Fibre Channel Guide

EXFO

EXPERTISE REACHING OUT

About EXFO

EXFO is the recognized expert in the global telecommunications industry through the design and manufacture of advanced and innovative test and measurement instruments. The Telecom Division, which represents our main business activity, offers fully integrated and complete test solutions to network service providers, system vendors and component manufacturers in approximately 70 countries. One of our strongest competitive advantages is our modular platform design, providing PC-based, Windows-centric test solutions that maximize technology reuse across several market segments. The Photonics and Life Sciences Division mainly leverages core telecom technologies to offer value-added solutions in high-tech industrial manufacturing and research sectors.

Table of Contents

1. Introduction to Fibre Channel

1. Introduction to Fibre Channel

1.1 Description

Fibre Channel[1] is a serialized protocol that was designed for storage-area-network communications. The concept of storage area networks (SAN) was created in order to integrate storage and network technologies for storage solutions, in response to today's enterprise needs. A SAN is a grouping of interconnected devices and servers using a common communication infrastructure, which capitalizes on the best of both the storage and network worlds. The *de facto* protocol for this infrastructure is Fibre Channel, an ANSI standard that allows for an inexpensive and scalable architecture.

Fibre Channel's principal trait is its fusion of storage and networking technologies, thus allowing communications with the following attributes (see Figure 1.1):

- High bandwidth
- Low latency
- High data integrity
- High connectivity
- Routing
- Large distances
- Management and troubleshooting tools

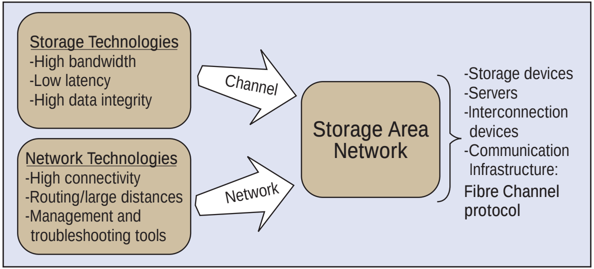


Figure 1.1: SAN Attributes

[1]Fibre Channel was originally spelled Fiber Channel and was designed to support fiber-optic cabling only. When copper support was added to its feature set some years ago, it was decided to change the spelling of "Fiber" to the European spelling, to reduce the association with fiber optics while maintaining the name recognition of the Fiber Channel technology.

The goal of Fibre Channel is to carry different types of traffic for applications that require the first-rate capabilities of storage *and* network technologies, such as IP, SCSI, iSCSI, FICON, ESCON, HiPPI, FDDI, ATM, etc.

1.2 History

The development of Fibre Channel first started in 1988, as a practical, inexpensive and scalable way of using fiber-optic cables to connect storage devices. It was approved as a standard by the American National Standards Institute (ANSI) in 1994.

Currently, the Fibre Channel standard is under the responsibility of Task Committee T11 of ANSI NCITS.

1.3 Data Rates and Interfaces

The medium for Fibre Channel is “Fibre”, a term that includes both fiber-optic and copper cabling. The distance over which Fibre Channel operates is dependent on the copper cabling, as well as on its corresponding transceiver.

Table 1.1 is a list of the currently supported Fibre Channel line rates.

Line Rate (Gb/s)	Throughput (MB/s)	T11 Specification Completion (Year)
1.0625	100	1996
2.125	200	2000
4.25	400	2002
8.5	800	2005
10.5	1200	2003

Table 1.1: Fibre Channel Transmission Rates

Fibre Channel data transfer bandwidth is dependent on clock rate and protocol overhead. The base clock rate for Fibre Channel is 1.0625 GHz, with 1 bit transmitted every clock cycle, which yields a rate of 1.0625 Gb/s. The throughput is calculated as follows:

With every Fibre Channel frame, which has a payload size of 2048 bytes, there are 120 bytes of overhead and 10 bits of data transferred for every byte of data, due to 8B/10B encoding (see Architecture section for details). Therefore, the effective data transfer rate is approximately 100 MB/s.

$$1.0625[Gb/s] \times \frac{2048[payload]}{2168[payload + overhead]} \times \frac{1[byte]}{10[codebits]} = 100.369 \text{ MB/s}$$

In the industry, both 1 Gb/s and 100 MB/s are used interchangeably to specify port or network transmission rates. The same holds true for 2 Gb/s, 4 Gb/s and 10 Gb/s rates and their MB/s counterparts.

Table 1.2 and 1.3 are lists of optical and electrical interfaces defined for Fibre Channel. The nomenclature for specifying Fibre Channel interfaces is as follows:

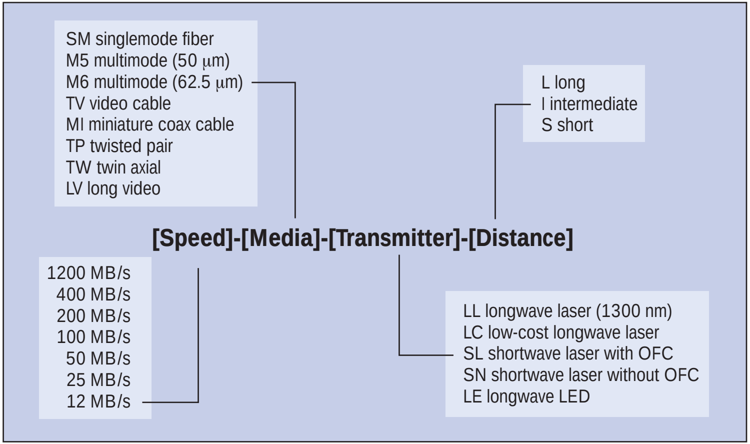


Figure 1.2: Fibre Channel Architecture

Data Rate (MB/s)	Bit Rate (MBaud)	Distance (meters)	FC-0 Code	Electrical Cable	Notes
100	1062.5	0 - 25	100-TV-EL-S	75 Ω Video	1
100	1062.5	0 - 10	100-MI-EL-S	75 Ω Mini Coax	2
50	531.25	0 - 50	50-TV-EL-S	75 Ω Video	1
50	531.25	0 - 15	50-MI-EL-S	75 Ω Mini Coax	2
25	265.625	0 - 75	25-TV-EL-S	75 Ω Video	1
25	265.625	0 - 50	25-TP-EL-S	150 Ω Shielded Twisted Pair	3
25	265.625	0 - 25	25-MI-EL-S	75 Ω Mini Coax	2
12.5	132.8125	0 - 100	12-TV-EL-S	75 Ω Video	1
12.5	132.8125	0 - 100	12-TP-EL-S	150 Ω Shielded Twisted Pair	3
12.5	132.8125	0 - 35	12-MI-EL-S	75 Ω Mini Coax	2

Table 1.2: Fibre Channel Electrical Interfaces

Note 1: The FC-PH standard specifies RG-6/U or RG-59/U coaxial video cable. Double-shielded versions of these cables will meet emission requirements. The recommended RG-6/U cable is Belden #1694A. The recommended RG-59/U cable is Belden #9209.

Note 2: The FC-PH standard specifies RG-179/U miniature coaxial cable. This cable will not meet emission requirements, but may be used inside a shielded enclosure. No other miniature coaxial cable has been recommended.

Note 3: The FC-PH standard specifies that shielded twisted pair cable shall conform to EIA/TIA568. The recommended shielded twisted pair cable is currently referred to as "IBM TYPE-1", a cable developed for IBM Token Ring.

Data Rate (MB/s)	Bit Rate (MBaud)	Distance (meters)	FC-0 Code	Fibre and Cable Type	Laser Type
400	4250	2 - 2000	400-SM-LL-I	Singlemode	1300 nm Longwave Laser ‡
400	4250	2 - 175	400-M5-SN-I	50 µm Multimode	780 nm Shortwave Laser ‡

Table 1.3: Fibre Channel Optical Interfaces *(Continued on page 7)*

*: The laser transmitter and receiver specifications for this link option are at a higher performance level to support the greater operating distance.
†: These are "alternative fibre cable plant".
‡: These do not include Open Fibre Control (OFC).

Data Rate (MB/s)	Bit Rate (MBaud)	Distance (meters)	FC-0 Code	Fibre and Cable Type	Laser Type
400	4250	2 - 50	400-M6-SN-I	62.5 µm Multimode †	780 nm Shortwave Laser ‡
200	2125	2 - 2000	200-SM-LL-I	Singlemode	1300 nm Longwave Laser ‡
200	2125	2 - 300	200-M5-SN-I	50 µm Multimode	780 nm Shortwave Laser ‡
200	2125	2 - 90	200-M6-SN-I	62.5 µm Multimode †	780 nm Shortwave Laser ‡
100	1062.5	2 - 10 000	100-SM-LL-L	Singlemode	1300 nm Longwave Laser *
100	1062.5	2 - 2000	100-SM-LL-I	Singlemode	1300 nm Longwave Laser
100	1062.5	2 - 500	100-M5-SL-I	50 µm Multimode	780 nm Shortwave Laser
100	1062.5	2 - 500	100-M5-SN-I	50 µm Multimode	780 nm Shortwave Laser ‡
100	1062.5	2 - 175	100-M6-SL-I	62.5 µm Multimode †	780 nm Shortwave Laser
100	1062.5	2 - 175	100-M6-SN-I	62.5 µm Multimode †	780 nm Shortwave Laser ‡
50	531.25	2 - 10 000	50-SM-LL-L	Singlemode	1300 nm Longwave Laser
50	531.25	2 - 1000	50-M5-SL-I	50 µm Multimode	780 nm Shortwave Laser
50	531.25	2 - 500	50-M6-LE-I	62.5 µm Multimode	1300 nm Longwave LED ‡
50	531.25	2 - 350	50-M6-SL-I	62.5 µm Multimode †	780 nm Shortwave Laser
25	265.625	2 - 10 000	25-SM-LL-L	Singlemode	1300 nm Longwave Laser *
25	265.625	2 - 2000	25-SM-LL-I	Singlemode	1300 nm Longwave Laser
25	265.625	2 - 2000	25-M5-SL-I	50 µm Multimode	780 nm Shortwave Laser
25	265.625	2 - 1500	25-M6-LE-I	62.5 µm Multimode	1300 nm Longwave LED ‡
25	265.625	2 - 700	25-M6-SL-I	62.5 µm Multimode †	780 nm Shortwave Laser
25	265.625	not defined	25-M5-LE-I	50 µm Multimode †	1300 nm Longwave LED ‡
12.5	132.8125	2 - 1500	12-M6-LE-I	62.5 µm Multimode	1300 nm Longwave LED ‡
12.5	132.8125	not defined	12-M5-LE-I	50 µm Multimode †	1300 nm Longwave LED ‡

Table 1.3: Fibre Channel Optical Interfaces (Continued)

1.4 Architecture

Figure 1.3 represents the Fibre Channel protocol stack.

- FC-0: This layer defines the physical link in the system, including the fiber, connectors, optical and electrical parameters for the different data rates.
- FC-1: This is the transmission protocol. The main functions are serial encoding/decoding, error control and special character management. The information transmitted over the fibre is carried with an 8B/10B encoding scheme in order to ensure that the transmission code is DC-balanced and that clock recovery is possible.

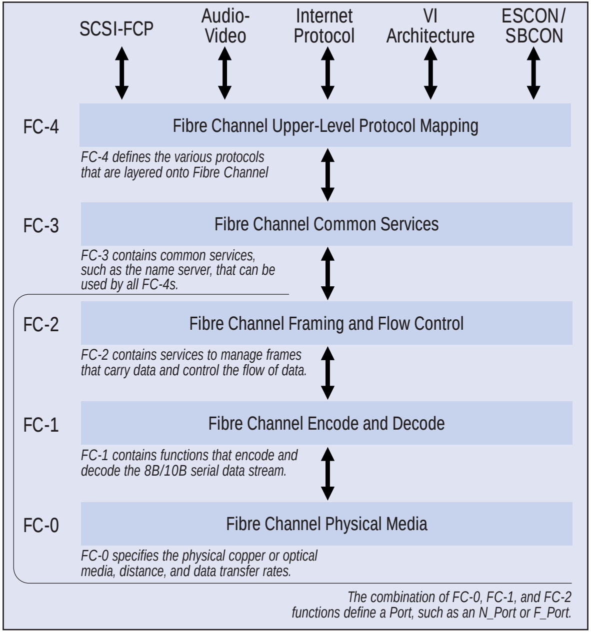


Figure 1.3: Fibre Channel Protocol Stack

- FC-2: This signaling protocol level serves as the transport mechanism of Fibre Channel and is similar to the transport layer in the OSI reference model. FC-2 defines framing rules, class-of-service management and sequence management by using the following building blocks (see table 1.4), which are detailed in the data flow section:

Ordered Set	Fibre Channel control information
Frame	Series of encoded transmission words
Sequence	Series of one or more frames
Exchange	Series of one or more non-concurrent sequences
Protocol	Set of frames transmitted for administrative purposes

Table 1.4: FC-2 Building Blocks

FC-2 is also where the login mechanism takes place (see *Login mechanism* section for details).

- FC-3: The FC-3 level is intended to provide the common services across multiple ports required for advanced features such as striping, hunt groups and multicast.
- FC-4: The highest level in the FC structure provides mapping to upper-layer protocols. These upper-layer protocols can be channel or network protocols.

1.4.1 Link management

Link management is the process of initializing Fibre Channel links and managing different states. The different states are listed in the following port transition chart (table 1.5).

ù

current \ next	Active	Link Recovery			Link Failure		Offline		
	AC	LR1	LR2	LR3	LF1	LF2	OL1	OL2	OL3
AC - Active sending data	Idle, R_RDY		LR	Err- LRR (4)	Err- NOS (1)	Err- Loss of sync(2) or signal(3)		OLS	
LR1 - LR Transmit sending LR		Idle, R_RDY	LR	LRR	Err- NOS (1)	Err- Loss of sync(2) or signal(3) timeout(1)		OLS	
LR2 - LR Receive sending LRR	Idle		LR	LR	Err- NOS (1)	Err- Loss of sync(2) or signal(3) timeout(1)		OLS	
LR3 - LRR Receive sending Idle			LR		Err- NOS (1)	Err- Loss of sync(2) or signal(3) timeout(1)		OLS	

Table 1.5: Fibre Channel Link Management States (Continued on page 11)

current \ next	Active	Link Recovery			Link Failure		Offline		
	AC	LR1	LR2	LR3	LF1	LF2	OL1	OL2	OL3
LF1 - NOS Receive sending OLS			LR		LRR, NOS, Idle	Err- Loss of sync(2) or signal(3) timeout(1)		OLS	
LF2 - NOS Transmit sending NOS					NOS	LR, LRR, Idle		OLS	
OL1 - OLS Transmit sending OLS			LR*		Err- NOS* (1)		LRR*, Idle*	OLS*	Loss of sync or signal*, timeout
OL2 - OLS Receive sending LR			LR	LRR	Err- NOS (1)			OLS, Idle	Loss of sync or signal, timeout
OL3 - Wait for OLS sending NOS					NOS	Err- LR, LRR		OLS	Idle, Loss of sync or signal

Table 1.5: Fibre Channel Link Management States (Continued)

[1]Note: Transitions marked by Err are invalid and trigger updates of a Link Error Status block; i.e., increment (1) link failure, (2) loss of synchronization, (3) loss of signal, or (4) primitive sequence protocol error counters in the LESB.

[2]Note: Transitions marked by * only occur as part of the link initialization protocol. In the online-to-offline protocol, the transitions are disabled, and the port may power down.

Link initialization is the first event that takes place when a Fibre Channel port receives a valid incoming signal (i.e., no more loss of signal). Figure 1.4 depicts the initialization process.

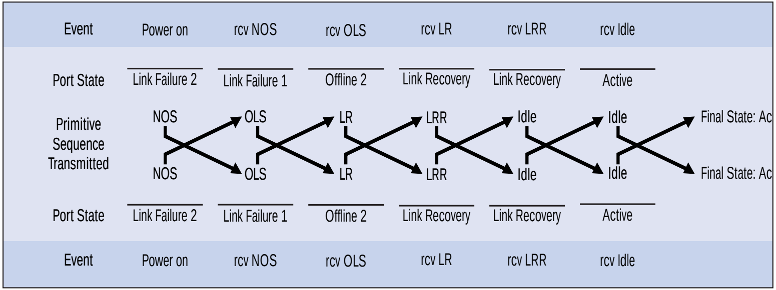


Figure 1.4: Link Initialization Sequence

1.4.2 Login mechanism

Fibre Channel has two different types of logins (see Addressing and Topologies sections for more information):

- Port login
- Fabric login

The **port login**, also known as PLOGI, is used to establish a session between two N_Ports (devices). During the port login, two N_Ports (devices) swap service parameters and introduce themselves to each other. This is necessary before any upper-level commands or operations can be performed.

The **fabric login**, also known as FLOGI, is carried out by the fabric-capable Fibre Channel device after it is attached to a fabric switch. Like PLOGI, FLOGI is an extended link service command that sets up a session between two participants. With FLOGI, a session is created between an N_Port or NL_Port and the switch. An N_Port sends a FLOGI frame with its node name, its N_Port name, and service parameters to a well-known address of 0xFFFFFE. The switch accepts the login and returns an acceptance (ACC) frame to the sender. If some of the service parameters requested by the N_Port or NL_Port are not supported, the switch sets the relevant bits in the RJT frame to indicate this. When the N_Port logs in, it uses a 24-bit port address of 0x000000. As a result, the fabric is able to assign the appropriate port address to that device, based on the domain-area-port address format; the new address is contained in the ACC response frame.

After FLOGI assigns the N_Port or public NL_Port its fabric address, the port must register with the SNS, using port login (PLOGI) at the address 0xFFFFFC. While the device may register values for some or all database objects, the most useful are the following:

- 24-bit port address
- 64-bit World Wide Node (WWN) name
- FC-4 protocols supported
- class-of-service parameters
- port type, such as N_Port or NL_Port

1.4.3 Flow control

Ports use "buffers" (memory) to temporarily store frames until they are assembled sequentially and delivered to the upper-layer protocol. The number of frames a port can store is referred to as a “buffer credit”.

N_Ports and F_Ports at each end of a link establish their buffer-to-buffer credit (BB_Credit) during login. N_Ports also establish end-to-end credit (EE_Credit) with each other.

Two counters are used to ensure that, during data transmission, a port does not send more frames than the buffer size of the receiving port before the receiving port indicates that it has processed the previous frame. The counters are BB_Credit_CNT and EE_Credit_CNT; during login, both are initialized at 0.

Whenever a port sends a frame, it augments BB_Credit_CNT and EE_Credit_CNT by 1. When it receives R_RDY from the adjacent port, it reduces BB_Credit_CNT by 1, and when it receives ACK from the destination port, it reduces EE_Credit_CNT by 1. If BB_Credit_CNT ever becomes equal to the BB_Credit, or if EE_Credit_CNT equals the receiving port's EE_Credit, the transmitting port ceases to send frames until the corresponding count is reduced. These flow control mechanisms are not universal, but vary according to class of service (see Classes of Service section).

Class of service	Flow control
1	EE
2	EE and BB
3	BB

Table 1.6: CoS vs. Flow Control

The relationship between flow control and class of service is shown in Table 1.6.

1.4.4 Data flow

Data hierarchy in Fibre Channel consists of the following building blocks:

- Frames: Series of transmission words used for transferring upper-level protocol data. There are data frames and link control frames (only with connection-oriented classes of service). Refer to the Framing section for details.
- Sequences: Unidirectional series of frames flowing from Sequence Initiator to Sequence Recipient. Sequences are identified by SEQ_ID, and the frames used are identified by SEQ_CNT.
- Exchanges: Series of sequences flowing from Exchange Originator to Exchange Responder. Exchanges are identified with X_ID.
- Protocol: Set of frames sent over one or more exchanges, transmitted for specific administrative purposes such as logins, port status retrieval, or exchange/sequence abortion.

Figure 1.5 is an example of the Frame/Sequence/Exchange hierarchy.

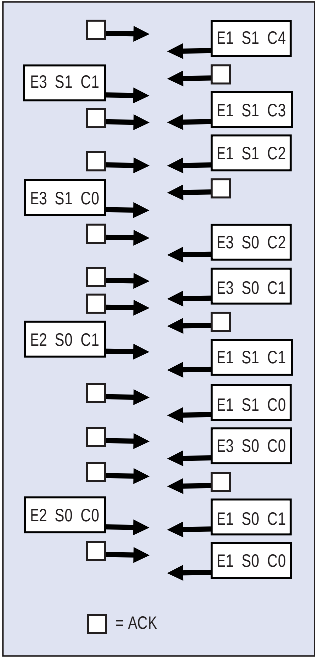


Figure 1.5: Frame/Sequence/Exchange Hierarchy

1.4.5 Framing

Framed FC-1

FC-1 frames are those that are generated at the FC-1 layer of the Fibre Channel stack. Framed FC-1 (see figure 1.6) has start-of-frame (SOF) and end-of-frame (EOF) primitives.

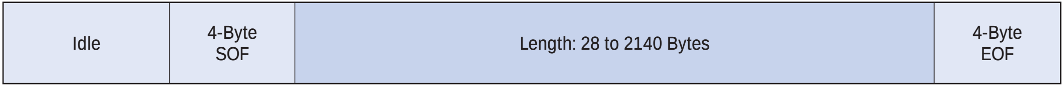


Figure 1.6: FC-1 Frame

Framed FC-2

FC-2 frames are generated and processed at the FC-2 layer of the Fibre Channel stack (see figure 1.7).

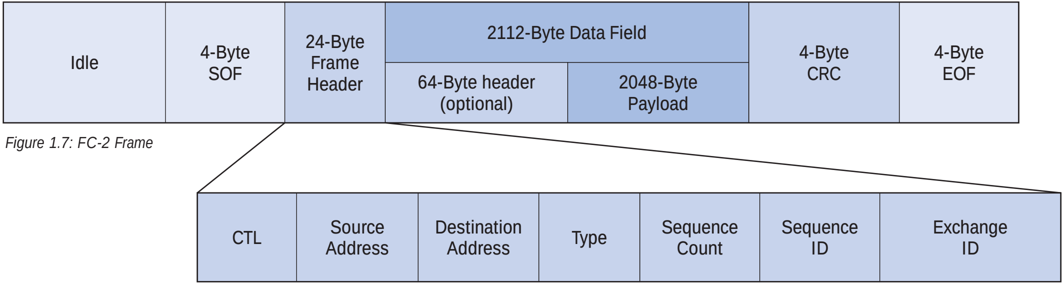


Figure 1.7: FC-2 Frame

These frames can form the following frame delimiter ordered sets (table 1.7):

Frame Delimiters			
Delimiter Function SOF = Start of Frame, EOF = End of Frame	Abbrev. RD	Beginning	Ordered Set
SOF Connect Class 1	SOFc1	Negative	K28.5 D21.5 D23.0 D23.0
SOF Initiate Class 1 SOF Normal Class 1	SOFi1 SOFn1	Negative Negative	K28.5 D21.5 D23.2 D23.2 K28.5 D21.5 D23.1 D23.1
SOF Initiate Class 2 SOF Normal Class 2	SOFi2 SOFn2	Negative Negative	K28.5 D21.5 D21.2 D21.2 K28.5 D21.5 D21.1 D21.1
SOF Initiate Class 3 SOF Normal Class 3 (Also SOFil, for AL initialization)	SOFi3 SOFn3	Negative Negative	K28.5 D21.5 D22.2 D22.2 K28.5 D21.5 D22.1 D22.1
SOF Fabric	SOFf	Negative	K28.5 D21.5 D24.2 D24.2
SOF Activate Class 4	SOFc4	Negative	K28.5 D21.5 D25.0 D25.0
SOF Initiate Class 4	SOFi4	Negative	K28.5 D21.5 D25.2 D25.2
SOF Normal Class 4	SOFn4	Negative	K28.5 D21.5 D25.1 D25.1
EOF Normal	EOFn	Negative Positive	K28.5 D21.4 D21.3 D21.3 K28.5 D21.5 D21.3 D21.3
EOF Terminate	EOFt	Negative Positive	K28.5 D21.4 D21.4 D21.4 K28.5 D21.5 D21.4 D21.4

Table 1.7: Frame Delimiters (Continued on page 17)

Frame Delimiters			
Delimiter Function SOF = Start of Frame, EOF = End of Frame	Abbrev. RD	Beginning	Ordered Set
EOF Disconnect-Terminate (Class 1) EOF Deactivate-Terminate (Class 4)	EOFdt	Negative Positive	K28.5 D21.4 D21.7 D21.7 K28.5 D21.5 D21.7 D21.7
EOF Remove-Terminate (Class 4)	EOFrt	Negative Positive	K28.5 D21.4 D25.4 D25.4 K28.5 D21.5 D25.4 D25.4
EOF Disconnect-Terminate-Invalid (Class 1) EOF Deactivate-Terminate-Invalid (Class 4)	EOFdti	Negative Positive	K28.5 D10.4 D21.4 D21.4 K28.5 D10.5 D21.4 D21.4
EOF Remove-Terminate-Invalid (Class 4)	EOFrti	Negative Positive	K28.5 D10.4 D25.4 D25.4 K28.5 D10.5 D25.4 D25.4
EOF Normal-Invalid	EOFni	Negative Positive	K28.5 D10.4 D21.6 D21.6 K28.5 D10.5 D21.6 D21.6
EOF Abort	EOFa	Negative Positive	K28.5 D21.4 D21.6 D21.6 K28.5 D21.5 D21.6 D21.6

Table 1.7: Frame Delimiters (Continued)

Primitive Signals			
Primitive Signal	Abbrev.	Beginning RD	Ordered Set
Idle	Idle	Negative	K28.5 D21.4 D21.5 D21.5
Receiver_Ready	R_RDY	Negative	K28.5 D21.4 D10.2 D10.2
Virtual Circuit Ready (Class 4)	VC_RDY	Negative	K28.5 D21.7 VC_ID VC_ID
Arbitrate	ARByx	Negative	K28.5 D20.4 y x
Arbitrate (val)	ARB(val)	Negative	K28.5 D20.4 val val
Open Full-Duplex	OPNyx	Negative	K28.5 D17.4 AL_PD AL_PS
Open Half-Duplex	OPNyy	Negative	K28.5 D17.4 AL_PD AL_PD
Open Broadcast Replicated	OPNyr	Negative	K28.5 D17.4 D31.7 D31.7
Open Selective Replicated	OPNfr	Negative	K28.5 D17.4 AL_PD D31.7
Close	CLS	Negative	K28.5 D5.4 D21.5 D21.5
Mark	MRKtx	Negative	K28.5 D31.2 MK_TP AL_PS
Dynamic Half-Duplex	DHD	Negative	K28.5 D10.4 D21.5 D21.5

Table 1.8: Primitive Signals

Primitive Sequences			
Primitive Sequence	Abbrev.	Beginning RD	Ordered Set
Offline	OLS	Negative	K28.5 D21.1 D10.4 D21.1
Meaning: Internal port failure. Transmitter may power down. Perform diagnostics or initialization. Receiver will ignore link errors or link failure. Response: LR			

Table 1.9: Primitive Sequences *(Continued on page 19)*

Primitive Sequences			
Primitive Sequence	Abbrev.	Beginning RD	Ordered Set
Not_Operational	NOS	Negative	K28.5 D21.2 D31.5 D5.2
Meaning: Link failure Response: OLS			
Link_Reset	LR	Negative	K28.5 D9.2 D31.5 D9.2
Meaning: Remove Class 1 Connections. Reset F_Port, or OLS Recognized Response: LRR			
Link_Reset_Response	LRR	Negative	K28.5 D21.1 D31.5 D9.2
Meaning: Link Reset Recognized Response: Idle			
Loop Initialization - F7, F7 no valid AL_PA	LIP(F7,F7)	Negative	K28.5 D21.0 D23.7 D23.7
Loop Initialization - F8,F7 loop failure, no valid AL_PA	LIP(F8,F7)	Negative	K28.5 D21.0 D24.7 D23.7
Loop Initialization - F7,x valid AL_PA	LIP(F7,x)	Negative	K28.5 D21.0 D23.7 AL_PS
Loop Initialization - F8,x loop failure, valid AL_PA	LIP(F8,x)	Negative	K28.5 D21.0 D24.7 AL_PS
Loop Initialization - Reset	LIPyx	Negative	K28.5 D21.0 AL_PD AL_PS
Loop Initialization - Reset All	LIPfx	Negative	K28.5 D21.0 D31.7 AL_PS
Loop Initialization - Reserved	LIPba	Negative	K28.5 D21.0 b a
Loop Port Enable	LPEyx	Negative	K28.5 D5.0 AL_PD AL_PS
Loop Port Enable All	LEPfx	Negative	K28.5 D5.0 D31.7 AL_PS
Loop Port Bypass	LPByx	Negative	K28.5 D9.0 AL_PD AL_PS
Loop Port Bypass	LPBfx	Negative	K28.5 D9.0 D31.7 AL_PS

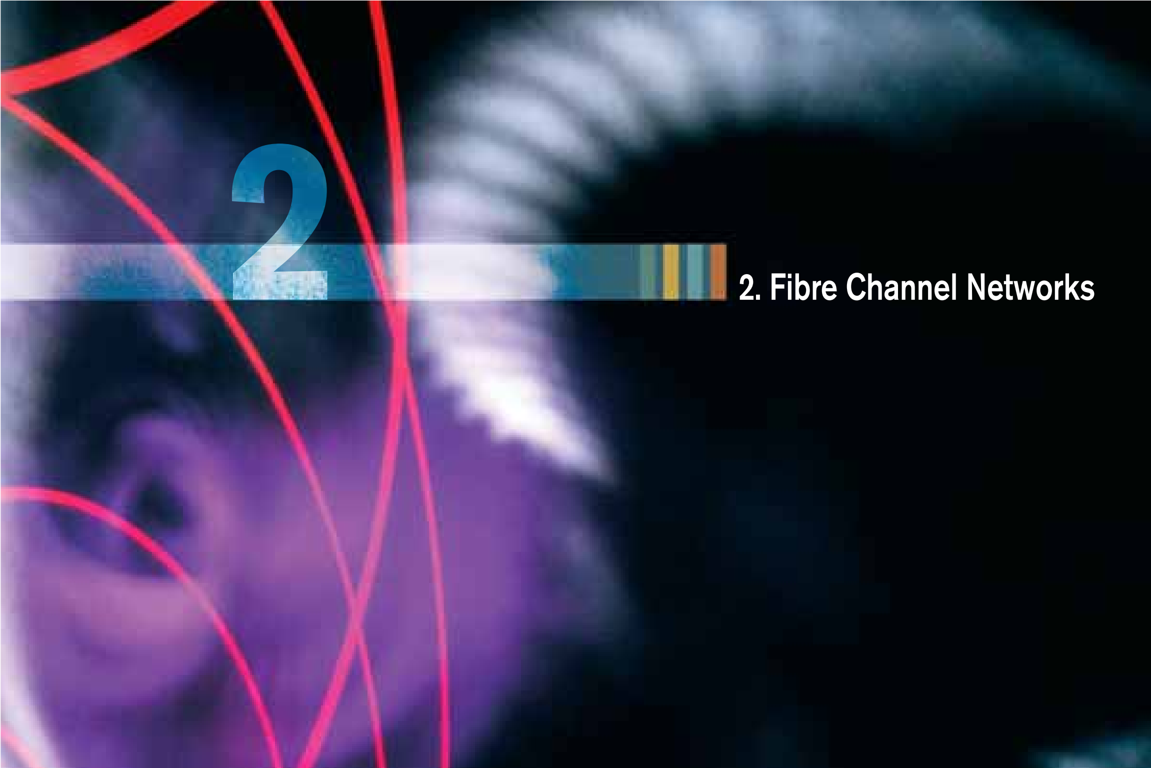
Table 1.9: Primitive Sequences (Continued)

1.5 Classes of Service

There are five classes of service defined for Fibre Channel.

- Class 1 is a dedicated channel between two connection devices. In this configuration, if a host and a device are connected, no other host can use that connection. The advantage of using service class 1 is speed and reliability.
- Class 2 is known as a "connectionless" service. It is a frame-switched link that guarantees delivery of packets from device to device and packet receipt acknowledgements.
- Class 3 is called unacknowledged connectionless service and is good for broadcasts. This configuration allows multiple transmissions to be sent across the Fibre Channel fabric to multiple devices.
- Class 4 is called "intermix", which creates a dedicated connection but also allows class 2 traffic to access the link. This method is very efficient and it allows for greater bandwidth because more than one connection can access the system at any time.
- Class 6 is dedicated to multicast. It differs from class 3 in that full channel bandwidth is guaranteed and the destination ports generate responses, which are aggregated in a single frame to the source port.

Class 3 is the most common class used. Data reliability is left up to the higher-level protocol mapped on the Fibre Channel protocol.



2. Fibre Channel Networks

2. Fibre Channel Networks

2.1 Topologies

Fibre Channel allows each application to use the topology that is most appropriate for its requirements by providing three different interconnection topologies:

- Point-to-point
- Arbitrated loop
- Switched fabric

The **point-to-point connection** (figure 2.1) is the least complex topology and is used when there are only two nodes and no expectation of expansion. The devices can use the link's total bandwidth, because the media is not shared.

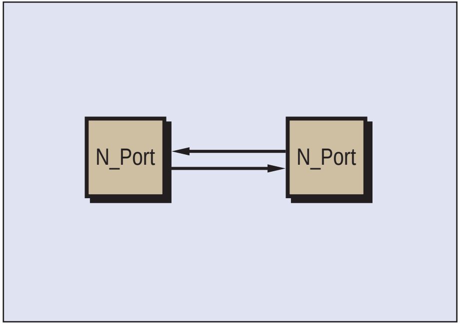


Figure 2.1: Point-to-Point

Fibre Channel Arbitrated Loop (FC-AL) (figure 2.2) is better for storage applications. It is a loop of up to 126 nodes (NL_Ports) that is managed as a shared bus. In this topology, a single connection is established between a sender and a receiver using arbitration protocol, and a data frame is transferred around the loop. When the communication between the two connected ports ends, the loop becomes available for arbitration and a new connection may be established. It is possible to configure loops with hubs to simplify connection management. The Fibre Channel standard supports a distance of up to 10 km for either of the above configurations, although loop size affects latency on the arbitrated loop configuration.

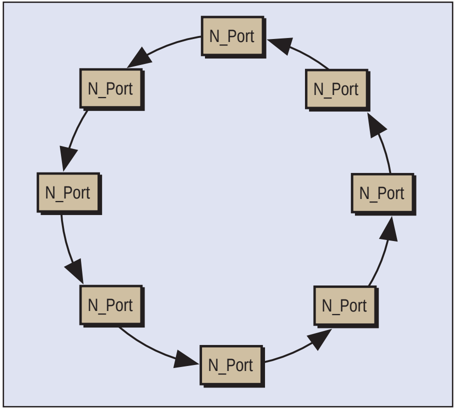


Figure 2.2: Loop

Fibre Channel Switched Fabric (FC-SW) is the third topology used in SAN implementations. It applies to directors that support the FC-SW standard; in other words, it is not limited to switches. A Fibre Channel fabric (see figure 2.3) consists of one or more fabric switches in a single, sometimes extended, configuration.

Unlike arbitrated loop implementations, with their shared bandwidth, switched fabrics provide full bandwidth per port. Adding a new device to an arbitrated loop further divides the shared bandwidth; in a switched fabric, in contrast, the addition of a new device or a new connection between existing devices increases the bandwidth: based on 2 Gb/s technology, an eight-port switch with three initiators and three targets can support three simultaneous 200 MB/s conversations, for a total of 600 MB/s throughput (1200 MB/s if full-duplex applications were available).

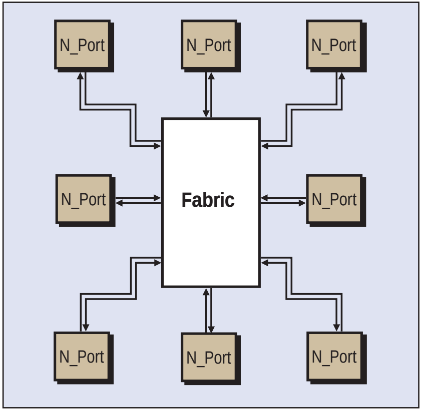


Figure 2.3: Fabric

2.1.1 Transport networks

One of the easiest ways to increase the strategic value of a SAN environment is to extend it across longer distances for disaster recovery and business continuance operations (see figure 2.4). In order to do so for distances longer than 10 km (the native Fibre Channel 10-km connectivity distance), a transport network running a protocol other than Fibre Channel must be used.

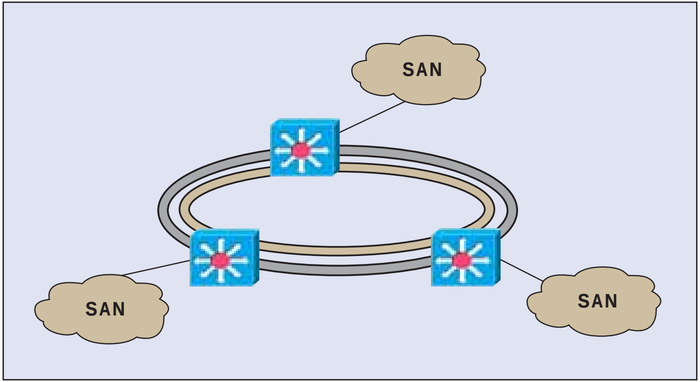


Figure 2.4: SAN Islands Connected through a Transport Network

2.1.2 Mapping

For each of the different transport mechanisms, particular mapping methods are used. Table 2.1 below lists the currently specified Fibre Channel mappings (the most deployed are Fibre Channel over SONET/SDH and Fibre Channel over DWDM networks).

Fibre Channel over DWDM/CWDM	DWDM/CWDM is a bit-rate-independent and protocol-independent transport system; therefore, there is no mapping mechanism. The Fibre Channel signal is simply carried on one of the wavelengths. This is done by a WDM mux/demux.
Fibre Channel over Dark Fiber	The mapping is similar to DWDM, since dark fiber is bit-rate-independent and protocol-independent.
Fibre Channel over SONET/SDH	Carrying Fibre Channel over SONET/SDH requires a framing mechanism that can adapt the FC signal to the SONET/SDH transport. This mechanism is Transparent GFP (GFP-T).
Fibre Channel over ATM	Carrying Fibre Channel over ATM requires encapsulating FC frames onto ATM frames.
Fibre Channel over IP	Fibre Channel over IP consists of encapsulating Fibre Channel frames in IP datagrams and forwarding to the destination. Class of service 1 is not supported, nor are primitive sequences.

Table 2.1: Fibre Channel Mappings

2.2 Equipment

Fibre Channel uses several specific types of equipment to create the different SAN islands, or topologies.

2.2.1 Interconnect devices

Fiber Channel interconnect devices consist of the following:

- hubs
- switching hubs
- switches
- directors

2.2.2 Translation devices

Translation devices connect the Fibre Channel network to outside networks or devices.

- An **HBA** or **Fibre Channel Adapter** is the most common translation device. It connects the Fibre Channel network to the server's host bus, which can be PCI or SBus.
- A **bridge** connects legacy SCSI or ESCON storage devices to the Fibre Channel network.
- An **adapter** connects FC to IP networks such as Ethernet or Token Ring.
- A **gateway** (also known as a router or director) interfaces with telecom networks, such as ATM or SONET.
- **Multifunction routers** connect multiple Fibre Channel ports to multiple protocols such as SCSI, ATM, or Ethernet.

2.2.3 Storage devices

Storage takes many forms and comes in a variety of configurations. The following are some storage devices and their respective capacities:

Storage devices:	Capacity measured in:
disk drives, BODs, and RAIDs	gigabytes
storage subsystems, tape drives	terabytes
tape libraries	petabytes

Table 2.2: Storage Device vs. Capacity

2.2.4 Servers

The **server** is the interface to IP networks, and the initiator in the Fibre Channel SAN. It interacts with the Fibre Channel fabric through the HBA. Microprocessors eliminate single points of failure utilizing single I/O buses or multiple host I/O buses. It is possible for servers to use multiple microprocessors and host I/O buses to put several instances of the operating system to use at the same time.

2.2.5 Port types

Fibre Channel specifies different types of ports, which depend on the type of topologies they can support and on the type of equipment they can connect to. The ports are listed in table 2.3.

E_Port	Expansion port A port is designated an E_Port when it is used as an inter switch expansion port to connect to the E_Port of another switch and enlarge the switch fabric.
F_Port	Fabric port that is not loop capable It is used to connect an N_Port point-to-point to a switch.
FL_Port	Loop-capable fabric port It is used to connect NL_Ports to the switch in a public loop configuration.
G_Port	Generic port that can operate as either an E_Port or an F_Port A port is defined as a G_Port after it is connected but has not received response to loop initialization or has not yet completed the link initialization procedure with the adjacent Fibre Channel device.
L_Port	Loop-capable node or switch port.
U_Port	Universal port; a more generic switch port than a G_Port It can operate as an E_Port, F_Port, or FL_Port. A port is defined as a U_Port when it is not connected or has not yet assumed a specific function in the fabric.
N_Port	Node port that is not loop capable It is used to connect an equipment port to the fabric.
NL_Port	Loop-capable node port It is used to connect an equipment port to the fabric in a loop configuration through an L_Port or FL_Port.
T_Port	Used previously as a mechanism for connecting directors together It has been largely replaced by the E_Port.

Table 2.3: Port Types

2.3 Addressing

In the Fibre Channel environment, each element has its own 64-bit address, known as a World Wide Node (WWN) name. However, if two WWN addresses were to be put into the same frame header, routing performance would be greatly affected, as this is too much data to process in the remaining 128 bits (16 bytes).

Therefore, a shorter identification scheme (24 bits), called port addressing, was implemented to produce a smaller frame header, and thus speed up the routing process. Thanks to this change, the Fibre Channel fabric is optimized for high-speed switching of frames, allowing up to 16 million addresses!

A 24-bit port address contains three different identifiers:

- Domain (bits 23 to 16): The domain identifies the switch itself, making it the most important data byte. One byte can include up to 239 available addresses (out of 256), so if a SAN environment has multiple interconnected switches, the domain number allows each switch to have a unique identifier.

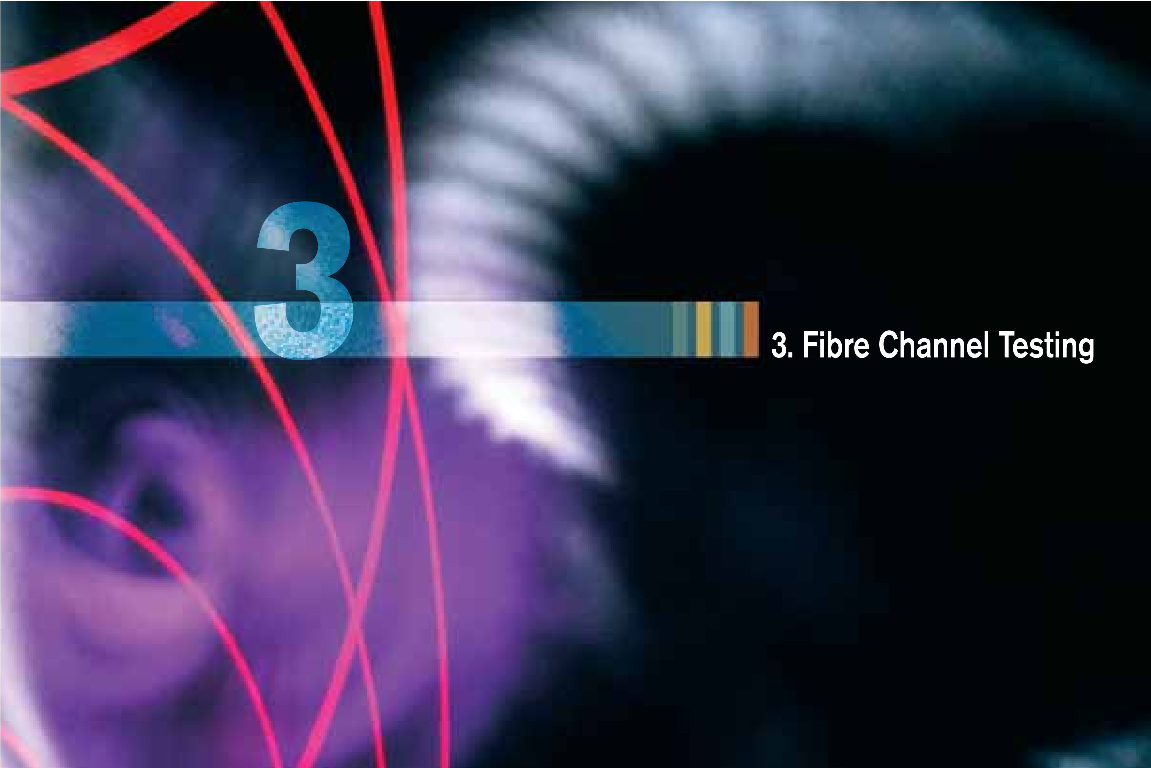
- Area (bits 15 to 08): The area number identifies the fabric ports. This field can assign addresses to up to 256 individual FL_Port's supporting loops, or group of F_Ports—for example, a card with multiple ports would be assigned a different area number for each group of ports.

- Port or arbitrated loop physical address AL_PA (bits 07 to 00): This last part identifies attached Fibre Channel ports (N_Ports and NL_Ports); again, up to 256 addresses can be assigned.

Under this addressing scheme, the topology itself assigns a unique 24-bit address to each port in the switched fabric (contrary to WWNs, which are assigned to manufacturers by the IEEE standards committee and are built into the device).

Although port addressing requires human intervention to change the scheme (from WWN), it is the switch itself that assigns and maintains the port addresses. When the WWN is first used to log onto a switch through a specific port, the switch assigns a new 24-bit address to that port; the switch also keeps track of the relation between the port address and the original WWN address of the device connected to that port.

Activated through an internal name server, the switch's naming function not only allows for dynamic port addressing that eliminates human error in address maintenance, but also allows more versatility in terms of add-ons, transfers, and any modifications made to the SAN.



3. Fibre Channel Testing

3. Fibre Channel Testing

Due to its stringent performance requirements, Fibre Channel requires extensive testing during deployment in order to assure the desired service level. The following sub-sections are the principal testing test scenarios that should be undertaken with a SAN test module (see www.exfo.com for product information).

3.1 Transport Networks

When troubleshooting or commissioning a network, the first test that needs to be performed is the validation of the network's transport path. This test can either be done locally with a remote loopback or using a test set at the remote location.

The test consists of sending a stream of traffic in order to measure the line quality and performance, as defined in the test below (see figure 3.1).

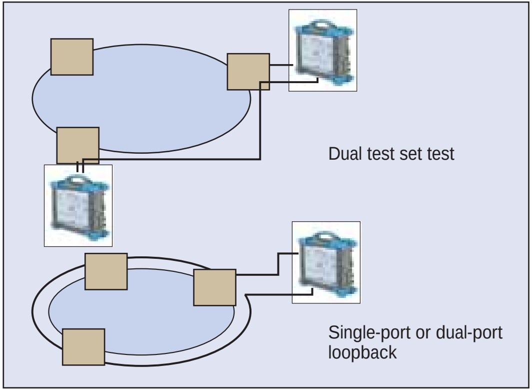


Figure 3.1: Test Topologies

Purpose:

- Initial evaluation of the transport network line quality

Test topologies:

- Single-port loopback
- Dual-port loopback
- Dual test set

Framing:

- FC-2, FC-1 and FC-0 (with and without SYNC)

Recommended tests:

- Transmit PRBS streams with different rate and frame size
- Analyze statistics and performance measurements

3.2 Login Testing

Doing transparent network testing in Fibre Channel has some limitations, since the test module cannot go through a fabric switch and must be connected directly to the Fibre Channel transport equipment (xWDM or SONET/SDH mux). With fabric switch login ability, some testers can connect to a remote location that can be accessed through a fabric network.

The login process will not only permit the unit to connect through a fabric, but it will also exchange some of the basic port characteristics (such as buffer-to-buffer credit and class of service) in order to efficiently transport the traffic through the network. Figures 3.2 and 3.3 depict login testing scenarios, the former in a SAN and the latter for a SAN-extended end-to-end network.

Login purpose:

- Connect to a fabric network in order to reach a remote device
- Exchange basic port characteristics

Topology supported:

- Dual test set

Expected results:

- Automatic detection of Port/Fabric login
- Login statuses (Successful Login, In progress, Failure and Logout)
- Remote buffer-to-buffer advertised credit

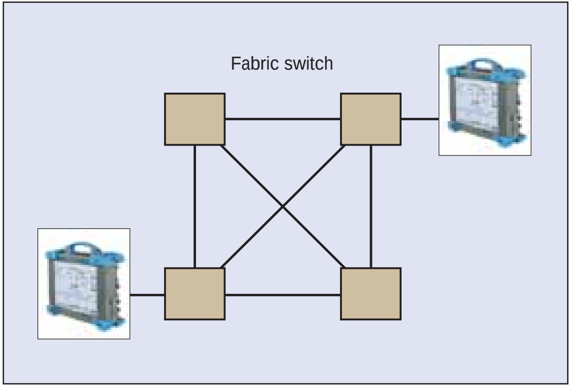


Figure 3.2: Login Testing with a Fabric

3.3 BER Testing

Bit-error-rate testing (also known as BERT) provides an objective measurement of the physical line performance by simulating the full range of frequency responses of the physical components of the network. It consists of sending a stream of frames with pseudo-random patterns (known as PRBS). The receiving side has the ability to decode the PRBS patterns and actually detect if any bits were inverted while being transported. The bit error rate is calculated by dividing the number of bits in error by the total number of bits received.

Some testers have the ability to generate a set of pre-defined PRBS values. Lower PRBS sequences can be used with lower line rates, and higher PRBS sequences are recommended for higher line rates. Units that can also generate specific payload patterns (xPAT) stimulate the physical line in order to detect possible jitter conditions.

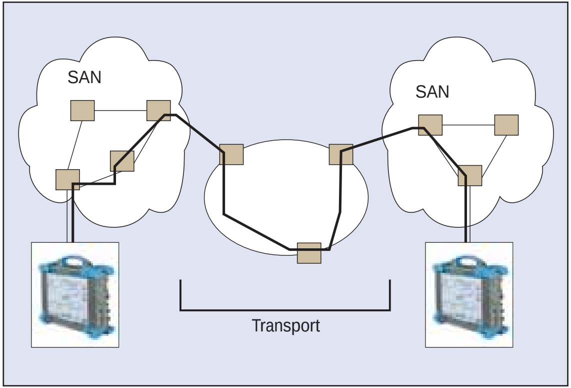


Figure 3.3: End-to-End Login Testing

Purpose:

- Validate the quality of the line (Fibre Channel BER should be < 10^{-9} or one bit error per billion bits)
- Validate the jitter response of the network with xPAT
- Observe the network response when too many errors are generated

Topology supported:

- Single-port loopback
- Dual-port loopback
- Dual test set (with remote or local port)

Patterns supported:

- PRBS from 10^{9-1} to 10^{31-1}
- CSPAT, CRPAT and CJTPAT (supported only with FC-1 frames)
- User-defined

Validation through error injection:

- Bit-error manual injection
- Bit-error automatic injection (rate or continuous)
- Pattern loss alarm generation

Related statistics:

- Bit-error count and rate
- Bit-mismatch count and rate
- Pattern loss alarm (when too many bit errors are detected)

3.4 Latency Measurements

Transmission of frames in a network is not instantaneous and is subject to multiple delays caused by the propagation delay in the fibre and by the processing time inside each piece of network equipment (such as storing the complete frame before forwarding it to the appropriate direction).

The total accumulation of delays between two end points is called latency. Some applications, such as real-time applications (VoIP or video conferencing) and storage area networks, are very sensitive to excess latency.

Some testers have the ability to measure the average latency between two end points.

Purpose:

- Measure the average latency between two end points

Topology supported:

- Single-port loopback (unidirectional latency)
- Dual-port loopback (unidirectional and round-trip latency)
- Dual test set (round-trip latency)

Framing mode:

- FC-1 or FC-2

Related latency statistics:

- Minimum
- Maximum
- Last
- Average
- Number of samples

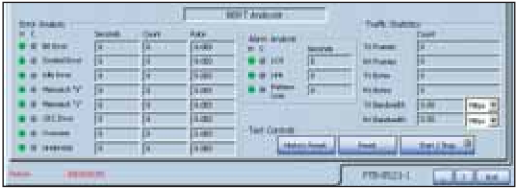


Figure 3.4: Sample GUI (BERT Analyzer)

Expected results:

- In SAN applications, the results are based on the network topology
- In real-time applications, round-trip latency should be below 250-300 ms

3.5 Buffer-to-Buffer Credit Estimation

Fibre Channel has a specific characteristic that is based on network topology (including distances) and on congestion in the pathway, enabling the regulation of traffic flow (see Flow Control section). This basically consists of transmitting an acknowledgement each time a frame is received. Each port accumulates the number of frames transmitted and expects an acknowledgement for each one.

With this mechanism, each port can specify the threshold of frames that can be transmitted without receiving a single acknowledgement. This threshold is known as the buffer-to-buffer credit.

A lower number of credits would not use the network at its full capacity because the transmitter would need to wait for acknowledgement even if the network bandwidth is under-utilized. A higher number of credits could create congestion somewhere in the network, increasing latency, and possibly discarding frames.

Purpose:

- Based on round-trip latency measurement, it provides an estimate of buffer-to-buffer credit for the flow-control mechanism.

Topology supported:

- Dual-port loopback
- Dual test set

Framing mode:

- FC-1 and FC-2

Related statistic:

- Estimate of buffer-to-buffer credit

How to use the estimate:

- The estimate value is used in the login process in order to indicate to the remote port, the amount of buffer-to-buffer credit you will be using. This way, the remote port will reserve sufficient memory to support this value.
- In an implicit login (manual login), the amount of buffer-to-buffer credit of the remote port must be entered manually at the local port.

3.6 Reported Alarms and Errors

A network is based on various equipment connected to each other through optical fiber, electrical wiring or radio waves. These items are not 100% guaranteed against failure or malfunctions. Overheating, electrical or electromagnetic interference, system errors, or any other misbehaviors of equipment, while not necessarily crashing the network, can still create errors in the stream of traffic.

An alarm is usually triggered when a fault is detected and added to the time or to the total number of faults detected. For example, a pattern loss alarm will be triggered if three consecutive errored words are detected.

Fibre Channel testers have the ability to detect numerous faults in the stream of traffic that are reported with three different metrics: count of errors, error in seconds and rate of error.

Purpose:

- Validate the quality of the line

Topology:

- Single-port loopback
- Dual-port loopback
- Dual test set

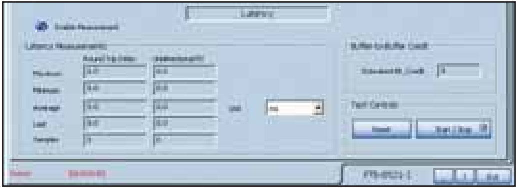


Figure 3.5: Sample GUI (Latency)

Framing mode:

- FC-2 and FC-1 provide pattern and frame errors
- FC-0 provides only pattern errors

Related statistics:

- Bit errors
- Mismatch 0/1 errors
- Symbol errors
- FCS errors
- Undersize and oversize errors

- Loss-of-signal alarm
- Loss-of-link alarm
- Pattern-loss alarm

Metrics:

- The count error provides the total number of errors detected throughout the test.
- The number of seconds in error provides the number of seconds during which at least one error was detected.
- The rate is the number of errors detected divided by the total number of bits, bytes or frames received throughout the test.

Interpretation of the results:

- The count error alone can be misleading. Its significance is provided by the rate metric.
- But even the count and rate alone have some limitations, since the distribution of the errors in time is unknown. If the errors are detected evenly in time, the count of seconds will be relatively high; if the errors are detected in bursts, the count of seconds in errors will be very small and could be attributed to faulty system behavior.

3.7 Reported Performance Statistics

Another way of monitoring the quality of the network is through performance measurements. Not only do they provide the throughput (or utilization) of the path, but they can also provide vital information on the quality of the line.

That is why SAN testers must follow ITU-T G.821 and G.826 performance measurement standards. Although these standards have been created for use with telecom traffic, they have also been recommanded for Fibre Channel traffic.

Purpose:

- Provide a fixed and well-known set of performance measurements

Topology:

- Single-port loopback
- Dual-port loopback
- Dual test set

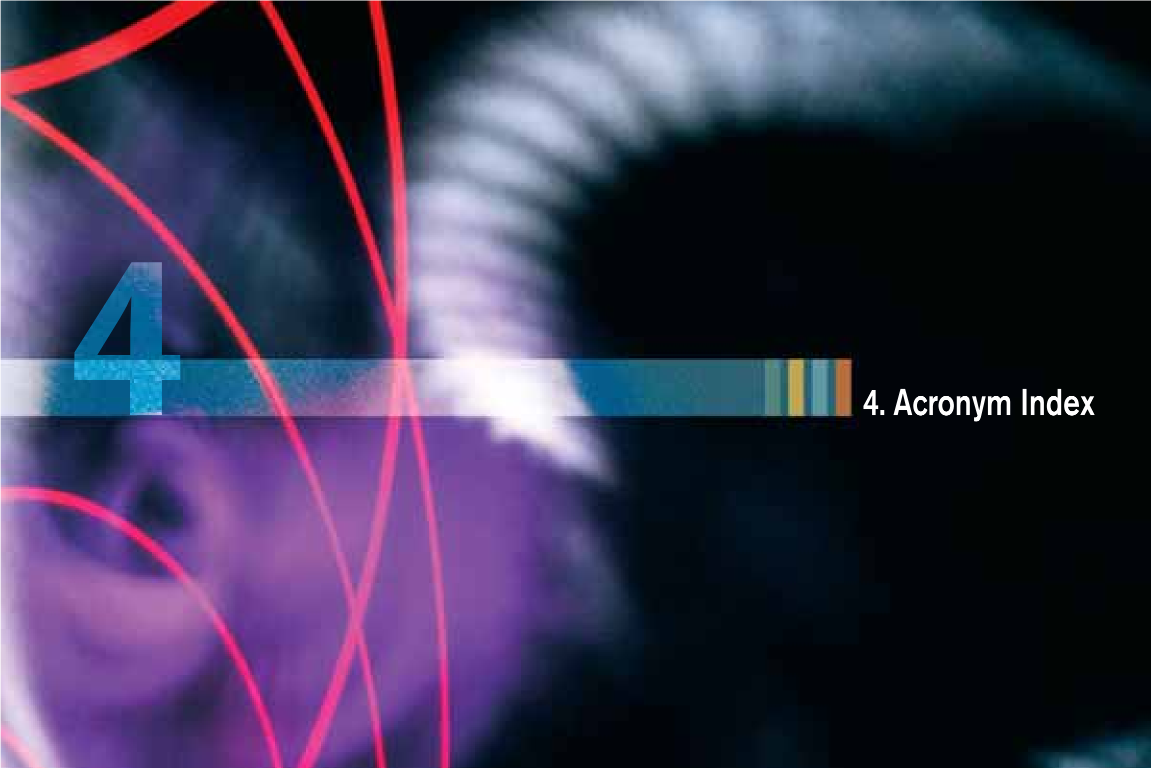
Framing mode:

- FC-2 and FC-1 only

Figure 3.6: Sample GUI (Performance Monitoring)

Related Statistics:

- Error Frame Second (EFS): number of seconds where at least one frame had an error
- Error Count (EC): total number of errors
- Error Second (ES): number of seconds where at least one error was detected
- Severely Errored Second (SES): number of seconds where multiple errors were detected
- Alarm Second (AS): number of seconds where at least one alarm was detected
- Unavailable Second (UAS): number of seconds where the path was unavailable
- Error Second Rate (ESR): the number of seconds in error divided by the total duration of the test (in seconds)
- Severely Errored Second Rate (SESR): the number of severely errored seconds divided by the total duration of the test (in seconds).



4. Acronym Index

A

ABTS . . . . Abort Sequence

ACC . . . . Accept

ACK . . . . Acknowledgement

ACT . . . . Activity

AL_PA . . . . Arbitrated-Loop Physical Address

ANSI . . . . American National Standards Institute

ARB . . . . Arbitrate

AS . . . . Available Second

ATM . . . . Asynchronous Transfer Mode

B

BB_Credit . . . . Buffer-to-Buffer Credit

BB_Credit_CNT . . . . Buffer-to-Buffer Credit Count

BBE . . . . Background Block Error

BBER . . . . Background Block Error Ratio

BER . . . . Bit Error Rate

BERT . . . . Bit-Error-Rate Test

BOD . . . . Block-Oriented Device

C

CJTPAT .... Compliant Jitter Tolerance Pattern

CLS .... Close

CRPAT .... Compliant Random Pattern

CRC .... Cyclic Redundancy Check

CS_CTL .... Class-Specific Control

CSPAT .... Compliant SSO Pattern

CTL .... Control

CWDM .... Coarse Wavelength-Division Multiplexing

D

D_ID .... Destination Identifier

DF_CTL .... Data Field Control

DHD .... Dynamic Half-Duplex

DWDM .... Dense Wavelength-Division Multiplexing

E

E_Port .... Expansion Port

EB .... Errored Block

EC .... Error Count

EE_Credit . . . . .End-to-End Credit

EE_Credit_CNT . . . .End-to-End Credit Count

EFS . . . . .Error-Free Second

EIA . . . . .Electronic Industries Alliance

EOF . . . . .End of Frame

ES . . . . .Errored Second

ESCON . . . . .Enterprise System Connection

ESR . . . . .Errored Second Ratio

F

F_CTL . . . . .Frame Control

F_Port . . . . .Fabric Port

FC . . . . .Fibre Channel

FC-0 . . . . .Fibre Channel, Layer 0

FC-1 . . . . .Fibre Channel, Layer 1

FC-2 . . . . .Fibre Channel, Layer 2

FC-AL . . . . .Fibre Channel Arbitrated Loop

FC-PH . . . . .Fibre Channel Physical

FC-SW . . . . .Fibre Channel Switched Fabric

FCS . . . . .Frame Check Sequence

FDDI . . . . .Fiber Distributed Data Interface

FICON . . . . .Fibre Connection

FL_Port . . . . .Loop-Capable Fabric Port

FLOGI . . . . .Fabric Login

FPGA . . . . .Field-Programmable Gate Array

f/s . . . . .Frame per Second

G

G_Port . . . . .Generic Port

Gb/s . . . . .Gigabit per Second

GFP . . . . .Generic Framing Procedure

H

HBA . . . . .Host Bus Adapter

HiPPI . . . . .High-Performance Parallel Interface

I

IBM . . . . .International Business Machines Corporation

ID . . . . .Identifier

IEEE . . . . .Institute of Electrical and Electronic Engineers
IFG . . . . .Inter-Frame Gap
IP . . . . .Internet Protocol
iSCSI . . . . .SCSI Protocol over TCP/IP
ISM . . . . .In-Service Monitoring
ISO . . . . .International Organization for Standardization

L

L_Port . . . . .Loop Port
LFn . . . . .Link Failure n
LIP . . . . .Loop Initialization
LOS . . . . .Loss of Signal
LPB . . . . .Loop-Port Bypass
LPE . . . . .Loop-Port Enable
LR . . . . .Link Reset
LRR . . . . .Link Reset Response
LRn . . . . .Link Recovery n

M

Mb . . . . .Megabit

MB . . . . .Megabyte

MBaud . . . . .Megabaud

Mb/s . . . . .Megabit per Second

MB/s . . . . .Megabyte per Second

MRKtx . . . . .Mark

N

N_Port . . . . .Node or Fibre Channel Port

NL_Port . . . . .Loop-Capable Node or Fibre Channel Port

NCITS . . . . .National Committee for Information Technology Standards

NE . . . . .Network Element

NOS . . . . .Not Operational

O

OFC . . . . .Open Fibre Control

OLn . . . . .Offline n

OLS . . . . .Offline

OSI . . . . .Open System Interconnection

OOSM ....Out-of-Service Monitoring

OPNyx ....Open Full-Duplex

OPNyy ....Open Half-Duplex

OPNyr ....Open Broadcast Replicated

OPNfr ....Open Selective Replicated

OX_ID ....Originator Identifier

P

PARAM ....Parameter

PCI ....Peripheral Component Interconnect (personal computer bus)

PHY ....Physical-Layer Device

PLOGI ....N_Port Login

PM ....Performance Monitoring

PRBS ....Pseudo Random Bit Sequence

PSP ....Primitive Sequence Protocol

R

RAID ....Redundant Array of Inexpensive Disks

R_CTL ....Routing Control

R_RDY ....Receiver Ready

RX . . . . .Receive

RX_ID . . . . .Responder Identifier

S

s . . . . .Second

S_ID . . . . .Source Identifier

SAN . . . . .Storage Area Network

SBus . . . . .Serial Bus

SCSI . . . . .Small Computer System Interface

SDH . . . . .Synchronous Digital Hierarchy

SEQ_CNT . . . . .Sequence Count

SEQ_ID . . . . .Sequence Identifier

SES . . . . .Severely Errored Second

SESR . . . . .Severely Errored Second Ratio

SFP . . . . .Small Form Factor Pluggable

SOF . . . . .Start of Frame

SONET . . . . .Synchronous Optical Network

SUI . . . . .Smart User Interface

SYNC . . . . .Synchronization

T

T_Port . . . . .Mechanism for connecting directors

TL_Port . . . . .Private-to-public bridging of switches or directors

TX . . . . .Transmit

TYPE . . . . .Data Structure Type

U

U_Port . . . . .Universal Port

UAS . . . . .Unavailable Second

ULP . . . . .Upper-Layer Protocol

V

VC_RDY . . . . .Virtual Circuit Ready (Class 4)

W

WWN . . . . .Worldwide Name

X

xPAT . . . . .CRPAT, CSPAT and CJTPAT

Notes

Notes

Acknowledgements

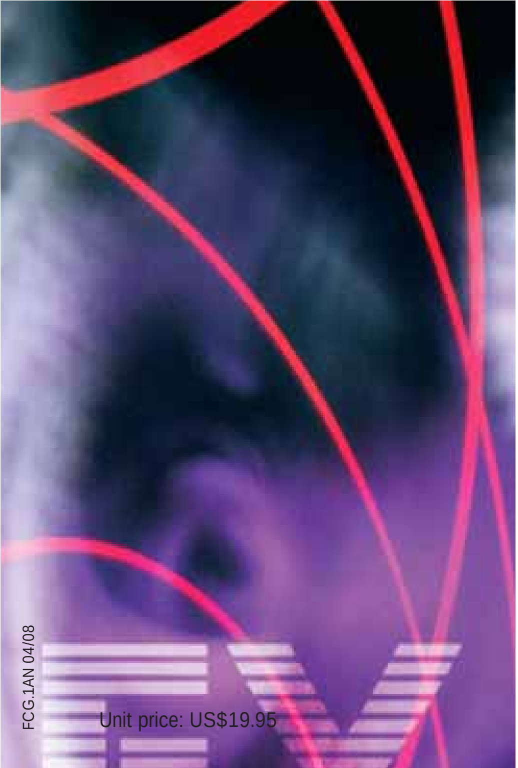
This guide would not have been possible without the enthusiasm and teamwork of EXFO staff, particularly the hard work and technical expertise of Mr. Raffi Tchakmakjian, Product Manager.

Printed and bound in Canada

ISBN 1-55342-003-9

Legal Deposit—National Library of Canada 2004
Legal Deposit—National Library of Quebec 2004

Fibre Channel Guide, 1st edition



For details on any of our products and services, or to download technical and application notes, visit our website at www.exfo.com.

EXFO

EXPERTISE REACHING OUT