Figure 18 shows how a single device or multiple devices can be placed on one hub port. The hub works like the SCSI bus in that all devices share the same bandwidth. Only one device has control of the loop at a given time.

Hubs offer several advantages over shared buses, including bypass circuitry, monitoring and management. Hubs provide bypass circuitry so that failed devices don't stop data traffic. Hubs might also have LEDs to indicate if the ports and devices are functioning properly. Advanced hubs may provide management services such as port diagnostics, ULP activity and interfaces to management applications. Hubs often consolidate several servers on a single shared loop.

Private Loop Hub

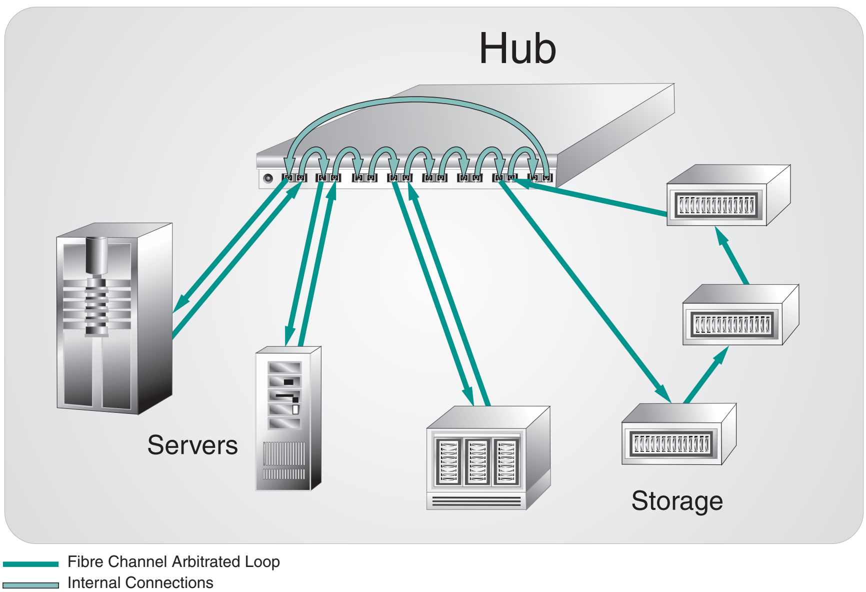


Figure 18: A hub connects up to 126 devices to a single loop. All of the devices share the loop and only one device can communicate over the loop at a given time. The arbitrated loop protocol has the same limitations as the SCSI bus, unless it is connected to a switch as in Figure 19. The storage devices on the right are connected to each other so that multiple storage devices can use a single hub port. This is how up to 126 devices can be connected to a single hub.

SWITCHING HUB

A switching hub differentiates itself from a hub by allowing multiple concurrent data transactions. A switching hub is still limited to 126 devices in a private loop configuration. Loop devices that have not logged into the fabric are known as private loop devices. To increase the scalability of arbitrated loop devices, a hub or switching hub can be connected to a switch, as shown in Figure 19.

After the hub or switching hub is connected to a switch, the configuration is known as a public loop. When the loop becomes public, loop devices can log into the fabric so that devices in the fabric can access the devices. When a loop device has logged into the fabric, the device is known as a public loop device. Public Loops greatly increases the accessibility of the devices on the loop.

Public Loop Switching Hub

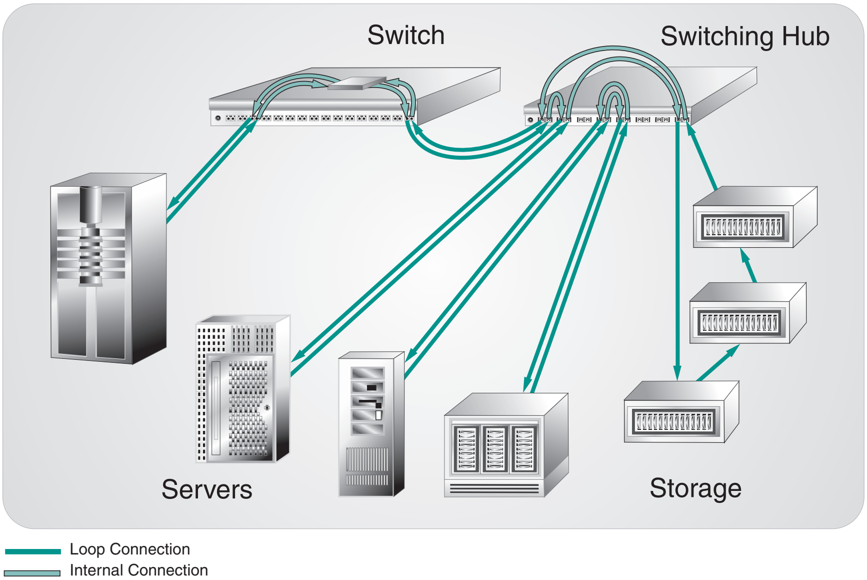


Figure 19: The switching hub allows multiple concurrent transactions of data. A switch is also connected to the switching hub to make it a public loop. The connection to the fabric allows any device on the fabric to communicate with the loop devices.

SWITCHES

The power of the switch results from its ability to carry multiple concurrent communications between any ports. There is no arbitration for a shared bus because the switching construct can support multiple concurrent communications. Each device can communicate with other devices without affecting the throughput of other devices. When more devices are attached to a switch, more bandwidth is enabled in the fabric.

Switches can also connect to other switches in a multi-switch fabric, as seen in Figure 20. Switched fabric architectures are highly scalable, since additional switches can be added to the fabric when more connectivity is needed. Beyond scalability, switched fabrics provide fabric services that help manage and configure the fabric as it scales. Scalability is a double-edged sword: as the fabric grows in power it becomes more difficult to manage. The fabric services overcome this problem by increasing manageability through the Name Server, zoning, and topology discovery.

Switch

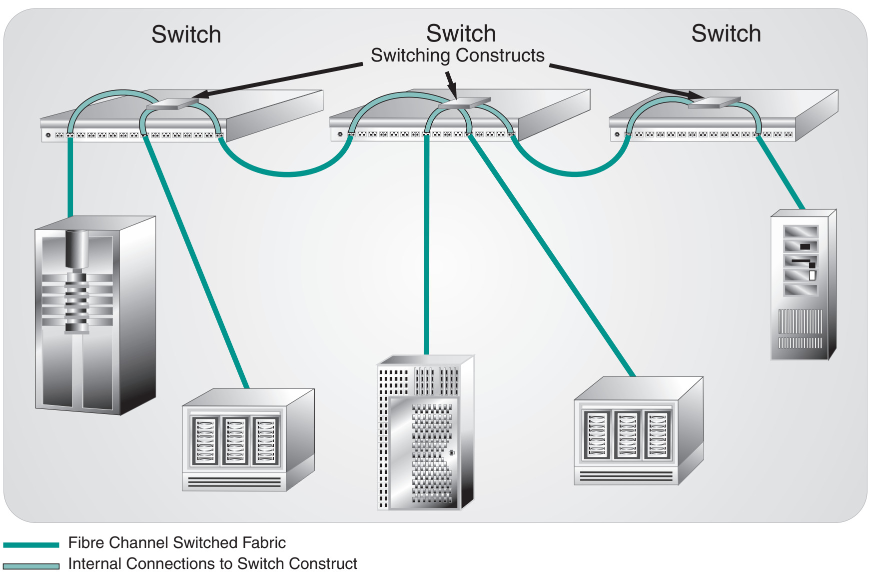


Figure 20: Switches can be connected to multiple switches or hubs to form very large fabrics. Fabric services help configure and control fabrics so that they can be optimized for a given implementation.

DIRECTORS

Directors are high-port count switches that have very high reliability, availability, and serviceability (RAS). The reliability of directors is usually above 99.999%. This correlates to less than five minutes of downtime per year. Directors support high availability by providing features like hot firmware upgrades and redundant logic. Directors are highly serviceable, allowing port cards, processors, fans, and other components to be easily replaced without affecting the other components of the director.

The RAS features of directors make them suitable for interconnecting switches to large storage subsystems and mainframe computers, as seen in Figure 21. Small switch implementations and departmental-level applications are often referred to as SAN islands. Directors are high port-count interconnect devices that are designed for the core of the SAN to connect multiple SAN islands. Directors are the heart of enterprise SANs.

Directors

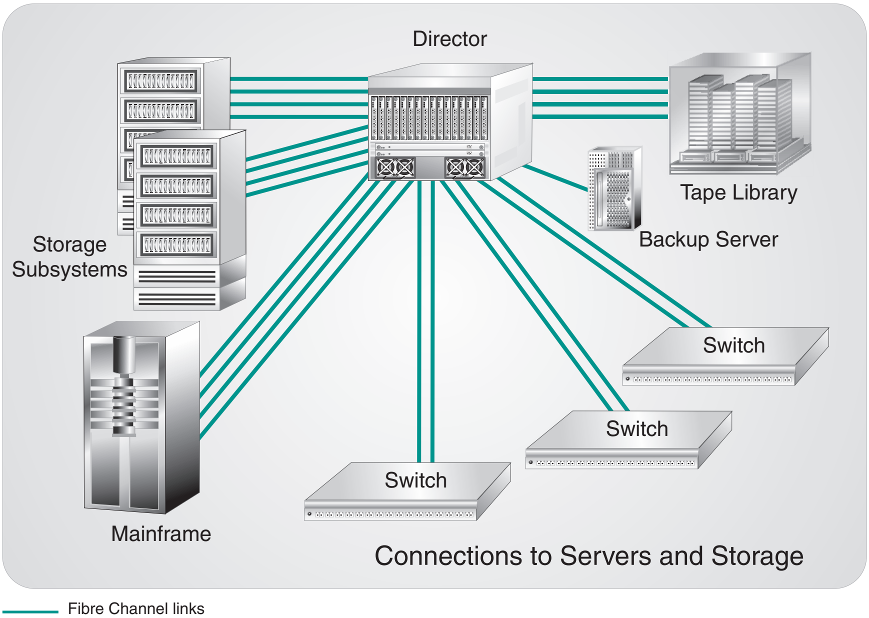


Figure 21: Directors are used at the core of the corporate SAN to interconnect multiple switches and pooled resources such as large storage subsystems, tape libraries and mainframes. Directors are designed to interconnect "SAN islands," which are SAN implementations at the departmental level.

WAN INTERCONNECT DEVICES

The first WAN interconnect devices were known as gateway devices. Gateways are interfaces to SONET or ATM networks. The link formed via a pair of gateways acts like a normal Fibre Channel link but may span continental distances. Gateways can only communicate with one other gateway.

FCIP devices map Fibre Channel to the Internet Protocol to allow a SAN to be connected to the Internet or to a private IP network. FCIP devices allow multiple concurrent connections to other FCIP devices. Since the FCIP device has a connection to the Internet, FCIP devices employ sophisticated security measures to ensure the integrity of the SAN.

WAN Interconnect Devices

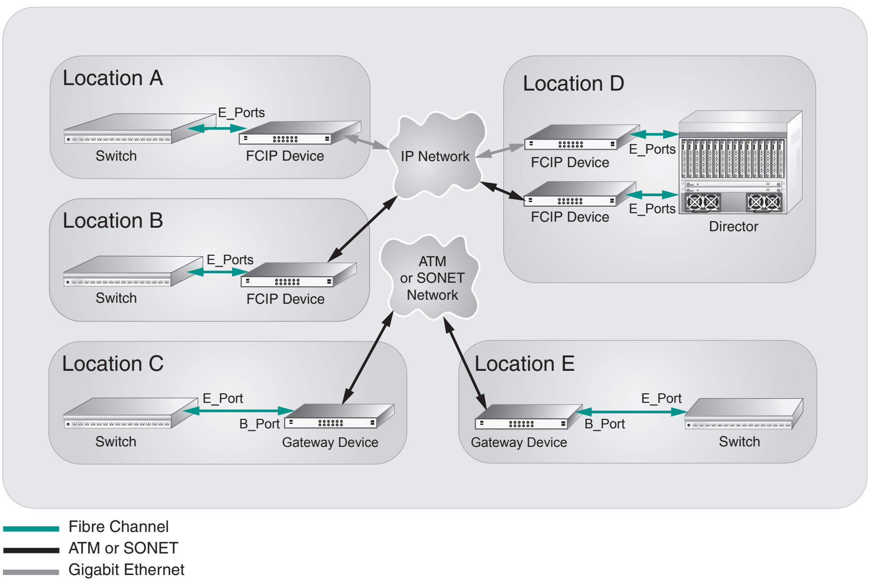


Figure 22: Gateway and FCIP devices connect SANs over global distances. Gateway devices use SONET or ATM networks to connect SANs, while FCIP devices connect SANs on private IP networks or on the Internet. Gateways have one-to-one connectivity, while FCIP devices allow one-to-many connectivity.

SUMMARY

The interconnect devices allow Fibre Channel to be deployed in multiple topologies and over long distances. The hub was the first interconnect device, connecting devices on a shared loop. The switching hub allows more than one device to communicate on the loop at one time but is still limited to only one active fabric connection. Switches allow connection between multiple devices simultaneously and allow scalability when one switch is connected to another. Directors offer high RAS and port counts to make them suitable for the core of corporate SANs. To overcome large distances, gateways and FCIP devices may be deployed to connect SANs across continents.

In the future, several of these devices will be integrated into a single device. FCIP devices will become blades on directors and provide 10 GFC connections across the Internet. Switches already deploy arbitrated loop on single switch ports so that multiple devices can connect to a single port. Higher port counts and faster speeds will continue to increase the power of interconnect devices.

- *Hubs*: Hubs share an arbitrated loop with up to 126 devices and allow only one device to communicate at a given time.
- *Switching Hubs*: Switching hubs allow up to 126 attached devices and allow multiple devices to communicate at a given time.
- *Switches*: Allowing multiple full-bandwidth connections between hundreds and thousands of devices, switches can be connected to other switches or directors to increase the size of the fabric.
- *Directors*: Designed to connect mainframes, storage subsystems, and tape libraries, directors connect SAN islands to form enterprise SANs.
- *WAN Interconnect Devices*: Gateways connect Fibre Channel SANs over ATM or SONET networks. FCIP devices connect multiple Fibre Channel SANs over the Internet or private IP networks.

[This is a blank page.]

Chapter 6

EDGE DEVICES

Edge devices, like storage subsystems and servers, are on the edge of the Fibre Channel fabric and use the fabric to communicate at gigabit speeds. Storage devices are edge devices that hold terabytes and petabytes of information and use controllers or bridges to interface with the Fibre Channel fabric. Servers or mainframes connect to the fabric via Host Bus Adapters (HBAs) and use the fabric to access storage devices. Translation devices connect Fibre Channel fabrics to SCSI devices and other networks. The Fibre Channel fabric connects these edge devices to create a SAN.

STORAGE DEVICES

Storage devices are incredibly powerful devices that store giga, tera, and petabytes of data. The basic building blocks of storage devices are shown in Figure 23. The disk drive is the basic electronic storage device that integrates the storage medium (the magnetized disk) and the drive. Disk drives are often placed in an enclosure and daisy chained together to form Just a Bunch of Disks (JBODs). The tape drive, on the other hand, has removable media in the form of tapes. These storage building blocks typically have a parallel SCSI interface, but many have been converted to a direct Fibre Channel interface.

RAIDs offer higher performance and reliability than an individual disk drive or JBOD. A controller is what distinguishes a RAID from a disk drive or JBOD and is the interface between the Fibre Channel fabric and the disk drives. The controller controls the storage devices and has cache memory so that multiple reads and writes can occur simultaneously. Figure 24 shows how the controller

Basic Storage Devices

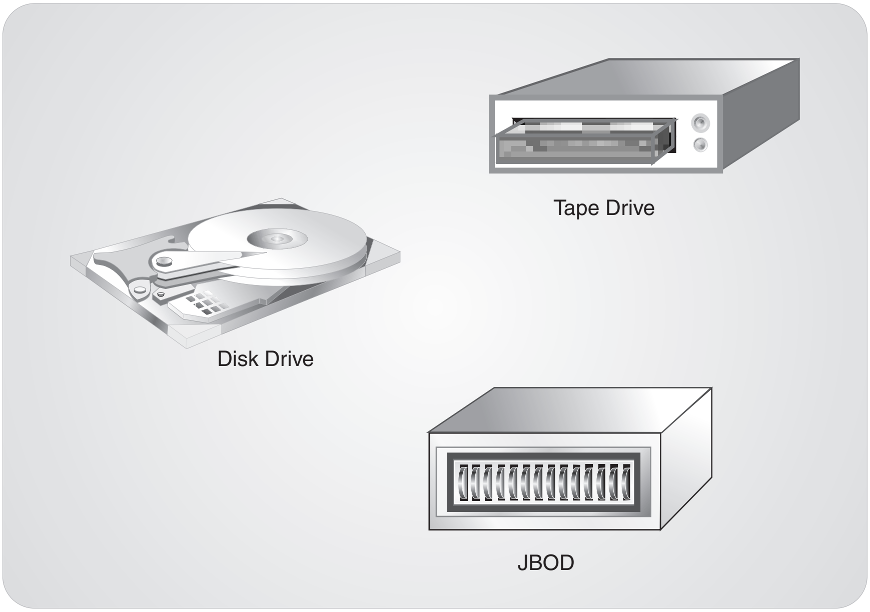


Figure 23: The basic storage devices are the disk drive and the tape drive. Disk drives store data on multiple platters or disks that are read by the magnetically sensitive heads. Tape drives store data on removable tapes that are ideal for archiving years of data.

may have more than one Fibre Channel connection and a backplane to connect to the disk drives.

Storage subsystems use the same basic building blocks of the RAID, but typically have an order of magnitude improvement in performance, capacity and availability over RAIDs. Storage subsystems use multiple controller cards, cabinets full of JBODs, cache memory boards and tens of transceivers, as seen in Figure 25. For redundancy and performance, storage subsystems may deploy a switched or redundant loop architecture to connect the various components. The cache module may have tens of gigabytes of RAM. Multiple adapters have tens of Fibre Channel ports to interface to the SAN. The controller processes millions of requests to the disk drives per second. The storage subsystems are the ultimate storage devices in data centers and act as the central repository for corporate data.

Storage Controller Architecture

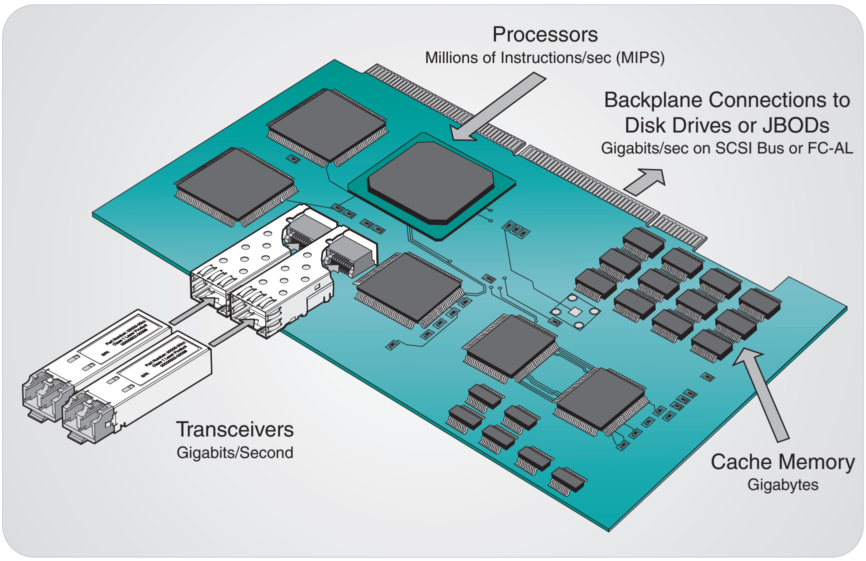


Figure 24: RAIDs use a storage controller to connect multiple disk drives or JBODs to the Fibre Channel fabric. The controller offers data protection and higher performance than an individual drive.

Storage Subsystem Structure

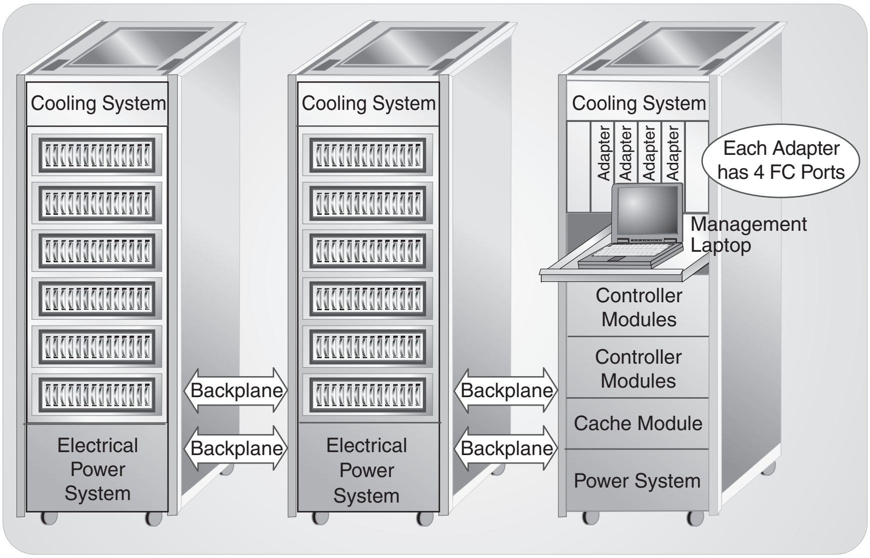


Figure 25: Storage subsystems offer performance, availability, and capacity that are an order of magnitude above RAIDs. Multiple cabinets of JBODs make storage subsystems the ultimate storage device. Redundant backplanes, controllers and adapters increase the RAS of the device.

Tape drives are typically installed in tape libraries, as shown in Figure 26. The tape library consists of several tape drives, tapes, a media changer and possibly a bridge. Fibre Channel tape drives can be directly connected to the fabric, while parallel SCSI-based tape drives need a bridge to connect to the SAN. The media changer may have its own Fibre Channel or SCSI connection, which controls the robotic tape handler. Large tape libraries have thousands of tapes in multiple silos. Small tape libraries may be rack-mounted devices that scale to meet any application.

Tape Library

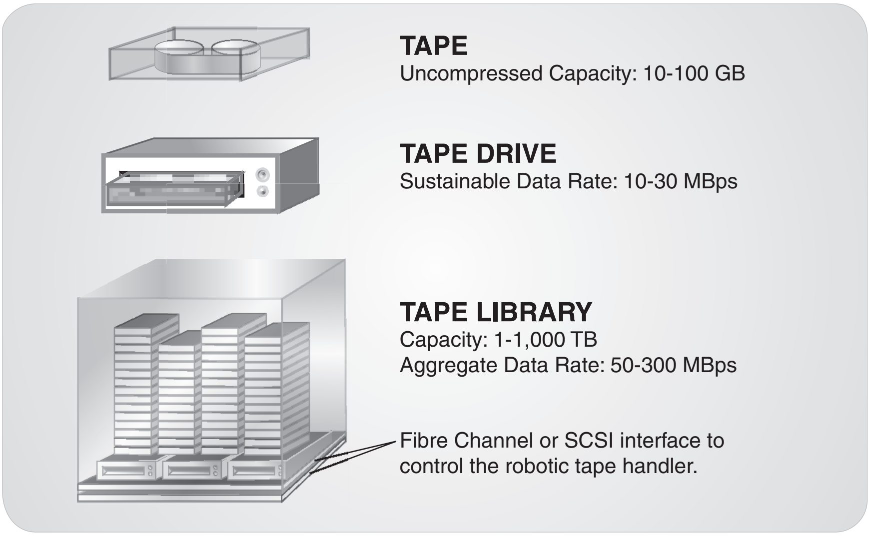


Figure 26: Tape libraries allow multiple tape drives to access hundreds of tapes that are loaded by robotic arms. Tapes are still the most common way to store massive amounts of data because they are cost effective and have long shelf lives.

HOST BUS ADAPTERS

HBAs provide the interface between the SAN and the host I/O bus of the computer (typically PCI or S-Bus). Multiple HBAs may be installed on a single workstation, server, or mainframe. Each HBA may have one, two, or four ports, so that the HBA can be connected to multiple SANs.

Since a single storage port may support multiple HBAs, HBAs are often the most common edge device in the SAN. Managing hundreds of HBAs from multiple vendors used to be problematic, so a common HBA Application Programming Interface (HBA API) was developed by the Storage Networking Industry Association (SNIA) to simplify management. This is another example of how the Fibre Channel community works together to increase manageability of SANs.

Host Bus Adapter

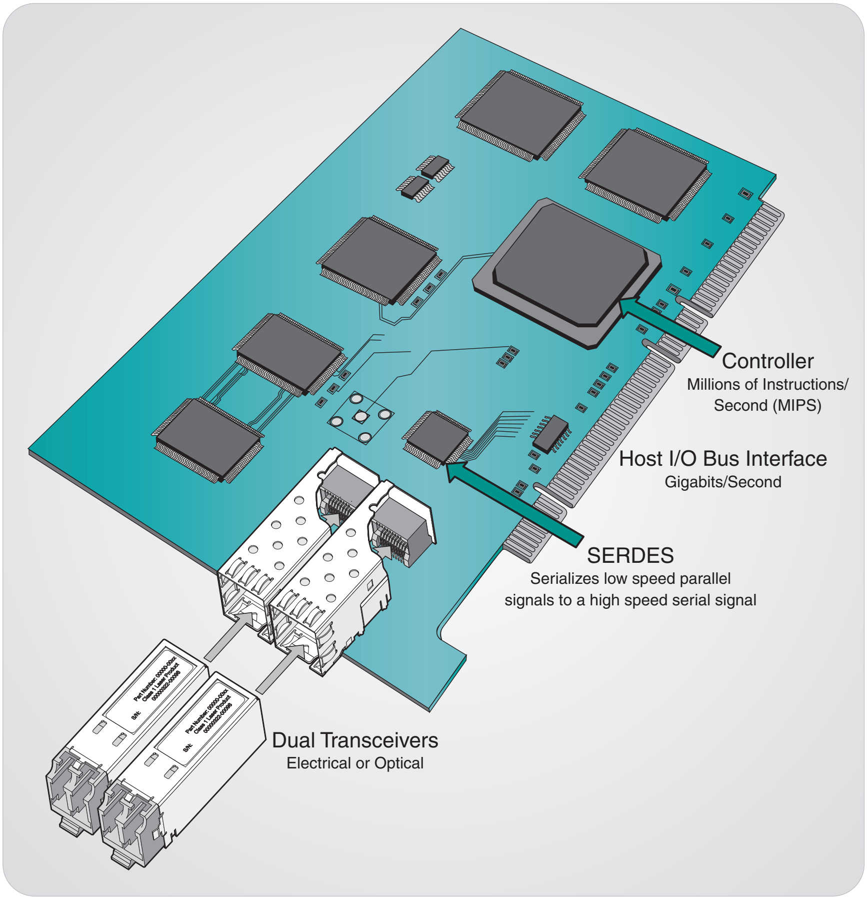


Figure 27: HBAs use SERDES and controllers to convert the Fibre Channel signals into signals for the host I/O bus on the computer. These prolific devices are the main interface between the storage world and the computing world.

TRANSLATION DEVICES

Translation devices connect the Fibre Channel network to outside networks or devices. The most common translation device is the HBA or Fibre Channel Adapter. The other common translation device is the storage controller that interfaces to the storage devices. A bridge (sometimes called an adapter) connects legacy SCSI or ESCON storage devices to the Fibre Channel network. An adapter usually connects Fibre Channel to IP networks such as Ethernet. Multifunction routers connect multiple Fibre Channel ports to multiple protocols such as SCSI, ATM, Infiniband, or Ethernet. Translation devices connect the Fibre Channel fabric to the hardware of the outside world.

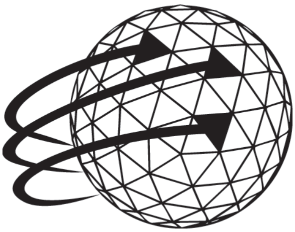
SUMMARY

The ubiquity of edge devices is the reason that Fibre Channel exists. Fibre Channel fabrics connect pools of storage devices to hundreds of servers. Servers use the fabric like an extended backplane to share valuable storage resources with other computing devices. Servers read and write terabytes of data at gigabit speeds to storage devices through the SAN.

Storage devices store data in disk drives or on tapes. Several disk drives are usually consolidated behind a controller that attaches to the SAN. The RAID is a common storage device that offers higher performance and increased reliability over individual disk drives. A storage subsystem incorporates multiple controllers, JBODs, and processors to increase the performance, capacity, and capability of the system. Advanced software on storage subsystems offers advanced services like remote mirroring, virtualization, and error recovery. Tape libraries have hundreds or thousands of tapes that can be robotically loaded into multiple tape drives. Fibre Channel connects these storage devices to computer resources via HBAs.

- *Storage Devices*: Combinations of disk drives and tape drives build powerful storage devices. Disk drives are combined to form JBODs. JBODs are used by RAIDs or storage subsystems to create high-capacity storage devices. Tape libraries archive petabytes of data for decades.

- *Host Bus Adapters*: HBAs are the interface between the host I/O bus and the SAN.
- *Translation Devices*: Translation devices are the interfaces between the Fibre Channel fabric and storage devices, servers, and other networks.



FABRIC SERVICES

Fabric services are the intelligent services provided by switched fabrics including the Name Server, Fabric Zone Server and Fabric Configuration Server. The Name Server stores attributes about ports and nodes and serves the data to querying devices and management applications. The Fabric Zone Server manages the zones which limit the information that the Name Server reports. The Fabric Configuration Server provides discovery of fabric topology and attributes to management applications. These three servers are the most widely deployed and used fabric services and are defined in Fibre Channel Generic Services 3 (FC-GS-3).

Storage Network Management (SNM) applications use fabric services to manage the Fibre Channel fabric. The fabric is divided into logical entities such as ports, nodes, and platforms which have a variety of attributes. When a port logs in to the fabric, the port registers various attributes that are stored in the fabric. SNM applications can then query the fabric to discover the attributes, capabilities, and topology of the fabric. Next, the SNM application configures zone sets, monitors the fabric, discovers the topology, and configures products.

COMMON TRANSPORT

An administrator typically accesses the fabric services via the SNM application resident on a management workstation. The management workstation may be attached to fabric devices in-band (over Fibre Channel) or out-of-band (via Ethernet). If the SNM application is manmanaging the fabric in-band, the SNM application accesses the fabric services via the Common Transport (CT) protocol. CT, shown in Figure 28, is similar to a ULP and is defined in FC-GS-3. The CT interface uses Fibre Channel Sequences to transfer Common Transport Information Units (CT_IUs) just like other ULPs. CT provides a standardized interface to the fabric services.

Fabric Services and Common Transport

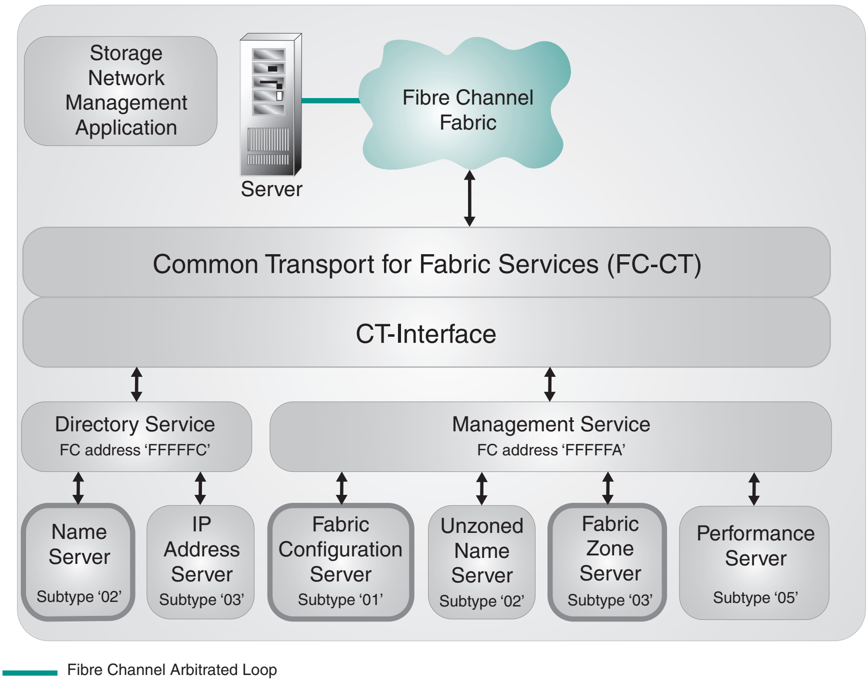


Figure 28: The CT protocol is used between SNM applications and the fabric services. The most widely used fabric services are the Name Server, Fabric Configuration Server, and Fabric Zone Server.

NAME SERVER

When a device logs in to the fabric, the ports must register with the Name Server. The Name Server is the main repository of information about the ports and nodes of edge devices, as seen in Figure 11. When the Name Server database changes, registered devices are notified by Registered State Change Notifications (RSCNs). The Name Server keeps track of which ports are currently available in the fabric and updates ports with RSCNs.

The Name Server is updated by numerous registration commands. Each port is referenced by its port identifier (Fibre Channel address), while node attributes are referenced by the node name. When the information in the Name Server changes, registered ports will receive an RSCN. After receiving an RSCN, the node (device) behind the port typically queries the Name Server to find out what devices are still available. The format of the RSCN identifies what type of event has occurred so that the node can direct its queries to the appropriate entity. For example, when a port logs in to the fabric, a port format RSCN is sent to devices in its zone so the device can send a PLOGI directly to the new device. The Name Server thus keeps the edge devices informed of the status of the other edge devices in the fabric.

FABRIC ZONE SERVER

To ensure security and dedicated resources in heterogeneous environments, an administrator creates zones and zone sets to restrict access to specified ports in the fabric. The Fabric Zone Server restricts the visibility of the Name Server for a given device.

A zone divides the fabric into groups of devices, as shown in Figure 29. Zone sets are groups of zones. Each zone set represents different configurations that optimize the fabric for certain functions.

For example, when the NT servers in Figure 29 are backed up, zone set 1 can be activated to enable a more efficient backup than when zone set 2 is active. Zones are usually configured by the WWNs of the attached devices. When zoning is based on WWNs, an attached device can be moved from one port to another, and the zoning is automatically maintained. Zoning enables heterogeneous environments and allows the fabric to be dynamically tuned for particular applications.

Zoning

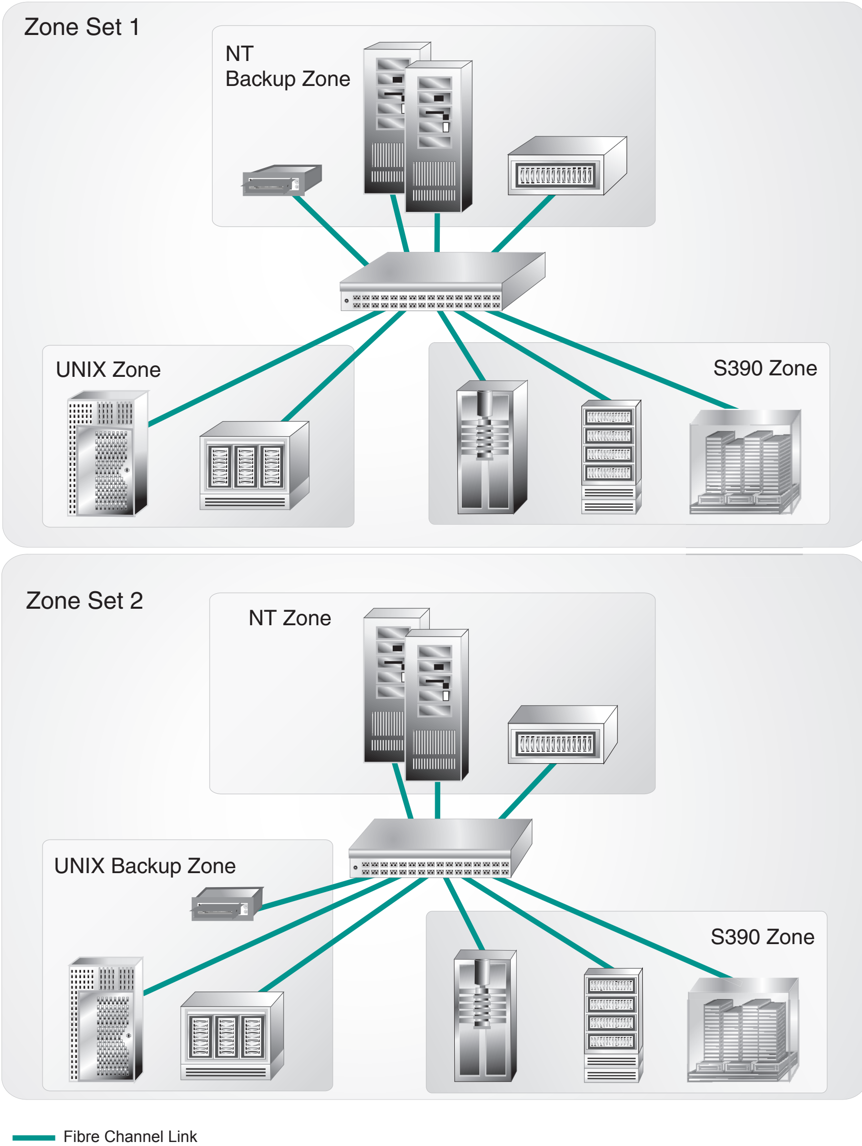


Figure 29: Zone sets are groups of zones that restrict traffic flow between various devices. Zone Set 1 is used when the NT servers are backed up and includes the tape drive. Zone Set 2 is used when the UNIX servers are backed up and includes the same tape drive.

Two types of zoning, hard and soft, are deployed in the fabric. Soft zoning is based on the information that the Name Server grants to the querying port. The Name Server limits the information about other ports in the fabric to those ports that are in the querying port's zone. The querying port sees only a limited view of the fabric based on the established zones. Hard zoning is managed identical to soft zoning, but prevents ports from communicating via hardware. The hard zone physically prevents frames from being routed through the fabric.

The Unzoned Name Server is a subtype of Management Services that allows an SNM application to discover all of the Name Server information without the limits of zoning. By using any of the commands defined for the Name Server, the Unzoned Name Server provides a powerful backdoor to the Name Server.

FABRIC CONFIGURATION SERVER

The Fabric Configuration Server delivers information about the interconnect elements (switches), switch ports, and platforms in the fabric. The Fabric Configuration Server attributes in Figure 30 are different from the Name Server attributes shown in Figure 11. The Fabric Configuration Server attributes define the elements of the fabric and the platforms that are entities above the nodes. The Name Server attributes define the nodes and ports that are attached to the fabric. The Configuration Server defines the physical configuration of the SAN.

The SNM application dynamically discovers the fabric topology via the Fabric Configuration Server. The Fabric Configuration Server is queried for topology information that enables SNM applications to form a dynamic topology map of the fabric. The fabric is a very interesting case study in distributed computing, as each switch must maintain and distribute information about its domain. The Fabric Configuration Server enables inband management of the fabric at unprecedented levels.

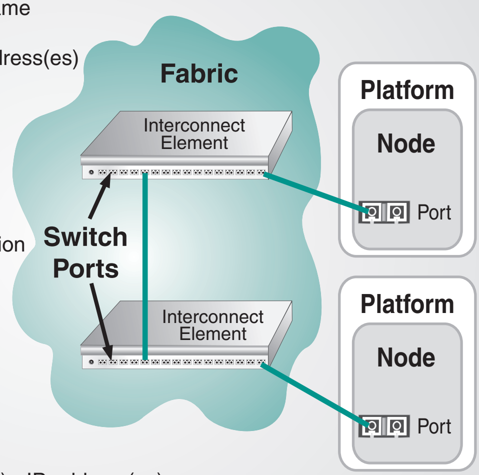
Configuration Server Elements

Interconnect Element Attributes:

Name - switch name
Type - predefined description
Domain Identifier - domain ID
Management Identifier - Domain Controller
Fabric Name - Principle Switch Name
Logical Name - Unique Text
Management Address(es) - IP address(es)
Information List - predefined text

Port Attributes:**

Name - WWN*
Type - predefined description
TX Type - predefined description
Module Type - predefined description
Physical Port Number - number
Attached Port Name(s) - WWN
Port State - predefined text



Platform Attributes:

Name - Unique Text
Type - predefined description
Node Name(s) - WWN
Platform Management Address(es) - IP address(es)
Platform Description - Unique Text
Platform Label - Unique Text
Platform Location - Unique Text

** **WWN** = World Wide Name that typically follows IEEE conventions.*
*** These port attributes are for the interconnect element ports, not edge device ports.*

Fibre Channel Link

Figure 30: The Fabric Configuration Server attributes define the interconnect elements and their ports. The topology of the fabric can thus be discovered by an SNM application. Platform information is also stored so that entities above the nodes can be addressed and described.

STORAGE NETWORK MANAGEMENT APPLICATIONS

SNM applications provide a single point of control for the fabric. Usually based on a simple browser interface, the SNM application is where administrators manage the fabric. From one console, the administrator can establish zoning, monitor the fabric, view the topology, and configure products.

Advanced features in SNM include policy-based management and SAN planning. Policy-based management uses administratively defined policies to manage one or more fabrics consistently. For example, alerts could be prioritized so that fault detection and recovery are improved. SAN planning software helps the administrator envision what the new SAN would look like with additional resources added. The SNM application might monitor port utilization statistics for months so that trends in data use could be visualized for the future. By looking at how data has been sent through the fabric, the administrator can better plan for the future. These advanced features will soon become commonplace in Fibre Channel.

SUMMARY

Fabric Services provide an effective, interoperable way to manage the fabric. As more fabrics utilize switches, storage and HBAs from multiple vendors, management of the SAN from one console is a crucial aspect of the deployment. SNM applications provide the single point of control for the fabric.

The standardization of the Fabric Services by ANSI is a huge advantage for interoperability and ease of management. Standards continue to be developed to facilitate device management within Fibre Channel, such as the HBA API and the Fabric Device Management Interface in FC-GS-4. The evolution of storage network management is a crucial aspect of the total-cost-of-ownership for SANs.

- *Common Transport*: SNM applications access the fabric services via CT-based commands.
- *Name Server*: The Name Server tracks the current state of the attached ports and nodes and reports changes to the registered ports.
- *Fabric Zone Server*: Zoning is controlled via the Fabric Zone Server to limit the number of devices that a device can access.
- *Fabric Configuration Server*: The Fabric Configuration Server keeps track of the fabric and the attached platforms to provide information on the topology of the fabric.

- *Storage Network Management Applications*: SNM applications provide a single point of control to all of the fabric services.

Chapter 8

FIBRE CHANNEL STANDARDS AND INTEROPERABILITY

ANSI fosters the creation of a wide variety of industry standards like Fibre Channel. The InterNational Committee for Information Technology Standards (INCITS) is an ANSI accredited entity with committees that develop IT standards. The T11 Technical Committee of INCITS develops the Fibre Channel standards. This chapter shows how the Fibre Channel standards are related.

Conformance to ANSI standards has direct impact on the interoperability of Fibre Channel products. Every Fibre Channel company tests for interoperability, which is required at some level for product sales. The FCIA is addressing interoperability with the SANmark program. The SANmark program encompasses explanations, requirements, and test suites to qualify products and systems to ensure that they comply with Fibre Channel Standards. The SANmark program is designed to increase the interoperability of Fibre Channel products.

FIBRE CHANNEL STANDARDS

Fibre Channel standards are primarily created by the INCITS T11 Technical Committee. A variety of standards are generated by the committee, which meets for a week every other month and has many interim meetings and teleconferences. Other organizations, like the INCITS T10 committee and IETF, also develop standards that apply to Fibre Channel.

The standards have been grouped into transport layer standards, which define the physical and signaling requirements of Fibre Channel; the topology standards, which define the interconnects and intelligence of Fibre Channel; the management standards, which define management interfaces; and the Upper Level Protocol mapping, which maps ULPs to Fibre Channel. In addition to standards, the T11 committees sponsor technical reports. The technical reports can specify requirements, define profiles, or act as implementation guides.

TRANSPORT LAYER STANDARDS

The basic functionality of Fibre Channel is described in the transport layer standards listed in Table 6. The transport layer standards specify all aspects of a point-to-point serial link, including cable requirements, transceiver types, link models, signal integrity, framing, signaling, and initialization. Fibre Channel - Framing and Signaling (FC-FS) has expanded the functionality of Fibre Channel to include many Extended Link Services, such as Process Login, State Change Registration, and Link Incident Reporting. The specification of the transport layer is so good that other gigabit network standards, like Gigabit Ethernet, have copied the basic transport layer of Fibre Channel.

Table 6. Transport Layer Standards

Standard	Description
Fibre Channel – Physical Interface	FC-PI defines the physical layer of Fibre Channel, including wiring, fiber optics, connectors, and transceivers. FC-PI specifies the next generation of physical layer requirements that were previously defined in FC-PH, FC-PH-2, and FC-PH-3.
Fibre Channel – Framing and Signaling	FC-FS defines the framing and signaling layers above the physical layer in Fibre Channel. FC-FS is the next generation of standards after FC-PH, FC-PH-2, and FC-PH-3.
Single Mode Longwave Laser – Vixelmedia	SM-LL-V is an addendum to FC-PI that defines a physical variant (a transceiver with a 1550 nm laser and singlemode fiber) that enables Fibre Channel links to span distances of over 50 kilometers.
Fibre Channel – 10 Gigabit	10 GFC defines the physical and signaling standards for Fibre Channel to operate at 10 Gbps.

TOPOLOGY STANDARDS

The transport layer standards have done an excellent job of specifying how one port communicates with another port. The topology standards outlined in Table 7 specify how multiple devices communicate in a switched fabric, an arbitrated loop, or a telecommunications network. FC-SW specifies how multiple switches communicate and distribute information about the fabric and the attached devices. FC-AL describes how devices can communicate in an arbitrated loop with or without a hub. FC-BB specifies three WAN interfaces that extend Fibre Channel fabrics via IP, ATM or SONET. FC-GS specifies the various services that a fabric provides to aid in the intelligence of the fabric. These fabric standards define the distributed architecture that allows communication across multiple interconnect devices in the fabric.

Table 7. Topology Standards

Standard	Description
Fibre Channel – Fabric Generic Requirements	FC-FG defines the minimum requirements for a topology-independent interconnecting fabric. This document laid the foundation for further standards development.
Fibre Channel – Framing and Signaling	FC-FS defines the framing and signaling layers above the physical layer in Fibre Channel. FC-FS is the next generation of standards after FC-PH, FC-PH-2, and FC-PH-3.
Fibre Channel – Switch Fabric	FC-SW describes the interaction and operation of multiple Fibre Channel switches through E_Ports, FL_Ports, and Switch Fabric Internal Link Services. The third generation of this document (FC-SW-3) will include distribution and negotiation of time-out values, multicasts, and security.
Fibre Channel – Arbitrated Loop	FC-AL specifies the signaling interface to allow communication via an arbitrated loop. The third generation of this document will enhance interoperability in arbitrated loops.
Fibre Channel – Backbone	FC-BB specifies the interface between Fibre Channel fabrics and SONET or ATM. FC-BB-2 extends Fibre Channel over IP networks with Fibre Channel over Intenet Protocol (FCIP).
Fibre Channel – Generic Services	FC-GS specifies the basic services of Fibre Channel, including Directory Services, Management Services, Time Services, Alias Services, and Key Distribution Services. The fourth generation of this standard (FC-GS-4) includes Performance Services, the Fabric Device Management Interface and more.

Table 7: The topology standards in combination with the transport layer standards define the three topologies in Fibre Channel. FC-FS defines the point-to-point topology, FC-SW defines the switched fabric topology, and FC-AL defines the arbitrated loop topology. The topologies are shown in Figure 31.

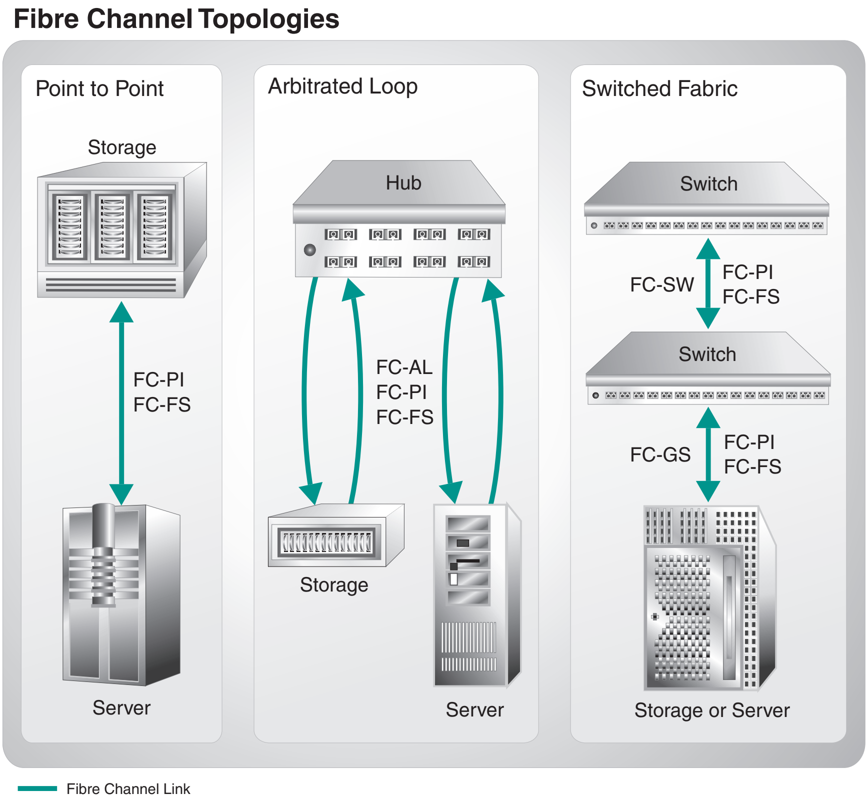


Figure 31: FC-PH defines the basic point-to-point topology that can connect two devices. Arbitrated loop, defined by FC-AL, joins several devices in one continuous loop. The switched fabric architecture, defined by FC-SW and FC-FS, allows highly scalable fabrics. The client interface to the fabric is defined in FC-GS.

MANAGEMENT STANDARDS

Management standards are evolving quickly to help IT administrators manage thousands of devices in large data centers. Table 8 shows the three standards that are being developed in the Storage Management Interfaces (SMI) Task Group of T11. The Application Programming Interfaces (APIs) and Management Information Bases (MIBs) are being standardized so that more management applications can take advantage of their use. The SMI Task Group encourages participation from more software management companies.

Table 8. Management Standards

Standard	Description
Fibre Channel – HBA API	FC-HBA defines an application programming interface for management applications to interface with HBAs from multiple vendors.
Fibre Channel – Switch API	FC-SWAPI defines an application programming interface for management applications to interface with switches from multiple vendors.
Fibre Channel – Management Information Base -FA	MIB-FA formalizes the MIB that was developed by the Fibre Alliance.

TECHNICAL REPORTS

Technical reports are created to explore a given topic or increase interoperability between Fibre Channel devices. Table 9 summarizes the main technical reports developed in the T11 committees. Profiles are technical reports that select specific options in the Fibre Channel standards to optimize the fabric for certain applications. Profiles apply to specific areas of Fibre Channel SANs, as shown in Figure 32.

Technical reports address more general problems that the Fibre Channel industry has observed. For example, high-speed jitter and signal specification is a specific aspect of Fibre Channel that needed to be investigated. The T11 committee started the Methodologies of Jitter Specification workgroup to study this crucial area in great detail. This T11 workgroup has studied many aspects of high-speed serial signals and turned their findings into an excellent standard that has been adopted by other standards organizations. Profiles and technical reports are excellent examples of collaborative efforts between some of the brightest and most experienced minds in storage networking.

Table 9. Fibre Channel Technical Reports

Standard	Description
Fibre Channel – Methodologies of Interconnects	FC-MI is a profile that addresses interoperability in loops, fabrics, and management applications. This comprehensive profile has direct input into the SANmark Program.
Fibre Channel – Device Attach	FC-DA is a profile that consolidates the previous profiles of FC-FLA, FC-PLDA, and FC-TAPE into one comprehensive document.
Fibre Channel – Fabric Loop Attach	FC-FLA specifies options regarding arbitrated loops that are attached to switch fabrics.
Fibre Channel – Private Loop SCSI Direct Attach	FC-PLDA specifies options regarding Fibre Channel, FCP, and SCSI-3 command sets for communication in private arbitrated loops.
Fibre Channel – Tape and Tape Medium Changers	FC-TAPE specifies options regarding Fibre Channel standards and SCSI devices that are required to operate streaming devices and medium changers in a public arbitrated loop environment.
Fibre Channel – Avionics Environment	FC-AE is a group of profiles which specify Fibre Channel options pertinent to commercial and military aerospace industries.
Fibre Channel – Methodologies of Jitter Specification -2	FC-MJS-2 enhances the jitter and signal specifications in FC-PH and FC-MJS.
Fibre Channel – Copper Interface Implementer's Guide	FC-CU describes the common practices and guidelines for implementing electric-based (copper) Fibre Channel connections.

SECURITY

One topic that is successfully being addressed in the ANSI committees is security. As SANs have grown across multiple data centers and more access points have become available through FCIP, the security of the SAN had to be maintained. A new security standard—Fibre Channel Security Protocols (FC-SP)—defines interoperable methods to secure the SAN at multiple levels.

The security points being addressed are authentication, authorization and encryption. Authentication uses digital secrets to prove that a device is not spoofing another device with a hijacked port name. Authentication can be between switches, end devices and management applications. FC-SP has decided that the required authentication mechanism will be Challenge Handshake

Authentication Protocol (CHAP). Authorization is another security measure that controls which devices are allowed to join the fabric with access control lists. Finally, when security is required at the highest levels, encryption of data ensures that intercepted data is rendered useless. The security of the fabric is of great importance and is supplemented by other security features adopted at the ULP and device level.

INTEROPERABILITY

Large investments of capital and time have been directed toward ensuring interoperability of Fibre Channel equipment. The largest investments have come from Fibre Channel companies that are continually testing products and product integration in SANs. One company has spent over a billion dollars on an interoperability lab. Many companies have formed businesses to test interoperability, and the University of New Hampshire's Interoperability Lab is devoted to testing the interoperability of Fibre Channel and other products.

The FCIA has developed the SANmark program to increase the interoperability of Fibre Channel devices, applications, and SANs. The FCIA works closely with T11 to meet the needs of the end user. The SANmark program is the evolutionary result of years of interoperability testing.

The SANmark program is designed to address all aspects of the SAN. Table 10 shows the array of Standards Conformance Documents (SCD) that the SANmark Committee addresses. The status of the SANmark program is continuously updated at www.fibrechannel.org. The low-level behaviors of Fibre Channel ports and products were addressed first, and growing levels of complexity will be addressed as the program continues. The SCDs were designed so that a variety of documents can be developed simultaneously.

The SANmark program also involves the development of test suites to carry out the tests specified in the SCDs. The documents and the tests are developed simultaneously so that the validity of the SCD and the interoperability of the products can be tested during development. The FCIA is committed to developing the most comprehensive interoperability program in storage networking.

Fibre Channel Profiles

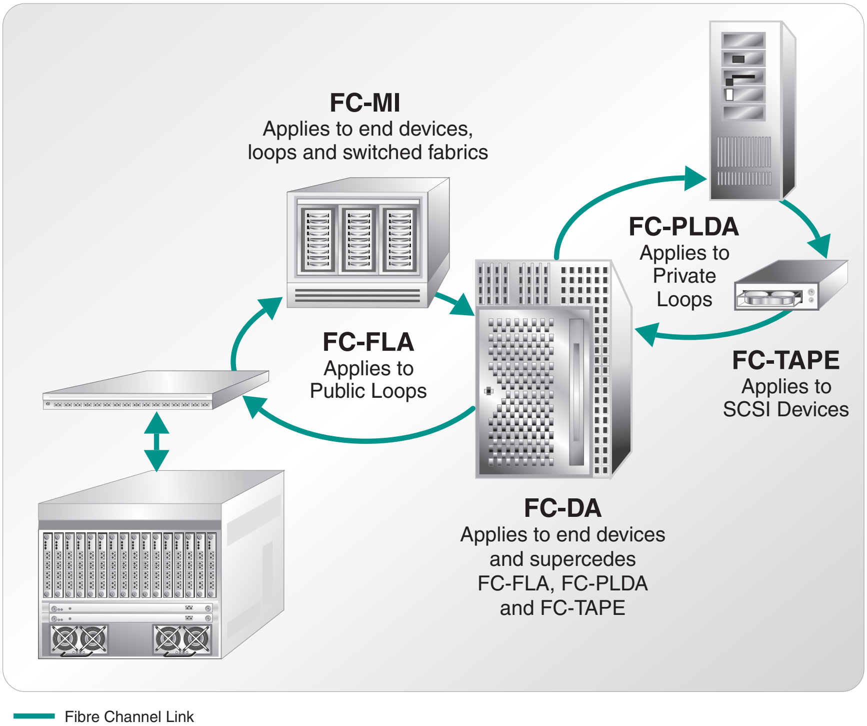


Figure 32: Profiles are technical reports that increase interoperability between devices. Profiles apply to specific aspects of Fibre Channel SANs so that solutions can be quickly deployed. FC-MI is the latest profile, encompassing all Fibre Channel topologies.

Table 10. Standard Conformance Documents (SCD)

Series	Description
0000	**Guiding Documents** – Describes program structure, the most current documents, and the direction of the program.
1000	**Physical and Mechanical** – Addresses cable plant, connectors, transceivers, signal integrity, hubs, and enclosures.
2000	**Port Behavior** – Addresses Nx_Ports, B_Ports, and multi-protocol ports.
3000	**Fabric Behavior** – Addresses E_Ports, Fx_Ports, distributed services, and FSPF routing protocol.
4000	**Management Functionality** – Addresses system and resource management, diagnostics, and middleware.
5000	**Fault Tolerance** – Addresses multi-path I/O, failover policies, robustness, and error recovery.
6000	**Security** – Addresses system security policies, authentication, encryption, and best practices.
7000	**Extended SAN Behavior** – Addresses distance connectivity, diagnostics, and troubleshooting for LAN, MAN, and WAN applications.
8000	**Applications and Solutions** – Addresses server clusters, copy services, backup, and high-level system services.

SUMMARY

ANSI standards have been developed to increase the performance and functionality of Fibre Channel. Standards range widely from the dimensions of the fiber-optic connectors to the management software interfaces. These standards help Fibre Channel products interoperate from the physical layer to the application.

The FCIA is committed to ensuring interoperability by developing the SANmark program, which addresses the various aspects of Fibre Channel SANs. The SANmark program is a collaborative effort by multiple system integrators and product vendors to ensure that SANs operate according to Fibre Channel standards and accepted practices.

- *Transport Layer Standards*: The transport layer standards define the physical, framing, and signaling standards for Fibre Channel links.
- *Topology Standards*: The topology standards define how interconnect devices and translation devices form fabrics and share information and connectivity.
- *Profiles and Technical Reports*: Various profiles and technical reports aid in the deployment of interoperable and robust SANs.
- *SANmark Program*: The SANmark program qualifies products, applications, and SANs to increase interoperability and high performance.

CONCLUSION

SANs based on Fibre Channel form the connection between clusters of servers and pools of storage. With DAS, servers have access to a limited amount of storage and storage is trapped behind a single server. Fibre Channel overcomes these limitations and enables sharing of powerful storage resources.

The Fibre Channel SAN also allows dynamic provisioning of resources. With advanced monitoring software, applications can dynamically allocate more servers or storage to an application when utilization reaches a specified threshold. For example, if an Internet video streaming application is redlining the five servers streaming a hot live video, the application could dynamically configure five more servers to double the streaming capacity. Likewise, if a storage device is nearing full capacity, storage virtualization could allocate more storage for the given application. Advanced application software optimizes utilization of SANs and their attached devices.

APPLICATION IS KING

Fibre Channel is the underlying technology that enables applications to scale to unprecedented levels. In the past, a single server used DAS to store data in a database or other application. As the database grew to hundreds of gigabytes or terabytes, the server not only had difficulty storing the massive amounts of data, but it also had difficulty serving the demand for that data. The distributive nature of the storage and servers attached to the SAN enables databases and applications to grow beyond the limitations of a single server or a few servers.

The advantage of Fibre Channel is that it facilitates the growth of applications to levels that have never existed before. The one area of computing that has scaled to massive proportions in the past is the mainframe environment. Fibre Channel built upon the foundations of the mainframe storage networks, Enterprise Systems Connection (ESCON), and utilized new technologies and protocols to form the most powerful information infrastructures ever created. Mainframes now use Fibre Channel for their storage networks and this same technology is used in open systems environments so Internet age applications can scale to global levels.

DYNAMIC RESOURCES

Fibre Channel is designed to accommodate the most fluid environments that are now experienced in today's data centers. Fibre Channel adapts to the changing applications environment easily. Zoning can be changed in an instant when different resources need to be used by different applications. For example, a server may only require a tape drive during backups at a specified time every night. A backup zone set could be activated at this time to offer dedicated resource for this special need. Fibre Channel enables the SAN to morph into various configurations as needed.

Variations in bandwidth demand can also be accommodated via Fibre Channel. Trunking is the best example of this. When a resource requires more bandwidth than is usually allocated to the resource, trunking allocates additional bandwidth on the fly to meet the needs of the resource. Paths on multiple interswitch links are optimized to allow full utilization of the available resources. As demand for data grows and shifts, Fibre Channel adapts to serve the changing demands of the data center.

FIBRE CHANNEL HARDWARE AND SOFTWARE

Fibre Channel hardware continues to grow in performance, interoperability, scale, and speed, while decreasing in size, power, latency, and cost. Firmware continues to grow in functionality, ease of use, and performance. Devices continue to incorporate more functionality by combining other devices such as routers, bridges, and switches into one product. The hardware continues to evolve, with large changes coming in 2003 when 10 GFC hits the market.

Fibre Channel software is external to devices and has a variety of uses, including management of the entire SAN, backup, and applications that span several devices. Some Fibre Channel companies are investing as much as 75% of their research and development money in software development. Software increases the level of automation and control of the hardware. As hardware becomes commoditized, software becomes the differentiator of products and companies.

FIBRE CHANNEL STANDARDS AND SANMARK

New generations of standards continue to be developed for the Fibre Channel industry. NCITS facilitates the main generation of Fibre Channel standards, and considerable efforts are being made by the IETF to connect remote Fibre Channel SANs. SNIA has a new interoperability testing facility in Colorado Springs, Colorado, that should see widespread use.

The SANmark program is intended to dispel any lingering interoperability concerns and will gain in acceptance and support. The program is designed to address a multitude of interoperability points and offers a way to specify behaviors that the industry wants to see, like trunking, fault tolerance, and data recovery.

LAST WORDS

Fibre Channel SANs are solving the real-world problems in IT today. The industry is scaling to support the exponential growth in data while the number of available IT staff is remaining nearly constant. Fibre Channel is a sound architecture that is the foundation for many development efforts. Fibre Channel has reached critical mass and continues to grow unabated for the foreseeable future.

- *Application Is King*: The Fibre Channel architecture opens the door for scalability, availability, and distance. Heterogeneous systems and multiple applications can run reliably on Fibre Channel SANs.
- *Dynamic Resources*: Fibre Channel adapts to the changing environment of the data center with zoning and trunking.
- *Fibre Channel Hardware and Software*: As certain as Moore's law, hardware continue to provide more for less, and software continues to grow in power and ease of use.
- *Fibre Channel Standards and SANmark*: Fibre Channel standards and SANmark continues to evolve to increase performance and interoperability.

Chapter 10

REFERENCES

BOOKS

1. Kembel, Robert, *A Comprehensive Introduction to Fibre Channel*, Northwest Learning Associates, Tucson, AZ, 2000.
2. Kembel, Robert, *Fibre Channel Switched Fabrics*, Northwest Learning Associates, Tucson, AZ, 2001.
3. Farley, Marc, *Building Storage Networks* – Second Edition, Osborne/McGraw-Hill, Berkeley, CA, 2001.
4. Clark, Tom, *Designing Storage Networks*, Addison Wesley Longman, Inc., Reading, MA, 1999.
5. *Fibre Channel Storage Area Networks*, Fibre Channel Industry Association, 2001.

WEBSITES

www.infostor.com – Infostor covers the complete storage networking market and has an excellent archive of past articles.

www.fibrechannel.org – The website for the Fibre Channel Industry Association has the latest news of the organization and its projects like SANmark.

www.searchstorage.com – Searchstorage covers the gamut of storage issues and has excellent interactive online events.

www.byteandswitch.com – Byte and Switch focuses on the actions of individual companies and offers breaking news on the storage industry.

www.t11.org – The T11 Technical Committee of NCITS develops most Fibre Channel standards and is an excellent resource center for released standards and standards in development.

www.fibrechannel-europe.com – The Fibre Channel Association of Europe features events specific to Europe and a variety of Fibre Channel topics.

www.storageperformance.com – The Storage Performance Council is defining a benchmark for storage subsystems.

http://lw.pennnet.com/home.cfm – The home of Lightwave magazine focuses on fiber-optic communication and has articles on fiber-optic applications like Fibre Channel SANs.

www.snia.org – The Storage Networking Industry Association covers many storage networking topics. SANs based on Fibre Channel are their main solution.

www.esj.com –The *Enterprise Systems Journal* covers enterprise systems and related issues, including SANs.

www.ncits.org – The National Committee for Information Technology Standards sponsors the T11 Technical Committee and many other standards committees.

www.wwpi.com – This is the website for Computer Technology News which covers many storage and computer topics.

ACRONYMS

AE – Avionics Environment
AL – Fibre Channel Arbitrated Loop
ANSI – American National Standards Institute
API – Application Programming Interface
ARP – Address Resolution Protocol
ASP – Application Service Provider
ATM – Asynchronous Transfer Mode
AV – Audio Visual
BB – BackBone
BSP – Business Service Provider
CIO – Chief Information Officer
CLS – Close Primitive Signal
CPU – Central Processing Unit
CT – Common Transport

CT_IU – Common Transport Information Unit
DA – Device Attach
DAS – Directly Attached Storage
E/O – Electrical/ Optical
ESCON – Enterprise Systems Connection
ERP – Enterprise Resource Planning
FC – Fibre Channel
FCIA – Fibre Channel Industry Association
FCP – Fibre Channel Protocol
FG – Fabric Generic
FICON – Fibre Connection
FLA – Fabric Loop Attach
FLOGI – Fabric Login
FS – Framing and Signaling
FSPF – Fabric Shortest Path First
GBIC – Gigabit Interface Converter
GFC – Gigabit Fibre Channel
GS – Generic Services
HBA – Host Bus Adapter
IEEE – Institute of Electrical and Electronic Engineers
IETF – Internet Engineering Task Force
INCITS – InterNational Committee for Information Technology Standards
I/O – Input / Output
IP – Internet Protocol
ISL – Inter-Switch Link
ISP – Internet Service Provider
IT – Information Technology
JBOD – Just a Bunch of Disks
LAN – Local Area Network
LPSM – Loop Port State Machine
MAN – Metropolitan Area Network
MI – Methodologies of Interconnects
MPEG – Motion Pictures Expert Group
MSA – Multi-Sourcing Agreement

NIC – Network Interface Card
NVOD – Near Video On Demand
OLTP – On-Line Transaction Processing
OLS – Offline Primitive Sequence
OPN – Open Primitive Sequence
PCI – Peripheral Component Interconnect
PH – PHysical
PI – Phyical Interface
PLDA – Private Loop SCSI Direct Attach
PLOGI – Port LOGIn
PRLI – PRocess LogIn
PSM – Port State Machine
RAID – Redundant Array of Independent Disks
RAM – Random Access Memory
RAS – Reliability, Availability and Serviceability
RSCN – Registered State Change Notification
SAD – Scalability, Availability and Distance
SAM – SCSI Architecture Model
SAN – Storage Area Network
SCSI – Small Computer Systems Interface
SERDES – SERializer DESerializer
SFF – Small Form Factor
SFP – Small Form Factor Pluggable
SNIA – Storage Netwroking Industry Association
SNM – Storage Network Management
SONET – Synchronous Optical Network
SSP – Storage Service Provider
SW – Switched Fabric
ULP – Upper Level Protocol
VI – Virtual Interface
VOD – Video On Demand
WAN – Wide Area Network
WDM – Wavelength Division Multiplexing
WWN – World Wide Name