Fibre Channel Functional Overview

Prior chapters have so far been dedicated to the fundamentals of the SCSI protocol and have placed much emphasis on the layered approach to distributed communications illustrated in Figure 1-1 (p. 1).

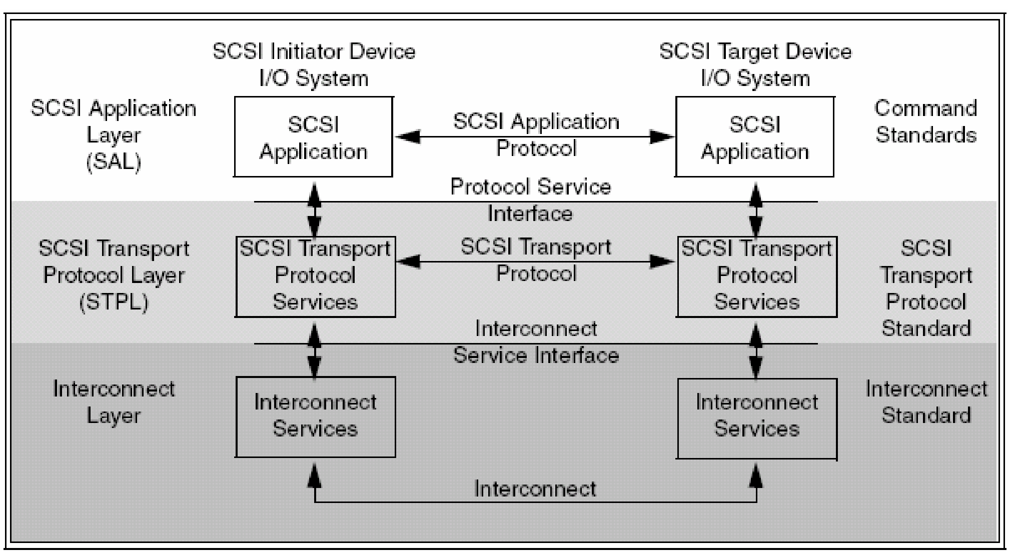

Figure 1-1: Layered SCSI Communication Model.

While the SCSI Application Layer (SAL) and the SCSI Transport Protocol Layer (STPL) are inherently part of the SCSI specification, the Interconnect Layer can be implemented by a variety of interconnect methods such as the SCSI Parallel Interface (SPI), Fibre Channel, InfiniBand or TCP/IP, to name the most prevalent. When Fibre Channel is used as an interconnect method for SCSI, the relationship between both protocol stacks is shown in Figure 1-2 (p. 2).

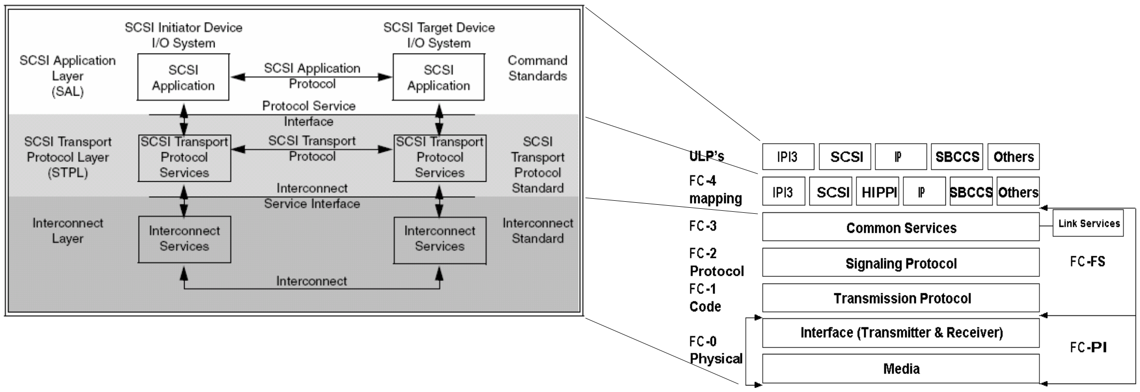

Figure 1-2: Relationship Between SCSI and FC Stacks.

Figure 3-2 reveals that Fibre Channel boasts a layered structure of its own in which various protocol functions are segregated into discrete levels. This layering is central to the Fibre Channel architecture and forms the basis for this chapter.

The previous chapter also introduces protocol mapping by describing how SCSI PDUs are grouped into IUs fit for transmission by using Fibre Channel sequence and exchange logical constructs. These constructs, along with the fundamental structure and capabilities of the Fibre Channel communications protocol, are presented in this chapter while highlighting key points which make Fibre Channel a choice interconnect for SCSI applications. Accordingly, this chapter also brings to a conclusion the protocol mapping introduction offered in **Error! Reference source not found.**: *Error! Reference source not found.* (p. **Error! Bookmark not defined.**).

1.1 Fibre Channel Concepts

While mostly used for SAN applications, Fibre Channel remains a stand-alone, networked communication protocol in its own right, and may serve purposes other than storage networking: FC-AE (Avionics Environment) is a T11 document which provides implementation guidance for ULPs used in airborne applications such as MIL-STD-1553, a command/response protocol designed for military use as a primary interface between aircraft and missiles. Remote Direct Memory Access (RDMA) and Virtual Interface (VI) are yet more examples of potential Fibre Channel applications.

> *Note: The spelling of fibre in "Fibre Channel" is used to emphasize the concept of a fabric resulting from interwoven threads. By comparison, the term "fiber optics" refers directly to the physical structure of a fiber.*

Fundamentally, Fibre Channel allows two or more *nodes* to communicate by sending information units (IUs) to each other. This is accomplished by fragmenting the IUs into *frames* which are then sent through a networked infrastructure. The main purpose of the infrastructure is to accept frames from transmitting nodes and deliver them to the intended recipient nodes. Nodes connect to the network via one or more Fibre Channel *port(s)* which implement the logical and physical interface to a *link*. This link represents the actual media and associated transceiver logic (collectively called the Link Control Facility, LCF) required to reliably transmit digital information at high data rates (typically 1 or 2 Gbps). Figure 1-3 (p. 3) leverages a previously introduced topology to illustrate the components of this high-level summary of Fibre Channel communications.

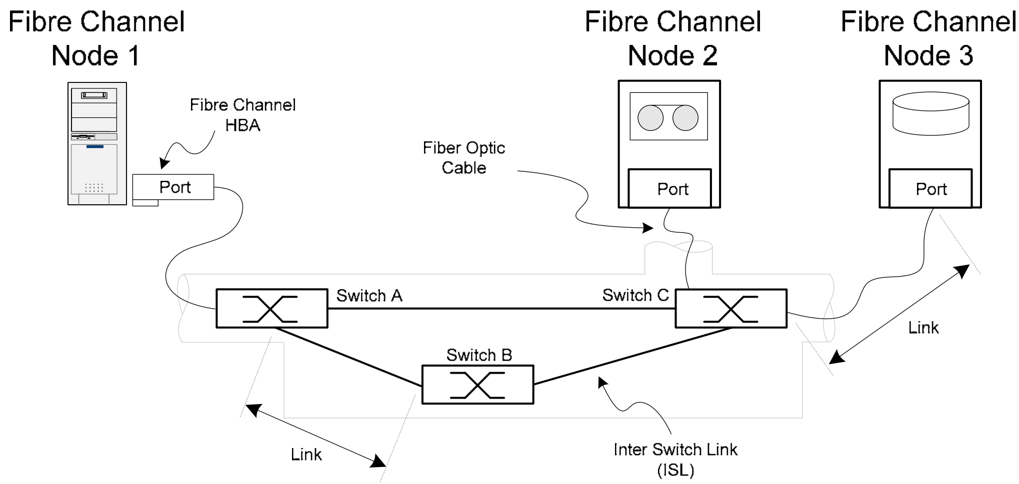

Figure 1-3: Fibre Channel Concepts

In the above scenario, Fibre Channel nodes are represented by a computer host and two storage devices. While storage devices usually boast built-in Fibre Channel ports, host computers require the addition of one or more *Host Bus Adapter(s)* to provide the required Fibre Channel port functionality. In this example, the ports interface to fiber optic media and provide the appropriate lasers to drive the optical signals across the link. While they are omitted for clarity in the figure above, Fibre Channel ports also exist on the switches which implement the *fabric* portion of this Fibre Channel network; this includes ports which connect switches to other switches by means of Inter Switch Links (ISLs). The topic of switch ports is elaborated as required in the following sections.

1.1.1 Fibre Channel Topologies and Port Types

The example depicted by Figure 1-3 shows Fibre Channel nodes connected to a core of three switches linked by a number of ISLs. Together, these interconnected switches implement a storage network through a *switch fabric* topology. Fibre Channel standards use the term *topology* to describe three supported methods for establishing a communication channel over a link between Fibre Channel ports:

- **point to point**: strictly speaking, the point to point topology does not realize a network as it can only be used to connect two nodes together, and is sometimes called back-to-back. As a result, storage applications which use this topology are usually referred as Direct Attached Storage (DAS), which mostly falls outside the scope of this book (point to point will be revisited when migration strategies are explored in Part III). Hence, the current section only covers the point to point topology to the extent that it must be acknowledged. An example of this topology is shown in Figure 1-4 (p. 4);

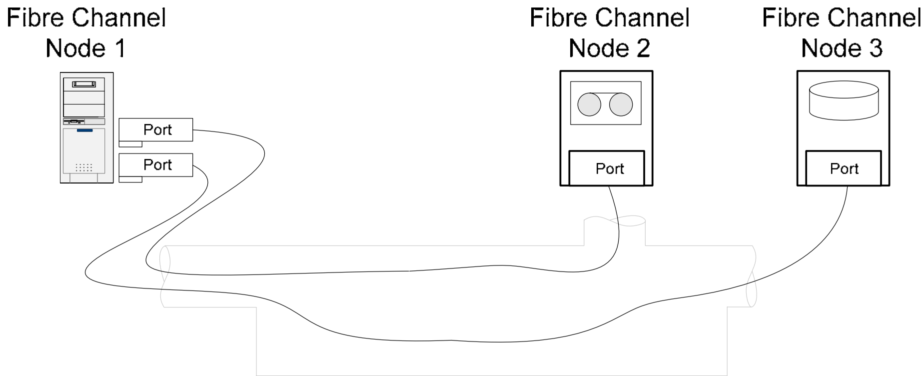

Figure 1-4: Point to Point Topology.

- **arbitrated loop** (private and public); while modern SANs favor switch fabric topologies (explained below), some storage devices may implement arbitrated loop interfaces. This topology will be covered in the context of practical applications but is not the main focus of this text. An example of arbitrated loop topology is shown in Figure 1-5 (p. 4);

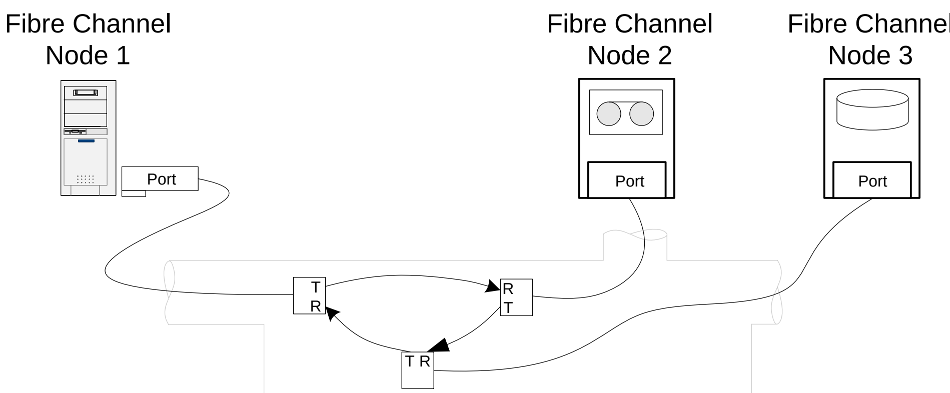

Figure 1-5: Arbitrated Loop Topology

- **switch fabric**: the switch fabric topology is the cornerstone of SAN technology and effectively delivers the "network" in Storage Area Networks. As the main focus of this book, the switch fabric topology is almost exclusively used throughout the following chapters. An example of switch fabric topology is shown in Figure 1-3 (p. 3).

Terminology note: the standards use the term *topology* to describe the attach mode between two ports. *Switch fabric* and *arbitrated loop* are two different topologies, but they can coexist in the same network. Therefore, the term *network* is used to describe a SAN in which multiple topologies are leveraged simultaneously; in the special case where every single SAN devices is N_Port (i.e. no arbitrated loop exists), then the terms *network* and *fabric* are synonymous.

When Fibre Channel ports are connected together, the underlying topology is discovered automatically. This feature highlights the fact that Fibre Channel ports may support one or more topologies. Fibre Channel adopts a standard nomenclature to distinguish the capabilities of a port, and also to indicate the current topology under which a particular port is operating.

For Fibre Channel **nodes**, ports are described as follows:

- **L_Port**: when a node port only supports an arbitrated loop topology, whether private, public or both, it is called an L_Port. L_Ports can connect to other L_Ports, to NL_Ports or to an FL_Port;
- **N_Port**: when a node port only supports a point to point or a switch fabric topology, it is called an N_Port. N_Ports can only connect to another N_Port (point-to-point topology) or to an F_Port (fabric topology).
- **NL_Port**: an NL_Port is an N_Port which also supports a public or private arbitrated loop and is currently operating in a loop topology;
- **Nx_Port**: an Nx_Port is an N_Port which also supports the public (but not private) arbitrated loop topology. When an Nx_Port is connected in a point to point topology or a fabric topology, it adopts the characteristics of an N_Port. When an Nx_Port is connected in a loop topology, it adopts the characteristics of a public NL_Port;

For Fibre Channel **switches**, ports are described as follows:

- **F_Port**: an F_Ports is a switch port which only supports a switch fabric topology, and can only be used to connect to N_Ports;
- **FL_Port**: an FL_Port is an F_Port which also supports an arbitrated loop and is currently operating in a loop topology. It is used to connect to L_Ports or NL_Ports;
- **Fx_Port**: an Fx_Port is a switch port that supports either a switch topology or an arbitrated loop. When connected to an N_Port, the Fx_Port adopts the characteristics of an F_Port. When connected to an L_Port or NL_Port, the Fx_Port adopts the characteristics of an FL_Port.
- **E_Port**: an E_Port is a expansion switch port and is used to connect to another E_Port on another switch, thereby creating an Inter-Switch Link (ISL). An E_Port can only connect to another E_Port (and sometimes a B_Port, if we choose to include it).

should we include B_Port?

Where Fibre Channel port references are made irrespective of type, the term **FC_Port** is used.

Figure 1-6 (p. 6) illustrates how this terminology applies to actual topologies. For the sake of this example, Switch A is assumed to have F_Port capability only, and Switch B is assumed to have FL_Port capability only (except for the two E_Ports shown). As a result, the Nx_Port of Node A can connect to switches A or B by adopting the port mode required by each individual switch. On the other hand, the port on Node B only supports L_Port: this allows it to connect to Switch B, but not to Switch A. Finally, switches A and B are seen to have established an ISL between themselves through the use of E_Ports.

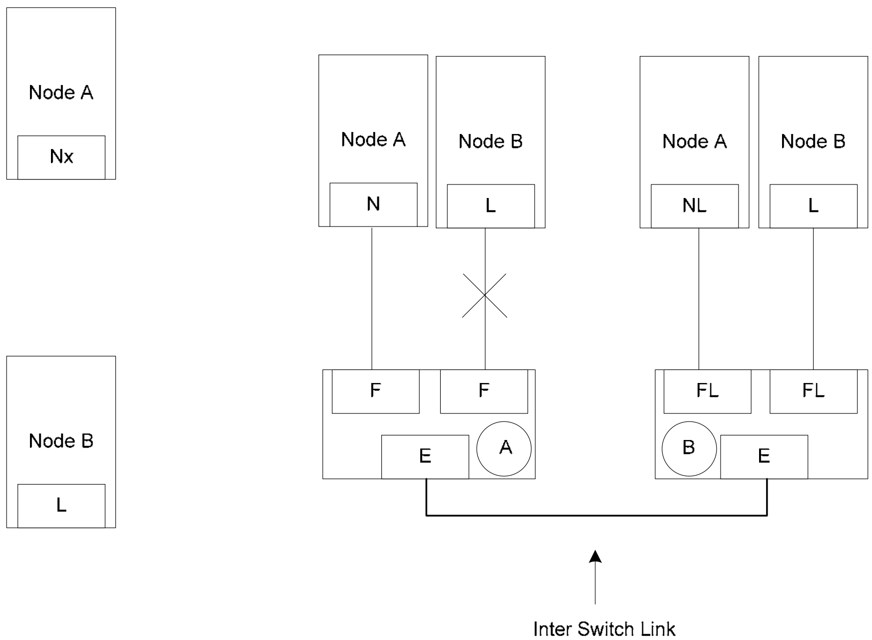

Figure 1-6: Port Types Depend on Topology

This summary of port nomenclature is by no means exhaustive and intends to clarify those types which are most commonly seen in practice. As switch vendors develop features which are not covered by the standards, there is a tendency to coin new port types to describe the associated functionality. An example of such practice can be seen in Cisco MDS 9000 Series of Fibre Channel switches where *trunking* ISL ports are called *TE_Ports*. These additional port types will be introduced in opportune time as they relate to the matter at hand.

1.1.2 Fibre Channel Nodes

Nodes are the communicating entities for which the Fibre Channel network exists. In the process of sending information units (IUs) to their peers, nodes are responsible for fragmenting each IU into a sequence of frames which they send into the network, and for extracting meaningful data from the network by reassembling each frame they receive into sequences, or IUs. In other words, nodes are responsible for managing communications with peer entities. This is illustrated in Figure 1-7 (p. 7) which presents a *node*-centric view of a Fibre Channel network. As a client of this network, Node 1 need only supply the network with a properly formatted frame which includes the address of the intended recipient; it is the responsibility of the network to deliver this frame to its destination, as shown by the dotted arrows. Nodes 1, 2 and 3 neither know nor care about what is inside this “cloud”, so long as frames are routed to the proper destination.

However, nodes must keep track of their conversations with other nodes – this is depicted by the "conversation records" shown within each node; this is where nodes keep track of the various IUs (sequences) which are usually part of an individual operation (exchange). Although sequences and exchanges will be explained shortly, their introduction at this point is meant to position these logical structures with respect to the Fibre Channel frame: sequence and exchange assignment is accomplished inside the host by setting correct values in the frame header, while the Fibre Channel frame proper is the only physical entity that actually travels in a network (i.e. the network processes frames, not sequences or exchanges). Note: fabric switches may implement additional intelligence to enhance fabric performance by making forwarding decisions based on these logical constructs.

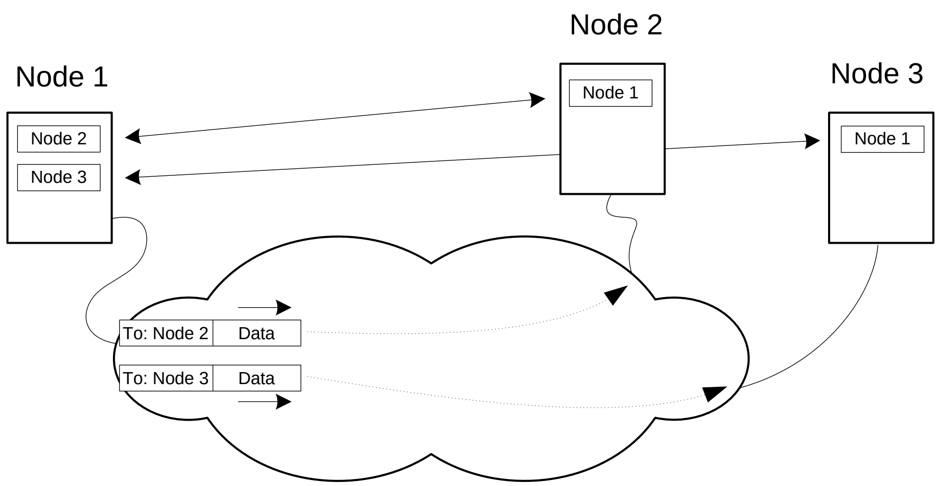

Figure 1-7: Fibre Channel nodes communicate by sending frames through a network.

1.1.3 The Fibre Channel Network

The network accepts Fibre Channel frames sent by transmitting nodes and routes them to the intended recipient based on the Fibre Channel address contained in the frame header. This network usually consists of one or more Fibre Channel switch(es) which are interconnected by Inter Switch Links (ISLs). In contrast to the previous illustration, Figure 1-8 (p. 8).shows a *network*-centric view of the same Fibre Channel fabric which, as revealed, consists of three switches connected by Inter-Switch Links (ISLs). The primary role of the three switches is to admit frames which are injected into the network by attached nodes and route these frames to their final destination by forwarding them to the next switch along the shortest path to the destination. This implies that the switches must know where each possible destination is located in the network, and that link or switch failures be detected so that alternate routes may be used when they exist. Finally, it should be observed that the network treats each frame independently and has no concept of any possible relationship between frames (theoretically – not in practice though). In other words, the network does not know or care about the fact that each individual frame may be carrying a fragment of a conversation between nodes.

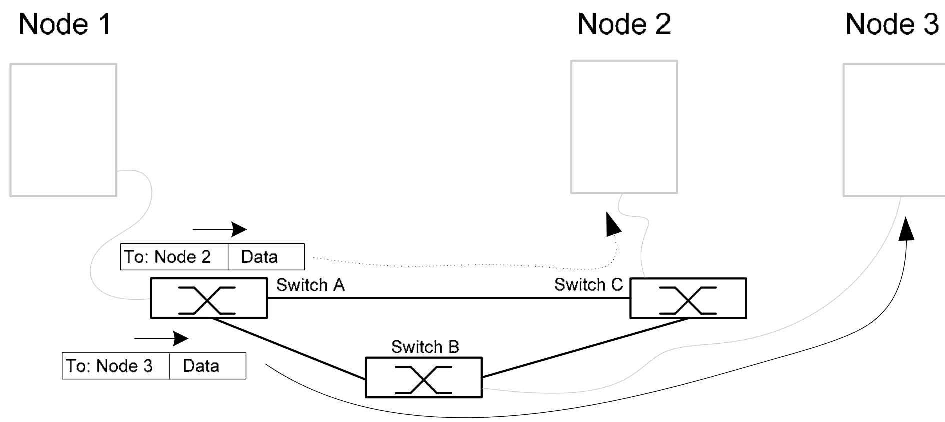

Figure 1-8: The Fibre Channel fabric routes frames to their destination.

The clear distinction between the client-centric and network-centric view of the Fibre Channel network is leveraged in this and upcoming chapters by discussing each topic independently: this chapter focuses on the client-centric view whereas Chapters 4 and 5 discuss network-related concepts. However, many components of the Fibre Channel architecture are common to both node and network components: FC-0, FC-1 and the framing portion of FC-2; these topics mainly relate to the communication model and are presented in the following section. FC-3 will be split in half: the common portion is described here, the node portion is described here, and the network portion is described in Chapter 5.

Need to standardize on where and how we want to introduce the standard text (FC-FS, FCP, etc) to which we refer in the respective chapters.

1.2 Fibre Channel Operations

The Fibre Channel communication protocol has thus far been described in terms of its steady state operation: under normal operating conditions, nodes spend 99% of their time sending and receiving frames, and network switches for the most part avail themselves of forwarding the frames to their respective destination nodes.

However, many preparatory tasks must be performed by the network before it is ready to conduct normal activities. For example, fabric switches must initialize ISLs between themselves and establish a structured addressing scheme along with an accompanying routing protocol. When zoning is configured, switches must merge their respective zone databases and agree on a common active zone set. For their part, end nodes must initialize the ports which connect them to the network in order to establish the connection topology, at which point a loop initialization (LIP, in the case of an arbitrated loop topology) or a fabric login (FLOGI) must be performed. These tasks share the common objective of establishing steady-state operating parameters between the end node and the switch. Exchanged parameters include the assignment of a Fibre Channel hardware address (FC_ID) and buffer-to-buffer credit distribution, to name just a few.

After fabric login, a node must register itself with the Fibre Channel Name Server (FCNS) by providing a record of its FC_ID, pWWN, nWWN, and FC-4 capabilities (FCP or IP, for example). Finally, the node can issue a query to the name server requesting database records of other registered nodes which meet a certain criteria. In the case of a SCSI initiator node, for example, the query might contain a request for the name server to return entries for which the capabilities field contains “scsi-fcp target”. The retrieval of this record (or records, if more than one meets the criteria) provides an initiator node with the information necessary to initiate a conversation with a target node. Specifically, WWNs are used for identification and the FC_ID is used to build the Fibre Channel frames which are ultimately sent in the network.

Once a node has identified other nodes with which it needs to communicate, there are two additional steps which must be completed before data may be allowed to flow. A node must perform a port login (PLOGI) to the Fibre Channel port of the destination node to exchange node-to-node parameters specific to this node pair. Once a port login is established, a node must finally login to the appropriate process (PRLI) with which it ultimately wants to converse. This “process” is the term used to describe the Upper Layer Protocol which will be invoked by the upcoming conversations. In the case of a SCSI node for example, the PRLI is performed by the initiator node against the SCSI-FCP process in the target node. The standards refer to this operation as “establishing an FC-4 image pair”.

While the operations described above are performed very few times in the lifecycle of a Fibre Channel node compared to steady-state data frame transfers, they are essential to the operation of the network and their failure will usually cripple the affected node(s). Operations such as FLOGI, PLOGI and PRLI are complex enough that they warrant a dedicated section and are formally presented in Chapter 4. Similarly, fabric-specific tasks such as ISL configuration, Domain_ID assignment and FSPF are given similar treatment in Chapter 5. The cursory description supplied above should allow the reader to complete this chapter with a minimal amount of pre-requisite information: when examples are used to explain how frames, sequences and exchanges are communicated between nodes, the reader will assume that nodes have previously performed the steps outlined above and trust that thorough explanations lie ahead.

1.3 The Fibre Channel Communication Model

Based on the previous discussion emphasizing the frame as the central communication unit, the ultimate intention of this section is to understand the structure of the frame and how it is transmitted from one port to another (i.e. we are not concerned about switching just yet).

The high level overview presented in the previous section emphasizes how the primary function of a Fibre Channel network is to transmit and receive frames. Nodes transmit each frame to a switch, which receives the frame and determines the next hop to the destination. If the destination is a node on the same switch, the frame is sent to the destination node port; if the destination is located on a different switch, the frame is sent to the next switch on the shortest path to the destination by transmitting it over the appropriate Inter-Switch Link (ISL). This section explores the various steps involved in the transmission and reception of Fibre Channel frames and in doing so, introduces the Fibre Channel Communication Model.

Most network standards follow a structured model that favors a layered approach to the distribution of various protocol functions. Fibre Channel is no exception, as can be seen in Figure 1-9 (p. 10). The leading advantage of a layered protocol is that it provides modularity: each layer is clearly defined and responsible for a set of predefined tasks; changes in one layer do not impact other layers. This simplifies protocol design, implementation and verification, and also facilitates protocol evolution through localized updates. When developing a communications protocol, structured layering of various functions must be carefully considered in the early design stages. In the case of Fibre Channel, the decisions leading to the layered structure shown in Figure 1-9 are driven by the specific tasks which need to be accomplished. In the "transmit" direction, each layer is characterized by:

- the input it receives from the layer directly above it; and,
- the output it supplies to the layer directly below.

Of course, the reverse is true in the "receive" direction: input from lower levels must be processed and supplied to the upper level.

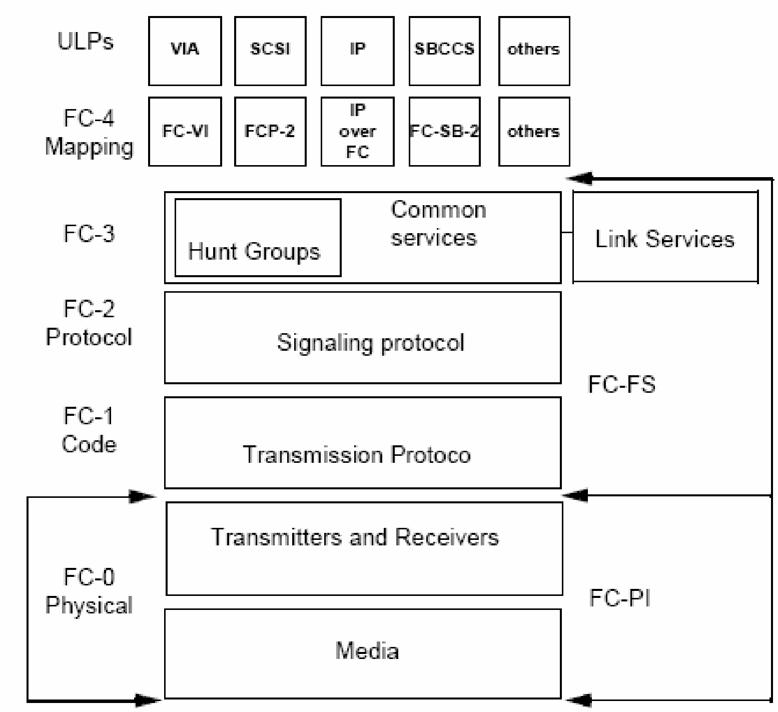

Figure 1-9: Layers of the Fibre Channel Protocol Suite

This collection of structured levels is called a communication model. While Figure 1-9 shows a complete picture containing every layer, the Fibre Channel standard documentation which describes each of these layers is broken into three distinct publications: FC-PI (Physical Interface) documents those aspects of Fibre Channel which relate to physical interfaces, which are collectively referred to as Fibre Channel level 0 (FC-0). FC-FS (Framing and Signaling) contains the specifications for serial transmission coding rules (FC-1) and also documents frame formats, logical structures and fabric services (FC-2). Finally, FC-LS (Link Services) defines how Fibre Channel manages link services – this level is also called FC-3, and contains additional functions besides link services. Note that FC-4 refers to protocol mapping and is included in Figure 1-9 for informational purposes only: protocol standards which define the FC-4 for a particular ULP are published by the same authorities which maintain the ULP in question. For example, the FC-4 for SCSI is defined in a document called FCP which is maintained by the SCSI Technical Committee T10 (and not by the Fibre Channel committee T11).

The following sections summarize the tasks performed at each layer by defining their input and their output. Advanced topics relating to each layer are covered in Part II of this book.

1.3.1 FC-0: The Physical Layer

The most fundamental communication task for any interconnect method consists in the transmission of a basic unit of digital data: the bit. FC-1 supplies FC-0 with a digital stream of bits which must be transmitted over some media via a specific type of transmitter/receiver compatible with the selected media. The ability to transmit a simple zero (0) or one (1) from a transmitter to a receiver is accomplished by the following steps:

- one or more supported media type(s) must be defined. Choice of media provides flexibility when deploying a physical infrastructure: for example, an electrical interface may provide a low-cost solution for short distances while fiber optics enable long range applications and higher transmission speeds. The media layer is also responsible for mechanically aligning the optical transmission fiber to the interface connector, defining requirements for connection types, operating ranges and insertion loss for optical and electrical cables. This is illustrated below in Figure 1-10 (p. 11).

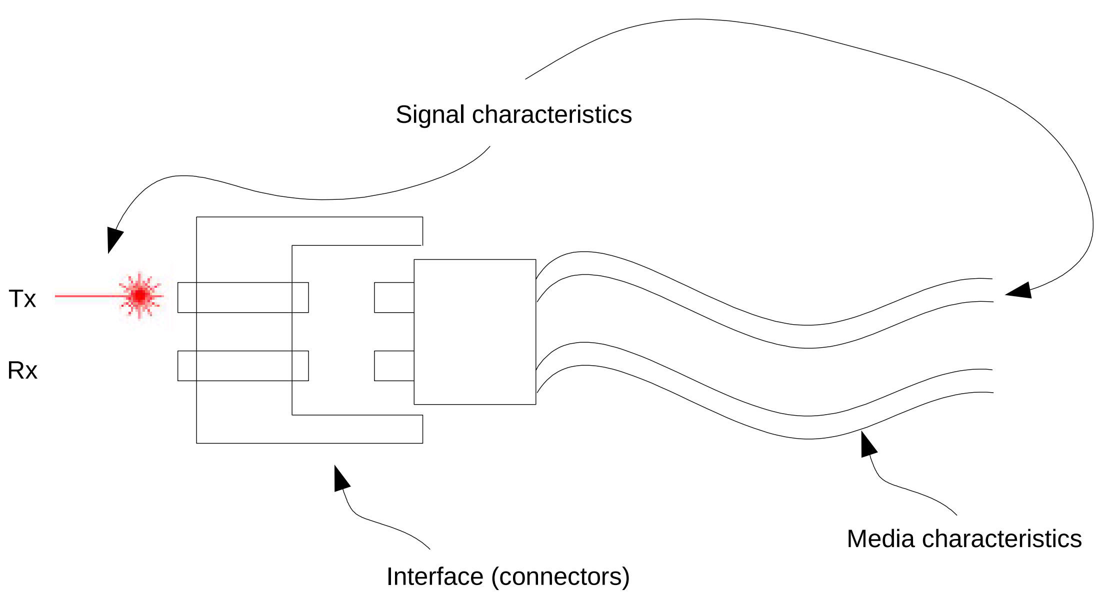

Figure 1-10: Sample FC-0 Media Type and Connector.

- the selected media is leveraged by selecting a transmission method which gives it the ability to carry signals representing binary values zero (0) or one (1) from a transmitting node to a receiving node within acceptable signal characteristics under minimum input signal power and extreme allowed conditions (EMI, RFI etc.). Conversely, a receiver must have the ability to receive these signals and output the associated binary level. In Fibre Channel optical communications, for example, the transmitter is a laser which is turned on to transmit a digital one (1) and switched off to transmit a digital zero (0) The receiver is an optical sensor which can detect the laser intensity level and translate that level to a digital one (1) or a digital zero (0), which is sent up one level to FC-1. As depicted in Figure 1-11 (p. 12), various laser power levels with different size and shape connectors may be used to interface with the fiber, depending on application requirements.

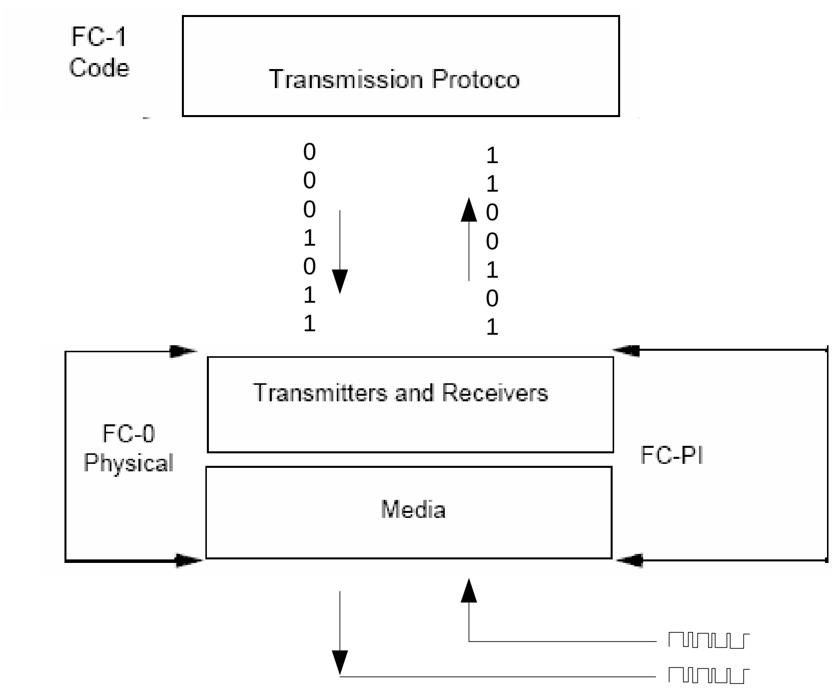

Figure 1-11: FC-0 Inputs and Outputs.

The fundamental characteristics described above are collectively referred to as Fibre Channel level 0 (FC-0), also known as the physical layer. A link, then, is formally defined as bidirectional physical media terminated at each end by a Link Control Facility. The LCF is contained within each Fibre Channel port and includes the transmitter and the receiver along with the appropriate control logic, and is responsible for managing the transmission and reception of data bits. This is illustrated below in Figure 1-12 (p. 13).

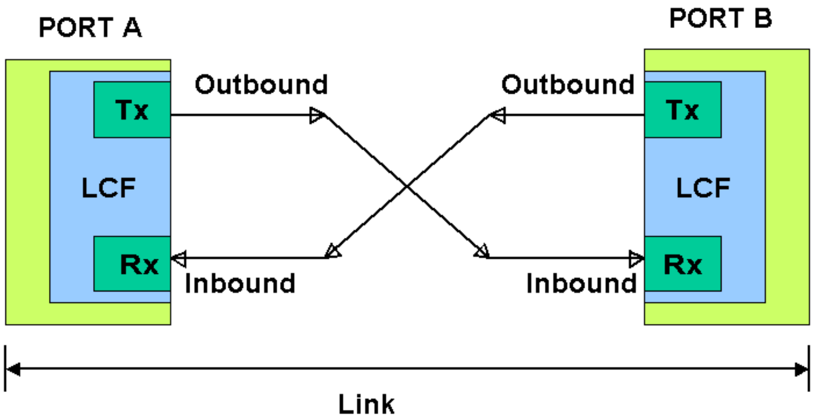

Figure 1-12: Transmitter and Receiver port of a Link

1.3.2 FC-1: Transmission Code

The FC-1 level accepts two fundamental transmission units from FC-2: four byte *data words*, used to carry FC-2 information, and four byte *ordered sets*, which are unique four byte patterns inserted in a stream of data words to signal special events such as the beginning of a frame or the allocation of a buffer to buffer credit. Data words and ordered sets are always four bytes in length and effectively constitute the smallest unit of information which can be sent over a Fibre Channel link. They are collectively known as Fibre Channel *transmission words*.

The main task to which FC-1 dedicates itself is the encoding of FC-2 transmission words into a format which is acceptable to FC-0 for transmission purposes. This "acceptability" is defined by such characteristics as high signal transition density, reasonable bit error detection capability, and DC Offset equalization.

This is accomplished by using 8B/10B encoding to convert FC-2 transmission words, which always consist of four bytes (32 bits), into four 10-bit characters (40 bits). For reasons which will be covered in Chapter 7, this coding scheme fulfills requirements a) and b) mentioned above. The resulting 40-bit pattern is then serialized and sent to FC-0 for transmission.

The reverse task must be accomplished at the receiving end: an incoming bit stream is supplied by FC-0 to FC-1, at which point the 40-bit patterns sent by the peer FC-1 layer must be recognized: each 10-bit character is decoded into its 8-bit equivalent, and the resulting four byte transmission word passed up to the FC-2 level.

The complete process implemented at the FC-1 level of the Fibre Channel communication model is illustrated in Figure 1-13 (p.14).

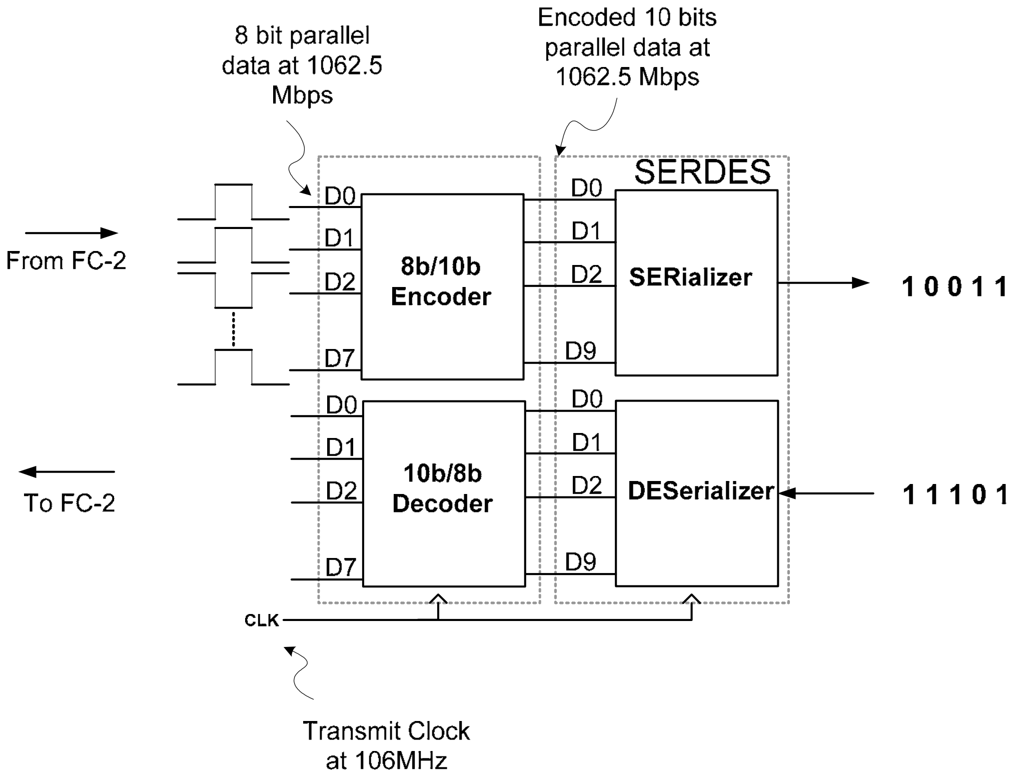

Figure 1-13: FC-1 Level Transmission Code.

1.3.3 FC-2: Framing, Flow Control and Class of Service

The advanced capabilities which uniquely distinguish Fibre Channel from other communication protocols are mostly made possible by the mechanisms defined at the FC-2 level of the communication model. Specifically, the following constructs and concepts are defined at the FC-2 level:

- the Fibre Channel frame;
- classes of service;
- flow control;
- sequences and exchanges;
- error detection.

This section is limited to framing as well as the interaction between FC-2 and FC-1. The logical constructs which enable Fibre Channel to perform all other tasks are discussed in section 1.3.4 (p. 19).

1.3.3.1 The Fibre Channel Frame

The fundamental unit of information manipulated by FC-2 is the Fibre Channel frame (this is not to be confused with the smallest unit of information which can be sent individually over a Fibre Channel link, which is the 8b/10b encoded version of a four-byte word, which amounts to 40 bits). This frame, shown in Figure 1-14 (p. 15), contains network- and node-related information in the frame header, and application data in the frame payload.

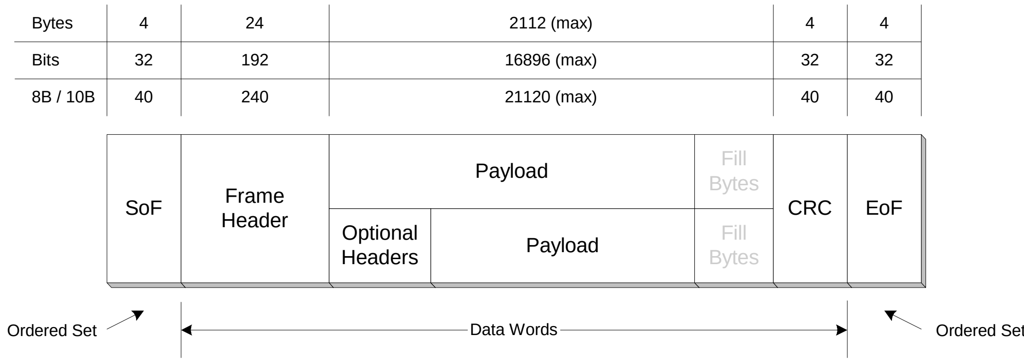

Figure 1-14: Fibre Channel frame format.

FC-2 always sends four-byte transmission words to the FC-1 layer. These transmission words must fall in one of two categories: data words and ordered sets. When FC-2 sends a complete frame to FC-1, it first sends the Start-of-Frame (SoF) ordered set, followed by the data words which constitute the frame.

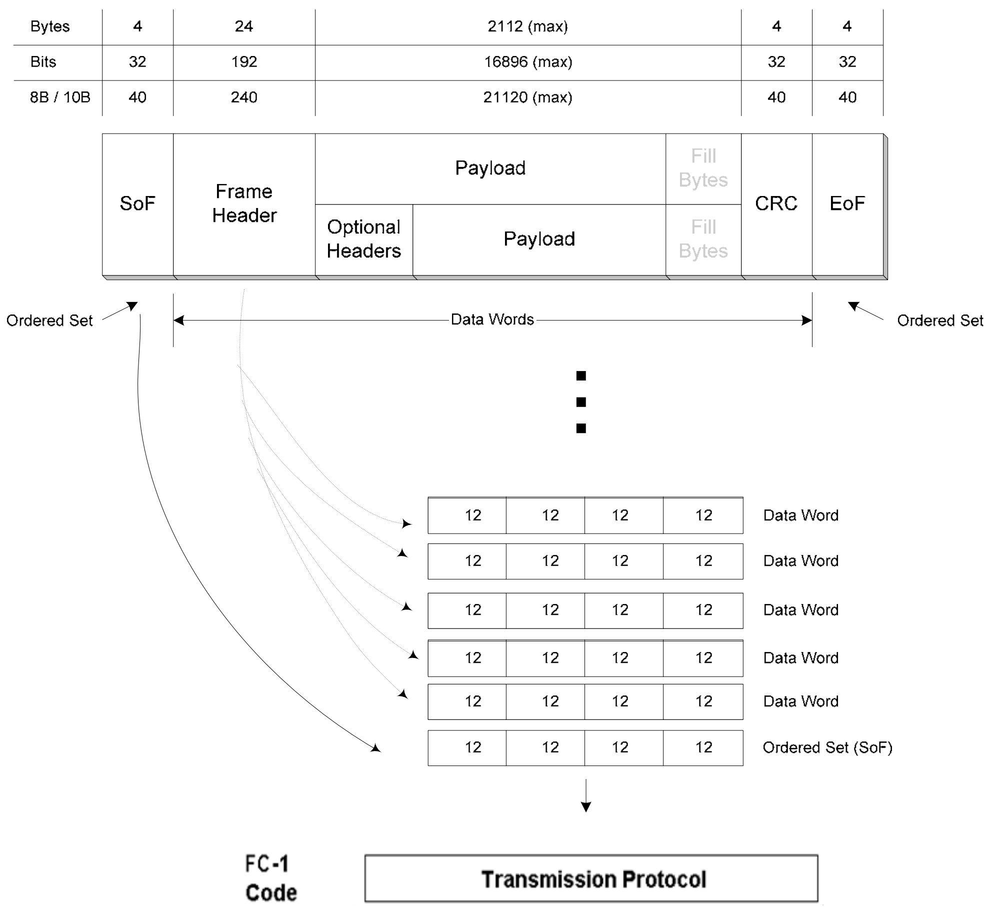

Figure 1-15: FC-2 supplies transmission words to FC-1.

Finally, the End-of-Frame (EoF) ordered set completes the transmission of a complete frame (this is not true – ordered sets are determined by FC-1 since they are disparitydependent. Rather, FC-2 must send “signals” which refer to a SoF or EoF – FC-1 must do the rest).

On the receiving end, the FC-1 layer passes received transmission words up to the FC-2 layer. When FC-2 receives an SoF ordered set, it immediately knows that subsequent transmission words can only be data words until the EoF ordered set is encountered; when this happens, the frame is known to have been completely received. FC-2 can start pre-processing the frame at any time after the SoF has been received since it can infer the position of the various header fields by simply counting byte positions, although the CRC should be checked for data integrity.

Introduce FT-0 and FT-1?

At this point, the reader should notice: FC frames are made up of ordered sets and data words. Ordered sets and data words are defined in FC-1, and are made up of digital bits, ones and zeroes which are transmitted by light impulses or electrical waves at level FC-0.

1.3.3.2 Classes of Service

Fibre Channel supports a feature called *class of service* which allows a node to request specific delivery characteristics from the network. This feature is introduced at this stage because the request for a specific class of service is implicitly contained in the SoF delimiter discussed in the previous paragraph. The SoF delimiter is an ordered set which precedes every frame sent into the network. The existence of multiple classes of service implies that many types of SoF exist depending on the requested class of service.

Fibre Channel supports six classes of service which are listed below. However, only three are in common use when it comes to storage networks; accordingly, this book only covers class 2, class 3 and class F.

- **Class 1**: class 1 is a connection-oriented service which is mainly characterized by the fact that the full link bandwidth is reserved and dedicated to a single session between two Nx_Ports. The term *connection-oriented* indicates that the network must configure a dedicated circuit in order to carry class 1 frames between the two ports. As required, this circuit may span many Fibre Channel switches and transit over ISLs, which implies that switches must collaborate in the establishment of this circuit by way of some switch-to-switch protocol. Dedicated circuits follow a single path through the network, which guarantees in-order delivery of frames and congestion-free communication. In addition to link-level flow control (buffer to buffer credit system), Class 1 also mandates end-to-end flow control (flow control is covered below on page 17).
- **Class 2**: class 2 is a connectionless service which leverages buffer-to-buffer and end-to-end flow control, in addition to confirmation of delivery (or non-delivery) of frames: each class 2 frame received by an Nx_Ports is acknowledged back to the sending port by using the Extended Link Service (ELS) The term *connectionless* means that the network treats each frame independently from every other frame and that no bandwidth is reserved for any particular conversation between Nx_Ports, which may lead to occasional congestion on busy ISLs. Same-destination frames may follow different paths through a fabric, potentially arriving out of order. Note that Fibre Channel does not place a formal

requirement for in-order delivery, although in practice this may cause problems and is usually avoided by switch vendors.

- **Class** 3: the near-totality of Fibre Channel ports deployed in modern SANs use class 3 service for storage traffic. Class 3 is a connectionless service which only supports the buffer-to-buffer credit system of flow control. As a result, class 3 does not guarantee in-order delivery, and no indication is returned to the sender of a frame which is lost due to congestion, corruption, or a busy port.
- **Class** 4: this class of service makes use of virtual circuits across a fabric-managed shared-bandwidth allocation protocol. Frames are guaranteed to be delivered in order, and an indication of delivery (or non-delivery) of frames is provided.
- **Class** 6: class 6 is a unidirectional class of service whereby a single stream of frames originating from an N_Port is transmitted to a fabric switch over a dedicated circuit (this implies a single session at full link bandwidth). The fabric replicates the stream of frames to a selected group of destinations over an equal amount of dedicated circuits. Class 6 is also known as the Multicast Service and requires a special entity in the fabric called a Multicast Server.
- **Class F:** class F frames are used exclusively by Fibre Channel switches when they communicate with each other over Inter-Switch Links for the purpose of configuring and maintaining the fabric. This topic is covered extensively in Chapter 5.

The Class of Service which applies to a particular frame is determined by the Start of Frame ordered set which must be supplied by the sending Nx_Port for every frame transmitted over a link. Since many classes of service are supported, the terminology summarized in Table 1-1 (p. 17) distinguishes a variety of delimiters:

Table 1-1: Fibre Channel Frame Delimiters

Data frame	Delimiters
Class 1	SOF_{c1}, SOF_{i1}, SOF_{n1}, EOF_{n}
Class 2	SOF_{i2}, SOF_{n2}, EOF_{n}
Class 3	SOF_{i3}, SOF_{n3}, EOF_{n}, EOF_{t}
Class 4	SOF_{c4}, SOF_{i4}, SOF_{n4}, EOF_{n}
Class 6	SOF_{c1}, SOF_{i1}, SOF_{n1}, EOF_{n}

Missing class F for now…

In summary, Nx_Ports select the required class of service for a frame by using the appropriate SoF delimiter at the beginning of the frame. Note that Table 1-1 indicates that various SoF delimiters exist even within the same class of service; this is required for sequence management and is covered in the next section. It should also be noted that Nx_Ports discover supported classes of service for fabric ports as well as peer Nx_Ports when they perform a fabric login (FLOGI) with the Fx_Port and a port login (PLOGI) with one or more peer Nx_Ports.

1.3.3.3 Flow Control

Flow control is a mechanism which ensures that a transmitter only send a frame when the receiver is ready to accept it. In Fibre Channel, two levels of flow control are defined: link-level flow control and end-to-end flow control. The difference is illustrated in Figure 1-16 (p. 18).

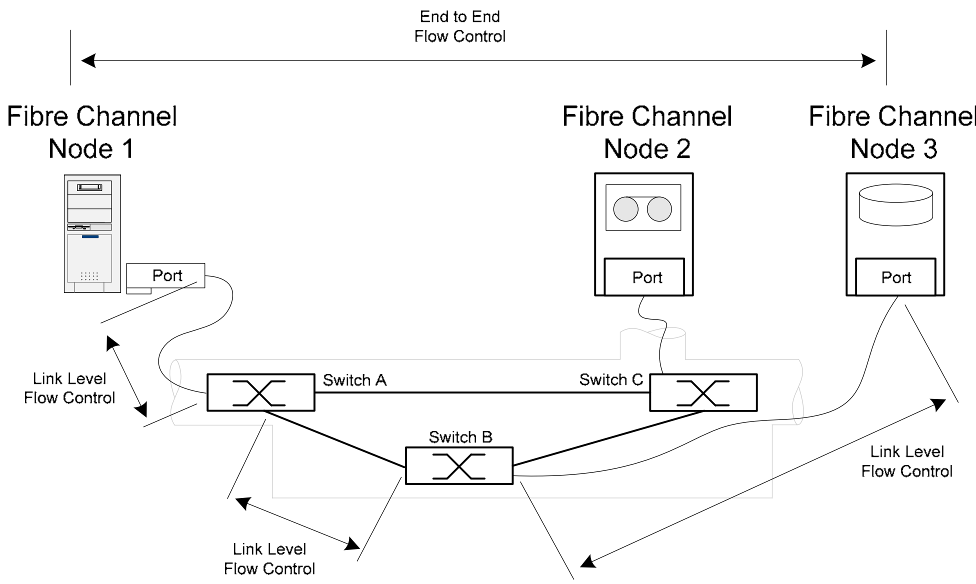

Figure 1-16: Link-level Flow Control and End to End Flow Control

Link-level flow control involves the FC_Ports at each end of a physical link through the use of Buffer to Buffer credits and is used by all classes of service (although only for connection establishment in classes 1 and 6), End to end flow control is performed between two Nx_Ports by use of ACKs and is used for all classes of service except class 3. As previously mentioned, modern SANs rely almost exclusively on class 3 traffic; accordingly, this section limits the discussion to link level flow control using buffer-to-buffer credits.

When Nx_Ports perform a fabric login (FLOGI), one of the numerous parameters exchanged with the Fx_Port during this operation is the initial allocation of buffer-to-buffer credits (BB_Credits). This allocation is bidirectional: the Nx_Port allocates a certain amount of credits to the Fx_Port, and the Fx_Port allocates a certain amount of credits (not necessarily the same) to the Nx_Port. Each individual credit constitutes a permission to send a single frame, after which the credit is considered to be spent: the FC_Port must decrement its BB_Credit counter by one. If a port exhausts its available BB_Credits, it is not allowed to send frames on the link until additional credits are received. Thus, FC_Ports perform flow control by granting credits at a rate which corresponds to their ability to receive frames. Credits are granted by sending a special transmission word (an ordered set) which indicates that one credit is being granted; this ordered set is called a Receiver Ready (R_RDY). Normally, receiving FC_Ports respond to every frame they accept by sending a R_RDY on the link from which the frame was received. When a sending FC_Port receives a R_RDY ordered set, it increments its BB_Credit counter by one.

1.3.4 FC-2: Sequences and Exchanges

The FC-2 layer is responsible for much more than just sending and receiving frames. The frame header fields must be correctly populated, sequences and exchanges must be tracked and error handling invoked in the case where abnormal situations occur. All these tasks are accomplished by the proper use of various fields in the frame header, or by using pre-defined ordered sets. While the Fibre Channel frame is a physical entity with a rigid format and clear boundaries, the remainder of Fibre Channel functions are performed by sending messages which consist of multiple frames related to each other by sequences and exchanges.

As previously mentioned, Fibre Channel nodes communicate by sending data frames to each other through a network topology. The format of the Fibre Channel frame is shown in Figure 1-17 (p. 19) and is seen to consist of three parts:

- the frame header contains a variety of fields which are used by Fibre Channel nodes and switches to process the frame;
- the data portion of the frame contains the actual payload to be carried across the network, along with payload-specific information and fill bytes, if required;

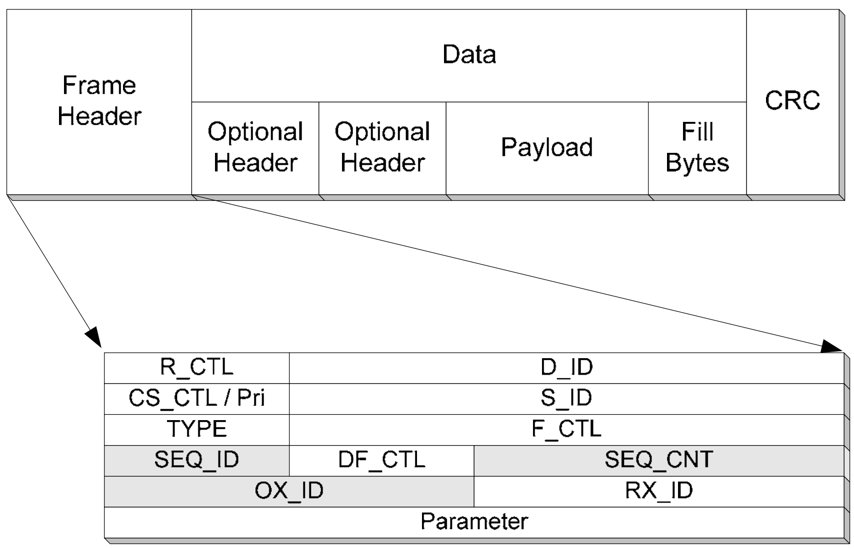

Figure 1-17: Fibre Channel frame format.

- the frame trailer consists of a Cyclic Redundancy Check (CRC) which is used to verify frame integrity.

While the data and trailer portions of the Fibre Channel frame are self-explanatory, the frame header is more complex as it contains a variety of fields which are used by nodes and fabric switches alike to process the frame at various stages of its journey from one node to the next. The three fields highlighted in Figure 1-17 are of particular interest at this juncture since they are used to implement one of Fibre Channel’s outstanding features: the ability to account for groups of related data frames, called sequences, and groups of related sequences, called exchanges.

Compared to traditional network protocols, Fibre Channel's unique ability to track sequences and exchanges results in a particular affinity for this protocol to carry transactional information such as SCSI command execution data. The relationship between frames, sequences and exchanges is shown in Figure 1-18 (p. 20).

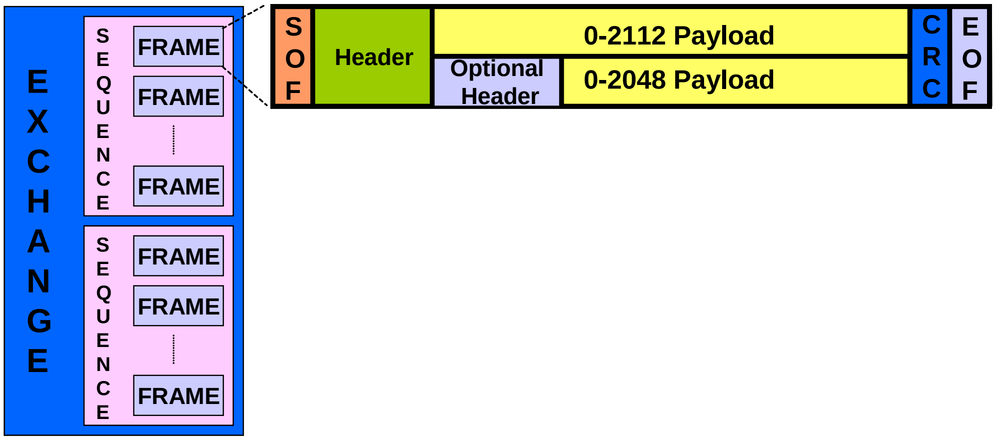

Figure 1-18: Frames, Sequences and Exchanges.

As can be seen above, an exchange consists of one or more sequence(s), and each sequence within an exchange contains one or more Fibre Channel frames. The association between frames, sequences and exchanges is defined within the header of each transmitted frame. This concept is illustrated in Figure 1-19 (p. 20).

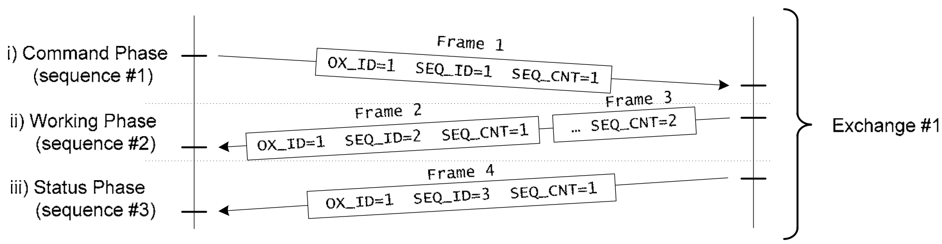

Figure 1-19: Fibre Channel frame headers contain information about sequence and exchange association.

This sample exchange between two nodes results in the transmission of three Fibre Channel sequences. The first sequence contains a single frame (Frame 1) and is transmitted by the *originator* to initiate the exchange. The originator populates the frame header with information that uniquely identifies the exchange (OX_ID=1) as well as the sequence (SEQ_ID=1) to which the frame belongs. Additionally, the ordering of the frame within the sequence is indicated by SEQ_CNT=1. As a result of receiving Sequence #1, the *responder* sends Sequence #2, which consists of two Fibre Channel frames. Again, the header within each frame contains information which associates it to the original exchange (OX_ID=1), but to a separate sequence (SEQ_ID=2). Finally, each frame within Sequence #2 is individually numbered (SEQ_CNT=1, SEQ_CNT=2). The third and final sequence of this exchange is again sent by the responder. Consisting of a single frame, the appropriate header fields indicate that this frame is the first (and last) frame of the third (and last) sequence of Exchange #1.

The following noteworthy characteristics apply to Fibre Channel exchanges and sequences:

- **sequences are unidirectional**: frames which belong to a sequence may only travel in a single direction. In other words (in conjunction with the following bullet), only the sequence initiator may transmit while holding the sequence initiative. In the example from Figure 1-19, this means that while sequence #2 is being sent by Node 2, Node 1 may not transmit frames as a part of this exchange; (is initiator-responder for the sequence defined relative to the direction of the sequence, or the direction of the exchange?)
- **concurrent sequences are not allowed within a single exchange**: within a particular exchange, only one sequence may be active at any time.

These characteristics imply that exchanges are inherently half-duplex: data frames related to an individual exchange may only travel in one direction at a time. However, a Fibre Channel node *may* have multiple concurrent exchanges active at the same time. In other words, exchanges can be multiplexed, which results in full duplex capability at the node level.

1.4 FCP Part II: Protocol Mapping Revisited

The protocol mapping function introduced in Chapter 2 (**Error! Reference source not found.**: ***Error! Reference source not found.***, p. **Error! Bookmark not defined.**) suggests that SCSI PDUs are grouped into related IUs which are then transmitted over the interconnect layer, as illustrated in Figure 1-20 (p. 21).

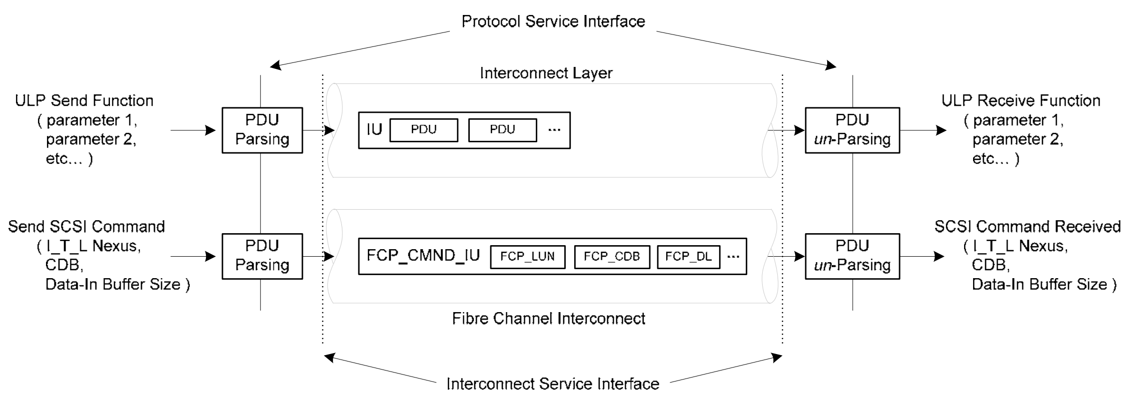

Figure 1-20: FCP IUs are transmitted over Fibre Channel.

When Fibre Channel is used as an interconnect method, SCSI PDUs are parsed into FCP IUs, and the resulting IUs are transmitted by leveraging those mechanisms made available to upper layer protocols (ULPs) by Fibre Channel. To summarize the previous section, these mechanisms are:

- **Fibre Channel Exchanges**: a collection of related half-duplex sequences can be tracked as a single exchange. In FCP parlance, sequences correspond to FCP_IUs while exchanges correspond to the half-duplex, bi-directional flow of IUs which ultimately results in the complete execution of a SCSI command;

- **Fibre Channel Sequences**: the succession of Fibre Channel frames which carry an FCP_IU is called a sequence;
- **Fibre Channel Frames**: the fundamental unit of data transmission is the Fibre Channel frame, which contains the fields necessary to associate each frame to a unique sequence within a unique exchange.

Generically, the execution of a SCSI READ command looks like the succession of IUs shown in Figure 1-21 (p. 22): an FCP_CMND_IU contains the command to be executed (READ), the data being read is returned in a FCP_DATA_IU, and a command status is conveyed to the initiator via the FCP_RSP_IU. The complete execution of a SCSI command is contained within a single Fibre Channel exchange while each FCP_IU within this exchange is transmitted by using a Fibre Channel sequence.

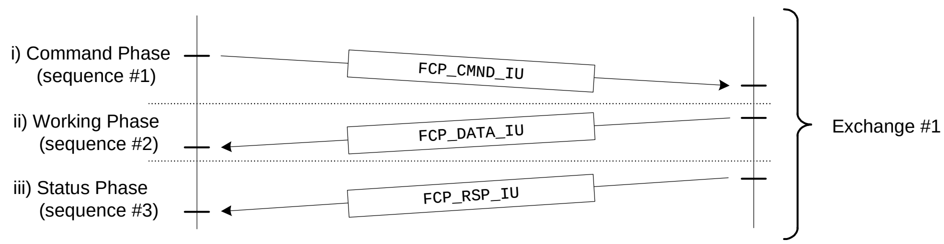

Figure 1-21: FCP_IUs

"On the wire", a sequence is transmitted as a succession of one or more Fibre Channel frames, and the group of sequences required to complete the execution of the SCSI command represents a Fibre Channel exchange. Therefore, the actual frames introduced in a Fibre Channel network to execute the SCSI command shown above might look like Figure 1-22 (p. 22).

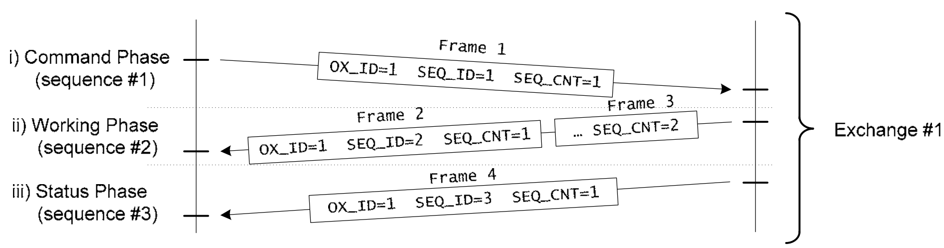

Figure 1-22: IUs Consist of One Or More Fibre Channel Frames.

The previous example demonstrates that a single exchange is required to execute this sample READ command (as is the case with all SCSI commands). The FCP_CMND_IU is sent by using a Fibre Channel sequence which contains a single frame, and the frame header contains the necessary information to associate the frame with the unique exchange (OX_ID=1) and sequence (SEQ_ID=1) to which it belongs. The FCP_DATA_IU consists of sequence #2 which, in this particular case, requires two Fibre Channel frames. Each frame within this sequence is tracked by the SEQ_CNT field contained within the header of each Fibre Channel frame. Finally, the FCP_RSP_IU is also a single-frame sequence (#3) which completes the Fibre Channel exchange and, by the same token, indicates the completion of the SCSI command execution.

While this overview associates Fibre Channel exchanges to SCSI operations and Fibre Channel sequences to FCP_IUs, it should be noted that this association is established through protocol mapping (specifically: by FCP). Exchanges and sequences are logical constructs which are intrinsically part of Fibre Channel and do not exist for the sole benefit of SCSI: indeed, every ULP that uses Fibre Channel as a transport protocol must use these fundamental structures to transmit data from one device to another since *any* Fibre Channel data transfer must be performed within the context of an exchange. For example, Fibre Channel may be used to carry Internet Protocol (IP) traffic, in which case every IP datagram must be protocol-mapped (via Fibre Channel Link Encapsulation, FC-LE) into a single sequence which may comprise one or more Fibre Channel frame(s). Furthermore, the half duplex nature of Fibre Channel sequences requires two concurrent exchanges (one in each direction) to realize the full-duplex capabilities of IP. This is but one example of how various ULP protocol mappings determine how to leverage sequences and exchanges when using Fibre Channel. In other words, protocol mapping is the glue between an ULP and Fibre Channel.

This independence between Fibre Channel and SCSI (and all other ULPs which use Fibre Channel, for that matter) should be kept in mind since the remainder of Part I is dedicated mainly to Fibre Channel and is meant to introduce the reader to key aspects of this communication protocol. Advanced Fibre Channel topics will be covered in Part II, Chapter need reference.

1.5 Fibre Channel Addressing

The very core function of a Fibre Channel SAN is to deliver every frame that enters the network to its intended recipient. It is therefore essential that all Fibre Channel ports be identified by some means such that frames may be addressed to them.

In Fibre Channel, every port that logs into the fabric (by way of FLOGI) is assigned a network-unique 24 bit *Fibre Channel identifier* (FC_ID) by the fabric switch. When a source port needs to send frames to a destination port, the source must populate the *Destination_ID* (D_ID) field of every Fibre Channel frame header with the FC_ID of the destination port such that Fibre Channel switches may inspect individual frames to determine the appropriate course of action. For example, Figure 1-23 (p. 24) illustrates a Fibre Channel frame transmitted by Node 1 entering Switch A. Since the D_ID field in the frame header contains the FC_ID of Node 3, Switch A forwards the frame towards Switch B.

> ***Note:*** *The exact procedure by which a switch determines* ***where*** *the destination lies in the network (and consequently determines the correct ISL over which to forward the frame) is accomplished through the use of a routing protocol called Fabric Shortest Path First (FSPF), which is covered in* ***Error! Reference source not found.***.

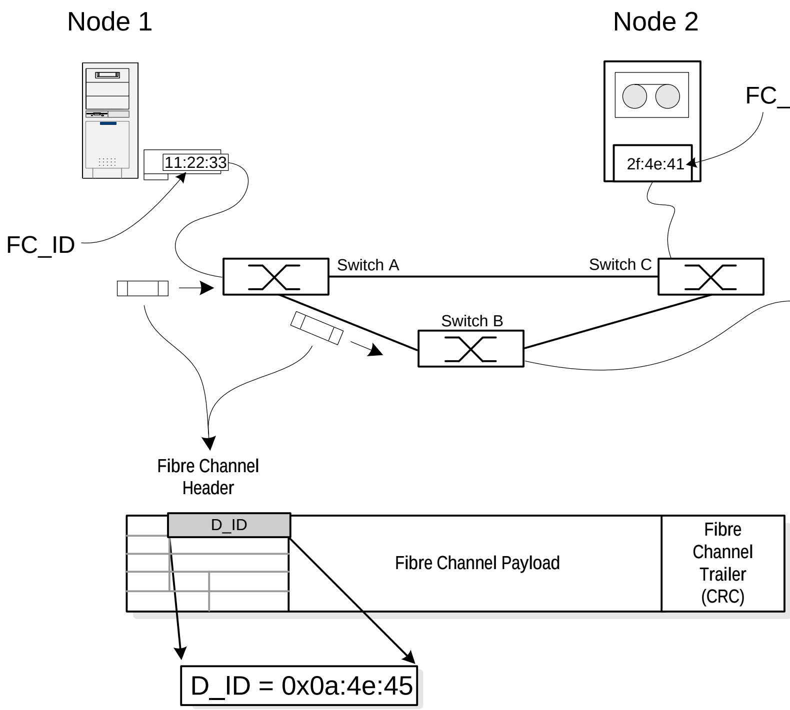

Figure 1-23: The Destination FC_ID Is Contained In The D_ID Field

1.5.1 Addressing Terminology

The generic term *addressing* is widely used in the networking industry to describe a variety of addressing methods. For example, a 48 bit Ethernet address of the form 0x00:0d:60:af 28:3d is also called a Medium Access Control (MAC) address, while Data Link Connection Identifiers (DLCIs) reflect a form of addressing used in Frame Relay networks. Not to be overlooked, the ubiquitous IP address has practically become a household term with the rising popularity of the Internet and is yet another example of the pervasive use of the word "addressing".

In Fibre Channel, addresses (or portions of addresses) are called *identifiers*, and are commonly designated by the suffix *_ID*. The Fibre Channel standards refer to the term *N_Port_ID* to designate N_Port addresses, and sometimes also uses the term *native address identifier*. In an effort to use a single term that applies to any topology, this text adopts the generic term *FC_ID* to describe all instances of 24 bit Fibre Channel addresses.

Fibre Channel identifiers are partitioned into three logical 8 bit fields. The standards use the terms *Domain_ID*, *Area_ID* and *Port_ID* to refer individually to each of these logical fields, as does this text.

Finally, FC_IDs are sometimes referred as *Destination_IDs* (D_IDs) and *Source_IDs* (S_IDs) when referring to their location in the Fibre Channel frame header.

1.5.2 FC_IDs Are Dynamic Hardware Addresses

When a Fibre Channel port is connected to a fabric switch, its FC_ID is initially undetermined. Through the FLOGI process, the fabric assigns a 24 bit network-unique address (the FC_ID) to each port that logs in (the exact process is covered in the next section). By virtue of the fact that it is *assigned*, the FC_ID is therefore dynamic in nature: the exact value of the FC_ID is determined by the switch at login time. As a result, logging out of the fabric and then logging in again at a later time may result in a completely different FC_ID assignment for a given port.

> ***Note:*** *The dynamic behavior of Fibre Channel address assignment may cause problems in some legacy operating systems. This problem (and its solution) are discussed in Part III, need reference.*

The generic term "hardware address" or "physical address" is sometimes used to characterize the nature of FC_IDs. This nomenclature refers to the fact that hardware adapters such as HBAs must examine each received frame to confirm that they are indeed the correct recipient. This verification is made by comparing the D_ID contained in the frame header to the FC_ID assigned to the Nx_Port during login. When the D_ID in a received frame header does not correspond to the assigned FC_ID for a particular Nx_Port, that frame is dropped and never "physically" enters the hardware adapter.

> ***Note:*** *Other terms that are used to describe hardware addresses include* ***Medium Access Control*** *(MAC – used in IEEE 802 standards), and* ***Layer 2 address****, used in reference to the layered OSI model for network communications.*

In summary, addresses that appear in the frame header of any frame-based networking technology are sometimes referred to as hardware addresses because it is these addresses that determine whether the frame should be admitted into the adapter hardware for further processing.

> ***Note****: The Ethernet/IP-savvy reader will mind the following distinctions between FC_IDs and MAC addresses: Ethernet MAC addresses and Fibre Channel identifiers are both considered hardware addresses because they appear in the data link layer frame header. Consequently, this is why some addresses, such as IP addresses, are not hardware addresses: indeed, there is no such thing as a Layer 2 IP frame – IP* ***datagrams*** *are always encapsulated within lower level protocols such as Ethernet, PPP, or… Fibre Channel. It is also worth noting that MAC addresses are usually burned into Ethernet adapters whereas FC_IDs are dynamically assigned by a switch; in this respect, it should be noted that the term "hardware address" does not relate to the fact that an address is burned in rather than dynamically allocated (or even assigned manually). This point is important because Fibre Channel HBAs have one or more burned-in World Wide Names (WWNs) in addition to their assigned FC_IDs. Although they are burned-in like MAC addresses, these WWNs are not addresses – they are* ***names****. WWNs are discussed further in the next section.*

1.5.3 FC_ID Format

The Fibre Channel identifier is a 24 bit number that is assigned to Nx_Ports as they log into a fabric switch. This FC_ID must be unique within the confines of a single fabric, which means that identical FC_IDs may exist in disjointed, independent fabrics. This possibility will be of considerable interest in the event that such fabrics should later be merged.

The FC_ID is logically divided in three parts, each being eight bits in length: the Domain_ID, the Area_ID and the Port_ID This is illustrated below in Figure 1-24 (p. 26).

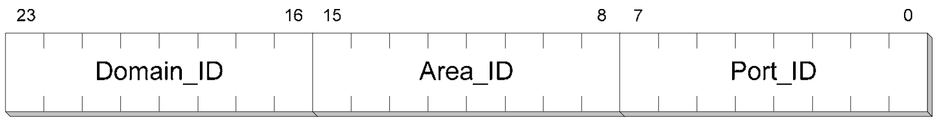

Figure 1-24: The FC_ID Contains Three Logical Fields

As will be covered in the following sections, these logical fields are leveraged in different ways according to the underlying topology (point-to-point, arbitrated loop and fabric), and various switch vendors follow slightly different rules for address assignment. While this should theoretically be of no consequence, there are instances when such discrepancies might cause problems, most notably during migrations and in multi-vendor fabrics. These issues are tackled in Part III need reference.

1.5.4 Fibre Channel Address Space Overview

The 24 bits in the FC_ID give way to 16777216 possible values, which collectively represent the complete Fibre Channel address space. Despite the fact that they are usually assigned dynamically, some FC_ID values are reserved for specific purposes, without regard for the topology in use. These Well Known Addresses (WKAs) and functional addresses are summarized below in Table 1-2 (p. 26).

Table 1-2: Summary Of Fibre Channel Address Space Partitioning

Domain	Area	Port	Description
00	00	00	Address used by N_port during FLOGI to request for address identifier
01-EF	00-FF	00	Address used by Fabric Loop port (FL_Port) connected to Public loop devices
01-EF	00-FF	00-FF	Address used by Fabric connected N_Ports and E_Ports
F0-FE	00-FF	00-FF	Reserved
FF	00-FA	00-FF	Reserved
FF	FB	00-FF	Reserved for Multicast Group ID
FF	FC	00	Reserved
FF	FC	01-EF	N_Port Identifier for Domain Controller
FF	FC	F0-FF	Reserved
FF	FD-FE	00-FF	Reserved
FF	FF	00-EF	Reserved
FF	FF	F0-F6	Reserved WKA
FF	FF	F7	WKA for key distribution service
FF	FF	F8	WKA for Alias Service
FF	FF	FA	WKA for Management Service
FF	FF	FB	WKA for Time Service
FF	FF	FC	WKA for Directory Service
FF	FF	FD	N_port Identifier for fabric controller
FF	FF	FE	N_port Identifier for fabric F_port
FF	FF	FF	Broadcast Address

While this table may seem complicated at first glance, it is presented here in its entirety to introduce the reader to the variety of address types that exist in support of Fibre

Channel. The current discussion focuses on the distinction between regular node port addresses and WKAs. The FC_ID breakdown into three logical fields is also covered in the following sections.

1.5.4.1 The Two Broad Categories of FC_IDs

The most notable use of FC_IDs in a Fibre Channel network is the assignment of dynamic identifiers to Nx_Ports as they log into the fabric. Once an FC_ID is acquired, nodes may communicate with other nodes by populating the D_ID field of the frame header with the identifier of the intended recipient, as illustrated in Figure 1-23. The FC_IDs used for this purpose are called *native address identifiers*, and constitute the most prevalent category of FC_IDs.

There is a fixed range of FC_IDs within the Fibre Channel address space that is dedicated to well known components of a Fibre Channel fabric. These components, such as the Fibre Channel Name Server (FCNS) or the Domain Controller, are assigned dedicated FC_IDs, called Well Known Addresses (WKAs), that never change. For example, the FCNS is always accessible through the FC_ID 0xff:ff:fc. The range of FC_IDs that are reserved for WKAs is identified by the shaded fields in Table 1-2.

1.5.4.2 FC_ID Allocation In A Fabric Topology

When an N_Port is initially connected to a fabric switch, it does not have a valid FC_ID with which to communicate with other fabric ports. By logging into the fabric, an N_Port implicitly requests that the fabric switch assign it a valid FC_ID. The meaningful fields of the resulting address are shown below in Figure 1-25 (p. 27).

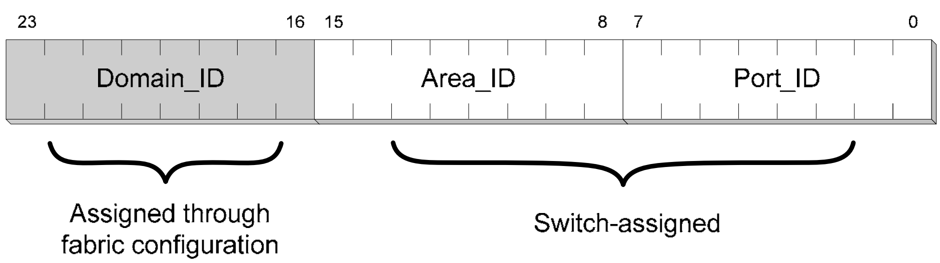

Figure 1-25: Fabric Topology FC_ID Format

When a switch assigns an FC_ID to a requesting N_Port, the Domain_ID field of the assigned address is always equal to the switch's own Domain_ID (in a stable, configured fabric, each switch is assigned a unique eight bit Domain_ID during fabric configuration – Chapter need reference covers this subject in detail). The value of the remaining 16 bits is left to the discretion of the assigning switch. As a result, there is no way to predict the FC_ID that will be assigned to any particular N_Port before the port in question actually logs into the fabric.

> *Note: While the previous statement is true, some legacy switches implement a semi-static address allocation scheme that links specific switch ports to pre-determined combinations of Area_ID and Port_ID values. As a result, some HBA vendors and operating systems rely on this semi-static assignment practice for internal binding purposes. This shortcoming and its implications are discussed in Part III, need reference.*

Figure 1-26 (p. 28) illustrates how the Domain_ID portion of the FC_ID (shaded in grey) remains fixed on a switch-per-switch basis.

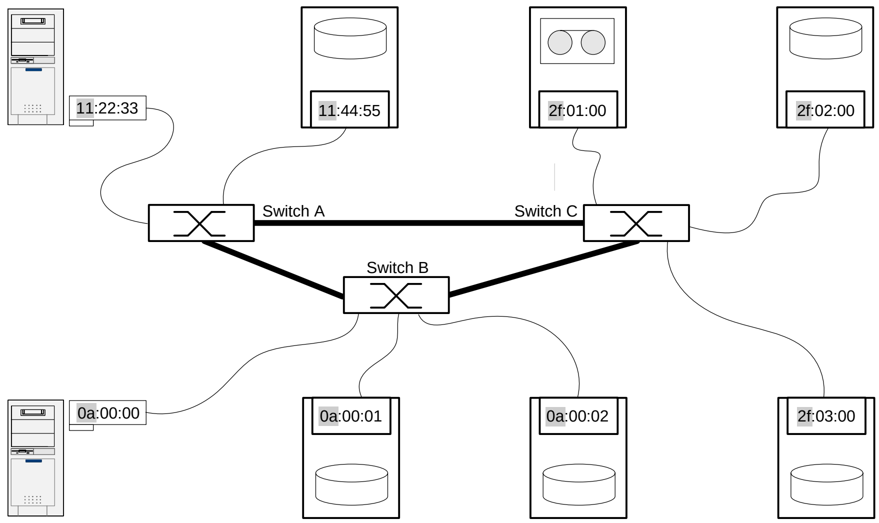

Figure 1-26: FC_ID Assignments In A Fabric Topology

For demonstration purposes, each switch in this example assigns FC_IDs in its own particular way:

- Switch A (Domain_ID = 11) sequentially assigns the values 22, 33, 44, etc. to Area_ID and Port_ID fields.
- Switch B (Domain_ID = 0a) keeps the Area_ID field equal to zero, and sequentially increments the Port_ID, starting at zero.
- Switch C (Domain_ID = 2f) keeps the Port_ID field equal to zero and sequentially increments the Area_ID field, starting at one.

While each switch assigns FC_IDs a bit differently, each method is perfectly valid, and the simultaneous use of these different assignment schemes is of no consequence to the proper operation of the fabric.

Note: While the previous statement is correct with respect to the standards, some vendors do not respect this requirement which causes their switches to mishandle some frames where FC_IDs are in a format that these vendors do not recognize. An example of this scenario is given in need reference.

1.5.4.3 Area_ID and Port_ID In Arbitrated Loop Topologies

On the basis of the previous example, there seems to be little justification for the existence of the Area_ID field. Indeed, were it not for the arbitrated loop topology, the FC_ID format would probably never have included this field. However, the possible coexistence of public loops and fabric ports within a single fabric introduces a situation where the addressing characteristics of the loop topology must be integrated within the fabric addressing hierarchy.

The simplest expression of a loop topology is the private loop. Private loop devices are characterized by the fact that they cannot communicate with devices that are not part of the same loop. When a private loop port is initialized, the loop initialization process results in the assignment of FC_IDs formatted as shown below in Figure 1-27 (p. 29).

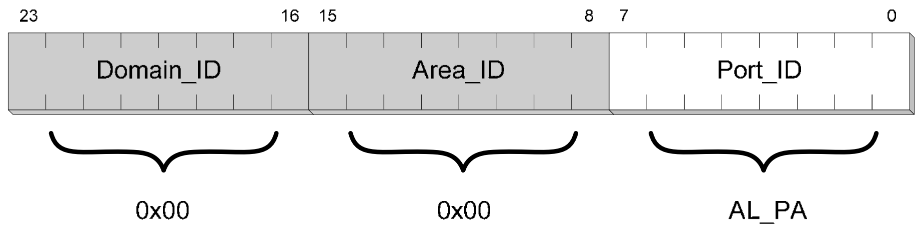

Figure 1-27: Private Loop FC_ID Format

The Domain_ID and the Area_ID fields are always set to zero in a private loop, and the Port_ID field adopts a value known as the Arbitrated Loop Physical Address (AL_PA). For reasons related to disparity considerations (which are covered in Chapter need reference), this field can only accept a subset of the 256 possible values that would otherwise be implied by its eight bits. This subset consists of 127 values contained within the range x00 to 0xef. For example, values 0x00, 0x01 and 0x02 are valid AL_PAs whereas 0x03 and 0x05 are not.

Note: While 0x00 is a valid AL_PA, it is only used by public loop FL_Ports and is never used by private L_Ports.

In the case of public loops, the address assignment scheme leverages both the Domain_ID and Area_ID fields of the FC_ID. The key feature of a public L_Port lies in its ability to communicate with Nx_Ports located anywhere within a Fibre Channel network. To this end, L_Ports must have a full 24 bit address, which they obtain by performing a fabric login (FLOGI). This login operation is performed by public loop ports after they acquire an AL_PA through normal loop initialization. In response to the FLOGI request, the Login Server returns a 24 bit address formatted as shown in Figure 1-28 (P. 29).

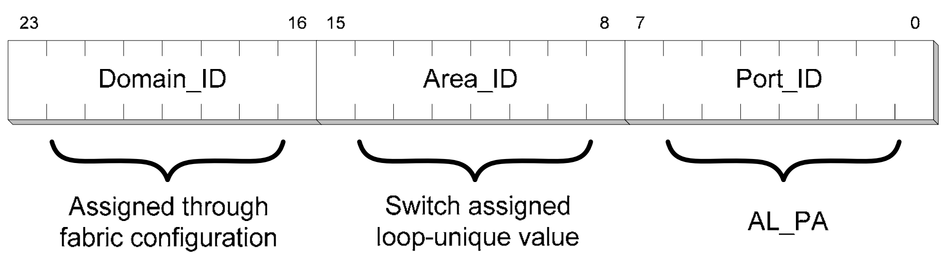

Figure 1-28: Public Loop FC_ID Format

Whereas the Domain_ID identifies a specific switch in a fabric, the Area_ID field is used to identify a particular physical loop within a switch. Accordingly, multiple public L_Ports located on the same physical loop will always be assigned the same Domain_ID and Area_ID, but will each have their own individual AL_PAs. This is shown in Figure 1-29 (p. 30): each loop-attached disk connected to Switch B has a Domain_ID equal to that of the switch (0x0a), but the Area_ID varies according to the specific loop to which each disk is connected.

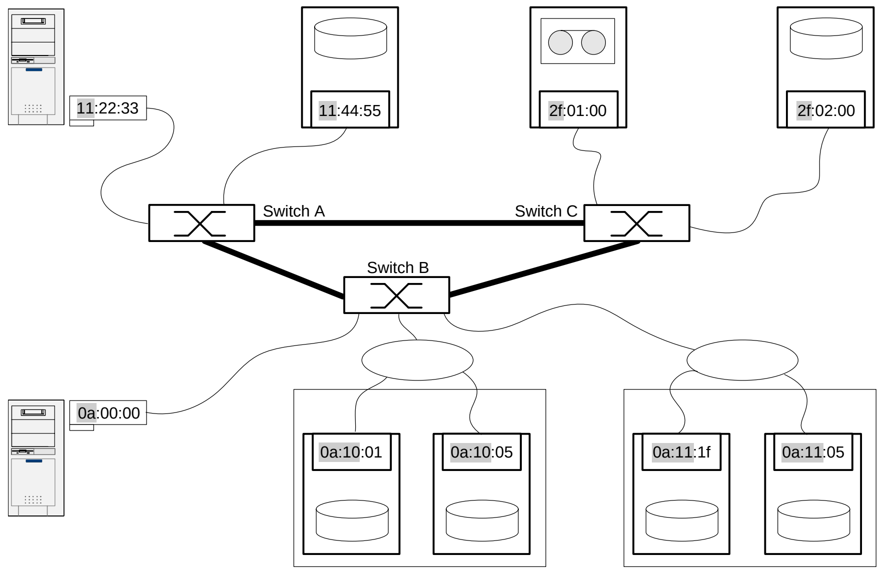

Figure 1-29: Public Loop Addressing Example

1.5.4.4 FC_ID Allocation: Point-To-Point

In a point-to-point topology, the FC_ID assignment is performed during the PLOGI stage by the N_Port that has the highest World Wide Name (this is discussed further in 0: *When a* fabric switch replies to an N_Port by sending an LS_ACC, it also registers the new N_Port in the Fibre Channel Name Server with a minimal amount of information, namely: FC_ID, pWWN and nWWN. However, N_Ports usually register themselves with the Name Server as well and, by doing so, usually include such detailed information as their protocol capabilities, symbolic node names, IP addresses, etc. A comprehensive section covering the Fibre Channel Name Server can be found at need reference.

Note: When an N_Port attempts to connect with the Name Server, it must direct a PLOGI operation to FC_ID 0xff ff fc, which is the WKA of the Name Server. PLOGI is explained in the next section.

When an N_Port that is also a SCSI initiator has successfully logged into a fabric, it usually attempts to discover fabric-attached N_Ports that are SCSI targets. In other words, it starts looking for storage. This is accomplished by querying the Name Server for a list of Nx_Ports that are logged in and, optionally, fulfill a particular criterion, such as “SCSI target”. This process is explained in need reference; for the purposes of this section, however, the reader may assume that a SCSI initiator already knows which Nx_Ports are SCSI targets. Indeed, an initiator N_Port will attempt a PLOGI with every discovered N_Port that is believed to be a SCSI target. In fact, some HBAs try to PLOGI into every single N_Port they learn about, initiator or target, and proceed to discover the port’s capabilities by using discovery capabilities provided by a subset of the Process Login (PRLI) Extended Link Service. This behavior may lead to problems that are exposed in need reference (i.e. single initiator zoning).

Port Logi (p. 42). The standards are not explicit about the exact range of addresses that can be assigned in a point-to-point scenario; it is reasonable to assume that any address is acceptable other than those reserved by Fibre Channel for special use (WWNs).

Therefore, point-to-point FC_IDs are in the range 0x00:00:01 to 0xEF:FF:FF, and no particular use is made of the three logical fields mentioned earlier (Domain_ID, Area_ID and Port_ID).

1.6 Fibre Channel World Wide Names (WWNs)

Most network protocols typically require that a fixed identity be associated to various entities such as nodes, ports, or other components. In Fibre Channel, this identification takes the form of a 64 bit value called a *World Wide Name* (WWN) and is leveraged to identify network entities and implement security policies. The WWN is a 64 bit value that is guaranteed to be "world unique" and is usually etched in silicon at the time a particular component is manufactured. It is important to understand that a WWN *is not an address*. Although the numerical format used for WWNs may look like a form of addressing, the reader will quickly dispel any such notion.

The example shown in Figure 1-30 (p. 31) demonstrates a typical use of WWNs. In this scenario, a particular target must restrict LUN access to a precise subset of initiators present within the fabric.

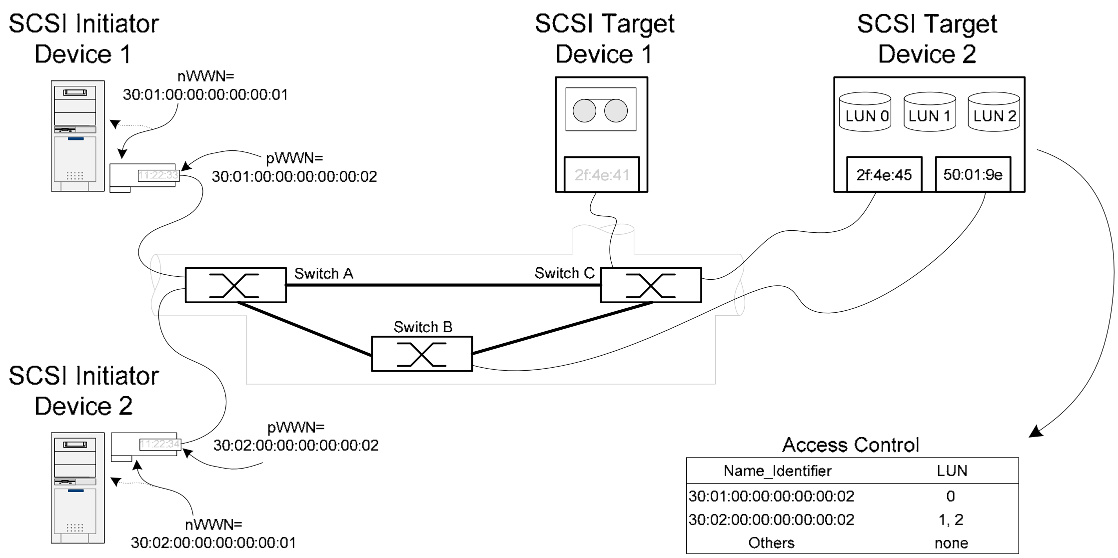

Figure 1-30: LUN Security Is Enforced Against WWNs

In this particular scenario, each initiator in Figure 1-30 needs to have its access restricted to a subset of available LUNs at the target ports. Through proper configuration, access control can be performed on the storage array such that only those initiator ports whose WWNs match an entry in the Access Control List (ACL) are granted access to the specified LUN(s). This example further demonstrates that such access control could not be performed reliably if hosts were identified by FC_ID instead of WWN: given that they are allowed to change over time, such identification would require the ACL in the storage array to be continuously updated to reflect a host's new FC_ID, which is clearly not manageable.

1.6.1 WWN Format

As previously mentioned, the length of a WWN is fixed at eight octets, or 64 bits. These 64 bits are divided in two basic fields: the first four bits form a field that is called the Name Assignment Authority field (NAA), while the field that consists of the remaining 60 bits contains the actual Name. This fundamental division is illustrated in Figure 1-31 (p. 32).

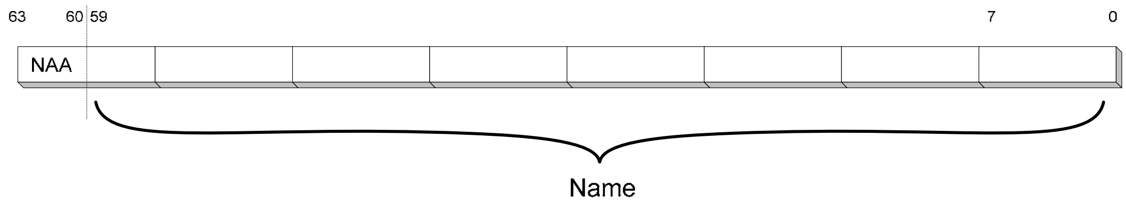

Figure 1-31: Fibre Channel WWN Format

The NAA field is required because more than one Name format is supported by the Fibre Channel naming standard. Each available format offers specific characteristics and varying scopes of uniqueness. Table 1-3 (p. 32) summarizes the various Name formats that correspond to all 16 possible NAA values.

Table 1-3: Name Formats According to NAA Value

Words 0, bits 31 - 28	NAA	Length
0000' b	Name not present	
'0001' b	IEEE 48-bit address	64
'0010' b	IEEE extended	64
'0011' b	Locally assigned	64
'0100' b	32-bit IP address	64
'0101' b	IEEE Registered	64
'0110' b	IEEE Registered Extended	128
'0111' b to '1011' b	Reserved	
'1100' b	EUI-64 Mapped	64
'1101' b	EUI-64 Mapped	64
'1110' b	EUI-64 Mapped	64
'1111' b	EUI-64 Mapped	64

In practice, only three Name formats are used in Fibre Channel networks. They are highlighted in Table 1-3 by a grey background, and correspond to NAA values 0x01, 0x02 and 0x05; accordingly, they are the only formats covered in the remainder of this text.

> ***Note:*** *It should be called to the reader's attention that the widespread use of the term World Wide Name (WWN) may lead one to believe that all Fibre Channel Names can be called WWNs. This is in fact not the case. The standards formally refer to the 64 bit value as a Name_Identifier. When the NAA for a particular Name_Identifier is equal to 0x1, 0x2, 0x5 or 0x6, it can then, and only then, be properly called a World Wide Name; this is because such formats include fields that are administratively maintained by Standards bodies which guarantee their "world uniqueness". As mentioned before, this is normally the case for all NAAs used in present-day Fibre Channel networks. However, it should be noted that the Fibre Channel Name_Identifiers used by many examples in this book start with 0x3 to properly reflect the fact that such names are concocted by the authors. As such, these names are not WWNs and are accordingly designated Fibre Channel Name_Identifiers.*

1.6.2 NAA = 0x1: IEEE 48 Bit Address

Some manufacturers elect to use the IEEE 48 Bit Address format to create Fibre Channel WWNs. When this is the case, the NAA field, which corresponds to the first four bits of the WWN, is set to 0b0001 (0x1). A Universal LAN MAC Address (ULA), which contains the manufacturer's Organizationally Unique Identifier (OUI), occupies the rightmost bits of the WWN, and all remaining bits are set to zero. This is illustrated in Figure 1-32 (p. 33).

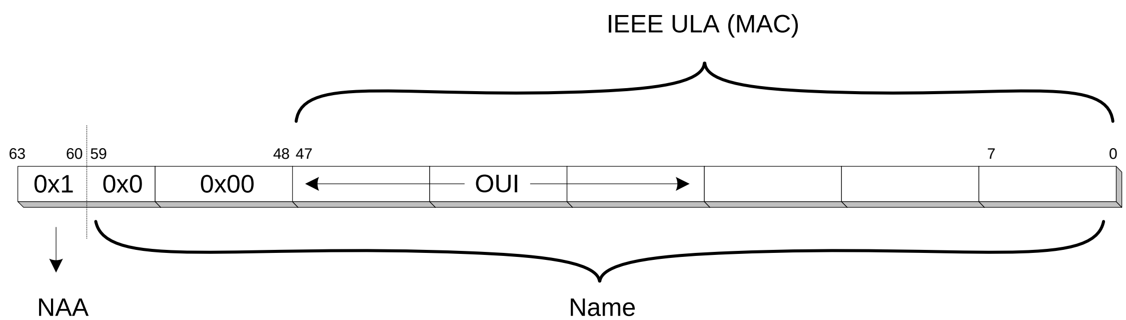

Figure 1-32: IEEE 48 Bit Address (NAA = 0x1)

A practical example of this format is demonstrated by the port WWN (pWWN) that is etched into the silicon of a particular Emulex Fibre Channel Host Bus Adapter, shown below in Figure 1-33 (p. 34).

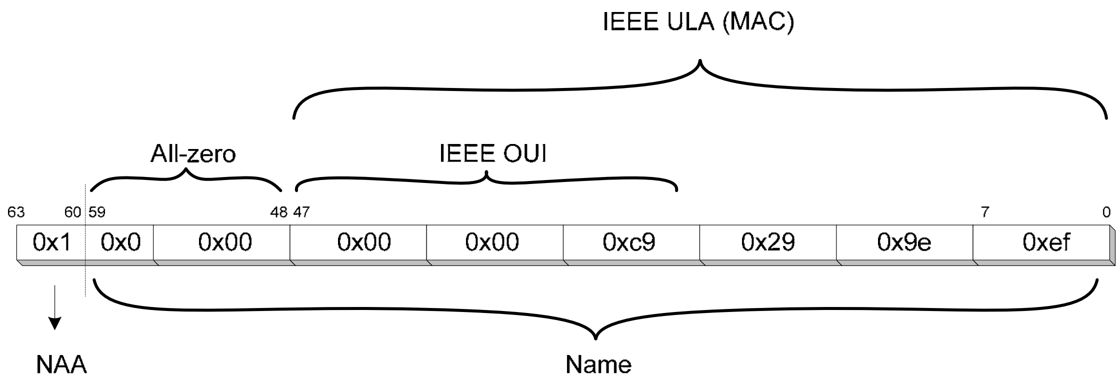

Figure 1-33: NAA = 0x1 Example: Emulex HBA pWWN

In this instance, one can observe that the Emulex's OUI is 00-00-c9, and that the number 29-9e-ef was exclusively assigned to this HBA by Emulex at the time of manufacture.

Note: *a detailed description of IEEE MAC address formats can be found in Appendix need reference; notably, the Ethernet-savvy reader can learn important distinctions between Ethernet MAC addresses, FC_IDs and Fibre Channel WWNs, and will also find a list of popular storage company OUIs.*

1.6.3 NAA = 0x2: IEEE Extended

When the NAA field is equal to 0b0010 (0x2), this indicates that the Name field complies with the IEEE Extended format. While similar to the IEEE 48 Bit Address format described in the previous section, the 12 previously unused bits may now be used by the manufacturer to identify a particular characteristic of the entity using this WWN format. Figure 1-34 (p. 34) illustrates this format as it applies to a Seagate Fibre Channel hard disk that was given a pWWN = 0x22:00:00:20:37:39:ab:f5 at the time of its manufacture.

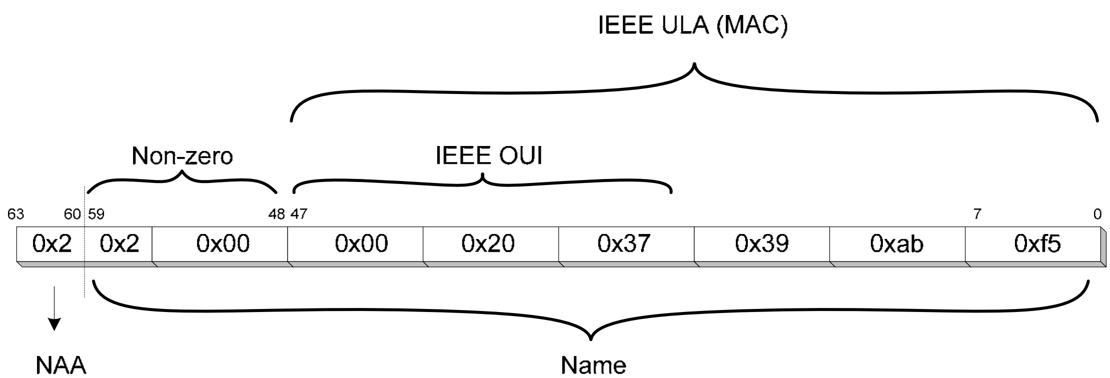

Figure 1-34: NAA = 0x2 Example: Seagate Fibre Channel Hard Disk Drive pWWN

Seagate's choice to use the IEEE Extended name format was their own; accordingly, the significance of 0x2 in the extended (non-zero) portion of the Company_ID is only known to them.

__Note:__ the term Company_ID is introduced at this stage to illustrate that Company_ID and Organizationally Unique Identifier (OUI) can be used interchangeably. A quick glance at the official list of OUIs maintained by the IEEE reveals that the difference is purely cosmetic: OUIs are represented as three hexadecimal-coded bytes separated by dashes (e.g. xx-yy-xx), while the corresponding Company_ID does away with the dashes and simply consists of a 24-bit hexadecimal number (e.g. xxyyzz). In practice, 802 committee standards use the term OUI while T11 committee texts use Company_ID, but both terms identify the same entity.

To better illustrate the practical use of the IEEE Extended format, the following example demonstrates how Cisco Systems assigns WWNs to each individual switch port using the IEEE Extended name format; effectively, each individual switch port is assigned a WWN with NAA = 0x2. The 12 non-zero bits from the previous example are used to number each individual port with respect to its physical location within the switch. In Figure 1-35 (p. 35), the "port index" field set to 0x001 indicates that this particular WWN is assigned to the first physical port installed on line card which is present in slot one in a particular chassis.

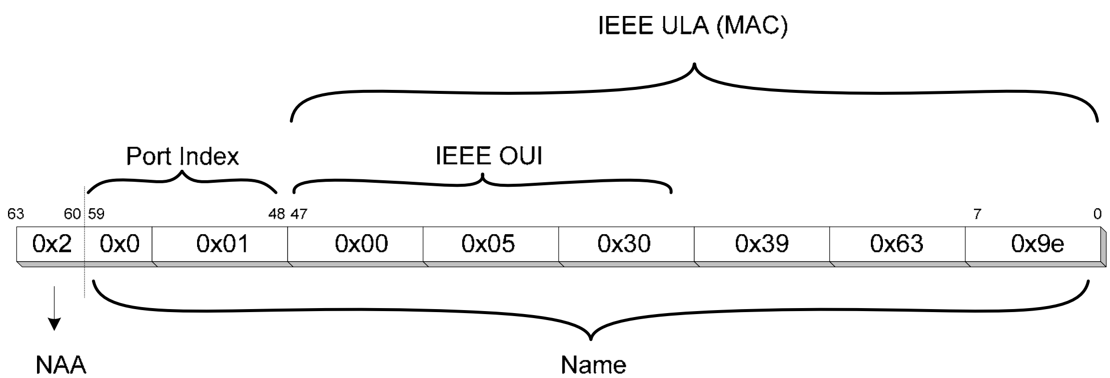

Figure 1-35: NAA = 0x2 Example: Cisco Systems WWN Assignment Scheme

The WWN allocation scheme in Cisco switches is such that each slot is assigned a pool of 64 available WWNs; these pools are distinguished by port index ranges (c.f. Figure 1-35). For example, the WWN of the second port in the third slot of the MDS 9509 chassis from the previous example would be in the form: 0x20:c2:00:05:30:39:63:9e. This is because the third slot is offset from the first slot by 3 × 64 WWNs, so the first port in slot three would be offset from the first port in slot one by 0xc0, which yields a port index of 0xc1. Being the next consecutive port on the same line card, port two would have a port index = 0xc2.

1.6.4 NAA = 0x3: Locally Assigned

While this subsection appears under a section titled *Fibre Channel World Wide Names (WWNs)*, it should be noted that Locally Assigned Name_Identifiers may not formally be called WWNs since local assignment cannot guarantee that a particular name is world-unique. With this understanding in mind, this format is nevertheless covered here to illustrate the use that is made of the Locally Assigned format in this book. Indeed, the reader may have already noticed that some WWNs that appear in various examples and figures have the NAA field set to 0x3, which indicates a Locally Assigned name. This format is illustrated below in Figure 1-36 (p. 36).

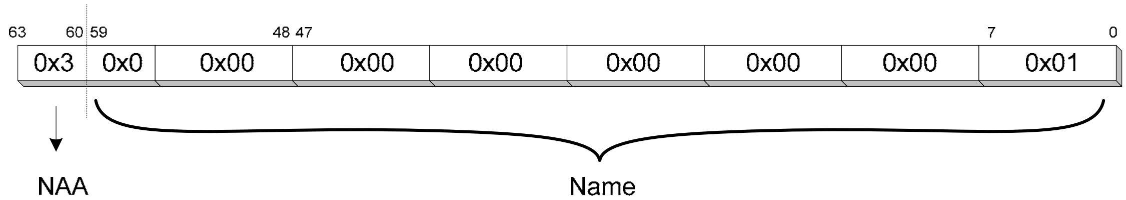

Figure 1-36: NAA = 0x3 Example: Locally Assigned Name_Identifier

The Locally Assigned format is leveraged in this book to improve the readability of potentially confusing Name_Identifiers. As a result, the Name_Identifiers used in Figure 1-30 (p. 31) are simple and descriptive; this would not have been the case if real-world WWNs had instead been used.

1.6.5 NAA = 0x5: IEEE Registered

Yet another WWN format supported by the Fibre Channel protocol, the IEEE Registered format is a slight variation on the IEEE Extended format covered previously. The IEEE OUI/ Company_ID occupies the bits immediately following the NAA, which is set to 0x5. This leaves bits 0-35 available for local assignment by the organization that owns the OUI. Figure 1-37 (p. 36) depicts a pWWN from an EMC Storage Array that is formatted as IEEE Registered.

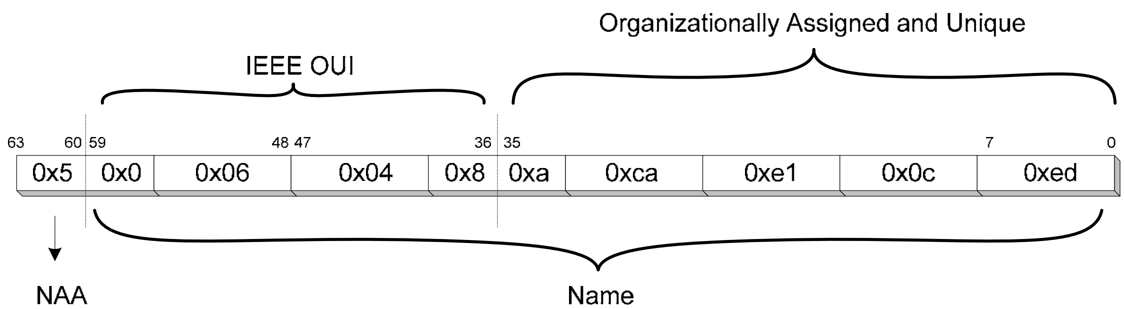

Figure 1-37: NAA = 0x5 Example: EMC Storage Array pWWN

There is no telling why EMC chooses to use the IEEE Registered format instead of the IEEE Extended format (nor is one better than the other – they both offer 36 bits of organizationally assignable bits, or Vendor Specified Identifier, VSID). From the perspective of a SAN administrator, the important thing is to understand where to find the OUI/Company_ID within the WWN when trying to identify a particular vendor. Section 1.6.7 (p. 37) summarizes the OUI position within various WWN formats.

1.6.6 Common Entities Identifiable By WWN

For the average SAN Administrator, the overwhelming majority of WWN encounters involve port WWNs (pWWNs), as they are commonly used to build zones, and also to build access control lists within storage arrays. In some cases, they are also used to enforce persistent binding (see need reference).

There are, however, many more entities within a SAN that are identified by WWNs. The following list cannot claim to be exhaustive since new entities may be added by vendors as they implement more features in their switches. For example, there was no such thing as a VSAN WWN (vWWN) before Cisco Systems invented this technology for their line of Fibre Channel switches. Furthermore, the standards bodies do not oversee the abbreviations, prefixes and suffixes that may be used to describe the application of a WWN to a particular entity. As an example, consider the term *Port World Wide Name*, which is often abbreviated as pWWN; some publications choose to call this a *World Wide Port Name*, and abbreviate it as *WWPN*. These terminological shortcomings should be kept in mind while reading Table 1-4 (p. 37).

Table 1-4: SAN Entities That Use WWNs As An Identifier

Entity	Abbreviation	Description
Host Port	pWWN	
Switch Port	pWWN (in context)	
Node	nWWN	
Switch	sWWN	
Fabric	fWWN	
VSAN	vWWN	
?		

1.6.7 Location Of The OUI/Company_ID

As a practical example of the various WWN formats that may be simultaneously at play, Table 1-5 (p. 38) shows the output of a command that summarizes the contents of the Fibre Channel Name Server database in a SAN which contains a variety of devices. This table graphically depicts the location of the OUI / Company_ID fields with respect to the format type (NAA value). OUIs are similarly positioned for NAA values 0x1 and 0x2, and NAA = 0x5 formats place the OUI in a different position (shifted 12 bits to the left).

Table 1-5: OUI / Company_ID Location For Various WWN Formats

NAA = 0x1, 0x2

Company	OUI	Company_ID
Cisco	00-0d-ec	000dec
Emulex	00-00-c9	0000c9
QLogic	00-e0-8b	00e08b
Seagate	00-20-37	0020-37

```
sjc7-9509-6# sh fcns database

VSAN 1:
--------------------------------------------------------------------------
FCID        TYPE             PWWN                   (VENDOR)     FC4-TYPE:FEATURE
--------------------------------------------------------------------------
0x7e0005    N      10:00:00:00:c9:29:9e:ef    (Emulex)     scsi-fcp:init
0x6a0005    N      10:00:00:00:c9:38:df:53    (Emulex)     scsi-fcp:init
0x630004    N      10:00:00:0d:ec:02:2d:c5    (Cisco)      ipfc
0x630300    N      21:00:00:e0:8b:04:5c:20    (Qlogic)     scsi-fcp:init
0x6a0200    N      21:00:00:e0:8b:05:48:e9    (Qlogic)     scsi-fcp:init
0x630100    N      21:00:00:e0:8b:0d:d4:c8    (Qlogic)     scsi-fcp:init
0x6a000b    N      22:00:00:0d:ec:02:2d:82    (Cisco)      virtual:volume_owner
0x630007    N      22:00:00:0d:ec:02:2d:c2    (Cisco)      virtual:volume_owner
0x6401b6    NL     22:00:00:20:37:18:1a:39    (Seagate)    scsi-fcp:target
0x6401c9    NL     22:00:00:20:37:38:8b:a7    (Seagate)    scsi-fcp:target
0x6401c6    NL     22:00:00:20:37:39:8c:ea    (Seagate)    scsi-fcp:target
0x6401b5    NL     22:00:00:20:37:39:ab:f5    (Seagate)    scsi-fcp:target
0x6401ba    NL     22:00:00:20:37:39:ae:1c    (Seagate)    scsi-fcp:target
0x6401b9    NL     22:00:00:20:37:46:06:ff    (Seagate)    scsi-fcp:target
0x6401c5    NL     22:00:00:20:37:46:07:df    (Seagate)    scsi-fcp:target
0x6401c7    NL     22:00:00:20:37:46:34:96    (Seagate)    scsi-fcp:target
0x6401bc    NL     22:00:00:20:37:9c:5b:5e    (Seagate)    scsi-fcp:target
0x6a000c    N      22:02:00:0d:ec:02:2d:82    (Cisco)      virtual:volume_owner
0x6a000d    N      22:03:00:0d:ec:02:2d:82    (Cisco)      virtual:volume_owner
0x6a0008    N      50:06:04:82:ca:e1:0c:d2    (EMC)        scsi-fcp    253
0x6a000a    N      50:06:04:82:ca:e1:0c:ec    (EMC)        scsi-fcp    253
0x6a0007    N      50:06:04:8a:ca:e1:0c:ed    (EMC)        scsi-fcp    253
0x6a0400    N      50:06:0b:00:00:14:63:12    (HP)         scsi-fcp
0x630400    N      50:06:0b:00:00:14:70:d8    (HP)         scsi-fcp:target
```

NAA = 0x5

Company	OUI	Company_ID
EMC	00-60-48	006048
HP	00-60-b0	0060b0

Note: The output shown above in Table 1-5 is sorted by WWN to group like-NAA values together. Normal switch output behavior actually sorts by FC_ID.

1.7 Fibre Channel Login Services

In the microseconds that follow a physical connection, the instant a Fibre Channel port detects light and synchronizes itself with its link partner, very little is known about the nature of the connection that has just been established. At the very least, a Fibre Channel node must determine if it is connected to another Nx_Port (point to point or arbitrated loop), or if it is connected to the Fx_Port of a switch (i.e. to a fabric). When this is the case, an N_Port must acquire an FC_ID and share such parameters as Supported Class of Service, BB_Credits, WWNs and Max Receive Field lengths, to name a few, with the switch to which it is connecting. This process establishes the communication parameters between the N_Port and the F_Port.

When the N_Port is a SCSI initiator, it attempts to discover if the fabric contains SCSI target N_Ports. When one or more target ports are discovered, the initiator must also establish communication parameters with the target port, much like it did when it initially connected to the fabric as described in the previous paragraph.

Finally, the actual SCSI process in the initiator needs to establish an “image pair” with the SCSI process in the target port before the first SCSI command may finally be issued.

The formal exchange of all the parameters discussed above is accomplished by invoking Fibre Channel Login Services. Specifically, the following three functions are defined as Extended Link Services:

- Fabric Login (FLOGI);
- Port Login (PLOGI);
- Process Login (PRLI).

As a subset of Extended Link Services, these functions are described in the standards document *Fibre Channel – Link Services (FC-LS)*. This section summarizes the basic functionality of the three functions listed above, and highlights important similarities between FLOGI and PLOGI when compared to PRLI.

1.7.1 FLOGI And PLOGI Login Service Parameters

Before delving into specifics of the various Login functions, it is important to understand that the parameters which are exchanged during FLOGI and PLOGI are very similar. So similar, in fact, that they share a common payload format which is used by both functions. A simplified view of this format is illustrated in Figure 1-38 (p. 39).

Common Service Parameters	WWNs	Class 1 and 6 Parameters	Class 2 Parameters	Class 3 Parameters	Class 4 Parameters	Vendor Version	Extended Login Parameters

Figure 1-38: FLOGI and PLOGI Payload Fields

When the FLOGI or PLOGI service is invoked, the complete set of fields shown above is appropriately populated and sent by the originator to the responder. The responder replies to the service invocation by sending a similar set of fields, populated with responder-specific values, back to the originator.

Each field shown in Figure 1-38 contains well-defined sub-fields which may or may not apply to the specific function being requested. This is well understood by the protocol, and such sub-fields are ignored when this is the case.

The following sections summarize the important sub-fields that are exchanged during FLOGI and PLOGI operations, and also deal with service-specific functions.

1.7.2 Fabric Login (FLOGI)

Once an Nx_Port has been connected, enabled, and has acquired bit and word synchronization, it must determine the nature of the remote port to which it is connected (i.e. Nx_Port or Fx_Port). As this book focuses on Fibre Channel fabrics, this section covers the events that follow a connection to a switch F_Port.

When an N_Port logs into a fabric, it initiates a Fibre Channel exchange and sends an FLOGI Extended Link Service request to the directly connected switch. The frame header fields indicates that the ELS is destined for D_ID = 0xff ff fe, which is the WKA of the local switch port; the S_ID is initially set to 0x00 00 00, which is an indication that that the N_Port is also requesting that the switch assign it a valid FC_ID.

The N_Port also includes the following key values in the FLOGI payload:

- Common Service Parameters
  - ⇒ BB_Credits
  - ⇒ Virtual Fabrics
  - ⇒ N_Port / F_Port
  - ⇒ Security (FC-SP)
  - ⇒ R_T_TOV (Receiver – Transmitter)
  - ⇒ Buffer to Buffer Receive Data Field Size
- WWNs
  - ⇒ pWWN
  - ⇒ nWWN
- Class 3 Parameters
  - ⇒ Class Valid
  - ⇒ Sequential Delivery

Note: The list of FLOGI parameters shown above, and included in Figure 1-39, is not exhaustive. Rather, it summarizes those parameters considered essential for a basic understanding of the login process.

The complete process is summarized in Figure 1-39 (p. 41).

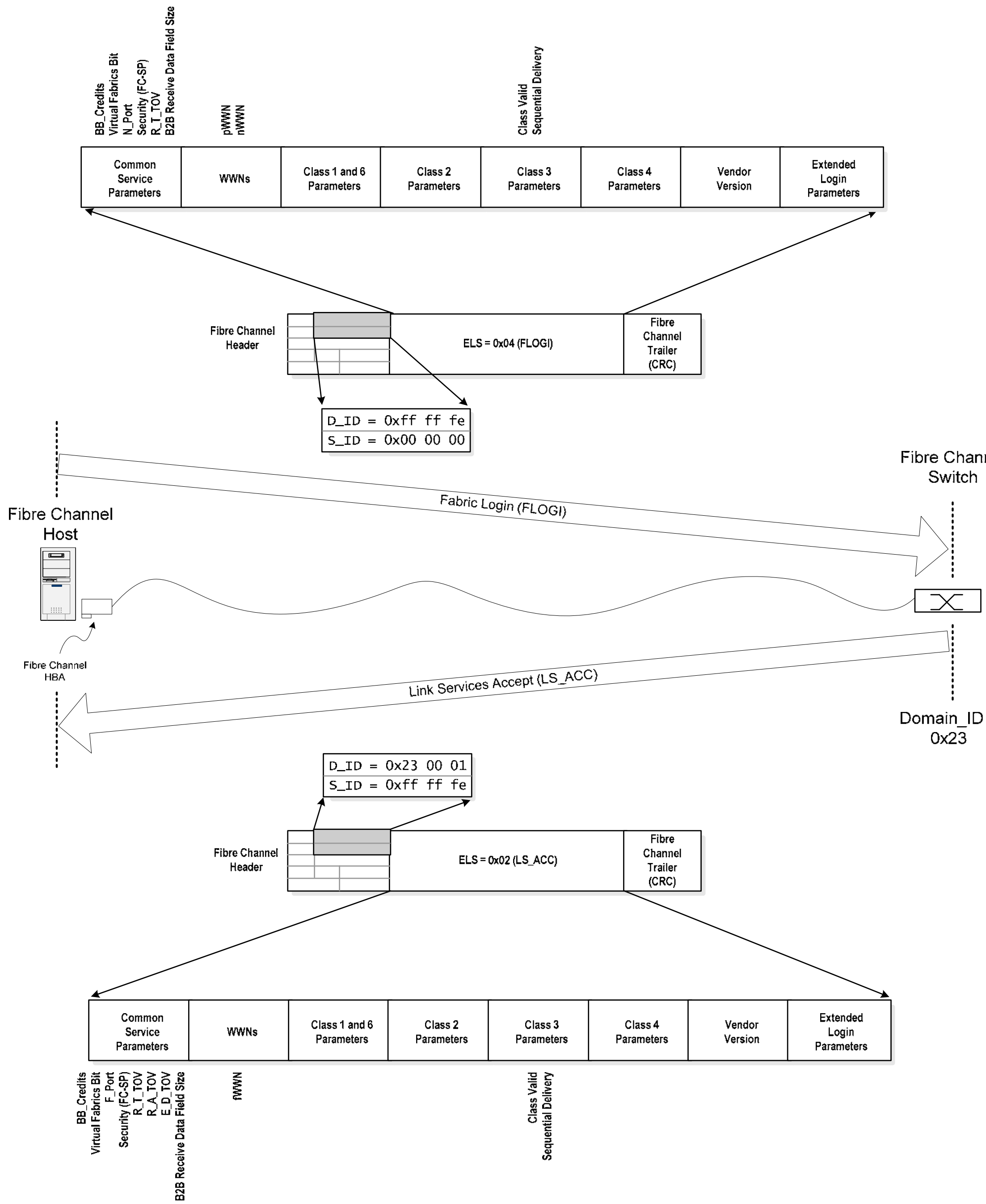

Figure 1-39: Fabric Login (FLOGI) Summary

Upon receiving the FLOGI ELS, the F_Port replies with a Link Service Accept (LS_ACC) sequence, which contains the same payload format as the FLOGI request. The D_ID field in the Fibre Channel frame header contains the FC_ID assigned by the switch to this N_Port. The F_Port populates the payload fields with values that apply to an LS_ACC response. Specifically:

- Common Service Parameters

- ⇒ BB_Credits
- ⇒ Virtual Fabrics
- ⇒ N_Port / F_Port
- ⇒ Security (FC-SP)
- ⇒ R_T_TOV (Receiver – Transmitter)
- ⇒ R_A_TOV (Resource Allocation)
- ⇒ E_D_TOV (Error Detection)
- ⇒ Buffer to Buffer Receive Data Field Size

- WWNs
  - ⇒ fWWN
- Class 3 Parameters
  - ⇒ Class Valid
  - ⇒ Sequential Delivery

When a fabric switch replies to an N_Port by sending an LS_ACC, it also registers the new N_Port in the Fibre Channel Name Server with a minimal amount of information, namely: FC_ID, pWWN and nWWN. However, N_Ports usually register themselves with the Name Server as well and, by doing so, usually include such detailed information as their protocol capabilities, symbolic node names, IP addresses, etc. A comprehensive section covering the Fibre Channel Name Server can be found at need reference.

Note: When an N_Port attempts to connect with the Name Server, it must direct a PLOGI operation to FC_ID 0xff ff fc, which is the WKA of the Name Server. PLOGI is explained in the next section.

When an N_Port that is also a SCSI initiator has successfully logged into a fabric, it usually attempts to discover fabric-attached N_Ports that are SCSI targets. In other words, it starts looking for storage. This is accomplished by querying the Name Server for a list of Nx_Ports that are logged in and, optionally, fulfill a particular criterion, such as “SCSI target”. This process is explained in need reference; for the purposes of this section, however, the reader may assume that a SCSI initiator already knows which Nx_Ports are SCSI targets. Indeed, an initiator N_Port will attempt a PLOGI with every discovered N_Port that is believed to be a SCSI target. In fact, some HBAs try to PLOGI into every single N_Port they learn about, initiator or target, and proceed to discover the port’s capabilities by using discovery capabilities provided by a subset of the Process Login (PRLI) Extended Link Service. This behavior may lead to problems that are exposed in need reference (i.e. single initiator zoning).

1.7.3 Port Login (PLOGI)

When an N_Port that is also a SCSI initiator discovers the existence of N_Ports which are believed to be SCSI targets, that initiator will attempt to log into the discovered target ports. As in the case of FLOGI, PLOGI is an Extended Link Service. The format of the payload, shown in Figure 1-38 (p. 39) is identical to that used with FLOGI, but as mentioned previously, only PLOGI-specific fields are used when the payload format is used to convey PLOGI information (note that some fields are used for both FLOGI and PLOGI operations).

The purpose of performing a PLOGI is to establish communication parameters with a specific N_Port peer. Therefore, an N_Port must maintain as many parameter sets as it has peers. For example, if an initiator logs into three different target ports, it must then maintain and abide by three separate sets of parameters.

Despite the fact that Figure 1-39 illustrates a host computer logging into a fabric switch, it should be clearly understood that storage ports must also perform a FLOGI to obtain their fabric operating parameters. Only initiators, however, attempt PLOGIs to other Nx_Ports. An exception to this rule is the name registration of target ports with the Fibre Channel name server. – a comprehensive treatment of this topic is offered in need reference.

When an N_Port logs into another N_Port, the D_ID field in the PLOGI ELS frame header contains the FC_ID of the destination port (indeed, the ultimate goal of the discovery process described earlier is to acquire this FC_ID so that it may subsequently be used in frame headers). The S_ID field in the frame header is set to the initiator's FC_ID, which was previously assigned by the FLOGI process.

In the PLOGI payload, the N_Port includes the following PLOGI-specific key values:

- Common Services Parameters
  - ⇒ BB_Credits
  - ⇒ N_Port / F_Port
  - ⇒ Security (FC-SP)
  - ⇒ R_T_TOV (Receiver – Transmitter)
  - ⇒ Buffer to Buffer Receive Data Field Size
  - ⇒ E_D_TOV
- WWNs
  - ⇒ pWWN
  - ⇒ nWWN
- Class 3 Parameters
  - ⇒ Class Valid
  - ⇒ Sequential Delivery

A summary of the Process Login execution is illustrated in Figure 1-40 (p. 44).

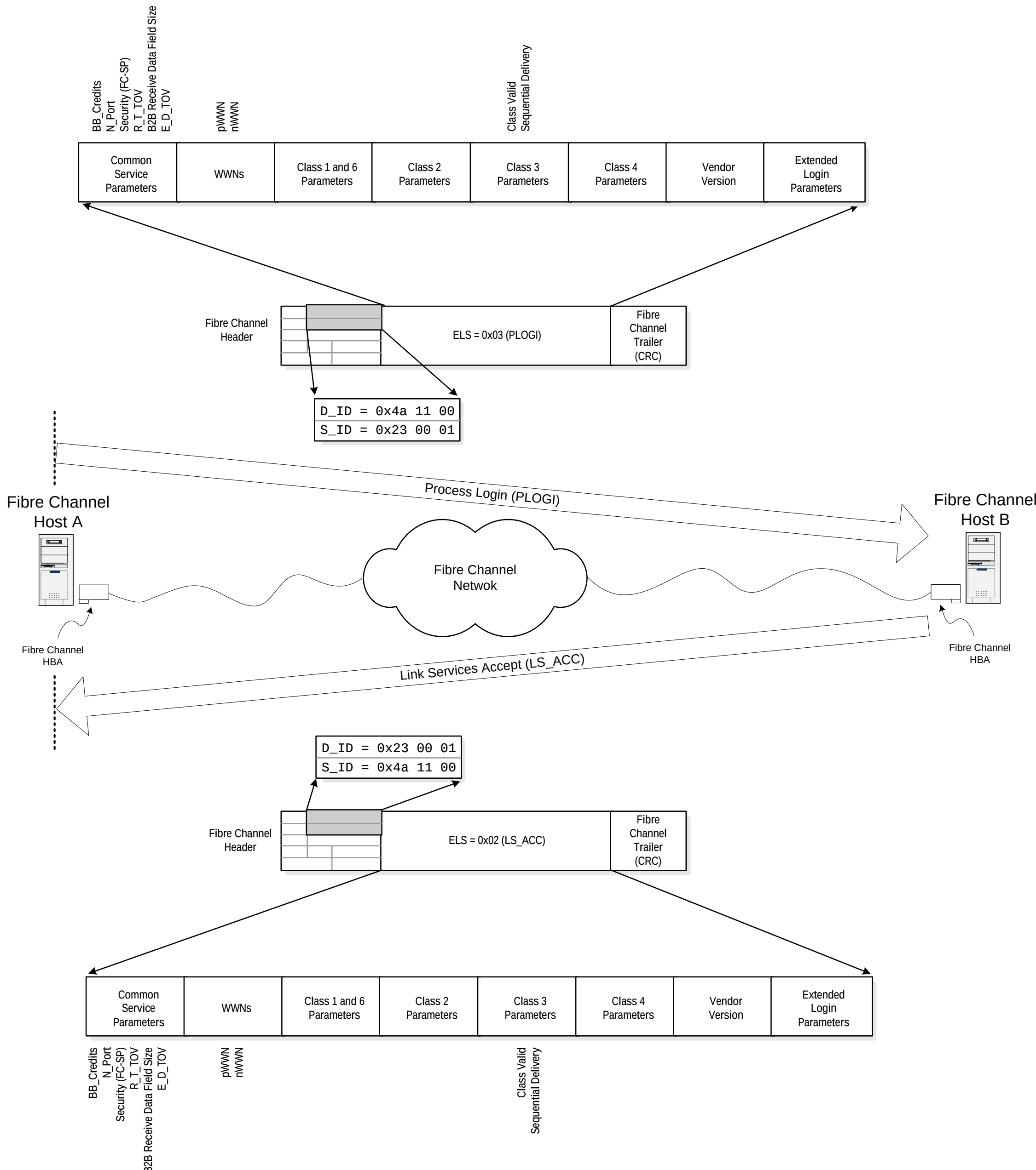

Figure 1-40: Port Login (PLOGI) Summary

The expected response to a PLOGI is similar to that received from an FLOGI request: the LS_ACC ELS is formatted according to the common payload format described previously, and PLOGI-specific fields are supplied in response to a Port Login request.

1.7.4 Process Login (PRLI)

Whereas the FLOGI and PLOGI operations share many similarities, the Process Login Extended Link Service (PRLI ELS) fundamentally differs in the following fashion: the payload of a PRLI is mostly determined by the upper layer protocol placing the process login request. In other words, a PRLI request is issued by the FC-4 protocol which is attempting to use the services provided by Fibre Channel to communicate with a peer protocol in a remote station. As it pertains to storage networks, this usually translates to a Process Login from the SCSI-FCP layer of an initiator into the SCSI-FCP layer of a Fibre Channel-attached target device. The basic format of the PRLI payload is shown in Figure 1-41 (p. 45).

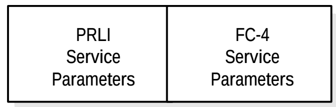

Figure 1-41: PRLI Payload Fields

The reader will recall that many FC-4 ULPs may use Fibre Channel to communicate with peer FC-4s in remote nodes. This is important when considering the payload of the PRLI request, which is essentially divided in two parts:

- A common portion defined in FC-LS, contains PRLI Service Parameters that are common to all PRLI requests. The fundamental parameters that are common to all FC-4 protocols are:
  - ⇒ Type Code (for example, SCSI-FCP = 0x08);
  - ⇒ Establish Image Pair.
- An FC-4 specific portion which is defined in the standard documents that govern the use of a particular ULP. The fundamental parameters that apply to the SCSI-FCP mapping are defined in the T10 Standard titled: *Fibre Channel Protocol for SCSI (SCI-FCP)*. They are:
  - ⇒ Initiator Function;
  - ⇒ Target Function.

***Note**: as in the case of FLOGI and PLOGI, the list of parameters shown above is partial, and only includes those parameters which are deemed essential to a clear understanding of basic Process Login functionality. A comprehensive description of all parameters is included in the advanced treatment of Process Login in **Error! Reference source not found.**, **Error! Reference source not found.**, (p. **Error! Bookmark not defined.**).*

A sample PRLI exchange is shown in Figure 1-42 (p. 46).

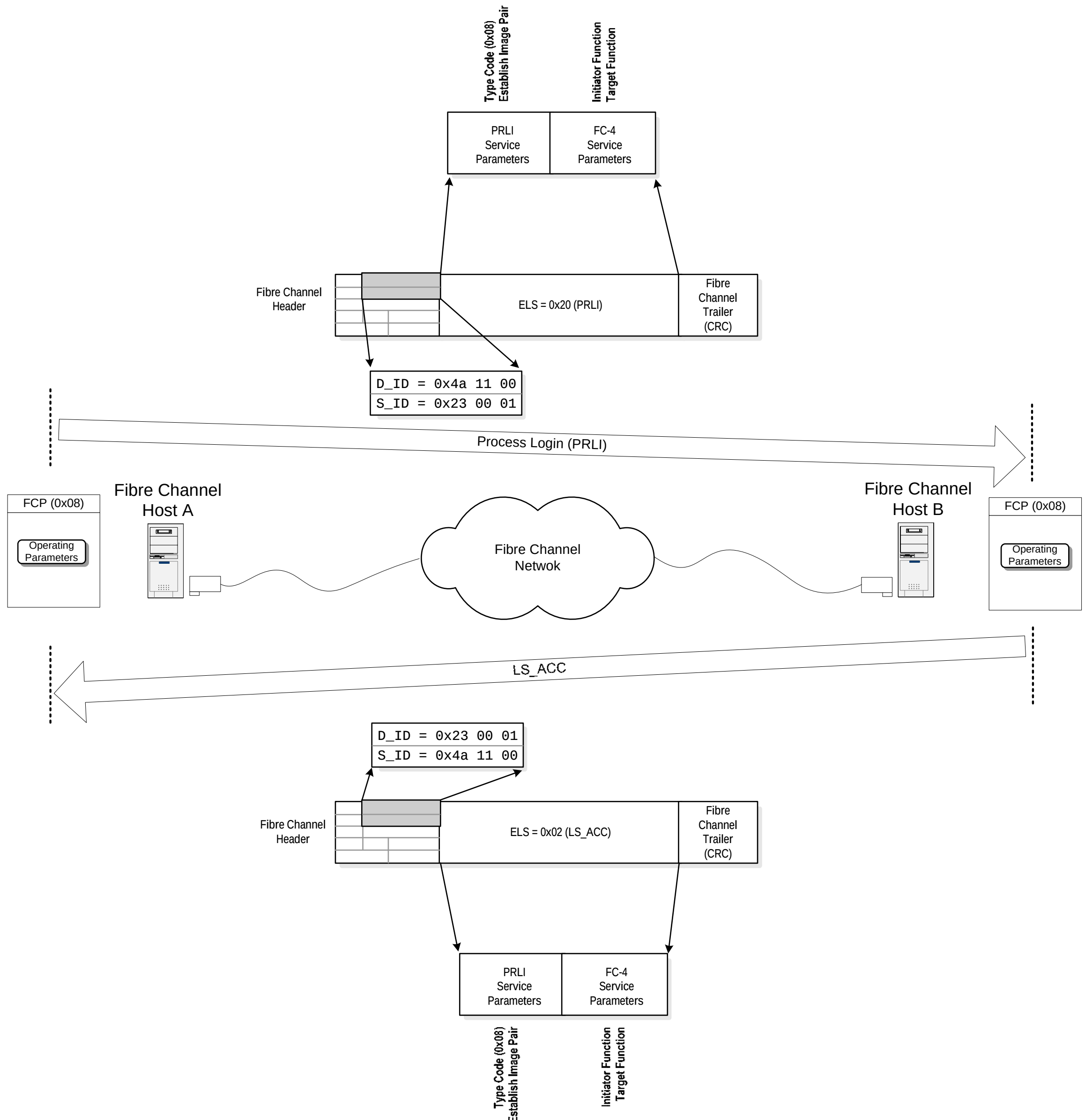

Figure 1-42: Process Login (PRLI) Summary

As seen previously, the LS_ACC ELS is used to acknowledge all Extended Link Service requests, including a PRLI.

At this juncture, the discovery capability offered by the PRLI function is of particular interest. Indeed, by setting the “Establish Image Pair” bit to zero (0), a Fibre Channel host indicates that it is not requesting an actual login into the peer process, but wishes to discover the values associated with certain FC-4 parameters. In the specific case of SCSI-FCP, it is common for initiators to issue PRLI requests to every single Fibre Channel host that is reported by the Name Server, and set the “Establish Image Pair” bit set to zero. By consulting the results of the “Target Function” bit returned in the LS_ACC, an initiator may then perform a full Process Login (with the “Establish Image Pair” bit set to one (1)) only with those devices which reported target capability.

1.8 Summary

This chapter covers four fundamental components of Fibre Channel:

- the Fibre Channel frame and the mechanisms by which it is transmitted over a physical link;
- the sequence and exchange logical structures used to track ULP IUs which are being exchanged by end nodes;
- the basic Fibre Channel addressing structure;
- the Fibre Channel Name function, which usually manifests itself as WWNs;
- basic concepts surrounding Login Services such as FLOGI, PLOGI and PRLI.