Copies of this document may be purchased from:
Global Engineering, 15 Inverness Way East,
Englewood, CO 80112-5704
Phone: (800) 854-7179 or (303) 792-2181 Fax: (303) 792-2192

X3.xxx-199x
X3T11/Project 959-D/Rev 3.0

FIBRE CHANNEL

SWITCH FABRIC

(FC-SW)

REV 3.0

X3 working draft proposed
American National Standard
for Information Technology

February 3, 1997

Secretariat:
Information Technology Industry Council

ABSTRACT:

NOTE:
This is a possible future draft proposed National Standard of Accredited Standards Committee X3. As such, this is not a completed document. The X3T11 Technical Committee or anyone else may modify this document as a result of comments received anytime, or during a future public review and its eventual approval as a Standard.

POINTS OF CONTACT:

Roger Cummings (X3T11 Chairman)
Distributed Processing Technology
140 Candace Drive
Maitland, FL 32751
Phone: (407) 830-5522 x348
Fax: (407) 260-5366
EMail: cummings_roger@dpt.com

Ed Grivna (X3T11 Vice Chairman)
Cypress Semiconductor
2401 East 86th Street
Bloomington, MN 55425
(612) xxx-xxxx
Fax: (612) 851-5087
E-Mail: elg@cypress.com

I. Dal Allan
(Fibre Channel Working Group Chairman)
ENDL
14426 Black Walnut Court
Saratoga, CA 95070
(408) 867-6630
Fax: (408) 867-2115
E-Mail: dal.allan@mcimail.com

Jeffrey Stai (Technical Editor)
Brocade Communications Systems, Inc.
15707 Rockfield Boulevard, Suite 215
Irvine, CA 92618
(714) 455-2908
Fax: (714) 455-9287
E-mail: stai@brocadesouth.com

Editor’s Notes, revision 3.0:

-

ANSI ®
X3.xxx-199x

draft proposed American National Standard
for Information Technology

Fibre Channel —
Switch Fabric (FC-SW)

Secretariat

Information Technology Industry Council

Approved ,199

American National Standards Institute, Inc.

Abstract

This report selects and restricts logical options from the Fibre Channel Physical and Signalling, Fibre Channel Protocol for SCSI, Fibre Channel Arbitrated Loop, Fibre Channel Switch, and Small Computer Systems Interface standards, such that any device complying with this report should interoperate. This report addresses options for devices that are both loop-attached to the fabric and direct-attach to the fabric.

American National Standard

Approval of an American National Standard requires verification by ANSI that the requirements for due process, consensus, and other criteria for approval have been met by the standards developer.

Consensus is established when, in the judgement of the ANSI Board of Standards Review, substantial agreement has been reached by directly and materially affected interests. Substantial agreement means much more than a simple majority, but not necessarily unanimity. Consensus requires that all views and objections be considered, and that a concerted effort be made towards their resolution.

The use of American National Standards is completely voluntary; their existence does not in any respect preclude anyone, whether he has approved the standards or not, from manufacturing, marketing, purchasing, or using products, processes, or procedures not conforming to the standards.

The American National Standards Institute does not develop standards and will in no circumstances give interpretation on any American National Standard. Moreover, no person shall have the right or authority to issue an interpretation of an American National Standard in the name of the American National Standards Institute. Requests for interpretations should be addressed to the secretariat or sponsor whose name appears on the title page of this standard.

CAUTION NOTICE: This American National Standard may be revised or withdrawn at any time. The procedures of the American National Standards Institute require that action be taken periodically to reaffirm, revise, or withdraw this standard. Purchasers of American National Standards may receive current information on all standards by calling or writing the American National Standards Institute.

Published by

American National Standards Institute
11 W. 42nd Street, New York, New York 10036

Printed in the United States of America

INSERT CODE HERE

Contents

Page

Page

Figures

Page

Page

Tables

Page

draft proposed American National Standard for Information Technology—

Fibre Channel — Switch Fabric (FC-SW)

1 Introduction and Scope

This American National Standard for FC-SW specifies tools and algorithms for interconnection and initialization of Fibre Channel switches to create a multi-switch Fibre Channel Fabric. This Standard defines an E_Port ("Expansion Port") that operates in a manner similar to an N_Port and F_Port, as defined in ANSI X3.230 FC-PH, with additional functionality provided for interconnecting switches.

This Standard also defines how ports that are capable of being an E_Port, F_Port, and/or FL_Port may discover and self-configure for their appropriate operating mode. Once a port establishes that it is connected to another switch and is operating as an E_Port, an address assignment algorithm is executed to allocate port addresses throughout the Fabric.

This Standard does not define credit models and management between E_Ports for the various Classes of Service other than Class F. Broadcast and multicast services are not defined. E_Ports conforming to this Standard support Class F, and also Class 1, Class 2, and/or Class 3; support for other Classes of Service are not defined by this Standard. The method by which routing of frames is established and effected is not described.

2 Normative references

The following Standards contain provisions which, through reference in the text, constitute provisions of this Standard. At the time of publication, the editions indicated were valid. All Standards are subject to revision, and parties to agreements based on this Standard are encouraged to investigate the possibility of applying the most recent editions of the Standards listed below.

Copies of the following documents can be obtained from ANSI: Approved ANSI Standards, approved and draft international and regional Standards (ISO, IEC, CEN/CENELEC, ITUT), and approved and draft foreign Standards (including BSI, JIS, and DIN). For further information, contact ANSI Customer Service Department at 212-642-4900 (phone), 212-302-1286 (fax) or via the World Wide Web at http://www.ansi.org.

Additional availability contact information is provided below as needed.

2.1 Approved references

[1] ANSI X3.230-1994, *Fibre Channel Physical and Signaling Interface (FC-PH)*.

[2] ANSI X3.272-1996, *Fibre Channel Arbitrated Loop (FC-AL).*

[3] ANSI X3.288-1996, *Fibre Channel - Generic Services (FC-GS).*

[4] ANSI X3.289-1996, *Fibre Channel - Fabric Generic (FC-FG).*

2.2 References under development

At the time of publication, the following referenced Standards were still under development. For information on the current status of the document, or regarding availability, contact the relevant Standards body or other organization as indicated.

NOTE – For more information on the current status of a document, contact the X3 Secretariat at the address listed in the front matter. To obtain copies of this document, contact Global Engineering at the address listed in the front matter, or the X3 Secretariat.

[5] ANSI X3.297-199x, *Fibre Channel - Physical and Signalling Interface-2 (FC-PH-2)*, X3T11/Project 901D/Rev 7.4

[6] ANSI X3.303-199x, *Fibre Channel - Physical and Signalling Interface-3 (FC-PH-3)*, X3T11/Project 1119D/Rev 9.0

[7] ANSI X3.xxx-199x, *Fibre Channel Arbitrated Loop (FC-AL-2),* X3T11/Project 1133D/Rev 5.2

[8] ANSI X3.xxx-199x, *Fibre Channel - Generic Services-2 (FC-GS-2)*, X3T11/Project 1134D/Rev xx

[9] ANSI X3.xxx-199x, *Fibre Channel - Fabric Loop Attachment (FC-FLA)*, X3T11/Project 1235DT/Rev 2.5

2.3 Other references

All of the following profiles are available from the Fibre Channel Association (FCA), 12407 MoPac Expressway North 100-357, P. O. Box 9700, Austin, TX 78758-9700; (800) 272-4618 (phone); or via e-mail, FCA-Info@amcc.com.

[10] FCSI-101*, FCSI Common FC-PH Feature Sets Used in Multiple Profiles*, Rev 3.1

[11] *FCA N_Port to F_Port Interoperability Profile,* Rev 1.0

3 Definitions and conventions

For FC-SW, the following definitions, conventions, abbreviations, acronyms, and symbols apply.

3.1 Definitions

3.1.1 address assignment: A process whereby addresses are dispensed to Switches and Switch Ports.

3.1.2 address identifier: As defined in FC-PH (see reference [1])**,** an unsigned 24-bit address value used to uniquely identify the source (S_ID) and destination (D_ID) of Fibre Channel frames.

3.1.3 Address Manager: A logical entity within a Switch which is responsible for address assignment.

3.1.4 Area: As defined in FC-FG (see reference [4]), the second level in a three-level addressing hierarchy.

3.1.5 Area Address Manager: A Switch which is responsible for address assignment to other Switches within a single Domain.

3.1.6 Area Identifier: As defined in FC-FG (see reference [4]), bits 15 through 8 of an address identifier.

3.1.7 byte: A group of eight bits.

3.1.8 Class F service: As defined in FC-FG (see reference [4]), a service which multiplexes frames at frame boundaries that is used for control and coordination of the internal behavior of the Fabric.

3.1.9 Class N service: A generic reference to a Class 1, Class 2, or Class 3 service, as defined in FC-PH (see reference [1]).

3.1.10 Domain: As defined in FC-FG (see reference [4]), the highest level in a three-level addressing hierarchy.

3.1.11 Domain Address Manager: A Principal Switch which is responsible for address assignment to other Switches outside of its Domain.

3.1.12 Domain Identifier: As defined in FC-FG (see reference [4]), bits 23 through 16 of an address identifier.

3.1.13 Domain_Map: A bitmap in which each bit corresponds to a Domain_ID value (see 6.2.3).

3.1.14 downstream Principal ISL: From the point of view of the local Switch, the downstream Principal ISL is the Principal ISL to which frames may be sent from the the Principal Switch to the destination Switch. All Principal ISLs on the Principal Switch are downstream Principal ISLs. A Switch that is not the Principal Switch may have zero or more downstream Principal ISLs.

3.1.15 E_Port: As defined in FC-FG (see reference [4]), a Fabric “Expansion” Port which attaches to another E_Port to create an Inter-Switch Link.

3.1.16 E_Port Identifier: An address identifier assigned to an E_Port.

3.1.17 E_Port_Name: A Name_Identifier which identifies an E_Port for identification purposes. The format of the name is specified in FC-PH. Each E_Port shall provide a unique E_Port_Name within the Fabric.

3.1.18 Error_Detect_Timeout value: A time constant defined in FC-PH. In this Standard, the recommended value of this time constant is 2 seconds.

3.1.19 F_Port: As defined in FC-PH (see reference [1]). In this Standard, an F_Port is assumed to always refer to a port to which non-loop N_Ports are attached to a Fabric, and does not include FL_Ports.

3.1.20 Fabric: As defined in FC-FG (see reference [4]), an entity which interconnects various Nx_Ports attached to it and is capable of routing frames using only the D_ID information in an FC-2 frame header.

3.1.21 Fabric Controller: 1. As defined in FC-FG (see reference [4]), the logical entity responsible for operation of the Fabric. 2. The entity at the well-known address hex ‘FF FF FD’.

3.1.22 Fabric Element: 1. As defined in FC-FG (see reference [4]), the smallest unit of a Fabric which meets the definition of a Fabric. From the point of view of an attached Nx_Port, a Fabric consisting of multiple Fabric Elements is indistinguishable from a Fabric consisting of a single Fabric Element.

3.1.23 Fabric F_Port: The entity at the well-known address hex ‘FF FF FE’. See reference [1].

3.1.24 FL_Port: An L_Port which is able to perform the function of an F_Port, attached via a link to one or more NL_Ports in an Arbitrated Loop topology (see FC-AL). The AL_PA of an FL_Port is hex’00’. In this Standard, an FL_Port is assumed to always refer to a port to which NL_Ports are attached to a Fabric, and does not include F_Ports.

3.1.25 Fx_Port: A Switch Port operating as an F_Port or FL_Port.

3.1.26 Fabric_Stability_Timeout value: A time constant used to detect inactivity during Fabric Configuration. The value of this time constant shall be 5 seconds.

3.1.27 Inter-Switch Link: A Link connecting the E_Port of one (local)Switch to the E_Port of another (remote) Switch.

3.1.28 Isolated: A condition in which it has been determined that no Class N traffic may be transmitted across an ISL.

3.1.29 L_Port: A port which contains Arbitrated Loop functions associated with the Arbitrated Loop topology.

3.1.30 Link: As defined in FC-PH.

3.1.31 local Switch: A Switch that can be reached without traversing any Inter-Switch Links.

3.1.32 Loop Fabric Address: An address identifier used to address a loop for purposes of loop management.

3.1.33 N_Port: As defined in FC-PH (see reference [1]). In this Standard, an N_Port is assumed to always refer to a direct Fabric-attached port, and does not include NL_Ports.

3.1.34 N_Port Identifier: An address identifier assigned to an N_Port.

3.1.35 Name_Identifier: As defined in FC-PH (see reference [1]), a 64-bit identifier.

3.1.36 NL_Port: An L_Port which is able to perform the function of an N_Port, attached via a link to one or more NL_Ports and zero or more FL_Ports in an Arbitrated Loop topology. In this Standard, an NL_Port is assumed to always refer to a loop-attached port, and does not include N_Ports.

3.1.37 Nx_Port: A Switch Port operating as an N_Port or NL_Port.

3.1.38 path: A route between a source and a destination.

3.1.39 path selection: A process whereby a path between a source and one or more destinations is discovered.

3.1.40 Port: 1. A generic reference to an N_Port, NL_Port, F_Port, FL_Port, or E_Port. 2. As defined in FC-FG (see reference [4]), the lowest level in a three-level addressing hierarchy.

3.1.41 Port Identifier: As defined in FC-FG (see reference [4]), bits 7 through 0 of an address identifier.

3.1.42 Port Mode: A generic reference to E_Port, F_Port or FL_Port operation.

3.1.43 Preferred Domain_ID: A Domain_ID previously granted to a Switch by the Domain Address Manager.

3.1.44 Principal ISL: An Inter-Switch Link that is used to communicate with the Principal Switch.

3.1.45 Principal Switch: A Switch which has been selected to perform certain duties.

3.1.46 remote Switch: A Switch that can be reached only by traversing one or more Inter-Switch Links.

3.1.47 Resource_Allocation_Timeout value: A time constant defined in FC-PH. In this Standard, the recommended value of this time constant is 10 seconds.

3.1.48 Router: An entity within a Switch responsible for routing of Class 2 and Class 3 frames.

3.1.49 routing: A process whereby the appropriate Switch Port(s) to deliver a Class 2 or Class 3 frame towards its destination is identified.

3.1.50 Switch: 1. A Fabric Element conforming to this Standard. 2. A member of the Fabric collective. Resistance is futile...

3.1.51 Switch_Name: A Name_Identifier which identifies a Switch for identification purposes. The format of the name is specified in FC-PH. Each Switch shall provide a unique Switch_Name within the Fabric.

3.1.52 Switch Port: An E_Port, F_Port, or FL_Port.

3.1.53 upstream Principal ISL: From the point of view of the local Switch, the upstream Principal ISL is the Principal ISL to which frames may be sent from the local Switch to the Principal Switch. A Switch that is not the Principal Switch always has exactly one upstream Principal ISL. The Principal Switch does not have an upstream Principal ISL.

3.2 Editorial conventions

In this Standard, a number of conditions, mechanisms, sequences, parameters, events, states, or similar terms that do not have their normal English meaning are printed with the following conventions:

- the first letter of each word in uppercase and the rest lowercase (e.g., Exchange, Class, etc.).

- a term consisting of multiple words, with the first letter of each word in uppercase and the rest lowercase, and each word separated form the other by an underscore (_) character. A word may consist of an acronym or abbreviation which would be printed in uppercase. (e.g., NL_Port, Transfer_Length, etc.).

– a term consisting of multiple words with all letters lowercase and each word separated form the other by a dash (-) character. A word may also consist of an acronym or abbreviation which would be printed in uppercase. (e.g., device-level, CUE-with-busy, etc.).

All terms and words not conforming to the conventions noted above have the normal technical English meanings.

Numbered items in this Standard do not represent any priority. Any priority is explicitly indicated.

In all of the figures, tables, and text of this Standard, the most significant bit of a binary quantity is shown on the left side. Exceptions to this convention are indicated in the appropriate sections.

The term “shall” is used to indicate a mandatory rule. If such a rule is not followed, the results are unpredictable unless indicated otherwise.

The fields or control bits which are not applicable shall be reset to zero.

If a field or a control bit in a frame is specified as not meaningful, the entity which receives the frame shall not check that field or control bit.

If a field or control bit is specified as reserved, it shall be filled with binary zeros by the source, and shall be ignored by the destination.

Temporary: Anything in “{ }” is an editor’s note indicating some unresolved issue.

3.2.1 Binary notation

Binary notation may be used to represent some fields. Single bit fields are represented using the binary values 0 and 1. For multiple bit fields, the binary value is enclosed in single quotation marks followed by the letter b. For example, a four-byte Process_Associator field containing a binary value may be represented as ‘00000000 11111111 10011000 11111010’b.

3.2.2 Hexadecimal notation

Hexadecimal notation may be used to represent some fields. When this is done, the value is en-closed in single quotation marks and preceded by the word hex. For example, a four-byte Process_ Associator field containing a binary value of ‘00000000 11111111 10011000 11111010’b is shown in hexadecimal format as hex’00 FF 98 FA’.

3.3 Abbreviations, acronyms, and symbols

Abbreviations and acronyms applicable to this International Standard are listed. Definitions of several of these items are included in 3.1. Abbreviations used that are not listed below are defined in FC-PH (see reference [1]).

3.3.1 Acronyms and abbreviations

AAM	Area Address Manager
AM	Address Manager
Area_ID	Area Identifier
BLS	Basic Link Service
DAM	Domain Address Manager
Domain_ID	Domain Identifier
E_D_TOV	Error_Detect_Timeout value

ELS	Extended Link Service
FC-AL	Fibre Channel Arbitrated Loop, reference [2]
FC-AL-2	Fibre Channel Arbitrated Loop-2, reference [7]
FC-FG	Fibre Channel - Fabric Generic, reference [4]
FC-FLA	Fibre Channel - Fabric Loop Attachment, reference [9]
FC-GS	Fibre Channel - Generic Services, reference [3]
FC-GS-2	Fibre Channel - Generic Services-2, reference [8]
FC-PH	Fibre Channel Physical and Signaling Interface, reference [1]
FC-PH-2	Fibre Channel Physical and Signaling Interface-2, reference [5]
FC-PH-3	Fibre Channel Physical and Signaling Interface-3, reference [6]
F_S_TOV	Fabric_Stability_Timeout value
ISL	Inter-Switch Link
IU	Information Unit
LAN	Local Area Network
LFA	Loop Fabric Address
Port_ID	Port Identifier
R_A_TOV	Resource_Allocation_Timeout value
SI	Sequence Initiative
SW_ACC	Switch Fabric Link Service Accept
SW_LS	Switch Fabric Link Service
SW_RJT	Switch Fabric Link Service Reject
ULP	Upper Level Protocol
WKA	Well-Known Address
WWN	World Wide Name

3.3.2 Symbols

Unless indicated otherwise, the following symbols have the listed meaning.

||	concatenation

4 Structure and Concepts

This clause provides an overview of a Switch-based Fabric.

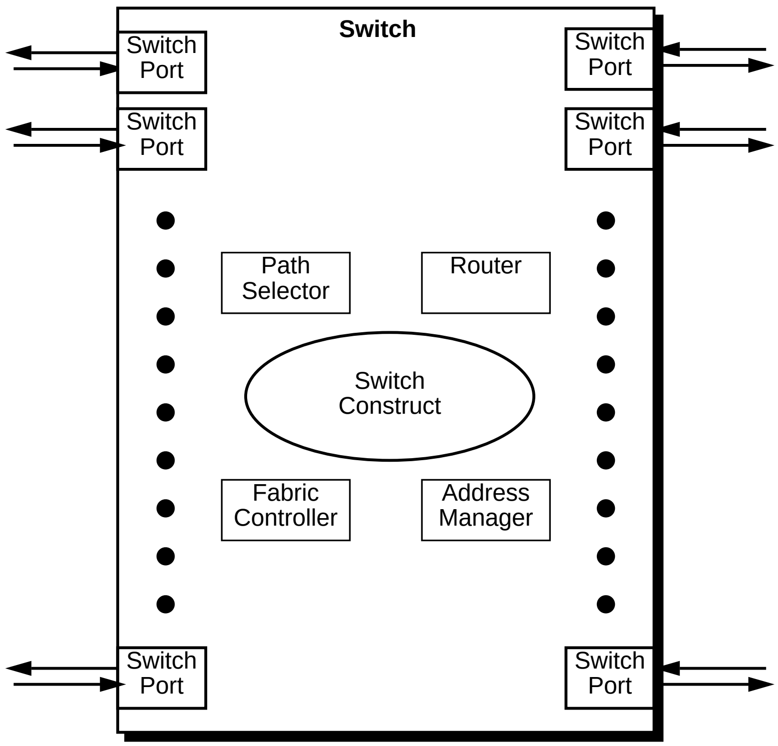
4.1 Fabric

A Fabric is a transport that provides switched interconnect between N_Ports. The general model of a Fibre Channel Fabric is defined in FC-FG, reference [4].

4.2 Switch

A Switch is the smallest entity that can function as a Switch-based Fibre Channel Fabric. Figure 1 illustrates the conceptual model of a Switch.

Figure 1 – Switch Model



A Switch is composed of the following major components:

- Three or more Switch Ports;
- a Switch Construct, capable of either multiplexed frame switching or circuit switching, or both;
- an Address Manager;
- a Path Selector, which performs path selection;
- a Router;

– and a Fabric Controller.

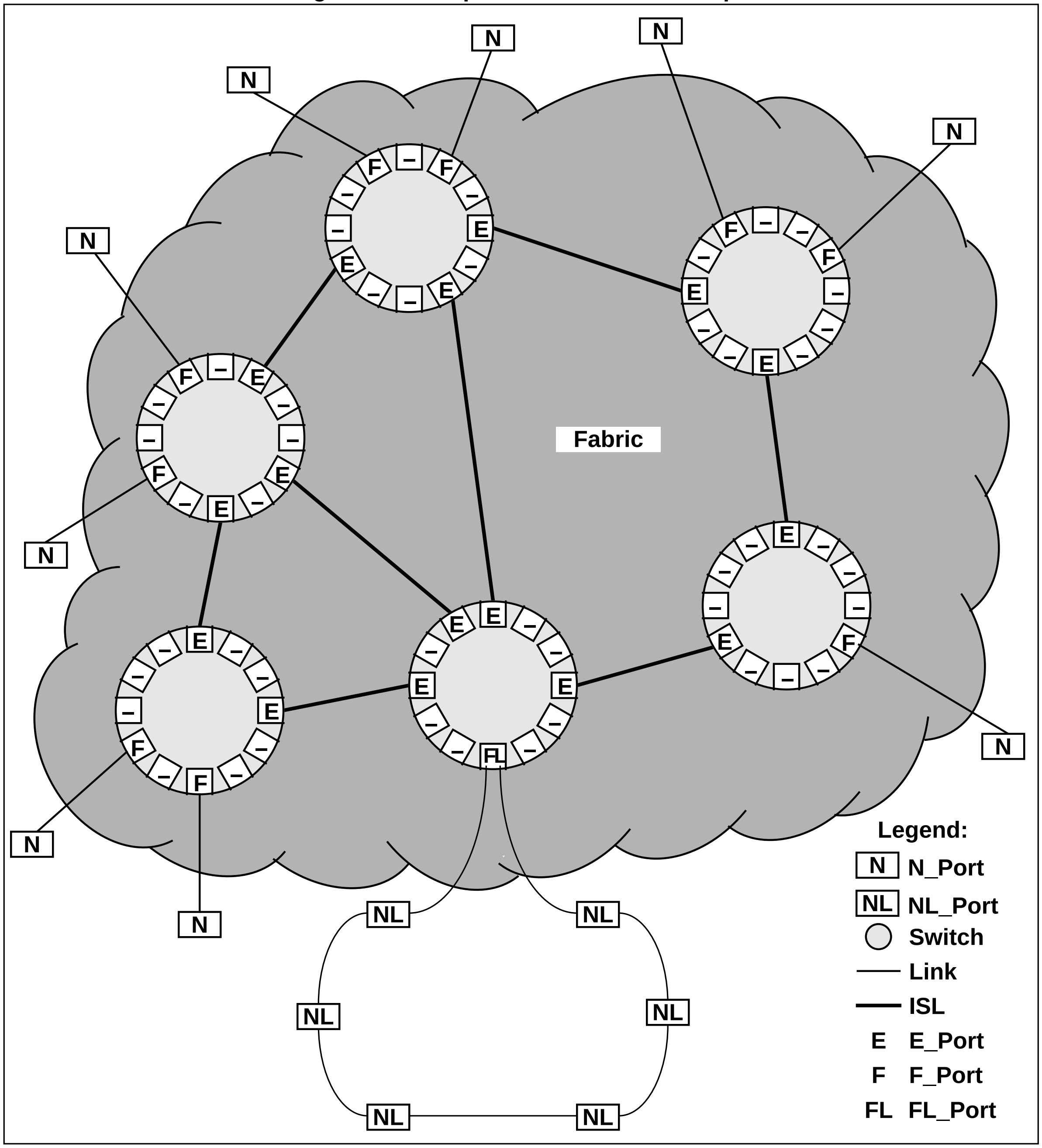
As defined, a Switch Port may be either an E_Port, an F_Port, or an FL_Port. A Switch Port that is capable of assuming more than one of these roles is called a multi-function Switch Port. Once a Switch Port assumes a role, via the Switch Port Initialization Procedure, it shall remain in that role until an event occurs that causes re-initialization.

The Link joining a pair of E_Ports is called an Inter-Switch Link (ISL). E_Ports conforming to this Standard use FC–PH compliant media, coding and data rates to form an ISL.

ISLs carry frames originating from the Node Ports and those generated within the Fabric. The frames generated within the Fabric serve as control, management and support for the Fabric.

Switches may be joined freely or in a structured fashion to form a larger Fabric, as illustrated in Figure 2.

Figure 2 – Multiple Switch Fabric Example



The structure of the Switch Construct in the Switch, as seen in figure 1, is undefined and beyond the scope of this Standard. It may support either or both circuit switching and multiplexed frame switching. It may be non-blocking, allowing concurrent operation of all possible combinations or it may be blocking, restricting operations. The Switch Construct may also contain redundancy, as may be required for high availability configurations.

The Address Manager is responsible for the assignment of addresses within some portion of the Fabric. Within the Switch, the Address Manager is responsible for acquiring a Domain and Area for the Switch, and allocating Port_IDs within the Domain and Area.

The Path Selector is a logical entity that establishes frame routing paths.

The Router is a logical entity that performs the routing of Class 2 and Class 3 frames to their final destination.

The Fabric Controller is a logical entity that performs the management of the Switch. The Fabric Controller has the characteristics of an N_Port, though it may or may not be attached to the Fabric via a Link.

4.3 Switch Topologies

Switch topologies are defined in FC-FG, reference [4].

4.4 Switching characteristics

Path, circuit, switching and frame routing within a Switch may occur synchronously or asynchronously to the current word alignment of the outbound fibre.

Synchronous switching guarantees retention of the established word alignment on the outbound fibre of the Switch Port. Asynchronous switching does not guarantee retention of word alignment on the outbound fibre of the Switch Port.

A Switch may employ either synchronous or asynchronous switching or a combination of the two (e.g., a Switch may use synchronous switching for Class F, Class 2 and Class 3, and asynchronous switching for Class 1). However, a Switch shall never mix the two within a given Class of Service.

A switching event occurs every time a connection less frame is transmitted and when a connection based service is established, suspended or terminated. Frame Intermixing and interjecting also constitute switching events.

4.4.1 Synchronous switching

Synchronous switching associated with connectionless frame routing and connection oriented Dedicated Connections or virtual connection Services shall guarantee the word alignment on the outbound fibre.

Switches shall ensure that synchronous switching only occurs between frames. Switches should use synchronous switching in support of Class 2, Class 3 and Class F service.

4.4.2 Asynchronous switching

Asynchronous switching may be performed any time Fill Words are being transmitted. Bit alignment and word alignment may be lost when an asynchronous switching event occurs. A recovery time that allows the attached Port time to regain synchronization shall be inserted before frame transmission resumes for the outbound fibre. Fill Words shall be transmitted during this recovery time. If conditions arise warranting transmission of a Primitive Sequence, then this should take precedence over transmission of Fill Words.

If a Switch or Node Port recognizes that it is linked to a Switch which employ asynchronous switching, and a permissible word realignment event occurs, then the Port may discount any resulting errors, i.e. not log errors resulting from the realignment event.

4.5 Switch Ports

A Switch shall have three or more Switch Ports. A Switch equipped only with F_Ports or FL_Ports forms a non-expandable Fabric. To be part of an expandable Fabric, a Switch shall incorporate at least one Switch Port capable of E_Port operation.

A Switch Port supports one or more of the following Port Modes: E_Port, F_Port, FL_Port. A Switch Port that is capable of supporting more than one Port Mode attempts to configure itself first as an FL_Port (as defined in FC-AL), then as an E_Port (as defined in in this Standard), and finally as an F_Port (as defined in FC-PH), depending on which of the three Port Modes are supported by the Switch Port.

The detailed procedure is described in 7.1.

4.5.1 F_Port

An F_Port is the point at which all frames originated by an N_Port enter the Fabric, and all frames destined for an N_Port exit the Fabric. An F_Port may also be the Fabric entry point for frames originated by an N_Port destined for an internal Fabric desitnation, such as the Fabric Controller. Similarly, an F_Port may also be the Fabric exit point for frames originated internal to the Fabric and destined for an N_Port. Frames shall not be communicated across a Link between an F_Port and anything other than an N_Port.

F_Ports are described in detail in 5.1.

4.5.2 FL_Port

An FL_Port is the point at which all frames originated by an NL_Port enter the Fabric, and all frames destined for an NL_Port exit the Fabric. An FL_Port may also be the Fabric entry point for frames originated by an NL_Port destined for an internal Fabric desitnation, such as the Fabric Controller. Similarly, an FL_Port may also be the Fabric exit point for frames originated internal to the Fabric and destined for an NL_Port. Frames shall not be communicated across a Link between an FL_Port and anything other than an NL_Port.

FL_Ports are described in detail in 5.2.

4.5.3 E_Port

An E_Port is the point at which frames pass between the Switches within the Fabric. Frames with a destination other than the local Switch or any N_Port or NL_Port attached to the local Switch exit the local Switch through an E_Port. Frames that enter a Switch via an E_Port are forwarded to a local destination, or are forwarded towards their ultimate destination via another E_Port. Frames shall not be communicated across a Link between an E_Port and anything other than an E_Port.

E_Ports are described in detail in 5.3.

4.6 Fabric Addressing

Switches use the address partitioning model described in FC-FG (Annex A), and as described below. The 24-bit address identifier is divided into three fields: Domain, Area, and Port, as shown in figure 3.

Figure 3 – Domain, Area, and Port Address Partitioning

23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
Domain_ID								Area_ID								Port_ID							
Address Identifier																							

A Domain is one or more Switches that have the same Domain_ID for all N_Ports and NL_Ports within or attached to those Switches, except for Well-Known Addresses. If there is more than one Switch in the Domain, any Switch within the Domain shall be directly connected via an ISL to at least one other Switch in the same Domain. {fyi, this rule is here for Area Controllers}

An Area_ID shall apply to either of the following:

– One or more N_Ports or E_Ports within and attached to a single Switch, except for Well-Known Addresses; or,

– an Arbitrated Loop of NL_Ports attached to a single FL_Port.

A single Arbitrated Loop shall have exactly one Area_ID.

A Port_ID shall apply to either of the following:

– a single N_Port or E_Port within a Domain/Area, except for Well-Known Addresses; or,

– the valid AL_PA of a single NL_Port or FL_Port on an Arbitrated Loop.

Address identifier values for this Standard are listed in table 1. Any value listed as Reserved is not meaningful within this Standard.

Table 1 – Address Identifier Values

Address Identifier (hex)			Description
Domain_ID	**Area_ID**	**Port_ID**	
00	00	00	Undefined (note 1)
00	00	AL_PA	E_Port: Reserved F_Port: Reserved FL_Port: Private Loop NL_Port (note 2)
00	00	non-AL_PA	Reserved
00	01 - FF	00 - FF	Reserved
01 - EF	00	00 - FF	Reserved
01 - EF	01 - FF	00	E_Port: E_Port Identifier (note 4) F_Port: N_Port Identifier (note 4) FL_Port: Loop Fabric Address (note 3)
01 - EF	01 - FF	AL_PA	E_Port: E_Port Identifier (note 4) F_Port: N_Port Identifier (note 4) FL_Port: N_Port Identifier for Public Loop NL_Port (note 3)

Table 1 – Address Identifier Values

Address Identifier (hex)			Description
Domain_ID	Area_ID	Port_ID	
01 - EF	01 - FF	non-AL_PA	E_Port: E_Port Identifier (note 4) F_Port: N_Port Identifier (note 4) FL_Port: Reserved
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 (note 5)
FF	FC	F0 - FF	Reserved
FF	FD - FE	00 - FF	Reserved
FF	FF	00 - EF	Reserved
FF	FF	F0 - FC	Well-Known Address (note 6)
FF	FF	FD	N_Port Identifier for Fabric Controller (note 7)
FF	FF	FE	N_Port Identifier for Fabric F_Port
FF	FF	FF	Well-Known Address (note 6)

Notes:

1 This value is used by an N_Port requesting an address identifier during FLOGI.

2 See FC-AL for a definition of AL_PA and FC-FLA for a definition of Private Loop and FL_Port operation with Private Loop devices.

3 See FC-FLA for the definition and use of Loop Fabric Address, and for a definition of Public Loop.

4 In FC-FG, the Area_ID range F0-FF is reserved for “Fabric Assisted Functions”, whatever that means.

5 A Domain Controller identifier may be used to address the Fabric Controller of a remote Switch that is not directly connected via an ISL to the originating Switch. The Port_ID field is set to the Domain_ID of the remote Switch.

6 The usage of Well-Known Addresses hex’FFFFF0’ through hex’FFFFFC’, and hex’FFFFFF’, are not defined by this Standard.

7 This address identifier has special usage depending on the originator. If the originator is an attached external N_Port or NL_Port (attached via an F_Port or FL_Port) then the destination of a frame sent to hex’FFFFFD’ is the Fabric Controller of the local Switch. If the originator is the Fabric Controller of the local Switch, then the destination of a frame sent to hex’FFFFFD’ via an ISL is the Fabric Controller of the remote Switch at the other end of the ISL.

4.7 Class F Service

Class F service is a connectionless service very similar to Class 2 that is used for internal control of the Fabric. Class F service as defined by this Standard differs in some ways from the definition in FC-FG. Class F service as used by this Standard is defined in 5.4.

4.8 Relationship Between this Standard and FC-FG

FC-FG defined the generic requirements for all Fabrics, independent of the specific type or topology. Many issues were appropriately left open for definition by later Fabric Standards specific to certain types and topologies.

In the process of defining the Switch Fabric, some items that were defined in FC-FG were found that required modification for use in this Standard.

In cases where this Standard and FC-FG conflict, this Standard shall take precedence.

5 Switch Ports

This clause defines the specific behaviors of all modes of Switch Port. Note that the models described below are defined for purposes of describing behavior. No implication is made as to whether the actual implementation of an element is in hardware or software. An element may be implemented on a per-Port basis, or may be a logical entity that is embodied in a single physical implementation shared by multiple ports.

5.1 F_Port Operation

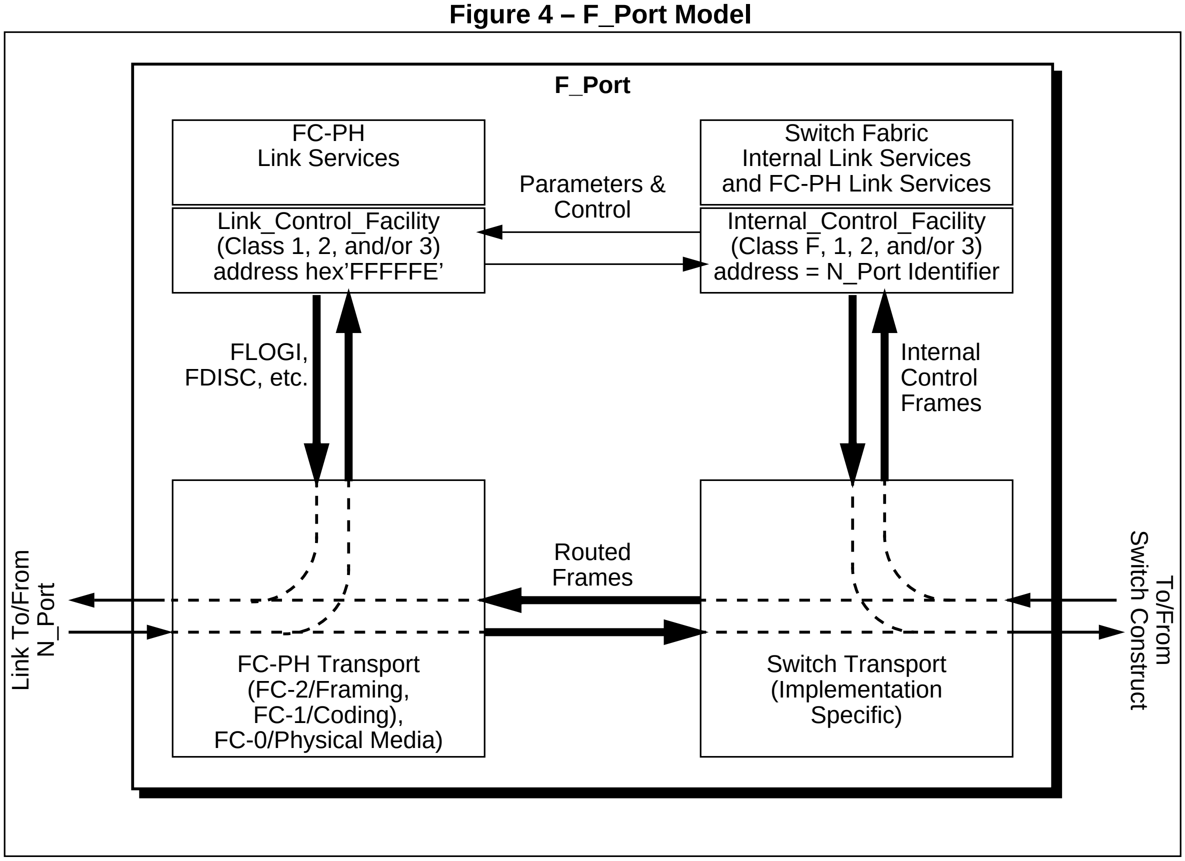
An F_Port is the point at which an external N_Port is attached to the Fabric. It normally functions as a conduit to the Fabric for frames transmitted by the N_Port, and as a conduit from the Fabric for frames destined for the N_Port.

An F_Port shall support one or more of the following Classes of service: Class 1 service, Class 2 service, Class 3 service. An F_Port shall not transmit Class F frames on its outbound fibre, nor shall an F_Port admit to the Fabric Class F frames or Primitive Sequences or Primitive Signals other than Idle received on its inbound fibre.

5.1.1 Model

The model of an F_Port is shown in figure 4.

Figure 4 – F_Port Model



An F_Port contains an FC-PH Transport element through which passes all frames and Primitives transferred across the Link to and from the N_Port. Frames received from the N_Port are either directed to the Switch Construct via the Switch Transport element, or directed to the

Link_Control_Facility.The Link_Control_Facility receives frames related to Link Services such as FLOGI, and transmits responses to those Link Service frames.

Frames received from the FC-PH Transport element that are destined for other ports are directed by the Switch Transport to the Switch Construct for further routing. Frames received from the Switch Construct by the Switch Transport are directed either to the FC-PH Transport for transmission to the N_Port, or to the Internal_Control_Facility. The Internal_Control_Facility receives frames related to Switch Fabric Internal Link Services, and transmits responses to those Internal Link Services frames. Information is passed between the Internal_Control_Facility and the Link_Control_Facility to effect the control and configuration of the Transport elements.

5.1.2 Link Behavior

The F_Port Link is used by Switches to transmit and receive frames with a single Node. A Link to an F_Port always connects to exactly one N_Port.

An F_Port Link follows the FC-0, FC-1, and FC-2 protocols defined for point-to-point Links as defined in FC-PH.

5.2 FL_Port Operation

An FL_Port is the point at which one or more external NL_Ports are attached to the Fabric. It normally functions as a conduit to the Fabric for frames transmitted by the attached NL_Ports, and as a conduit from the Fabric for frames destined for the attached NL_Ports.

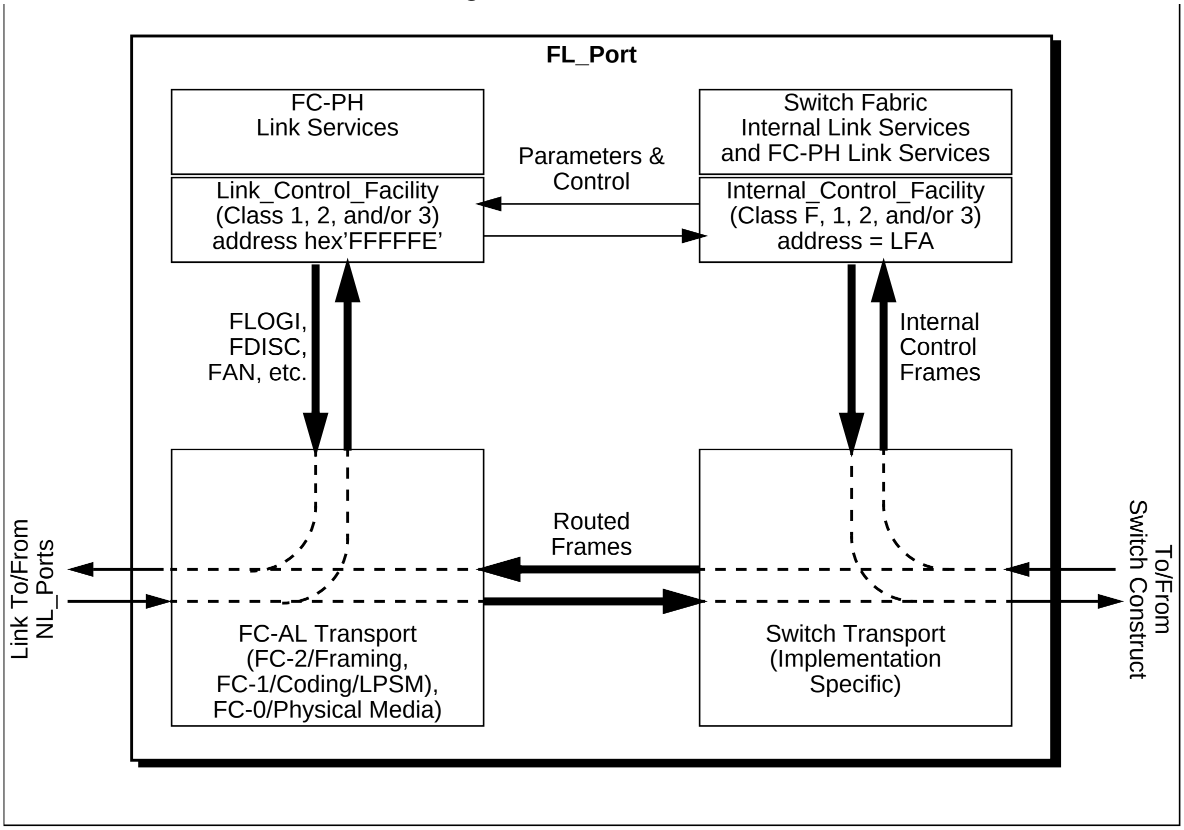
An FL_Port shall support one or more of the following Classes of service: Class 1 service, Class 2 service, Class 3 service. An FL_Port shall not transmit Class F frames on its outbound fibre, nor shall an FL_Port admit to the Fabric Class F frames or Primitive Sequences or Primitive Signals other than Idle received on its inbound fibre.

An FL_Port that conforms to this Standard shall conform to the FL_Port requirements defined in FC-FLA (reference [9]). {I bet I can’t say this, can I?}

5.2.1 Model

The model of an FL_Port is shown in figure 5.

Figure 5 – FL_Port Model



An FL_Port contains an FC-AL Transport element through which passes all frames and Primitives transferred across the Link to and from the multiple NL_Ports. Frames received from the NL_Ports are either directed to the Switch Construct via the Switch Transport element, or directed to the Link_Control_Facility.The Link_Control_Facility receives frames related to Link Services such as FLOGI, and transmits responses to those Link Service frames. The Link_Control_Facility also transmits and receives Loop Initialization Sequences and transmits the FAN ELS.

Frames received from the FC-AL Transport element that are destined for other ports are directed by the Switch Transport to the Switch Construct for further routing. Frames received from the Switch Construct by the Switch Transport are directed either to the FC-AL Transport for transmission to the destination NL_Port, or to the Internal_Control_Facility. The Internal_Control_Facility receives frames related to Switch Fabric Internal Link Services and Loop management Extended Link Services (see FC-FLA), and transmits responses to those Link Services frames. Information is passed between the Internal_Control_Facility and the Link_Control_Facility to effect the control and configuration of the Transport elements.

5.2.2 Link Behavior

The FL_Port Link is used by Switches to transmit and receive frames with multiple Nodes. A Link to an FL_Port connects to one or more NL_Ports.

An FL_Port Link follows the FC-0, FC-1, and FC-2 protocols defined in FC-PH, with the additional Arbitrated Loop protocols defined in FC-AL.

5.3 E_Port Operation

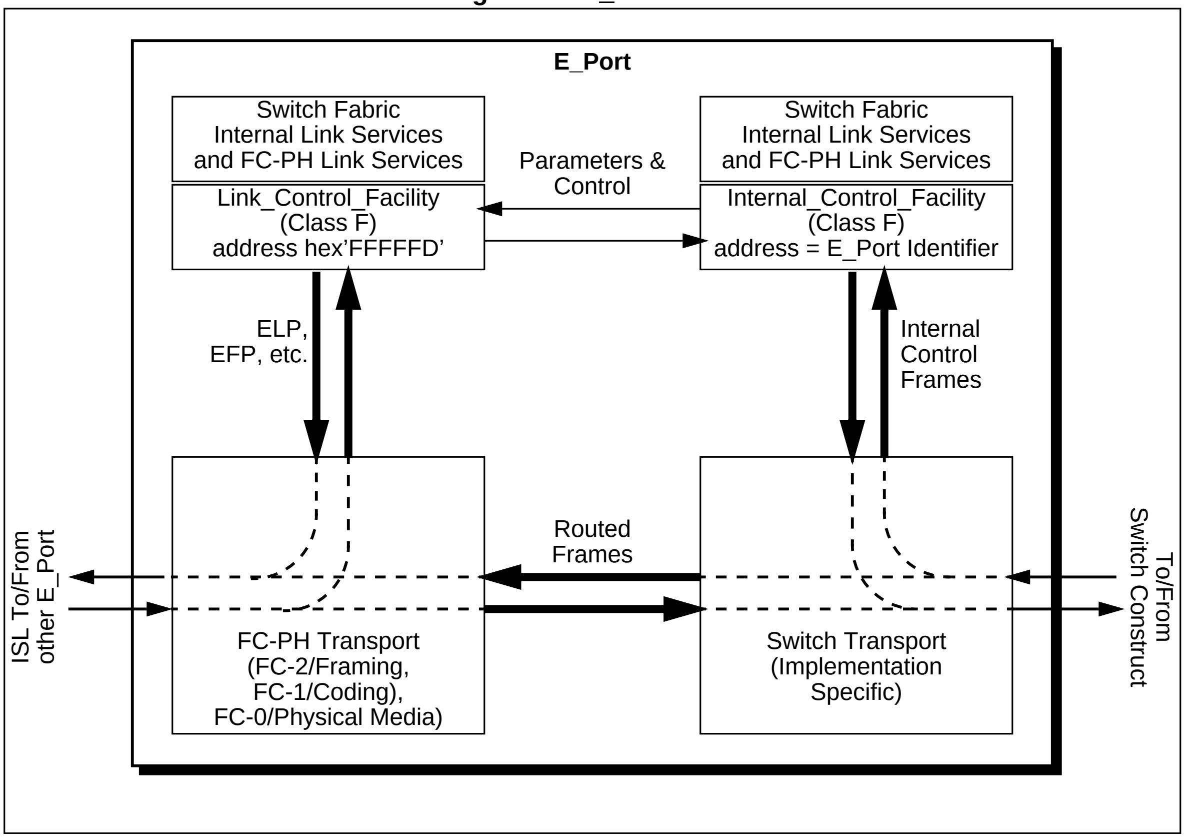
An E_Port is the point at which a Switch is connected to another Switch to create a Fabric. It normally functions as a conduit between the Switches for frames destined for remote N_Ports and NL_Ports. An E_Port is also used to carry frames between Switches for purposes of configuring and maintaining the Fabric.

An E_Port shall support the Class F service. An E_Port shall also support one or more of the following Classes of service: Class 1 service, Class 2 service, Class 3 service. An E_Port shall not admit to the Fabric Primitive Sequences or Primitive Signals other than Idle received on its inbound fibre.

5.3.1 Model

The model of an E_Port is shown in figure 4.

Figure 6 – E_Port Model



An E_Port contains an FC-PH Transport element through which passes all frames and Primitives transferred across the Link to and from the other E_Port. Frames received from the other E_Port are either directed to the Switch Construct via the Switch Transport element, or directed to the Link_Control_Facility. The Link_Control_Facility receives frames related to Switch Fabric Internal Link Services such as ELP, and transmits responses to those Link Service frames.

Frames received from the FC-PH Transport element that are destined for other ports are directed by the Switch Transport to the Switch Construct for further routing. Frames received from the Switch Construct by the Switch Transport are directed either to the FC-PH Transport for transmission to the other E_Port, or to the Internal_Control_Facility. The Internal_Control_Facility receives frames related to Switch Fabric Internal Link Services, and transmits responses to those Internal Link Services frames.

Information is passed between the Internal_Control_Facility and the Link_Control_Facility to effect the control and configuration of the Transport elements.

5.3.2 Inter-Switch Link Behavior

Inter-Switch Links (ISLs) are used by Switches to transmit and receive frames with other Switches. An ISL always connects exactly one E_Port on a Switch to exactly one E_Port on another Switch.

An ISL follows the FC-0, FC-1, and FC-2 protocols defined for point-to-point Links as defined in FC-PH, with the exception that Class F frames are allowed to transit the Link, as defined in FC-FG. The use of R_RDY shall be restricted to the management of buffer-to-buffer flow control of Class F frames on the ISL prior to the completion of the exchange of Link parameters (see 6.2.2 and 7.1); an alternate method of buffer-to-buffer flow control may defined in that process. Flow control of Class N frames shall be managed by other means not defined in this Standard.

NOTE – It is expected that the various flow control models will be defined by Profile.

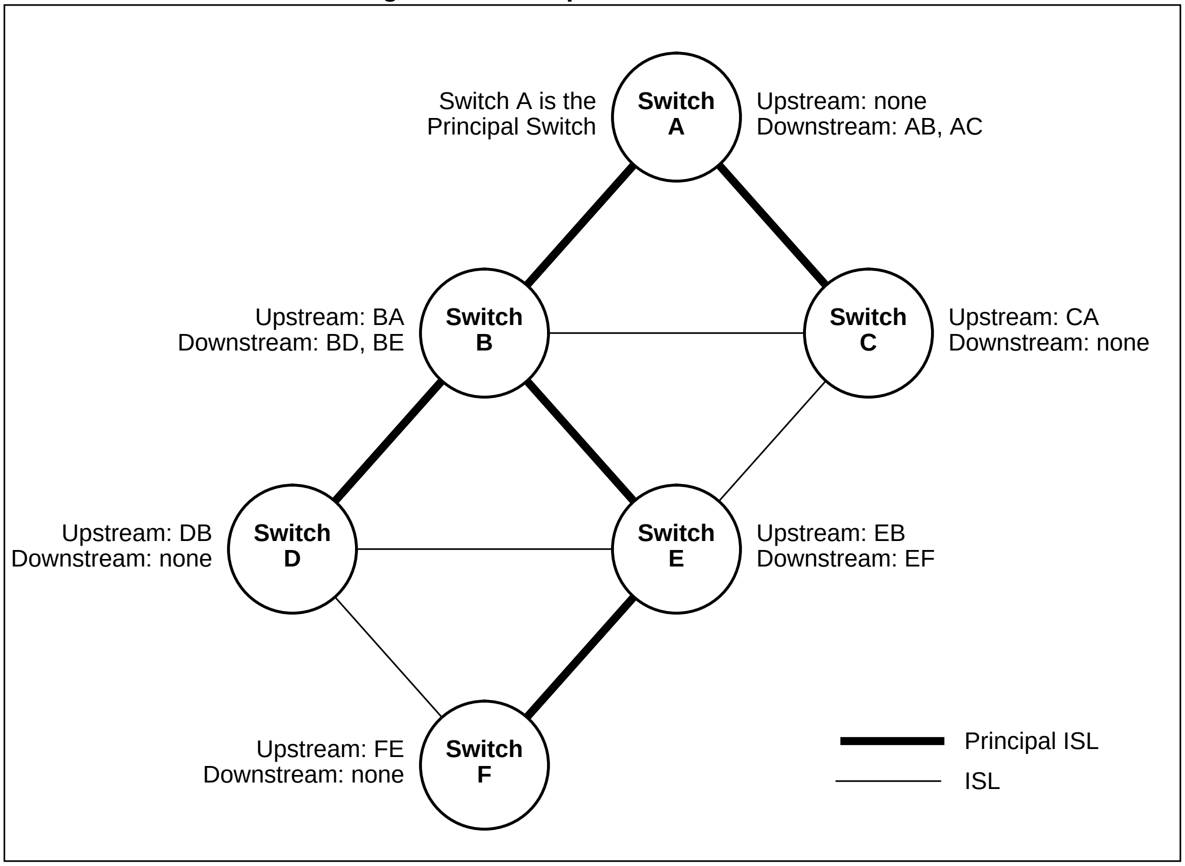
For purposes of defining and maintaining the Fabric Configuration, an ISL may be designated as a Principal ISL. The Principal ISL is a path that is used during configuration and address assignment to route Class F configuration frames, and is therefore a known path between two Switches. If a Principal ISL is lost, there may be no other available paths between the two affected Switches, so as a result the Fabric Configuration is possibly broken and must be rebuilt (by issuing the BF SW_ILS, see 6.2.9). If a non-Principal ISL is lost, at least one other path is known to be available between the Switches (i.e., the Principal ISL), therefore the lost ISL can be resolved via a routing change.

A Switch discovers the Principal ISL(s) during the process of Principal Switch Selection (see 7.2) and Address Distribution (see 7.3). During this process, the Switch identifies two kinds of Principal ISL. The Principal ISL that leads towards the Principal Switch is called the upstream Principal ISL. All frames from the Switch to the Principal Switch are sent via the upstream Principal ISL. The Principal Switch has no upstream Principal ISL; all other Switches have exactly one upstream Principal ISL.

A Principal ISL that leads away from the Principal Switch is called the downstream Principal ISL. Any frame sent by the Switch to another Switch as a result of a frame received on the upstream Principal ISL is sent via the downstream Principal ISL that leads towards that Switch. The Principal Switch may have one or more downstream Principal ISLs; all other Switches may have zero or more downstream Principal ISLs.

Principal ISLs are further illustrated in figure 7.

Figure 7 – Principal Inter-Switch Links



5.4 Class F Service

Class F Service is a connectionless service with notification of non-delivery between E_Ports, used for control, coordination, and configuration of the Fabric. Class F Service is defined by this Standard for use by Switches communicating across Inter-Switch Links. This definition of Class F for Inter-Switch Links supercedes the definitions of Class F for Inter-Element Links in FC-FG.

5.4.1 Class F Function

A Class F Service is requested by an E_Port on a frame by frame basis. The Fabric routes the frame to the destination E_Port. If the E_Port transmits consecutive frames to multiple destinations, the Fabric demultiplexes them to the requested destinations. Class F delimiters are used to indicate the requested service and to initiate and terminate one or more Sequences as described in FC-PH.

5.4.2 Class F Rules

To provide Class F Service, the transmitting and receiving E_Ports and the Fabric shall obey the following rules:

a) Except for some Switch Fabric Internal Link Service protocols, an E_Port is required to have exchanged Link parameters (see 6.2.2 and 7.1) with the associated destination with which it intends to communicate **(Login)**.

b) The Fabric routes the frames without establishing a Dedicated Connection between communicating E_Ports. To obtain Class F service, the E_Port shall use Class F delimiters as defined in 5.4.3. **(Connectionless service)**

c) An E_Port is allowed to send consecutive frames to one or more destinations. This enables an E_Port to demultiplex multiple Sequences to a single or multiple destinations concurrently. **(demultiplexing)**

d) A given E_Port may receive consecutive frames from different sources. Each source is allowed to send consecutive frames for one or more Sequences. **(multiplexing)**

e) A destination E_Port shall provide an acknowledgment to the source for each valid Data frame received. The destination E_Port shall use ACK_1 for the acknowledgment. If a Switch is unable to deliver the ACK_1 frame, the Switch shall return an F_BSY or F_RJT. **(Acknowledgment)**

f) The Sequence Initiator shall increment the SEQ_CNT field of each successive frame transmitted within a Sequence. However, the Switches may not guarantee delivery to the destination in the same order of transmission. **(non-sequential delivery)**

g) An E_Port may originate multiple Exchanges and initiate multiple Sequences with one or more E_Ports. The E_Port originating an Exchange shall assign an X_ID unique to the Originator called OX_ ID and the Responder of the Exchange shall assign an X_ID unique to the responder called RX_ID. The value of OX_ ID or RX_ID is unique to a given E_Port. The Sequence Initiator shall assign a SEQ_ID, for each Sequence it initiates, which is unique to the Sequence Initiator and the respective Sequence Recipient pair while the Sequence is Open. **(concurrent Exchanges and Sequences)**

h) Each E_Port exercises buffer-to-buffer flow control with the E_Port to which it is directly attached. End-to-end flow control is performed by communicating E_Ports. ACK_1 frames are used to perform end-to-end flow control and R_RDY is used for buffer-to-buffer flow control. **(dual flow control)**

i) If a Switch is unable to deliver the frame to the destination E_Port, then the source is notified of each frame not delivered by an F_BSY or F_RJT frame with corresponding D_ ID, S_ID, OX_ID, RX_ID, SEQ_ID, and SEQ_CNT from the Switch. The source is also notified of valid frames busied or rejected by the destination E_Port by P_BSY or P_RJT. **(non-delivery)**

j) A busy or reject may be issued by an intermediate E_Port or the destination E_Port with a valid reason code. **(busy/reject)**

k) If a Class F Data frame is busied, the sender shall retransmit the busied frame up to the ability of the sender to retry, including zero. **(retransmit)**

l) The Credit established during the ELP protocol by interchanging Link Parameters shall be honored. Class F shall not share Credit with any other Class of service. **(Credit)**

m) Effective transfer rate between any given E_Port pair is dependent upon the number of E_Ports a given E_Port is demultiplexing to and multiplexing from. **(bandwidth)**

n) Frames within a Sequence are tracked on a Sequence_Qualifier and SEQ_CNT basis. **(tracking)**

o) An E_Port shall be able to recognize SOF delimiters for Class F, Class 1, Class 2, and Class 3 service, whether or not all Classes of service are supported by the Port. An E_Port shall accept frames for all FC-PH service Classes. **(invalid Class)**

p) An E_Port receiving a Vendor Unique Class F frame may discard the frame without notification. A Vendor Unique Class F frame is indicated by an R_CTL field value of {hex'FF' or hex'F0' or hex'0F' depending on how you read FC-FG 6.9.1}. **(vendor unique)**

q) An E_Port shall support insertion of Class F frames onto an established Class 1 Dedicated Connection. However, this insertion shall not cause loss of any Class 1 frames. A Switch may abort (**EOFa**) or discard an Intermixed Class 2 or Class 3 frame in progress if its transmission of a Class F frame interferes. A Switch shall not abort an Inserted Class F frame. **(Class F intermix)**

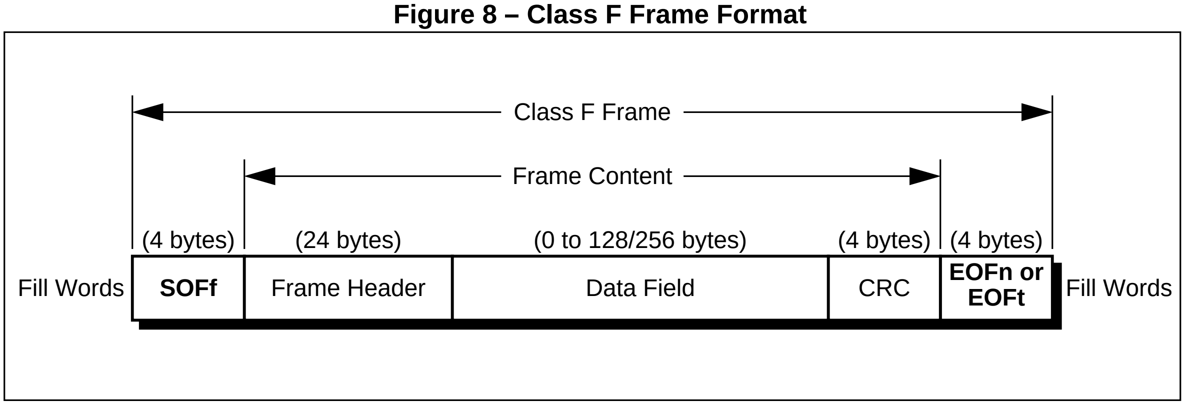
r) An E_Port shall use R_RDY and FC-PH buffer-to-buffer flow control with the E_Port to which it is directly attached, until after the exchange of Link parameters(see 6.2.2 and 7.1). The BB_Credit prior to the exchange of Link parameters shall be 1. E_Ports may agree to use an alternate buffer-to-buffer credit model for Class F following the exchange of Link parameters. **(alternate credit models)**

5.4.3 Class F Frame Format

Class F frames shall use the Frame_Header defined in Clause 18 of FC-PH. The Class F frame format is shown in figure 8. The Start_of_Frame Fabric (**SOFf**) delimiter shall precede the frame content of all Class F frames. The Data Field size of all Class F frames shall be less than or equal to {128/256} bytes. All Class F frames shall include the CRC defined in Clause 17 of FC-PH. The End_of_Frame Normal (**EOFn**) delimiter shall immediately follow the CRC of all normally completed Class F Data frames and all normally completed Class F Link_Control frames except the last frame of a Sequence. The End_of_Frame Terminate (**EOFt**) delimiter shall immediately follow the CRC of all Class F Link_Control frames that indicate the last frame of a Sequence which is normally terminated. A Class F frame is preceded and followed by the fill words appropriate to the Port Mode.

An E_Port or Switch may invalidate or discard without notification any incorrectly formed Class F frame, or any Class F frame with a code violation or CRC error.

Figure 8 – Class F Frame Format



5.4.4 Class F Flow Control

Class F service uses both buffer-to-buffer and end-to-end flow controls. R_RDY is used for buffer-to-buffer flow control. R_RDY is transmitted by the E_Port at one end of the ISL, to the E_ Port at the other end of the ISL, to indicate that a buffer is available for further frame reception by the first E_Port. This process operates in both directions on the ISL.

ACK_1 frames are used to perform end-to-end flow control. ACK_1 frames shall begin with an **SOFf** delimiter. The ACK_1 frame shall be terminated by an **EOFn** or **EOFt** delimiter. The ACK_0 and ACK_N Link Control frame shall not be used for Class F service.

6 Switch Fabric Services

This clause describes services provided for use by and with Switch Fabrics.

6.1 Switch Fabric Extended Link Services

{do we have any? or maybe the FAN, LINIT, etc. get 'standardized' here...?}

6.2 Switch Fabric Internal Link Services (SW_ILS)

This clause describes Link Services that operate internal to the Fabric between Switches. All SW_ILS frames shall be transmitted using the FT-1 frame format via the Class F service. The following defines the header fields of all SW_ILS frames:

- R_CTL: This field shall be set to hex'02' for all request frames, and to hex'03' for all reply frames.
- CS_CTL: This field shall be set to hex'00'.
- D_ID and S_ID: Set as indicated for the specific SW_ILS.
- TYPE: This field shall be set to hex'22', indicating Fibre Channel Fabric Switch Services.

All other fields shall be set as appropriate according to the rules defined in FC-PH.

The first word in the payload specifies the Command Code. The Command Codes are summarized in table 2.

Table 2 – SW_ILS Command Codes

Encoded Value (hex)	Description	Abbr.
01 00 00 00	Switch Fabric Internal Link Service Reject	SW_RJT
02 00 00 00	Switch Fabric Internal Link Service Accept	SW_ACC
10 00 00 00	Exchange Link Parameters	ELP
11 00 00 00	Exchange Fabric Parameters	EFP
12 00 00 00	Announce Address Identifier	AAI
13 00 00 00	Request Domain_ID	RDI
14 00 00 00	Hello	HLO
15 00 00 00	Link State Update	LSU
16 00 00 00	Link State Acknowledge	LSA
17 00 00 00	Build Fabric	BF
18 00 00 00	Reconfigure Fabric	RCF
20 00 00 00	Disconnect Class 1 Connection	DSCN
21 00 00 00	Detect Queued Class 1 Connection Request Deadlock	LOOPD
others	Reserved	
70 00 00 00 to 7F 00 00 00	Vendor Unique	

6.2.1 Switch Fabric Internal Link Service Reject (SW_RJT)

The Switch Fabric Internal Link Service Reject shall notify the transmitter of an SW_ILS request that the SW_ILS request Sequence has been rejected. A four-byte reason code shall be contained in the Data_Field. SW_RJT may be transmitted for a variety of conditions which may be unique to a specific SW_ILS request.

Protocol: SW_RJT may be sent as a reply Sequence to any SW_ILS request.

Format: FT–1

Addressing: The S_ID field shall be set to the value of the D_ID field in the SW_ILS request. The D_ID field shall be set to the value of the S_ID field in the SW_ILS request.

Payload: The format of the SW_RJT reply Payload is shown in table 3.

Table 3 – SW_RJT Payload

Item	Size Bytes
hex ‘01 00 00 00’	4
Reserved	1
Reason Code	1
Reason Code Explanation	1
Vendor Unique	1

Reason Code: The Reason Codes are summarized in table 4.

Table 4 – SW_RJT Reason Codes

Encoded Value (Bits 23-16)	Description
0000 0001	Invalid SW_ILS command code
0000 0010	Invalid revision level
0000 0011	Logical error
0000 0100	Invalid payload size
0000 0101	Logical busy
0000 0111	Protocol error
0000 1001	Unable to perform command request
0000 1011	Command not supported
others	Reserved
1111 1111	Vendor Unique error

Invalid SW_ILS command code: The Command Code is not recognized by the recipient.

Invalid revision level: The recipient does not support the specified revision level.

Logical error: The request identified by the Command Code and the Payload content is invalid or logically inconsistent for the conditions present.

Invalid payload size: The size of the Payload is inconsistent with the Command Code and/or any Length fields in the Payload.

Logical busy: The recipient is busy and is unable to process the request at this time.

Protocol error: An error has been detected that violates the protocol.

Unable to perform command request: The recipient cannot perform the request.

Command not supported: The command code is not supported by the recipient.

Vendor Unique Error: The Vendor Unique field indicates the error condition.

Reason Code Explanation: The Reason Code Explanation is summarized in table 5.

Table 5 – SW_RJT Reason Code Explanation

Encoded Value (Bits 15-8)	Description
0000 0000	No additional explanation
0000 0001	Class F Service Parameter error
0000 0011	Class N Service Parameter error
0000 0100	Unknown Switch Profile code
0000 0101	Invalid Switch Profile Parameters
0000 1101	Invalid Port_Name
0000 1110	invalid Switch_Name
0000 1111	R_A_TOV or E_D_TOV mismatch
0001 0000	Invalid Domain_Map
0001 1001	Command already in progress
0010 1001	Insufficient resources available
0010 1010	Domain_ID not available
0010 1100	Request not supported
	{anything else? esp. for ELP?}
others	Reserved

Vendor Unique: This field is valid when the Reason Code indicates a Vendor Unique error.

6.2.2 Exchange Link Parameters (ELP)

The Exchange Link Parameters Switch Fabric Internal Link Service requests the exchange of Link Parameters between two E_Ports connected via an ISL. The exchange of Link Parameters establishes the operating environment between the two E_Ports, and the capabilities of the Switches that are connected by the E_Ports. When an ELP is received by an E_Port, any Active or Open Class F Sequences between the two E_Ports, and any Dedicated Connections, shall be abnormally terminated prior to transmission of the SW_ACC reply Sequence.

Use of the ELP SW_ILS for Switch Port initialization is described in 7.1. Other uses of ELP are not defined by this Standard.

Protocol:

Exchange Link Parameters (ELP) request Sequence
Accept (SW_ACC) Reply Sequence

Format: FT–1

Addressing: For use in Switch Port initialization, the S_ID field shall be set to hex'FFFFFD', indicating the Fabric Controller of the originating Switch; the D_ID field shall be set to hex'FFFFFD', indicating the Fabric Controller of the destination Switch.

Payload: The format of the ELP request Payload is shown in table 6.

Table 6 – ELP Request Payload

Item	Size Bytes
hex '10 00 00 00'	4
Revision	1
Reserved	3
R_A_TOV	4
E_D_TOV	4
Requester E_Port_Name	8
Requester Switch_Name	8
Class F Service Parameters	16
Class 1 E_Port Parameters	4
Class 2 E_Port Parameters	4
Class 3 E_Port Parameters	4
Reserved	20
Switch Profile ID	2
Switch Profile Parameter Length (N)	2
Switch Profile-Specific Parameters	N

Revision: This field denotes the revision of the protocol. The first revision has the value of 1.

R_A_TOV: This field shall be set to the value of R_A_TOV required by the Switch.

E_D_TOV: This field shall be set to the value of E_D_TOV required by the Switch.

NOTE – The Value of R_A_TOV and E_D_TOV may be established by Profile or other means.

E_Port_Name: The E_Port_Name is an eight-byte field which identifies an E_Port for identification purposes. The format of the name is specified in FC-PH. Each E_Port shall provide a unique E_Port_Name within the Fabric.

Switch_Name: The Switch_Name is an eight-byte field which identifies a Switch for identification purposes. The format of the name is specified in FC-PH. Each Switch shall provide a unique Switch_Name within the Fabric.

Class F Service Parameters: This field contains the E_Port Class F Service Parameters. The format of the Parameters is shown in table 7.

Table 7 – E_Port Class F Service Parameters

<table>
<tr><th>Word</th><th>31</th><th>30</th><th>29</th><th>28</th><th>27</th><th>26</th><th>25</th><th>24</th><th>23</th><th>22</th><th>21</th><th>20</th><th>19</th><th>18</th><th>17</th><th>16</th><th>15</th><th>14</th><th>13</th><th>12</th><th>11</th><th>10</th><th>9</th><th>8</th><th>7</th><th>6</th><th>5</th><th>4</th><th>3</th><th>2</th><th>1</th><th>0</th></tr>
<tr><td>0</td><td>VAL</td><td colspan="15">Reserved</td><td colspan="16">Reserved</td></tr>
<tr><td>1</td><td colspan="2">R</td><td>XII</td><td colspan="13">Reserved</td><td colspan="16">Receive Data Field Size</td></tr>
<tr><td>2</td><td colspan="16">Concurrent Sequences</td><td colspan="16">End-to-End Credit</td></tr>
<tr><td>3</td><td colspan="16">Open Sequences per Exchange</td><td colspan="16">Reserved</td></tr>
</table>

The Class F Service Parameters are defined as follows:

– VAL (Class Valid): This bit shall be set to one.

– XII (X_ID Interlock): This bit when one indicates that the E_Port supplying this parameter requires that an interlock be used during X_ID assignment in Class F. In X_ID assignment, the Sequence Initiator shall set the Recipient X_ID value to hex'FFFF' in the first Data frame of a Sequence, and the Recipient shall supply its X_ID in the ACK frame corresponding to the first Data frame of a Sequence. The Sequence Initiator shall not transmit additional frames until the corresponding ACK is received. Following reception of the ACK, the Sequence Initiator continues transmission of the Sequence using both assigned X_ID values.

– Receive Data Field Size: This field shall specify the largest Data Field size in bytes for an FT-1 frame that can be received by the E_Port supplying the Parameters as a Sequence Recipient for a Class F frame. This field shall be set to {128/256/?}.

– Concurrent Sequences: This field shall specify the number of Sequence Status Blocks provided by the E_Port supplying the Parameters for tracking the progress of a Sequence as a Sequence

Recipient. The maximum number of Concurrent Sequences that can be specified is 255. A value of zero in this field is reserved. In Class F, the value of SEQ_ID shall range from 0 to 255, independent of the value in this field. An E_Port is allowed to respond with P_BSY to a frame initiating a new Sequence if E_Port resources are not available.

- End-to-End Credit: End-to-end credit is the maximum number of Class F Data frames which can be transmitted by an E_Port without receipt of accompanying ACK or Link_Response frames. The minimum value of end-to-end credit is one. The end-to-end credit field specified is associated with the number of buffers available for holding the Data_Field of a Class F frame and processing the contents of that Data_Field by the E_Port supplying the Parameters. Bit 15 of this field shall be set to zero. A value of zero for this field is reserved.

- Open Sequences per Exchange: The value of the Open Sequences per Exchange shall specify the maximum number of Sequences that can be Open at one time at the Recipient between a pair of E_Ports for one Exchange. This value plus two shall specify the number of instances of Sequence Status that shall be maintained by the Recipient for a single Exchange in the Exchange Status Block. This value is used for Exchange and Sequence tracking. The value in this field limits the link facility resources required for error detection and recovery (see FC-FG).

Class N E_Port Parameters: E_Port Parameters indicate that the E_Port is capable of transporting the indicated Class of Service, and the conditions under which it can transport the Class. One word of the ELP Payload is allocated for each Class.

Class 1 E_Port Parameters: This field contains the Class 1 E_Port Parameters. The format of the Parameters is shown in table 8.

Table 8 – Class 1 E_Port Parameters

Word	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
0	VAL	MIX	XPS	LKS	Reserved												Reserved															

The Class 1 E_Port Parameters are defined as follows:

- VAL (Class Valid): This bit is set to one if the E_Port supports Class 1. If this bit is zero, all other Class 1 E_Port Parameters shall be invalid.

- MIX (Intermix): This bit is set to one if the E_Port can perform Intermix as defined in FC-PH. Intermix shall be functional only if both E_Ports indicate support for this feature.

- XPS (Transparent Mode Stacked Connect Request): This bit is set to one if the E_Port can perform Transparent Mode Stacked Connect Requests as defined in FC-PH. Transparent Mode Stacked Connect Requests shall be functional only if both E_Ports indicate support for this feature. A Switch shall not indicate support for both XPS and LKS.

- LKS (Lock-down Mode Stacked Connect Request): This bit is set to one if the E_Port can perform Lock-down Mode Stacked Connect Requests as defined in FC-PH. Lock-down Mode Stacked Connect Requests shall be functional only if both E_Ports indicate support for this feature. A Switch shall not indicate support for both XPS and LKS.

Class 2 E_Port Parameters: This field contains the Class 2 E_Port Parameters. The format of the Parameters is shown in table 9.

Table 9 – Class 2 E_Port Parameters

Word	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
0	VAL	SEQ	Reserved														Receive Data Field Size															

The Class 2 E_Port Parameters are defined as follows:

- VAL (Class Valid): This bit shall be set to one if the E_Port supports Class 2. If this bit is zero, all other Class 2 E_Port Parameters shall be invalid.

- SEQ (Sequential Delivery): If this bit is set to one by an E_Port, it is indicating that the Switch is able to guarantee sequential delivery (as defined in FC-PH) of Class 2 frames. Sequential Delivery shall be functional only if both E_Ports indicate support for this feature.

- Receive Data Field Size: This field shall specify the largest Data Field size in bytes for an FT-1 frame that can be received by the E_Port supplying the Parameters for a Class 2 frame. Values less than 256 or greater than 2112 are invalid. Values shall be a multiple of four bytes.

Class 3 E_Port Parameters: This field contains the Class 3 E_Port Parameters. The format of the Parameters is shown in table 10.

Table 10 – Class 3 E_Port Parameters

Word	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
0	VAL	SEQ	Reserved														Receive Data Field Size															

The Class 3 E_Port Parameters are defined as follows:

- VAL (Class Valid): This bit shall be set to one if the E_Port supports Class 3. If this bit is zero, all other Class 3 E_Port Parameters shall be invalid.

- SEQ (Sequential Delivery): If this bit is set to one by an E_Port, it is indicating that the Switch is able to guarantee sequential delivery (as defined in FC-PH) of Class 3 frames. Sequential Delivery shall be functional only if both E_Ports indicate support for this feature.

- Receive Data Field Size: This field shall specify the largest Data Field size in bytes for an FT-1 frame that can be received by the E_Port supplying the Parameters for a Class 3 frame. Values less than 256 or greater than 2112 are invalid. Values shall be a multiple of four bytes.

Switch Profile ID: This field indicates an ID code which specifies the Switch Profile supported by the E_Port. Values of hex'0000' and hex'FFFF' are reserved. Values of hex'8000' through hex'FFFE' are Vendor Unique. All other values are reserved for future Profiles.

Switch Profile Parameter Length: This field specifies the length in bytes of the Switch Profile-Specific Parameters that follow. Values shall be a multiple of four. A value of zero indicates no parameters follow.

Switch Profile-Specific Parameters: These parameters contain Switch Profile-Specific information used to configure the ISL.

NOTE – Different switch implementations may use different methods for managing flow control of user frames across an ISL. These parameters are intended to provide a switch-specific way to indicate these flow control parameters. Consult the appropriate Switch Profile for more information.

Reply Switch Fabric Internal Link Service Sequence:

Service Reject (SW_RJT)
Signifies the rejection of the ELP command
Accept (SW_ACC)
Signifies acceptance of the ELP request.
– Accept Payload

Payload: The format of the ELP Accept Payload is shown in table 11.

Table 11 – ELP Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4
Revision	1
Reserved	3
R_A_TOV	4
E_D_TOV	4
Responder E_Port_Name	8
Responder Switch_Name	8
Class F Service Parameters	16
Class 1 E_Port Parameters	4
Class 2 E_Port Parameters	4
Class 3 E_Port Parameters	4
Reserved	20
Switch Profile ID	2
Switch Profile Parameter Length (N)	2
Switch Profile-Specific Parameters	N

The fields in table 11 are the same as defined for table 6.

6.2.3 Exchange Fabric Parameters (EFP)

The Exchange Fabric Parameters Switch Fabric Internal Link Service requests the exchange of Fabric Parameters between two E_Ports connected via an ISL. The exchange of Fabric Parameters is used to establish the address allocation within the Fabric. When an E_Port receives EFP from another E_Port, all Active or Open Class F Sequences and Dedicated Connections shall be unaffected.

Use of the EFP SW_ILS for Fabric Configuration is described in 7.2 and 7.3. Other uses of EFP are not defined by this Standard.

Protocol:

Exchange Fabric Parameters (EFP) request Sequence
Accept (SW_ACC) Reply Sequence

Format: FT–1

Addressing: For use in Fabric Configuration, the S_ID field shall be set to hex’FFFFFD’, indicating the Fabric Controller of the originating Switch. The D_ID field shall be set to hex’FFFFFD’, indicating the Fabric Controller of the destination Switch.

Payload: The format of the EFP request Payload is shown in table 12.

Table 12 – EFP Request Payload

Item	Size Bytes
hex ‘11 00 00 00’	4
Reserved	3
Principal Switch_Priority	1
Principal Switch_Name	8
Current Allocated Domain_Map	32

Principal Switch_Priority: This field shall specify the priority level of the Switch that the transmitting Switch believes is the Principal Switch. Values for this field are summarized in table 13.

Table 13 – Switch_Priority Field Values

Value (hex)	Description
00	Reserved
01	The Switch was the Principal Switch prior to sending or receiving BF. (note 1)
02 to FE	Higher to lower priority values. (note 2)
FF	The Switch is not capable of acting as a Principal Switch.

Notes:

1 This allows the same Switch to become Principal Switch if it is still part of the Fabric after sending and/or receiving the Build Fabric SW_ILS.

2 The Switch_Priority value for a given Switch is established by means not defined by this Standard.

Principal Switch_Name: This field shall specify the Switch_Name of the Switch that the transmitting Switch believes is the Principal Switch.

Current Allocated Domain_Map: This field shall contain 8 words that constitute a bitmap of Domain_IDs that have been allocated within the Fabric. Bit 0 of Word 0 set to one indicates that Domain_ID=hex’00’ has been allocated, Bit 1 of Word 0 set to one indicates that Domain_ID=hex’01’ has been allocated, and so forth. The bits corresponding to Domain_IDs hex’00’ and hex’F0’ through hex’FF’ shall always be set to zero.

Reply Switch Fabric Internal Link Service Sequence:

Service Reject (SW_RJT)
Signifies the rejection of the EFP command
Accept (SW_ACC)
Signifies acceptance of the EFP request.
– Accept Payload

Payload: The format of the EFP Accept Payload is shown in table 14.

Table 14 – EFP Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4
Reserved	3
Principal Switch_Priority	1
Principal Switch_Name	8
Current Allocated Domain_Map	32

The fields in table 14 are the same as defined for table 12.

6.2.4 Announce Address Identifier (AAI)

The Announce Address Identifier Switch Fabric Internal Link Service communicates the address identifier of the E_Port to another E_Port. This communication establishes that the E_Port has been assigned an address identifier, and that the Recipient may request an address identifier from the Originating E_Port.

Use of the AAI SW_ILS for Fabric Configuration is described in 7.3. Other uses of AAI are not defined by this Standard.

Protocol:

Announce Address Identifier (AAI) request Sequence
Accept (SW_ACC) Reply Sequence

Format: FT–1

Addressing: For use in Fabric Configuration, the S_ID field shall be set to hex’FFFFFD’, indicating the Fabric Controller of the originating Switch. The D_ID field shall be set to hex’FFFFFD’, indicating the Fabric Controller of the destination Switch.

Payload: The format of the AAI request Payload is shown in table 15.

Table 15 – AAI Request Payload

Item	Size Bytes
hex ‘12 00 00 00’	4
Switch_Name	8
Reserved	1
Address identifier	3

Switch_Name: This field shall contain the Switch_Name of the Originating E_Port.

Address identifier: This field shall contain the address identifier of the Originating E_Port.

Reply Switch Fabric Internal Link Service Sequence:

Service Reject (SW_RJT)
Signifies the rejection of the AAI command
Accept (SW_ACC)
Signifies acceptance of the AAI request.
– Accept Payload

Payload: The format of the AAI Accept Payload is shown in table 16.

Table 16 – AAI Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4

6.2.5 Request Domain_ID (RDI)

The Request Domain_ID Switch Fabric Internal Link Service is sent by a Switch to request a Domain_ID from the Domain Address Manager. RDI shall not be sent by a Switch unless the Switch has received an AAI SW_ILS since the last reconfiguration event.

Use of the RDI SW_ILS for Fabric Configuration is described in 7.3. Other uses of RDI are not defined by this Standard.

Protocol:

Request Domain_ID (RDI) request Sequence
Accept (SW_ACC) Reply Sequence

Format: FT–1

Addressing: For use in Fabric Configuration, the S_ID field shall be set to hex'FFFFFD', indicating the Fabric Controller of the originating Switch. The D_ID field shall be set to hex'FFFFFD', indicating the Fabric Controller of the destination Switch.

Payload: The format of the RDI request Payload is shown in table 17.

Table 17 – RDI Request Payload

Item	Size Bytes
hex '13 00 00 00'	4
Requesting Switch_Name	8
Reserved	3
Requested Domain_ID	1

Requesting Switch_Name: This field specifies the Switch_Name of the Switch requesting a Domain_ID.

Requested Domain_ID: This field shall contain the requested Domain_ID of the Switch requesting a Domain_ID. This field is set to either the Preferred Domain_ID if it is available, or zero.

Reply Switch Fabric Internal Link Service Sequence:

Service Reject (SW_RJT)
Signifies the rejection of the RDI command
Accept (SW_ACC)
Signifies acceptance of the RDI request.
– Accept Payload

Payload: The format of the RDI accept Payload is shown in table 18.

Table 18 – RDI Accept Payload

Item	Size Bytes
hex '02 00 00 00'	4
Requesting Switch_Name	8
Reserved	3
Granted Domain_ID	1

Requesting Switch_Name: This field specifies the Switch_Name of the Switch requesting a Domain_ID.

Granted Domain_ID: This field shall contain the Domain_ID granted by the Domain Address Manager to the requesting Switch.

6.2.6 Hello (HLO)

The Hello Switch Fabric Internal Link Service is used to periodically poll a remote Switch to ensure that it is still part of the Fabric. {more TBD}

Use of the HLO SW_ILS for Fabric Configuration {will be} described in 7.2. Other uses of HLO are not defined by this Standard.

Protocol:

Hello (HLO) request Sequence
Accept (SW_ACC) Reply Sequence

Format: FT–1

Addressing: The S_ID field shall be set to {TBD}. The D_ID field shall be set to {TBD}.

Payload: The format of the HLO request Payload is shown in table 19.

Table 19 – HLO Request Payload

Item	Size Bytes
hex ‘14 00 00 00’	4
Reserved	1
Address identifier of local Switch	3
Reserved	1
Address identifier of remote Switch	3

Address identifier of local Switch: This field shall contain the address identifier of the local Switch.

Address identifier of remote Switch: This field shall contain the address identifier of the remote Switch.

Reply Switch Fabric Internal Link Service Sequence:

Service Reject (SW_RJT)
Signifies the rejection of the HLO command
Accept (SW_ACC)
Signifies acceptance of the HLO request.
– Accept Payload

Payload: The format of the HLO accept Payload is shown in table 20.

Table 20 – HLO Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4
Reserved	1
Address identifier of local Switch	3
Reserved	1
Address identifier of remote Switch	3

6.2.7 Link State Update (LSU)

The Link State Update Switch Fabric Internal Link Service is used to establish routing.... {more TBD}

Protocol:

Link State Update (LSU) request Sequence
Accept (SW_ACC) Reply Sequence

Format: FT–1

Addressing: The S_ID field shall be set to {TBD}. The D_ID field shall be set to {TBD}.

Payload: The format of the LSU request Payload is shown in table 19.

Table 21 – LSU Request Payload

Item	Size Bytes
hex ‘15 00 00 00’	4
{TBD}	

Field: This field shall contain {TBD}.

Reply Switch Fabric Internal Link Service Sequence:

Service Reject (SW_RJT)
  Signifies the rejection of the LSU command
Accept (SW_ACC)
  Signifies acceptance of the LSU request.
– Accept Payload

Payload: The format of the LSU accept Payload is shown in table 20.

Table 22 – LSU Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4
{TBD}	

6.2.8 Link State Acknowledge (LSU)

The Link State Acknowledge Switch Fabric Internal Link Service is used to establish routing.... {more TBD}

Protocol:

Link State Acknowledge (LSA) request Sequence
Accept (SW_ACC) Reply Sequence

Format: FT–1

Addressing: The S_ID field shall be set to {TBD}. The D_ID field shall be set to {TBD}.

Payload: The format of the LSA request Payload is shown in table 19.

Table 23 – LSA Request Payload

Item	Size Bytes
hex ‘16 00 00 00’	4
{TBD}	

Field: This field shall contain {TBD}.

Reply Switch Fabric Internal Link Service Sequence:

Service Reject (SW_RJT)
Signifies the rejection of the LSA command
Accept (SW_ACC)
Signifies acceptance of the LSA request.
– Accept Payload

Payload: The format of the LSA accept Payload is shown in table 20.

Table 24 – LSA Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4
{TBD}	

6.2.9 Build Fabric (BF)

The Build Fabric Switch Fabric Internal Link Service requests a non-dispruptive reconfiguration of the entire Fabric. Fabric Configuration is performed as described in clause 7.

NOTE – Since the RCF causes a complete reconfiguration of the Fabric, and may cause addresses allocated to a Switch to change, the RCF SW_ILS should be used with caution. The BF SW_ILS allows the Fabric to attempt reconfiguration without loss of or change of address. Examples of situations in which BF is appropriate include a loss of a Principal ISL (Link Failure or Offline), or when two Fabrics are joined.

The transmission or reception of BF shall not of itself cause the loss of Class N frames, or cause a busy response to any Class N frames. Active or Open Class F Sequences between the two E_Ports, and any Dedicated Connections, shall not be abnormally terminated.

Use of the BF SW_ILS for Fabric Configuration is described in 7.2 and 7.3. Other uses of BF are not defined by this Standard.

Protocol:

Build Fabric (BF) request Sequence
Accept (SW_ACC) Reply Sequence

Format: FT–1

Addressing: For use in Fabric Configuration, the S_ID field shall be set to hex’FFFFFD’, indicating the Fabric Controller of the originating Switch. The D_ID field shall be set to hex’FFFFFD’, indicating the Fabric Controller of the destination Switch.

Payload: The format of the BF request Payload is shown in table 25.

Table 25 – BF Request Payload

Item	Size Bytes
hex ‘17 00 00 00’	4

Reply Switch Fabric Internal Link Service Sequence:

Service Reject (SW_RJT)
  Signifies the rejection of the BF command
Accept (SW_ACC)
  Signifies acceptance of the BF request.

– Accept Payload

Payload: The format of the BF accept Payload is shown in table 26.

Table 26 – BF Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4

6.2.10 Reconfigure Fabric (RCF)

The Reconfigure Fabric Switch Fabric Internal Link Service requests a disruptive reconfiguration of the entire Fabric. Fabric Configuration is performed as described in clause 7.

NOTE – Since the RCF causes a complete reconfiguration of the Fabric, and may cause addresses allocated to a Switch to change, this SW_ILS should be used with caution. Examples of situations in which RCF is appropriate include detection of overlapped Domains, or the failure of a Fabric Reconfiguration initiated by a BF.

When an RCF is transmitted by an E_Port, any Active or Open Class F Sequences between the two E_Ports, and any Dedicated Connections, shall be abnormally terminated. Also, all Class N frames shall be discarded, and all Dedicated Connections shall be abnormally abnormally terminated.

When an RCF is received by an E_Port, any Active or Open Class F Sequences between the two E_Ports, and any Dedicated Connections, shall be abnormally terminated prior to transmission of the SW_ACC reply Sequence. Also, all Class N frames shall be discarded, and all Dedicated Connections shall be abnormally abnormally terminated prior to transmission of the SW_ACC reply Sequence.

Use of the RCF SW_ILS for Fabric Configuration is described in 7.2 and 7.3. Other uses of RCF are not defined by this Standard.

Protocol:

Reconfigure Fabric (RCF) request Sequence
Accept (SW_ACC) Reply Sequence

Format: FT–1

Addressing: For use in Fabric Configuration, the S_ID field shall be set to hex’FFFFFD’, indicating the Fabric Controller of the originating Switch. The D_ID field shall be set to hex’FFFFFD’, indicating the Fabric Controller of the destination Switch.

Payload: The format of the RCF request Payload is shown in table 27.

Table 27 – RCF Request Payload

Item	Size Bytes
hex ‘18 00 00 00’	4

Reply Switch Fabric Internal Link Service Sequence:

Service Reject (SW_RJT)
Signifies the rejection of the RCF command
Accept (SW_ACC)
Signifies acceptance of the RCF request.
– Accept Payload

Payload: The format of the RCF accept Payload is shown in table 28.

Table 28 – RCF Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4

6.2.11 Disconnect Class 1 Connection (DSCN)

The Disconnect Class 1 Connection Switch Fabric Internal Link Service requests that the receiving E_Port end a Class 1 Connection. This SW_ILS is used for error recovery only.

Protocol:

Disconnect Class 1 Connection (DSCN) request Sequence
Accept (SW_ACC) Reply Sequence

Format: FT–1

Addressing: The S_ID field shall be set to the address identifier of the sending E_Port. The D_ID field shall be set to the address identifier of the destination E_Port.

Payload: The format of the DSCN request Payload is shown in table 29.

Table 29 – DSCN Request Payload

Item	Size Bytes
hex ‘20 00 00 00’	4
Reserved	3
Reason code for disconnect	1

Reply Switch Fabric Internal Link Service Sequence:

Service Reject (SW_RJT)
Signifies the rejection of the DSCN command
Accept (SW_ACC)
Signifies acceptance of the DSCN request.
– Accept Payload

Payload: The format of the DSCN accept Payload is shown in table 30.

Table 30 – DSCN Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4

6.2.12 Detect Queued Class 1 Connection Request Deadlock (LOOPD)

The Detect Queued Class 1 Connection Request Deadlock Switch Fabric Internal Link Service is used to check for possible deadlocks caused by Connection requests being queued at the destination E_Port (Camp-On). For example, if a connection request from port A is queued at port B, a request from port B is queued at port C, and a request from port C is queued at port A, a deadlock has occurred.

A LOOPD SW_ILS is originated when a Camp-On connection is queued. The LOOPD follows the path of pending Connection requests until the path is broken, or the LOOPD gets back to the original sender. If the LOOPD gets back to the original sender, a deadlock has occurred. The Switch shall busy one of the pending Connection requests to break the deadlock.

Protocol:

Detect Queued Class 1 Connection Request Deadlock (LOOPD) request Sequence
Accept (SW_ACC) Reply Sequence

Format: FT–1

Addressing: The S_ID field shall be set to the address identifier of the sending E_Port. The D_ID field shall be set to the address identifier of the destination E_Port.

Payload: The format of the LOOPD request Payload is shown in table 31.

Table 31 – LOOPD Request Payload

Item	Size Bytes
hex ‘21 00 00 00’	4
Reserved	1
Address identifier of originating E_Port	3

Reply Switch Fabric Internal Link Service Sequence:

Service Reject (SW_RJT)
  Signifies the rejection of the LOOPD command
Accept (SW_ACC)
  Signifies acceptance of the LOOPD request.
– Accept Payload

Payload: The format of the LOOPD accept Payload is shown in table 32.

Table 32 – LOOPD Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4

7 Fabric Configuration

The Fabric Configuration process enables a Switch Port to determine its operating mode, exchange operating parameters, and provides for distribution of addresses. This process is summarized in table 33.

Table 33 – Fabric Configuration Summary

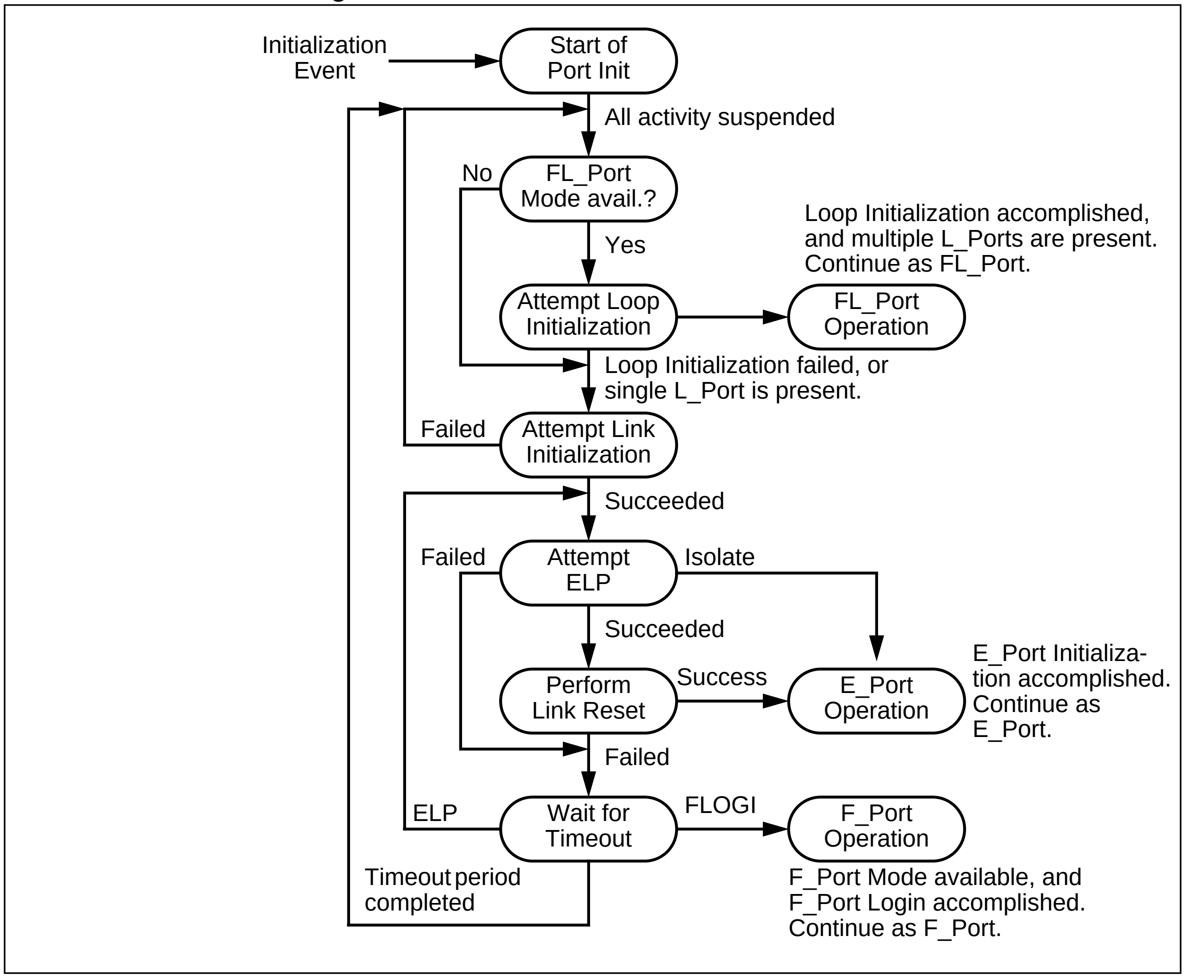
Step	Starting State	Process	Ending State
1. Establish Link Parameters and Switch Port operating mode	Switch Port has achieved word synch.	The Switch Port attempts to discover whether it is an FL_Port, an E_Port or an F_Port.	Switch Port mode is known. If a Port is an E_Port, Link Parameters have been exchanged and Credit has been initialized.
2. Select Principal Switch	BF or RCF SW_ILS transmitted or received.	Switch_Names are exchanged over all ISLs to select a Principal Switch, which becomes the Domain Address Manager.	The Principal Switch is selected.
3. Domain_ID Distribution	Domain Address Manager has been selected.	Switches request a Domain_ID from the Domain Address Manager.	All Switches have a Domain_ID.

{Area Address distribution, TBD. should look like Domain dist.}

7.1 Switch Port Initialization

Switch Ports shall initialize as detailed below. Figure 9 shows a schematic of the process to illustrate the flow. If the figure is different than the text, the text shall apply. Note also that this flow assumes that a Switch Port is capable of at least E_Port operation; either E/F/FL_Port, E/F_Port, E/FL_Port, or E_Port. Initialization of Switch Ports that are F/FL_Port, FL_Port, or F_Port is defined in FC-PH and FC-AL.

Figure 9 – Switch Port Mode Initialization Flow



a) **Start of Switch Port Initialization**. Switch Port initialization begins whenever an Initialization Event occurs. An Initialization Event is defined as either: a power-on reset condition; or, a transition to Link Offline, as defined in FC-PH; or, a loss of word synchronization; or, a failure to successfully complete a prior initialization attempt. When an Initialization Event occurs, all activity on the Switch Port is suspended until the Initialization is complete. Go to **step (b)**.

b) **FL_Port Mode Available?**. If the Switch Port is FL_Port-capable, go to **step (c)**. Otherwise, go to **step (d)**.

c) **Attempt Loop Initialization**. An FL_Port-capable Switch Port attempts Loop Initialization (as defined in FC-AL clause 10). If the Loop Initialization succeeds (the FL_Port transitions from the OPEN_INIT state to the MONITORING state), and the resulting AL_PA bitmap generated during the LISA Loop Initialization Sequence indicates more than one L_Port (other than the Switch Port) is attached, the Switch Port shall go to **step (h)**. If the Switch Port had attempted Loop Initialization at least once before and succeeded, but then attempted Link Initialization at least once and failed, the Switch Port may go to **step (h)**. Otherwise, go to **step (d)**.

d) **Attempt Link Initialization**. In this step, if the Switch Port is FL_Port-capable, and it has detected only one attached L_Port (NL_Port or FL_Port), attempting to establish a point-to-point Link is appropriate, and is necessary for detecting an attached E_Port. The Switch Port shall attempt Link Initialization as defined in FC-PH. If the Link Initialization succeeds, proceed to **step (e)**. Otherwise, the Switch Port shall return to **step (b)** and retry the initialization.

e) **Attempt to Exchange Link Parameters**. The Switch Port shall originate an ELP SW_ILS request Sequence (see 6.2.2). Table 34 below defines the responses and actions to an ELP request for the originating E_Port.

Table 34 – Responses to ELP Request for Originating E_Port

Response to ELP	Indication	Originating E_Port Action
1. R_RDY	Request received at destination	Wait E_D_TOV for response frame
2. ACK_1	Request received at destination	Wait E_D_TOV for response frame
3. SW_ACC	Destination E_Port received and processed request	Send ACK_1, continue configuration with **step (f)**
4. F_BSY or P_BSY	Destination is busy	Retry (note 1)
5. F_RJT or P_RJT	The frame is not acceptable	Respond accordingly (note 3)
6. ELP (rcvd E_Port_Name > own E_Port_Name)	Both E_Ports sent ELP at the same time	Send SW_ACC, continue configuration with **step (f)** (see Figure 10 for an example of this response)
7. ELP (rcvd E_Port_Name < own E_Port_Name)	Both E_Ports sent ELP at the same time	Send SW_RJT (note 2) (see Figure 10 for an example of this response)
8. ELP (rcvd E_Port_Name = own E_Port_Name)	E_Port output is looped back to input	Remove loopback condition
9. SW_RJT	Reason code/explanation: - Command already in progress - Logical busy - other	 - send SW_ACC (note 3) - retry (note 1) - respond accordingly

Table 34 – Responses to ELP Request for Originating E_Port

Response to ELP	Indication	Originating E_Port Action
10. FLOGI	Destination is an N_Port	Respond accordingly (note 3)
11. any other frame	Could be anything	Discard frame and retry (note 1)
12. E_D_TOV expires	Destination is busy; or, ELP, SW_ACC, ACK_1 frame lost; or, destination is not an E_Port	Retry (note 1)

Notes:

1 The retry is performed following a timeout period, as defined in **step (g)** below.

2 The Reason Code shall be “Unable to perform command request” with an Reason Explanation of “Command already in progress”.

3 Response is defined in FC-PH. A retry may be appropriate.

4 The SW_ACC is sent for the other ELP Exchange in progress, as described in Response #6.

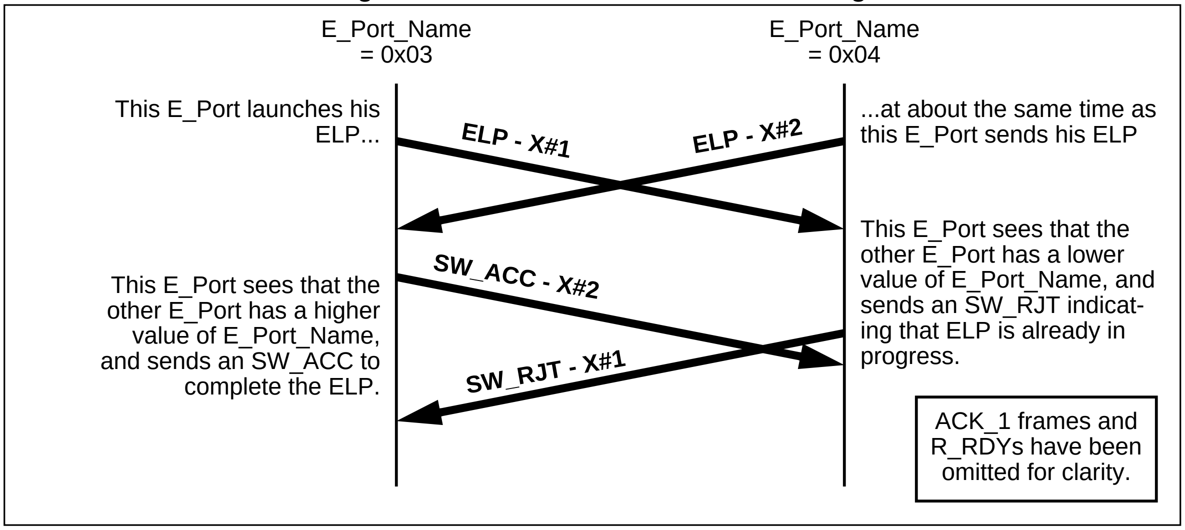
The originating E_Port shall consider the exchange of Link Parameters complete when it has received the SW_ACC and has transmitted the ACK_1 for the SW_ACC or SW_RJT reply Sequence. The responding E_Port shall consider the exchange of Link Parameters complete when it has received the ACK_1 for the SW_ACC or SW_RJT. The exchange of Link Parameters shall be considered successful if the reply to the ELP is an SW_ACC, and both E_Ports agree that the parameters exchanged are acceptable. If the exchange of Link Parameters is successful, the Switch Port shall go to **step (f)**. If the responding E_Port does not agree that the parameters are acceptable, it shall return an SW_RJT reply Sequence indicating the reason for the disagreement, and wait for the originating E_Port to initiate another ELP request Sequence. If the originating E_Port does not agree that the parameters in the SW_ACC are acceptable, or it receives an SW_RJT indicating the parameters in the ELP request were not acceptable to the responding E_Port, it may:

1) originate a new ELP request Sequence with modified parameters; or,

2) go to **step (i)** and operate as an Isolated E_Port (see 7.4); or,

3) perform the Link Offline protocol as defined in FC-PH and go to **step (g)** and retry the initialization.

Figure 10 – Simultaneous ELP Processing



f) **Perform Link Reset**. Following the successful completion of ELP, the value of buffer-to-buffer and end-to-end Class F Credit are initialized. In order to initialize the Profile-specific Credit parameters, the Switch Port that originated the successful ELP SW_ILS shall attempt the Link Reset protocol as defined in FC-PH. If the Link Reset succeeds, go to **step (i)**. Otherwise, go to **step (g)**.

NOTE – The re-initialization of Link credit is necessary since the Profile-Specific parameters in the ELP Payload are intended to communicate Link credit parameters for a specific credit model. The Link Reset is the common method defined by FC-PH for establishing a known credit state.

g) **Wait**. The Switch Port shall wait for R_A_TOV before retrying the ELP SW_ILS. If during the timeout period a FLOGI ELS (as defined in FC-PH) is received by the Switch Port, and F_Port Mode is available, the Switch Port shall go to **step (j)**; if F_Port Mode is not available, ignore the FLOGI. If during the timeout period an ELP SW_ILS is received by the Switch Port, the Switch Port shall go to **step (e)**. Otherwise, after the timeout period has expired, go back to **step (b)**.

h) **Initialize as an FL_Port**. The Switch Port has detected a functional Arbitrated Loop, populated with more than one other L_Port. The Switch Port shall continue to operate as an FL_Port until the next Initialization Event.

i) **Initialize as an E_Port**. The Switch Port has completed the exchange of Link Parameters with another E_Port. If the Link Parameters exchanged were not acceptable, then the E_Port shall become Isolated and not continue with Fabric Configuration, as described in 7.4. If the Link Parameters exchanged were acceptable, then the E_Port shall participate in the next phase of Fabric Configuration, described in 7.2. In either case, the Switch Port shall continue to operate as an E_Port until the next Initialization Event.

j) **Initialize as an F_Port**. The Switch Port has detected an attached N_Port. The Switch Port shall continue to operate as an F_Port until the next Initialization Event.

7.2 Principal Switch Selection

A Principal Switch shall be selected whenever at least one Inter-Switch Link is established. The selection process chooses a Principal Switch, which is then designated as the Domain Address Manager. The behavior of a Switch during this process is as follows:

– A Switch may request a Fabric Reconfiguration at any time by transmitting a BF or an RCF request Sequence. Unless warranted by current conditions, a Switch shall always first attempt a non-disruptive Fabric Reconfiguration by sending BF request Sequence. The recommended uses of BF and RCF are summarized in table 35.

Table 35 – Recommended BF and RCF Usage Summary

Event	BF or RCF Response
A Principal ISL experiences Link Failure or a transition to Offline	BF
A configured Fabric is joined to another configured Fabric, and they do not overlap	BF
An unconfigured Switch or Fabric is joined to a configured Fabric	neither (see below)
A configured Fabric is joined to another configured Fabric, and an overlap is detected	RCF
Reconfiguration caused by BF fails for any reason	RCF

– If the Switch is attempting a non-disruptive Fabric Reconfiguration, the Switch shall transmit a BF request Sequence on all E_Ports that the Switch has not yet received a BF request. The Switch shall respond appropriately to any BF request Sequence received on any E_Port, and shall not transmit a BF request Sequence on any E_Port from which a BF request Sequence is received. Any Class F frames other than RCF requests and the associated SW_ACC and ACK_1 frames shall receive an F_BSY response, with a Reason Code of “The Fabric is busy”.

– If the Switch is attempting a disruptive Fabric Reconfiguration, the Switch shall transmit an RCF request Sequence on all E_Ports that the Switch has not yet received an RCF request. The Switch shall respond appropriately to any RCF request Sequence received on any E_Port, and shall not transmit an RCF request Sequence on any E_Port from which an RCF request Sequence is received.

– If a Switch receives an RCF request Sequence while it is in the process of attempting a non-disruptive Fabric Reconfiguration, it shall stop the non-disruptive Fabric Reconfiguration and begin processing RCF requests as described above. Any Active or Open BF Sequences shall be abnormally terminated.

– A Switch that is not yet configured (for example, after initial power-on) is not required to transmit BF or RCF. It may instead transmit an EFP SW_ILS to all initialized E_Ports to determine if the Switch is attached to a configured Fabric.

– The Switch shall wait for twice F_S_TOV following the completion of the last BF or RCF Exchange before originating an EFP request Sequence.

– The Switch shall process all EFP Payloads based on the information available at the time of processing. A Switch may receive an EFP Payload either by receiving an EFP request Sequence at an E_Port, or by receiving at an E_Port an SW_ACC reply Sequence in response to an EFP request Sequence.

- If the Switch had a Domain_ID prior to the start of a non-disruptive Fabric Reconfiguration, then the Domain_Map shall be set to the value in the last EFP request Sequence Payload received from the Domain Address Manager. If the Switch did not have a Domain_ID, or if the Switch is performing a disruptive Fabric Reconfiguration (RCF), then all bits in the Domain_Map shall be zero (“zero Domain_Map”).

- The Switch shall retain a Switch_Priority||Switch_Name value that it believes is the lowest in the Fabric. This value is initialized at the start of Fabric Reconfiguration (caused by BF or RCF) to the Switch’s value of Switch_Priority||Switch_Name. After the Switch is configured, it retains as the lowest value the Switch_Priority||Switch_Name of the Principal Switch.

- If the Switch receives in an EFP Payload a non-zero Domain_Map and the Switch has a zero Domain_Map, then the Switch shall retain the received Switch_Priority||Switch_Name as the new value. The Switch shall also note from which E_Port it received the new value, for potential use as the upstream Principal ISL during address distribution.

- If the Switch receives in an EFP Payload a zero Domain_Map and the Switch has a non-zero Domain_Map (i.e., it has received a Domain_ID), the Switch retains its current lowest Switch_Priority||Switch_Name value as the lowest value (without comparing with the received value).

- If the Switch receives in an EFP Payload a zero Domain_Map and the Switch has a zero Domain_Map, and the received Switch_Priority||Switch_Name is lower than its current retained value, it discards the old value and retains the new value. The Switch shall also note from which E_Port it received the new value, for potential use as the upstream Principal ISL during address distribution.

- The Switch shall communicate its retained Switch_Priority||Switch_Name to all E_Ports that it has not yet communicated that value. The Switch shall accomplish this either by originating a new EFP request Sequence, or by an SW_ACC reply Sequence to a received EFP request.

- If the switch receives a new lower value of Switch_Priority||Switch_Name before it has had a chance to communicate a prior lower value to all other E_Ports, it shall not attempt to communicate the prior value, and shall instead attempt to communicate the new value. The Switch shall not abort or otherwise abnormally terminate an existing EFP Exchange originated by the Switch for the sole reason of the value of Switch_Priority||Switch_Name being adjusted lower prior to the completion of the Exchange.

- The Switch shall always return the lowest known value of Switch_Priority||Switch_Name in a SW_ACC reply Sequence to an EFP request Sequence.

- If the Domain_Map of the Switch is non-zero, and the Domain_Map in a received EFP Payload is non-zero, and if no corresponding bits are set to one in both Domain_Maps, then the E_Port shall request a non-disruptive Fabric Configuration, as described above.

- If the Domain_Map of the Switch is non-zero, and the Domain_Map in a received EFP Payload is non-zero, and if any corresponding bits are set to one in both Domain_Maps, then the E_Port shall not continue with Fabric Configuration, and shall become Isolated, as described in 7.4.

- If the retained value of Switch_Priority||Switch_Name does not change for twice F_S_TOV, and if the retained value of Switch_Priority is equal to 0xFF, then there is no Switch capable of becoming a Principal Switch. The Switch shall cause all E_Ports to become Isolated, as described in 7.4.

- If the retained value of Switch_Priority||Switch_Name does not change for twice F_S_TOV, and if the retained value of the Switch_Priority||Switch_Name is equal to the value of the Switch, then the Switch has become the Principal Switch.

- If the Switch receives an AAI request Sequence, then a Principal Switch has been selected. The Switch shall request a Domain_ID as described in 7.3.

- The Switch shall continue to process and generate EFP requests as appropriate until it either: determines that it has become the Principal Switch; or, it determines it has become Isolated from all other Switches; or, it receives a BF or RCF request (which restarts the selection process); or, it {times out}; or, it receives an AAI request Sequence.

At the completion of this process, all Switches other than the DAM shall retain knowledge of the E_Port through which was received the lowest value of Switch_Priority||Switch_Name. This E_Port is the start of the first ISL in the path to the DAM for the Switch; this ISL is called the upstream Principal ISL.

7.3 Address Distribution

Once a Domain Address Manager has been selected, Switches may request a Domain_ID. The DAM shall assign all Domain_IDs. All other Switches shall request Domain_IDs from the DAM.

7.3.1 Domain_ID Distribution by the DAM

The DAM shall conduct Domain_ID distribution as follows:

- At the completion of Principal Switch Selection, the Principal Switch shall assume the role of DAM. The Principal Switch shall set its Switch_Priority value to hex'01'. The Principal Switch shall clear all bits in its Domain_Map to zero.

- The DAM shall then grant itself a Domain_ID from the pool of available Domain_IDs. This pool is maintained by the DAM. If the DAM had a specific Domain_ID prior to the Reconfiguration Event, it shall grant itself that Domain_ID, if it is available.

- The DAM shall then transmit an AAI SW_ILS request Sequence via all E_Ports. After receiving the SW_ACC reply, the DAM may receive one or more RDI SW_ILS request Sequences via one or more of the E_Ports.

- When the DAM receives an RDI SW_ILS request Sequence with a non-zero requested Domain_ID, in the absence of any error condition preventing it, it shall allocate the requested Domain_ID to the requesting Switch, if available. If the requested Domain_ID is not available or is zero, it shall grant an available Domain_ID to the requesting Switch. This Domain_ID is communicated to the Switch by transmitting the SW_ACC reply Sequence via the E_Port on which the corresponding RDI request Sequence was received. {should DAM send RCF if preferred not grant-able and BF started things?}

- The DAM shall not grant the same Domain_ID to more than one requesting Switch.

- If the DAM receives an RDI request for the same requested Domain_ID as it granted to that Switch in a previous RDI request received after DAM Selection, it shall not be considered an error; the DAM shall grant the Domain_ID to the Switch. If a Switch that has already been granted a Domain_ID transmits a request to the DAM for a different Domain_ID, the DAM shall transmit BF or RCF, as appropriate.

– If the DAM receives an RDI request and no Domain_IDs are available, the DAM shall return SW_RJT with a reason/explanation of: “Unable to perform command request”, “Domain_ID not available“.

– All Principal ISLs via which the DAM receives RDI requests shall be downstream Principal ISLs.

– Each time the DAM grants a Domain_ID to a Switch (including itself), it shall transmit an EFP SW_ILS request Sequence via all E_Ports, with each bit in the Domain_Map corresponding to a granted Domain_ID set to one.

7.3.2 Domain_ID Requests by the Switches

The Switches shall request a Domain_ID as follows:

– At the completion of Principal Switch Selection, the Switch receives the AAI SW_ILS request Sequence via the upstream Principal ISL, and shall shall reply to the request with the appropriate SW_ACC or other response. An AAI request Sequence received on any other E_Port shall be replied to with the appropriate SW_ACC or other response, but shall otherwise be ignored. The AAI request received via the upstream Principal ISL is the indication that the DAM has assigned a Domain_ID to all Switches between the DAM and the Switch receiving the AAI request.

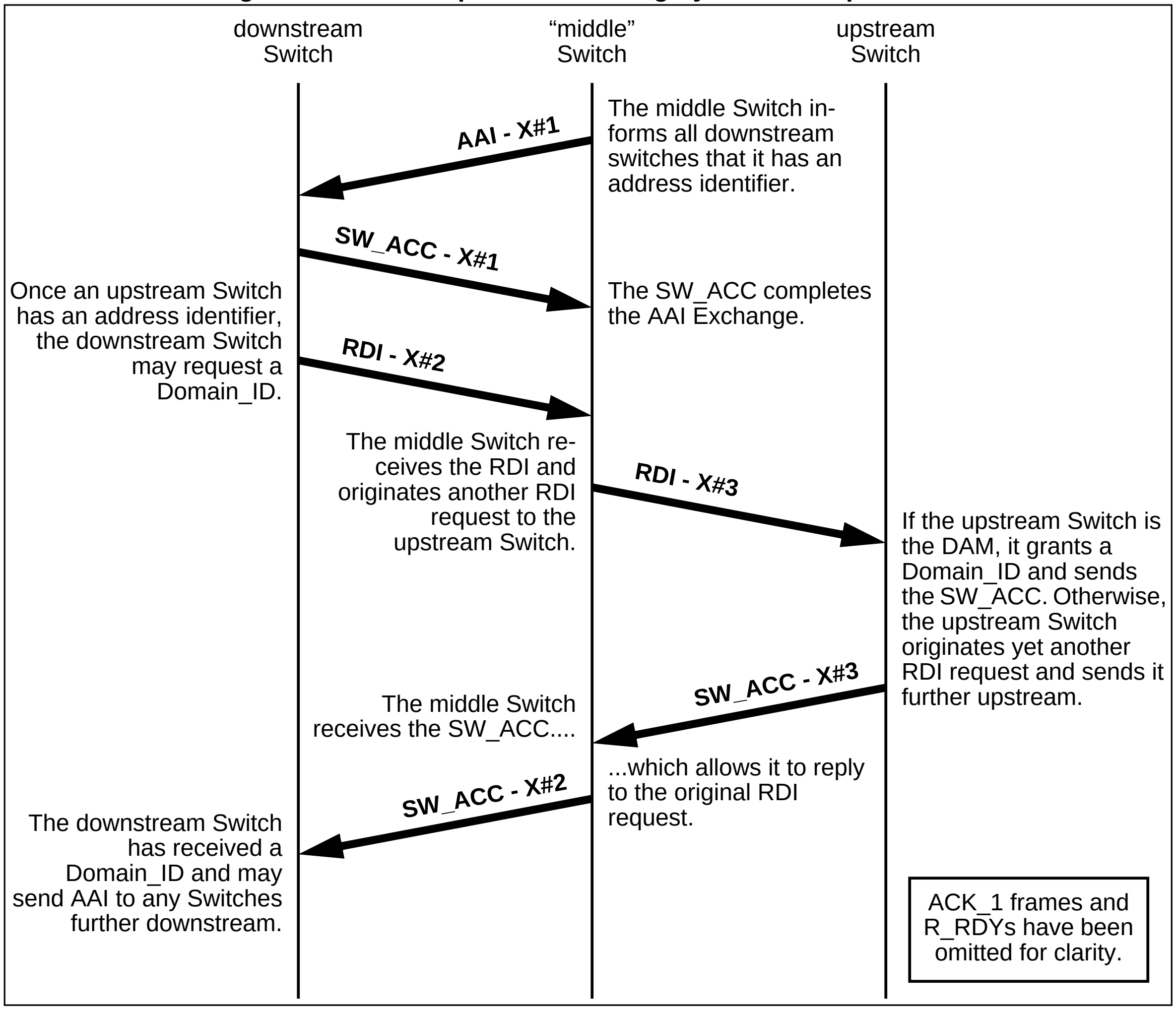
– After transmitting an SW_ACC reply to the AAI request, the Switch shall transmit an RDI request Sequence via the upstream Principal ISL. When the Switch receives the reply SW_ACC to the RDI request, it shall assign address identifiers to all Ports within its Domain as appropriate.

– After the Switch is granted a Domain_ID, it shall then transmit an AAI SW_ILS request Sequence via all E_Ports other than the Principal ISL. After receiving the SW_ACC reply, the Switch may receive one or more RDI SW_ILS request Sequences from one or more of the E_Ports.

– All Principal ISLs via which the Switch receives RDI requests shall be downstream Principal ISLs.

– When the Switch receives an RDI request Sequence from one of its E_Ports, it shall originate an RDI request Sequence with the same Payload via its upstream Principal ISL. When the reply SW_ACC is received via the upstream Principal ISL, it shall transmit an SW_ACC reply Sequence via the downstream Principal ISL on which the initial request was received. An example of this process is illustrated in Figure 11.

Figure 11 – RDI Request Processing by non-Principal Switch



7.4 E_Port and Fabric Isolation

An E_Port connected via an Inter-Switch Link to another E_Port may determine that it cannot communicate with the other E_Port for one of the reasons listed below.

– The two E_Ports have incompatible Link Parameter requirements. For example, if one Switch has an E_D_TOV setting different than another, Class 2 frames sent by an N_Port on one Switch may not receive timely F_BSY responses from the other Switch.

– The two E_Ports are a new Link between two existing Fabrics, and the Domain_ID allocations in each Fabric overlap. For example, if each existing Fabric had allocated Domain_ID hex’44’ to a Switch, one Switch would have to give up its Preferred Domain_ID to reconfigure; this could cause a major disruption to current traffic.

– The two E_Ports are a Link between Switches that are not capable of performing the DAM function, and are each also not attached via an ISL to any other Switch capable of performing the DAM function. Since no Switch can allocate Domain_IDs, no Class N frames can be sent between the Switches.

When any of the above conditions occurs, the E_Port shall Isolate itself from the other E_Port. Appropriate Class F frames may be communicated between Isolated E_Ports, but no routing of Class N frames shall occur across the ISL. {class F BB flow?}

If it is still desired to create a single Fabric via Isolated E_Ports, a Switch may override the Isolated condition by originating an RCF SW_ILS request Sequence via the appropriate ISL. The RCF shall force the selection of a single DAM from within the previously Isolated Fabrics.

Annex A
(informative)

Broadcast and Multicast Operation for Switch Fabrics

This annex defines additional services and requirements for Switch Fabrics that support Broadcast and Multicast, as defined in FC-PH-2. These additions are not intended to be comprehensive; rather, they express the current direction of the standardization effort at the time this standard was completed.

A.1 Multicast Group ID

The address identifier range hex'FFFB00' through hex'FFFBFF' will be used as multicast group identifiers...

other stuff...

Payload: The format of the XYZ request Payload is shown in table A.1.

Table A.1 – XYZ Payload

Item	Size Bytes
hex ‘xx 00 00 00’	4

Annex A
(informative)

Link Switches

This annex defines the Link Switch....

A.1 Extended Stuff

The following new stuff...

Payload: The format of the PDQ request Payload is shown in table A.2.

Table A.2 – PDQ Payload

Item	Size Bytes
hex ‘yy 00 00 00’	4

A

F

L

N

S