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

TR X3.xxx-199x
X3T11/Project xxxT/Rev 1.1

FIBRE CHANNEL

PUBLIC LOOP PROFILE

(FC-PLP)

REV 1.1

X3 working draft proposed
Technical Report

May 22, 1996

Secretariat:
Information Technology Industry Council

ABSTRACT:

NOTE:
This is a possible future draft proposed Technical Report of Accredited Standards Committee X3. As such, this is not a completed report. 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 technical report.

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

Carl Zeitler (X3T11 Vice Chairman)
IBM Corporation - AWD, MS 9440
11400 Burnet Road
Austin, TX 78758
(512) 838-1797
Fax: (512) 838-1852
E-Mail: zeitler@ausvm6.vnet.ibm.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 Avenue
Irvine, CA 92718
(714) 455-2908
Fax: (714) 455-9287
E-mail: stai@brocadesys.com

Editor’s Notes, revision 1.1:

- Extensive changes throughout to tighten up the specs and ‘de-SCSI’ the generic port clauses, as a result of the May 1-2 working meeting.
- Clause 4 now has some explanation for all this stuff.
- The FL_Port section has been reworked to separate the fabric function from the other services provided by the servers at “well-known addresses”.
- The “N_Port-to-F_PortInteroperability Profile” has been incorporated, including F_RJT stuff.
- The concept of keeping NL_Port and FL_Port in separate chapters seems somewhat at odds with the notion of listing originator and recipient requirements in one table in cases where the one can only be the originator and the other only the recipient. This tends to make those tables a little odd. Topic for discussion May 29-30.
- I have merged what used to be clauses 8, 9, 10 into clause 8 to make one big SCSI clause. I have made one pass through the SCSI stuff to de-private and de-class-3 it. Go see how well/badly I did.
- Initialization still needs work in each main clause 5,6,7. So does Annex B.
- There was much discussion at the end of the May 1-2 meeting about Annex A, and some big changes were proposed. In an effort to capture those ideas and present them for discussion, I have duplicated all Annex A clauses associated with Loop Change and edited the changes into the duplicate. There was also discussion about whether these services should be directed to the Loop base Address directly, or to the Fabric Controller (x’FFFFFD’). Topic for discussion May 29-30.

Editor’s Notes, revision 1.0:

- I have broken out NL_Ports, FL_Ports, and N_Ports into separate chapters for now. I think that will help with the discussion process. As we proceed, we will certainly discover commonalities that will allow collapsing sections.
- The following individuals contributed to this draft report:
  - Stuart Berman (Emulex)
  - Jim Coomes (Seagate)
  - Tom Harrington (LSI Logic)
  - Kumar Malavalli (Brocade)
  - Charles Monia (DEC)
  - Wayne Rickard (Gadzoox)
  - Pak-Lung Seto (Quantum)
  - Robert Snively (Sun)
  - Jeff Stai (Brocade)
  - Bent Stoevhase (Brocade)

ANSI ®
TR X3.xxx-199x

draft proposed X3 Technical Report

Fibre Channel —
Public Loop Profile (FC-PLP)

Secretariat

Information Technology Industry Council

Approved ,199

American National Standards Institute, Inc.

Abstract

This profile 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 the profile should interoperate. Small Form Factor document SFF-8045 defines the native device connector for disk devices.

X3's Technical Report Series

This Technical Report is one in a series produced by the American National Standards Committee, X3, Information Technology. The secretariat for X3 is held by the Computer and Business Equipment Manufacturers Association (CBEMA), 1250 Eye Street, NW Suite 200, Washington DC 20005.

As a by-product of the standards development process and the resources of knowledge devoted to it, X3 from time to time produces Technical Reports. Such Technical Reports are not standards, nor are they intended to be used as such.

X3 Technical Reports are produced in some cases to disseminate the technical and logical concepts reflected in standards already published or under development. In other cases, they derive from studies in areas where it is found premature to develop a standard due to a still changing technology, or inappropriate to develop a rigorous standard due to the existence of a number of viable options, the choice of which depends on the user's particular requirements. These Technical Reports, thus, provide guidelines, the use of which can result in greater consistency and coherence of information processing systems.

When the draft Technical Report is completed, the Technical Committee approval process is the same as for a draft standard. Processing by X3 is also similar to that for a draft standard.

PATENT STATEMENT

CAUTION: The developers of this Technical Report have requested that holder's of patents that may be required for the implementation of the Technical Report, disclose such patents to the publisher. However, neither the developers nor the publisher have undertaken a patent search in order to identify which, if any, patents may apply to this Technical Report.

As of the date of publication of this Technical Report and following calls for the identification of patents that may be required for the implementation of the Technical Report, no such claims have been made. No further patent search is conducted by the developer or the publisher in respect to any Technical Report it processes. No representation is made or implied that licenses are not required to avoid infringement in the use of this Technical Report.

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

Page

x

Page

draft proposed X3 Technical Report
for Information Technology—

Fibre Channel —
Public Loop Profile (FC-PLP)

1 Introduction and Scope

This Technical Report specifies Fibre Channel options between Public NL_Ports and FL_Ports, between Public NL_Ports and other Public NL_Ports (whether on the same or different loops), and between Public NL_Ports and N_Ports. This report also specifies SCSI and IP options for communication across a fabric between NL_Ports and N_Ports.

This report is intended to serve as an implementation guide whose primary objective is to maximize the liklihood of interoperability between conforming implementations. This specification prohibits and requires features which are neither prohibited nor required (i.e., are optional) in ANSI standards.

A second objective of this Technical Report is to simplify implementations and their associated documentation, testing, and support requirements. This means that there will be some optional features which are not mutually exclusive, but are still prohibited or required solely for the purpose of this simplification.

Internal characteristics of conformant implementations are not defined by this Technical Report. This Technical Report incorporates features from the standards described below. Where needed, changes are or have been proposed to the appropriate ANSI X3 technical committees to ensure this Technical Report remains a strict subset of ANSI standards.

2 Normative references

The following American National Standards contain provisions which, through reference constitute provisions of FC-PLP. 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. Members of IEC and ISO maintain registers of currently valid International Standards. ANSI performs a similar function for American National Standards.

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

[2] ANSI X3.xxx-199x, *Fibre Channel - Physical and Signalling Interface-2 (FC-PH-2)*, X3T11/Project 901D/Rev 7.3

[3] ANSI X3.xxx-199x, *Fibre Channel - Physical and Signalling Interface-3 (FC-PH-3)*, X3T11/Project 1119D/Rev 8.6

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

[5] ANSI X3.xxx-199x, *Fibre Channel Arbitrated Loop (FC-AL-2),* X3T11/Project 1133D/Rev 5.0

[6] ANSI X3.288-199x, *Fibre Channel - Generic Services (FC-GS)*, X3T11/Project 1050D/Rev 3.0

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

[8] ANSI X3.289-199x, *Fibre Channel - Fabric Generic (FC-FG)*, X3T11/Project 958D/Rev 3.3

[9] ANSI X3.xxx-199x, *Fibre Channel - Switched Fabric (FC-SW)*, X3T11/Project 959D/Rev 2.0

[10] ANSI TR.xxx-199x, *Fibre Channel - Private Loop Direct Attach (FC-PLDA)*, X3T11/Project xxxxT/Rev 1.2

[11] ANSI X3.131-1994, *Small Computer Systems Interface - 2 (SCSI-2)*

[12] ANSI X3.270-1996, *SCSI-3 Architecture Model (SAM)*

[13] ANSI X3.269-1996, *Fibre Channel Protocol for SCSI (FCP)*

[14] ANSI X3.xxx-199x, *SCSI-3 Primary Commands (SPC),* X3T10/Project 995D/Rev 4

[15] ANSI X3.xxx-199x, *SCSI-3 Block Commands (SBC),* X3T10/Project 996D/Rev 0

[16] ANSI X3.xxx-199x, *SCSI-3 Controller Commands (SCC)*, X3T10/Project 1047D/Rev 5

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

[18] *FCSI SCSI Profile*, Rev 2.2

3 Definitions and conventions

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

3.1 Definitions

3.1.1 Available_Receive_Buffers: The number of buffers in a receiving port which are available for receiving frames at link rate. Equal to the largest number of R_RDYs an NL_port can issue immediately upon transmission or receipt of an OPN. BB Flow control rules keep this number greater than or equal to the Available BB_Credit variable in a transmitting port. The number of available receive buffers must be greater than or equal to Login_BB_Credit upon receipt of any OPN.

3.1.2 Byte: A group of eight bits. Bytes are packed four per 32-bit word, or eight per 64-bit word.

3.1.3 Collation: The action by an FL_Port of collecting frames for delivery at a later time to a Loop Device.

3.1.4 F_Port: As defined in FC-PH. In this report, an F_Port is assumed to always refer to a port to which non-loop N_Ports are attached, and does not include FL_Ports.

3.1.5 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. The AL_PA of an FL_Port is x'00'.

3.1.6 Frame Bundling: A process whereby the Fabric collects frames for a source and destination pair before attempting to deliver the collected frames within a single Loop Tenancy to the destination NL_Port.

3.1.7 Frame Piling: A process whereby the Fabric collects frames for a destination from any source before attempting to deliver the collected frames within a single Loop Tenancy to the destination NL_Port.

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

3.1.9 Link Service Facilitator: The entity at the well-known address hex ‘FF FF FE’. (This is called the “Fabric F_Port” in FC-PH.)

3.1.10 LIS_HOLD_TIME: The maximum time allowed for each node to forward a loop initialization sequence frame.

3.1.11 Local Loop: The Physical Loop to which an L_Port is directly attached.

3.1.12 Loop Base Address: The 24-bit Port ID of the entity residing at the lowest address on an Arbitrated Loop. On a Public Loop, this is effectively the address of the entity at AL_PA hex ‘00’, which is the FL_Port.

3.1.13 Loop Device: A device with at least one NL_Port.

3.1.14 Loop_ID: Loop_IDs are 7-bit values numbered contiguously from 0 to 126 decimal and represent the 127 legal addresses on a loop (not all of the 256 possible AL_PAs are used in FC-AL for reasons related to running disparity).

3.1.15 Loop Tenancy: The period of time beginning when a port wins arbitration and ending when it receives a CLS in response to it’s own CLS, or forwards a CLS transmitted to it.

3.1.16 N_Port: As defined in FC-PH. In this report, an N_Port is assumed to always refer to a direct fabric-attached port, and does not include NL_Ports.

3.1.17 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.

3.1.18 OPN Originator: The port on an Arbitrated Loop that sent the OPN primitive.

3.1.19 Physical Loop: An Arbitrated Loop as defined in the FC-AL standard. It is useful to note that a Physical Loop can be thought of as a set of devices and the interconnects between them arranged such that any FC-PH or FC-AL frame or primitive may be sent from one device, pass through every device once, and then return to the sending device. What is important is the loop functionality, rather than the nature of the physical interconnect and routing between the devices. {editor’s note: this definition is intended to allow for some of the clever interconnects being talked about, without getting hung up on what ‘physical’ means. Help me out here!}

3.1.20 PL_TIME: Physical Loop Time, the maximum time allowed for a signal to propagate around the maximum size loop permitted by this document.

3.1.21 Preferred Address: On FC-AL, the AL_PA which an NL_Port attempts to acquire first during loop initialization. {more TBD}

3.1.22 Previously Acquired Address: {this definition may change and be moved to a clause on initialization} This address only has meaning during loop initialization. During initialization, it is the AL_PA which was in use prior to receipt of LIP. Immediately following power-on and between the time a loop initialization completes and the next one begins, an NL_Port has no previously acquired address.

3.1.23 Private loop device: A device with only private NL_Ports.

3.1.24 Private NL_Port: An NL_Port which is observing the rules of private loop behavior (see clause 4.1).

3.1.25 Public loop device: A device with at least one public NL_Port.

3.1.26 Public NL_Port: An NL_Port which can observe the rules of either public or private loop behavior (see clause 4.1). A public NL_Port may have open Exchanges with both private and public NL_Ports concurrently.

3.1.27 Remote Loop: The Physical Loop separate from the Local Loop to which the other communicating NL_Port is directly attached. A Remote Loop can only be reached across the Fabric.

3.1.28 Sequence Collation: A process whereby the Fabric collects all frames for a Sequence before attempting to deliver the collected frames within a single Loop Tenancy to the destination NL_Port.

3.1.29 Tenancy Collation: A process whereby the Fabric collects all frames for a Loop Tenancy from an NL_Port to an NL_Port on a different Loopbefore attempting to deliver the collected frames within a single Loop Tenancy to the destination NL_Port.

3.2 Editorial conventions

In this Technical Report, 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 Technical Report 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 note: open issues and notes are indicated in squiggly braces {}.

Hexadecimal notation

Hexadecimal notation is used to represent fields. For example, a four-byte Process_Associator field containing a binary value of 00000000 11111111 10011000 11111010 is shown in hexadecimal format as x’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.

3.3.1 Acronyms and abbreviations

A_W_TOV	Arbitration_Wait_Timeout value
BLS	Basic Link Service
E_D_TOV	Error_Detect_Timeout value
ELS	Extended Link Service
FC-AL	Fibre Channel Arbitrated Loop, reference [4]
FC-AL-2	Fibre Channel Arbitrated Loop-2, reference [5]
FC-FG	Fibre Channel - Fabric Generic, reference [8]
FC-GS	Fibre Channel - Generic Services, reference [6]
FC-GS-2	Fibre Channel - Generic Services-2, reference [7]
FC-PH	Fibre Channel Physical and Signaling Interface, reference [1]
FC-PLDA	Fibre Channel - Public Loop Direct Attach, reference [10]
FC-SW	Fibre Channel - Switched Fabric, reference [9]
FCP	Fibre Channel Protocol, reference []
FCSI	Fibre Channel Systems Initiative
IP	Internet Protocol
IU	Information Unit
FRU	Field Replaceable Unit
LAN	Local Area Network
LIS_HOLD_TIME	Loop Initialization Sequence Hold Time
LLC	Logical Link Control
L_M_TOV	Loop_Master_Timeout value
MAC	Media Access Control
NFS	Network File System or Network File Server
PL_TIME	Physical Loop Delay
R_A_TOV	Resource_Allocation_Timeout value
RR_TOV	Resource_Recovery_Timeout value
R_T_TOV	Resource_Allocation_Timeout value
SCSI	Small Computer System Interface
SI	Sequence Initiative
ULP	Upper Level Protocol
ULP_TOV	Upper_Layer_Protocol_Timeout value
WAN	Wide Area Network
WWN	World Wide Name

3.3.2 Symbols

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

||	concatenation

3.4 Applicability and Use of This Technical Report

If a feature is Prohibited, it means that it shall not be used between compliant implementations. It does not mean that an implementation may not use that feature to communicate with non-compliant implementations. However, interoperability is not guaranteed if prohibited features are used, or if required features are not implemented.

This Technical Report does not prohibit the implementation of features, only their use between compliant implementations. The only features required or prohibited are those which have been determined to result in non-interoperability if implemented differently.

The tables in the following clauses list features described in the various standards specific to the operations described in the clause. These tables indicate whether the feature is required, prohibited, or optional for compliance with this report. Features which are not listed do not affect interoperability of public loop devices.

The following legend is used for table entries in these clauses:

‘n’ the parameter shall be set to this value
‘P’ the implementation is Prohibited from using the specified feature
‘R’ the implementation is Required to support the specified feature
‘I’ the implementation may Invoke the specified feature
‘A’ use of the specified feature is Allowed (optional)
‘X’ this parameter has no required value
‘-’ this parameter is not meaningful

4 Structure and Concepts

This clause provides an overview of a public loop environment which allows NL_Ports and N_Ports to communicate across a Fabric without regard to the details of the topology. This environment also provides for efficient use of the loop topology by the FL_Port.

4.1 Summary of Private versus Public Loop Behaviors

The definition of Public and Private Loop behavior is as follows:

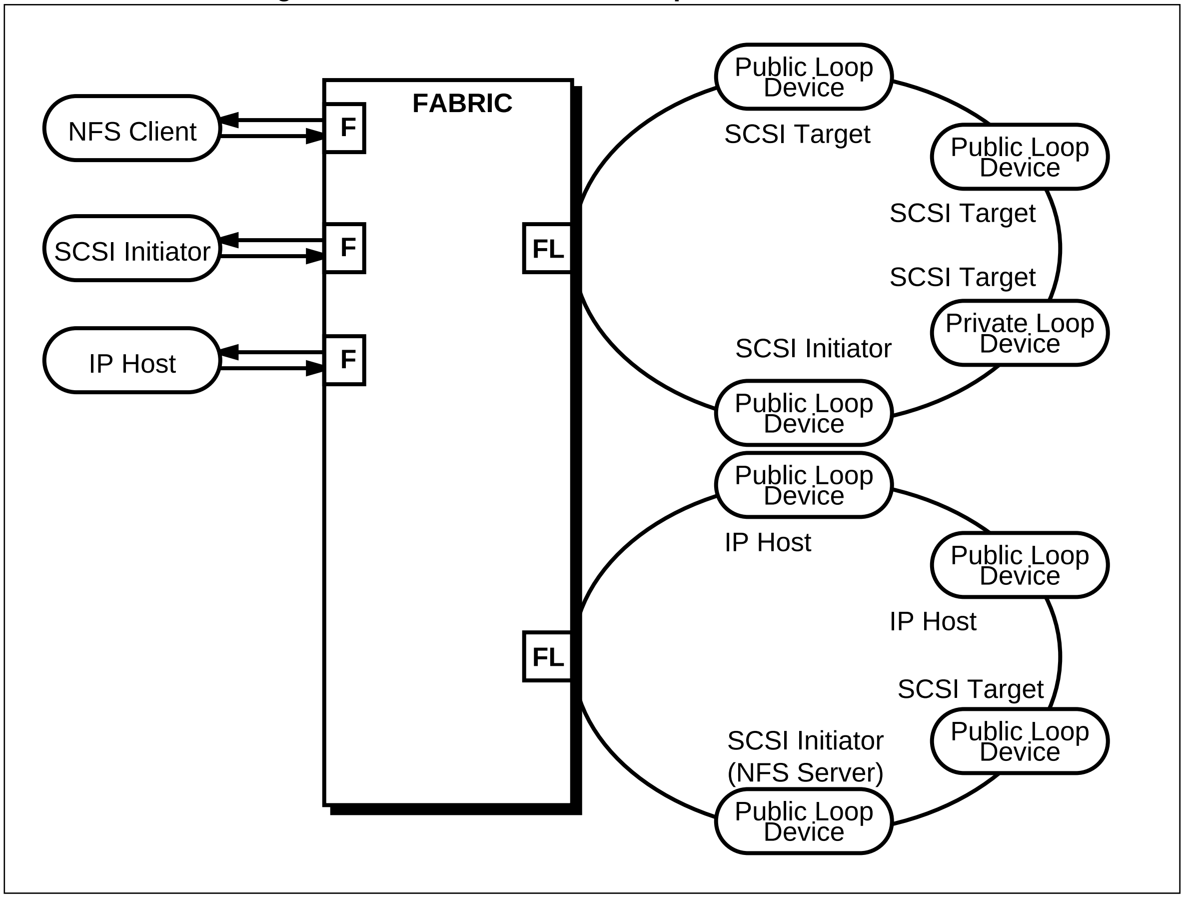
Table 1 – Private vs Public Loop Behavior

Behavior	Private Loop Behavior	Public Loop Behavior
Domain + Area of device’s NL_Port ID=0	Required	Prohibited
NL_Port may Open AL_PA=0	Prohibited	Required
FL_Port may respond to AL_PA=0	Prohibited	Required
FL_Port may Open device	Prohibited	Required
F/NL_Port may Open device	Prohibited	Required
Public NL_Port may Open device	Required	Required

A port which **only** exhibits Private Loop behavior is called a Private NL_Port. A port which exhibits Public behavior is called a Public NL_Port, even though it may communicate with Private NL_Ports. For example, a Public NL_Port in Figure 1 may be an NFS server which communicates with NFS Clients residing directly on the fabric using IP, and with local SCSI Targets on the same loop using FCP.

Devices with only private NL_Ports are called private loop devices. A private loop device is prohibited from providing or requesting fabric services. Devices with at least one Public NL_Port are called public loop devices.

Figure 1 – Private and Public Loop Device Coexistence



4.2 Private Loop and Public Loop Addressing

A 24-bit NL_Port native address identifier is divided into three 8 bit fields. In order of bit significance these are: Domain, Area, and Port. The AL_PA of a Private NL_Port corresponds to its Port address. No two ports on the same Local Loop shall share the same AL_PA. All ports on the same Local Loop shall have the same Domain+Area.

NOTE – The FC-PLDA report states that the Domain+Area of a private NL_Port is zero, and that a private NL_Port shall accept frames from NL_Ports which have any value for their Domain+Area in the S_ID of the frame header.

The following table summarizes how S_ID and D_ID and OPNs shall be created in various situations. The S_ID and D_ID are created by the sender of the frame, and the sender performs an OPN on the source loop. The FL_Port performs the OPN on the destination loop; this OPN is received by the frame recipient. If both source and destination are on the same Local Loop, the OPN is performed by the frame sender only.

Table 2 – Private and Public Loop Addressing

Frame sent from	Frame received by	OPN AL_PD		S_ID		D_ID	
		source loop	dest. loop	23:8	7:0	23:8	7:0
Public NL_Port	Fabric-Attached N_Port	x’00’	n/a	Local D&A	Source AL_PA	N_Port ID	
Fabric-Attached N_Port	Public NL_Port	n/a	D_ID 7:0	N_Port ID		D&A	Dest. AL_PA
Local Public NL_Port	Remote Public NL_Port	x’00’	D_ID 7:0	Local D&A	Source AL_PA	Remote D&A	Dest. AL_PA
Remote Public NL_Port	Local Public NL_Port	x’00’	D_ID 7:0	Remote D&A	Source AL_PA	Local D&A	Dest. AL_PA
Local Public NL_Port	Local Public NL_Port	D_ID 7:0		Local D&A	Source AL_PA	Local D&A	Dest. AL_PA
Local Public NL_Port	Local Private NL_Port	D_ID 7:0		Local D&A	Source AL_PA	x’0000’	Dest. AL_PA
Local Private NL_Port	Local Public NL_Port	D_ID 7:0		x’0000’	Source AL_PA	Remote D&A	Dest. AL_PA
Local Private NL_Port	Local Private NL_Port	D_ID 7:0		x’0000’	Source AL_PA	x’0000’	Dest. AL_PA

Notes:

“D&A” refers to the Domain and Area; “Local” means the Domain and Area on the Local Loop, “Remote” means the Domain and Area of the Remote Loop.

4.3 FL_Port Operation

The FL_Port is the gateway to the world for the Public NL_Ports on its Local Loop, allowing NL_Ports to act as if they were N_Ports and interact with other N_Ports attached to the fabric.

NOTE – For purposes of this report we restrict the meaning of N_Port to “direct Fabric-attached N_Port”. If a device is loop-attached, it is an NL_Port. When we refer to a characteristic that applies to either, we will explicitly refer to both. This usage also applies to F_Ports and FL_Ports. See clauses 3.1.4, 3.1.5, 3.1.16, and 3.1.17.

An FL_Port has a responsibility to use the loop efficiently; i.e., delivering one frame per Loop Tenancy is probably not very efficient. Therefore, the FL_Port is expected to collect (collate) frames using an appropriate algorithm, and to attempt to deliver the collected frames to the destination within a single Loop Tenancy.

The FL_Port must also manage the initialization and control of the Local Loop to the extent that a N_Port can interoperate with the NL_Ports on the Local Loop as if those NL_Ports were direct fabric-attached devices. The FL_Port must perform LIPs and AL_PA assignments transparently to the rest of the fabric. N_Ports and NL_Ports do discovery of other N_Ports and NL_Ports via the Directory Server.

However, the fabric may provide optional Extended Link Services so that an N_Port may exercise more control over a Remote Loop.

4.3.1 Collation of Frames

Frames may be collated using a variety of different algorithms. This section will provide a survey of interesting methods identified during the development of this report. The following table summarizes these methods:

Table 3 – Frame Collation Algorithms

Collation Method	Collate Frames based on...			
	D_ID	S_ID	SEQ_ID	OPN-CLS at Source
Sequence Collation	X	X	X (1)	-
Tenancy Collation	X	X	-	X (2)
Frame Bundling	X	X	-	-
Frame Piling	X	-	-	-

Notes:

1 SEQ_ID is unique for each D_ID/S_ID pair, as per FC-PH clause 18.6.

2 N_Ports cannot source Tenancy Collation because there is currently no way for an N_Port to define the OPN/CLS boundaries.

Frame Piling is the most general case, and most closely resembles N_Port to N_Port behavior, with the added latency introduced by the accumulation process and other loop-induced latencies. Frames are accumulated from all sources and delivered to the destination in its next turn.

Sequence Collation is interesting in that it allows the fabric to present complete sequences to the destination, if the destination requires it or is helped by it. This requires more buffering by the fabric, possibly resulting in greater latency. Also, the length of Sequence that can be collated is limited by the size of the available buffer.

Frame Bundling is a general case of Sequence Collation that accumulates all frames from a single source, independent of SEQ_ID. Frame Bundling does not seem very useful, except that it can act in a similar manner to Sequence Collation if the NL_Port destination login with the source allows only one concurrent Sequence.

Tenancy Collation sounds interesting at first glance but is probably not truly useful. The general consensus was that an L_Port should be constructed to be independent of Loop Tenancy behavior, mostly because a CLS can occur at any time. Also, the burden for the fabric is great when it must track OPN-CLS behavior across fabric elements; and, if the source generates frames for multiple destinations in one Loop Tenancy, the fabric must create a separate OPN-CLS indication at each destination.

4.3.2 Delivery of Frames

The FL_Port, after accumulating frames, then delivers the frames to the destination. There is a trade-off between the latency induced by accumulation and the inefficiency induced by delivering too few frames per Loop Tenancy. Also, the FL_Port must ensure that frames get delivered in a timely manner, even if the preferred “trigger” is not acheived. The following lists some interesting trigger mechanisms for frame delivery:

– Frame delivery may be triggered when some number of frames have been accumulated using a collation algorithm. The disadvantage of this trigger is that the source may send fewer frames than the trigger level and have no more to send. This issue can be overcome by coupling this trigger with another more deterministic trigger. (Frame Piling; Frame Bundling)

- Delivery can be triggered when one or more Sequences have been accumulated. (Sequence Collation)
- Delivery can be triggered when a complete Loop Tenancy has been accumulated. (Tenancy Collation)
- Delivery can be triggered when the accumulated frames have become “old”. A time stamp can be recorded for each collection of frames. When the time stamp passes an age threshold, delivery is triggered.
- Delivery can be triggered if the fabric is becoming congested, and by delivering the accumulated frames, the fabric can relieve the congestion without losing or busying frames.
- Delivery can be triggered if a Basic Link Service or Extended Link Service frame is accumulated. This is based on the observation that BLS and ELS Sequences are used for initialization and error recovery, and should be passed along without delay. Also in many instances, after sending a BLS or ELS, the source will send no more frames to that destination until the BLS or ELS has completed.
- Delivery can be triggered whenever the loop is not busy and the FL_Port has frames to send. “Loop not busy” is defined as no devices actively arbitrating for the loop.

4.3.3 Fabric Services

A fabric may provide additional Extended Link Services that allow an N_Port or NL_Port to manage and control a Remote Loop. These services (documented in annex A), are summarized below:

- Loop Initialize (LINIT): Allows an N_Port or NL_Port to request that a LIP (of any type) be sourced by the FL_Port on a Remote Loop.
- Loop Port Control (LPC): Allows an N_Port or NL_Port to request that LPEs and LPBs (of any type) be sourced by the FL_Port on a Remote Loop.
- Loop Status (LSTS): Allows an N_Port or NL_Port to determine the current condition of a Remote Loop.
- Loop Change Subscribe (LCS): An N_Port or NL_Port uses this service to request that the fabric alert the N_Port or NL_Port to any change in state of any Remote Loop.
- Loop Change Unsubscribe (LCU): Used by an N_Port or NL_Port to cancel an LCS request.
- Loop Change Notification (LCN): Sent from the fabric to indicate a change in state of a Remote Loop.

4.4 NL_Port Operation

{TBD}

4.5 N_Port Operation with NL_Ports

The intent of this Technical Report is to provide a Fabric environment such that a direct fabric-attached N_Port need not be aware that the other N_Port with which it is communicating is in fact a loop-attached NL_Port.

Some tools have been provided within this report that can facilitate the management of Public NL_Ports in error recovery and initialization. These tools should be considered optional enhancements, and are not required for interoperation of compliant implementations.

4.6 Document Structure

This report is structured such that the features of NL_Ports, FL_Ports, and N_Ports are separate from the features of the Upper Layer Protocols. This structure will allow the use of compliant ports with ULPs not covered by this report. This has been done to increase the extensibility of Public Loop implementations.

5 NL_Port Common Feature Sets

The tables in this clause list features described in the Fibre Channel Physical and Signaling Interface standard (FC-PH) and in the Fibre Channel Arbitrated Loop (FC-AL), specific to the operation of NL_Ports. These tables indicate whether the feature is required, prohibited, or optional for compliance with this Technical Report. Features which are not listed do not affect interoperability of public loop devices.

The legend for table entries used in this clause may be found in clause 3.4 on page 5.

5.1 FLOGI Request Parameters for NL_Ports

This clause specifies the features used by an NL_Port when sending a FLOGI request. Table 4 lists the features used and parameters requested by NL_Ports with usage defined by this document. See clause 6.1 on page 25 for the requirements for the FLOGI ACC response.

Table 4 – FLOGI Features and Parameters for NL_Ports

Feature/Parameter	NL_Port	Notes
FLOGI S_ID (hex value)	x'0000'||AL_PA]	1
Class of service for FLOGI	3	
FC-PH Version Highest Version Lowest Version	 X x'20'	
BB_Credit (min)	0	
Valid Vendor Version Level	0	2
N_Port/F_Port	0	
Alternate BB_Credit Management	1	
Maximum BB Receive Data Field Size (min)	X	3
Supported Classes of Service: Class 1 Class 2 Class 3 Class 4 and up	 P A R P	
Registration with Simple Name Server at well-known address 'FFFFFC' hex after completion of FLOGI	R	4

Notes:

1 The lower 8-bits of the S_ID is set to the AL_PA value of the NL_Port, which the NL_Port received during loop initialization. The fabric retains this value for the lower 8-bits of the assigned ID, and uses this value to generate the AL_PD when it returns the FLOGI ACC.

2 Profile versions are communicated via the Report Node Capabilities (RNC) ELS. The Valid Vendor Version Level shall not be set to 1 for the purposes of communicating support of this Technical Report.

3 This field is not meaningful in the FLOGI request.

4 A Public NL_Port shall register its: Native Address Identifier; Port_Name; Node_Name; supported Classes or Service; supported FC-4 types; and its IP address if it supports IP. The Simple Name Server will be documented as part of FC-GS-2. {NOTE: there will be a section on initialization and this feature will be moved to that section.}

5.2 PLOGI Parameters for NL_Ports

This clause specifies the parameters sent by an NL_Port to an NL_Port or N_Port for a PLOGI request or PLOGI ACC. Table 5 lists the parameters requested by NL_Ports with usage defined by this document.

Table 5 – PLOGI Parameters for NL_Ports

Parameter	NL_Port	Notes
Class of service for PLOGI and PLOGI ACC	3	1
Supported Classes of Service: Class 1 Class 2 Class 3 Class 4 and up	 P A R P	2
Validation of Classes of service other than Class 3 via ADISC when communicating outside the Local Loop	R	3

Notes:

1. If an NL_Port receives a PLOGI in Class 2, it shall return a P_RJT with an Action Code of “Rejected frame may be retryable” and a Reason Code of “Class not supported”. The NL_Port shall tolerate this behavior; the expectation is that an older N_Port is trying Class 2 first before trying Class 3.
2. An NL_Port shall not advertise support for a Class of Service that is not supported by the fabric (as indicated by FLOGI ACC) when sending a PLOGI or PLOGI ACC to any N_Port or NL_Port that is not attached to the Local Loop.
3. Completion of PLOGI indicates the Classes of Service that the N_Ports and/or NL_Ports may use in communications between them. Completion of PLOGI also indicates that all fabric elements between the ports support Class 3 traffic (since PLOGI is performed in Class 3 and the frames were delivered by the fabric). However, completion of PLOGI does not indicate that all fabric elements between the ports can support delivery of Class 2. After completion of PLOGI, an NL_Port that wishes to use Class 2 with another N_Port or NL_Port outside the Local Loop shall send ADISC in Class 2 and receive ADISC ACC in Class 2 before it may issue other Class 2 Sequences. {More on this as we nail down initialization for NL_Ports. Yes, Bob, I still have your flowchart...;-)}

5.2.1 PLOGI Common Service Parameters for NL_Ports

Table 6 lists PLOGI Common Service Parameters for NL_Ports with usage defined by this document.

Table 6 – PLOGI Common Service Parameters for NL_Ports

Common Service Parameter	NL_Port	Notes
FC-PH Version Highest Version Lowest Version	 X x'20'	
BB_Credit (min)	0	
Continously Increasing Relative Offset	1	
Random Relative Offset	0	
Valid Vendor Version Level	0	1
N_Port/F_Port	0	
Alternate BB_Credit Management	1	
BB Receive Data Field Size (min)	256	
Total Concurrent Sequences (min)	1	
Relative Offset by Information Category	R	
E_D_TOV	-	2

Notes:

1 Profile versions are communicated via the Report Node Capabilities (RNC) ELS. The Valid Vendor Version Level shall not be set to 1 for the purposes of communicating support of this Technical Report.

2 FC-AL is not a point-point topology, therefore E_D_TOV in login does not apply. See clause 10 for rules regarding use of various timers and timeouts.

5.2.2 PLOGI Class 2 Service Parameters

Table 7 lists PLOGI Class 2 Service Parameters with usage defined by this document.

Table 7 – PLOGI Class 2 Service Parameters for NL_Ports

Class 2 Service Parameter	NL_Port	Notes
X_ID Reassignment	00	
Initial Process Associator Required (value = 11) All other values	 P A	
ACK Initiator Capability ACK_0 ACK_N	 X 0	
ACK Generation Assistance	X	1
ACK Recipient Capability ACK_0 ACK_N	 X 0	
Initiator Data compression capable	0	
Initiator Data compression History buffer size	-	
X_ID Interlock	X	
Recipient Error Policy Supported	00	
Recipient Categories per Sequence	00	
Recipient Data compression capable	0	
Maximum Receive data field size (min)	256	
Concurrent Sequences (min)	1	
End to End Credit (min)	1	
Open Sequences per Exchange (min)	1	

Notes:

1 ACK generation assistance is possibly useful for ACK_0 generation.

5.2.3 PLOGI Class 3 Service Parameters

Table 8 lists PLOGI Class 3 Service Parameters with usage defined by this document.

Table 8 – PLOGI Class 3 Service Parameters for NL_Ports

Class 3 Service Parameter	NL_Port	Notes
Initial Process Associator Required (value = 11) All other values	 P A	
Initiator Data compression capable	0	
Initiator Data compression History buffer size	X	
Recipient Error Policy Supported	00	
Recipient Categories per Sequence=00 (one)	R	
Recipient Data compression capable	0	
Maximum Receive data field size (min)	256	
Concurrent Sequences (min)	1	
Open Sequences per Exchange (min)	1	
Notes: 1		

5.3 Other NL_Port FC-2 Features

Table 9 lists other FC-2 features for NL_Ports with usage defined by this document.

Table 9 – Other FC-2 Features for NL_Ports

Feature	NL_Port Originator	NL_Port Recipient	Notes
X_ID reassignment	P	A	
X_ID invalidation	P	A	
R_CTL Routing Bits FC-4 Device_Data frame Extended Link_Data frame Basic Link_ Data frame Link_Control frame	 R R R I	 R R R A	 1
Nonzero Continue Sequence Condition values	P	-	
Ignore nonzero Continue Sequence values	-	R	
X_ID Interlock in Class 2	A	A	
ACK Assist in Class 2	A	A	2
Sequence Retransmission	?	?	
Abort Sequence Condition	?	?	
Optional Headers (all)	P	A	
Payload size when frame is not the last frame of a sequence	(size MOD 4) == 0	(size MOD 4) == 0	3
Node Name Format IEEE IEEE Extended	 A A	 A A	 4
N_Port Name Format IEEE IEEE Extended	 A A	 A A	 4

Notes:

1 All Link Control frames are allowed for Class 2, except for ACK_N which is prohibited by this profile.

2 Values of 01 and 11 are allowed based on the result of login. A value of 10 (for ACK_N) is prohibited by this profile.

3 As per FC-PH, the size of a frame that is not the last frame of an Information Category must be a multiple of four. A frame length of zero is allowed.

4 All 64-bits of the Node Name and N_Port Name shall be significant.

{reaction to F_BSY and P_BSY in Class 2?}

5.4 Basic Link Services

All NL_Ports are prohibited from transmitting NOP {?} and RMC. NL_Ports must to be able to transmit ABTS and to accept ABTS.

5.5 FC-AL Features

Table 10 lists Fibre Channel Arbitrated Loop features for NL_Ports with usage defined by this document.

Table 10 – FC-AL Features for NL_Ports

Feature	NL_Port Originator	NL_Port Recipient	Notes
Open Full Duplex Open NL_Port on local loop Open FL_Port	 I R	 R -	
Open Half Duplex Open NL_Port on local loop Open FL_Port	 I P?	 R -	
Unfairness	A	R	1
Transfer mode (use of TRANSFER loop state)	A	R	1
LILP/LIRP	R	R	2
Multicast via Selective Replicate	P	P	
Broadcast via Broadcast Replicate	P	P	3
Alternate BB_Credit model	R	R	4
Login_BB_Credit=0	R	R	5
Login_BB_Credit>0	I	A	5
LPEyx/LPByx/LPEfx	I	A	6
LIP (F7, F7) and (F7, AL_PS) (F8, F7) and (F8, AL_PS) (AL_PD, AL_PS)	 I I I	 R R P	 7 8 9
MRK	P	R	10

Notes:

1 NL_Ports shall tolerate the use of Unfairness and Transfer Mode by other L_Ports. They don't have to like it, but they shall tolerate it!

2 Some legacy devices that do not implement LILP and LIRP would otherwise be compliant with this profile.

3 NL_Ports may use Broadcast Replicate to communicate on a Local Loop, however, the FL_Port shall not receive nor originate OPNfr. There is no currently defined method for discovering support for OPNfr, therefore the FL_Port cannot rely on it to deliver broadcast frames. {open issue}

4 Alternate BB_Credit management is mandatory in FC-AL.

5 NL_Ports shall interoperate with NL_Ports and FL_Ports which advertise any Login_BB_Credit. A recipient of frames may login with Login_BB_Credit>0, but the originator of frames is not required to take advantage of it.

6 LPEfx is useful for resetting the bypass circuits of NL_Ports which have been bypassed and have lost their AL_PAs due to a LIP (and therefore cannot be enabled using an addressed LPEyx).

7 This LIP may be issued by an NL_Port to request an AL_PA if it has none, or if it is unable to win arbitration within A_W_TOV. NL_Port response to a LIP is described in clause 5.6 on page 21.

8 This LIP may be issued by an NL_Port which detects Loss of Signal or Loss of Synchronization.

9 An NL_Port may invoke this LIP to reset a Private NL_Port. A Public NL_Port shall not receive this LIP as a Target Reset function. The NL_Port responds to this LIP in the same manner as a LIP(F7, F7).

10 An NL_Port that receives MRK shall attempt to forward the MRK. FC-AL states that an L_Port may remove a MRK if necessary for clock skew management.

5.6 NL_Port Initialization Features

NL_Ports shall initialize as follows:

a) Whenever an NL_Port receives any LIP, it shall begin initialization. The NL_Port shall suspend all open Exchanges with all other N_Ports and NL_Ports.

b) If a LIP(x'F8', X) is received, the NL_Port shall be implicitly logged out from the fabric.

c) An NL_Port shall compete to become Loop Master as defined in FC-AL. (The NL_Port is expected to lose to an FL_Port. If it doesn't then the loop does not conform to this report.)

d) If the NL_Port had an AL_PA prior to the LIP, and the NL_Port had completed FLOGI, then the NL_Port shall attempt to acquire its Fabric-assigned AL_PA during the LIFA Loop Initialization Sequence. If the NL_Port had an AL_PA prior to the LIP, and the NL_Port had not completed FLOGI, then the NL_Port shall attempt to acquire its previously-assigned AL_PA during the LIPA Loop Initialization Sequence. Otherwise, or if the NL_Port could not acquire an AL_PA during LIFA or LIPA, the NL_Port shall attempt to acquire an AL_PA during the LIHA or LISA Loop Initialization Sequences.

e) An NL_Port shall implicitly log out with the Fabric, and explicitly LOGO and clear all open Exchanges with all N_Ports and NL_Ports, if the L_bit was set to one in any Loop Initialization Sequence, or if the NL_Port was not able to acquire its AL_PA during LIFA or LIPA. {more cases?}

f) After completion of Loop Initialization, the NL_Port shall perform a FLOGI if required by a previous step. If the N_Port_ID received in the FLOGI ACC is different than the N_Port_ID of the NL_Port prior to Loop Initialization, the NL_Port shall explicitly LOGO and clear all open Exchanges with all N_Ports and NL_Ports.

g) The NL_Port shall perform PLOGI implicitly or explicitly with the Directory Server at x'FFFFFC' and register at least the following with the Simple Name Server: Native Address Identifier; Port_Name; Node_Name; supported Classes of Service; supported FC-4 types; and its IP address if it supports IP.

h) The NL_Port shall perform {some sort of discovery of other N_Ports and NL_Ports using the Simple Name Server} and perform PLOGI with those N_Ports and/or NL_Ports. If the NL_Port wishes to use Class 2 to communicate with an N_Port or NL_Port, and FLOGI and PLOGI with that port establish that it is possible, the NL_Port shall originate an ADISC ELS with that port in Class 2. If the ADISC ACC is received successfully, the NL_Port may use Class 2 with the other port; otherwise, Class 2 may be assumed to not be available on that path.

i) If the NL_Port still has open Exchanges from before the start of initialization, it shall authenticate the Exchanges. For each N_Port or NL_Port for which the NL_Port has an open Exchange, the NL_Port shall originate an ADISC ELS. If the N_Port_ID||Port_Name||Node_Name do not match a logged-in N_Port or NL_Port, the NL_Port shall originate LOGO to that port. If the N_Port_ID||Port_Name||Node_Name does match a logged-in N_Port or NL_Port, the NL_Port may resume the associated Exchange(s).

5.6.1 Responses to Link Service Frames Before PLOGI

Table 11 summarizes the frame that shall be generated as a result of receiving different Link Service frames when the recipient port is not currently logged in with the sending port. The entries in this table only apply to Link Services and Sequences associated with Link Services, not FCP-level frames.

Table 11 – Responses to Link Service Frames from Logged Out Ports

Frame Received	N_Port or NL_Port Not Logged In (PLOGI)	N_Port or NL_Port Logged In (PLOGI)
ABTS	Discard and send LOGO	BA_ACC, BA_RJT (note 2)
ADISC	Discard and send LOGO	ACC (note 1), LS_RJT
LOGO	ACC	ACC
PDISC	Discard and send LOGO	ACC (note 1), LS_RJT
PLOGI	ACC	ACC
PRLI	Discard and send LOGO	ACC
PRLO	Discard and send LOGO	ACC (note 3)
RLS	Discard and send LOGO	ACC
RRQ	Discard and send LOGO	ACC

Notes:

1 all three identifiers: N_Port ID, N_Port Name, and Node Name must match a logged-in port for ACC to be returned.

2 BA_ACC or BA_RJT as specified in clause 8.12.2

3 if no PRLI, set the reason code to “image pair does not exist”

5.6.2 Responses to ULP-Level Frames

Table 12 summarizes the frame that shall be generated as a result of receiving different frames associated with a ULP-level Sequence when the recipient port is either currently not logged in with the sending port, or is logged in but has not received a Process Login from the sending port.

Table 12 – Responses to ULP-Level Frames from Logged Out Ports

Frame Received	No PLOGI	PLOGI but no PRLI
ABTS	Discard and send LOGO	Discard and send PRLO
ULP frame	Discard and send LOGO	Discard and send PRLO
RRQ	Discard and send LOGO	Discard and send PRLO

5.7 Extended Link Services

Table 13 lists Extended Link Services with usage defined by this document. Devices which receive requests for Extended Link Services which are unsupported shall return LS_RJT with a reason code "Command not supported."

Table 13 – Extended Link Services for NL_Ports

Feature	NL_Port Originator	NL_Port Recipient	Notes
ABTX	P	P	
ADISC	R	R	1
ADVC	P	P	
ECHO	P	P	
ESTC	P	P	
ESTS	P	P	
FDISC	I	P	2
FLOGI	R	P	2
LOGO	R	R	
PDISC	I	R	1
PLOGI	R	R	
PRLI	R	R	
PRLI Common Service Parameters	P	P	
Single Service Parameter page per PRLI request	R	R	
Multiple Service Parameter pages per PRLI request	P	A	
ACC contains only those pages specified in PRLI	R	R	
Accept Response code="Command executed"	R	R	
PRLO	R	R	
RCS	P	P	
RES	P	P	
RLS	I	A	3
RNC (previously called RVU)	R	R	
RRQ	R	R	
RSI	P	P	
RSS	P	P	
RTV	P	P	
TEST	P	P	
TPRLO	R	R	

Notes:

1 ADISC is the preferred method to authenticate addressing following loop initialization.

2 Public NL_Ports do not receive FDISC or FLOGI requests.

3 Not all fields in the LESB may be supported (see vendor-specific documentation). RLS can be a diagnostic tool for isolating link degradation in loop hub or backplane topologies. Using RLS, diagnostic software can use a combination of self-directed test frames, LPB, and LESB polling to determine which device or link segment is causing link errors.

6 Fabric and FL_Port Feature Sets

The tables in this clause list features described in the Fibre Channel Physical and Signaling Interface standard (FC-PH) and in the Fibre Channel Arbitrated Loop (FC-AL), specific to the operation of Fabrics and FL_Ports. These tables indicate whether the feature is required, prohibited, or optional for compliance with this specification. Features which are not listed do not affect interoperability of public loop devices.

The legend for table entries used in this clause may be found in clause 3.4 on page 5.

6.1 FLOGI ACC Parameters to NL_Ports

This clause specifies the FLOGI ACC parameters returned by an FL_Port to an NL_Port to complete login. Table 14 lists the features used and parameters returned by FL_Ports with usage defined by this document. See clause 5.1 on page 14 for the requirements for the FLOGI request.

Table 14 – FLOGI ACC Features and Parameters for FL_Ports

Feature/Parameter	FL_Port	Notes
Class of service for FLOGI ACC	3	1
Returned D_ID in FLOGI frame header	ddaa||AL_PA	2
Maximum size of FLOGI ACC frame	256	
FC-PH Version Highest Version Lowest Version	 X x'20'	
BB_Credit (min)	0	
Valid Vendor Version Level	0	3
N_Port/F_Port	1	
Alternate BB_Credit Management	1	
Maximum BB Receive Data Field Size (min)	256	4
R_A_TOV	r_a_tov	5
E_D_TOV	e_d_tov	5
Supported Classes of Service: Class 1 Class 2 Class 3 Class 4 and up	 P A? R P	 6 7
Sequential Delivery bit Class 2 Class 3	 1 1	

Notes:

1 If a fabric receives a request for FLOGI to x'FFFFFE' in any Class of Service other than Class 3, it shall issue a F_RJT. The F_RJT shall begin with SOFnX and end with EOFt. The Action Code shall be “Rejected frame is not retryable”, and the Reason Code shall be “Class not supported by entity at x'FFFFFE'”.

2 The fabric shall build the assigned ID by using the requestor's Port_ID (AL_PA) for the low order byte (the AL_PA was transferred in the FLOGI request as the low order byte of the S_ID). The fabric may freely assign the Domain_ID (dd) and Area_ID (aa). The value of ‘ddaa’ shall be the same for all Public NL_Ports on a given loop.

3 Profile versions are communicated via the Report Node Capabilities (RNC) ELS. The Valid Vendor Version Level shall not be set to 1 for the purposes of communicating support of this Technical Report.

4 Public loop devices shall be capable of receiving 256-byte frames. FC-AL initialization requires NL_Ports to be able to receive 132-byte frames. There is no restriction on the *maximum maximum* frame size (up to the limits specified by FC-PH).

5 See clause 10 on page 66 for the timer values.

6 Any SOFc1 frame received by an FL_Port shall be shall be rejected. The F_RJT shall begin with SOFn1 and end with EOFdt. The Action Code shall be “Rejected frame is not retryable”, and the Reason Code shall be “Class not supported”. All SOFn1 and SOFi1 frames shall be discarded by the FL_Port.

7 May be required for fabrics - {OPEN ISSUE}

6.1.1 Completion of Fabric Login (FLOGI)

{NOTE: By mandating Class 3 for FLOGI, we may have a little problem. FC-PH 23.3.4.1 and 23.3.4.2 describe how the N_Port and F_Port may decide that login is completed. For Class 3, we have:

– N_Port completes FLOGI when it transmits the R_RDY in response to the FLOGI ACC.

– F_Port completes FLOGI when it receives the R_RDY in response to the FLOGI ACC.

The problem is that most loop designs (as I understand them) will NOT send that R_RDY if the CLS follows the ACC frame (please correct me if I am wrong!) because there is no need to balance BB_Credit. And if the NL_Port does not have any more buffer space available, it also will not send the R_RDY. If the R_RDY is not sent, that means that technically the FLOGI does not complete.

Now, if this is truly a problem, I can see a few ways to deal with it:

a) Mandate that the CLS not be sent following a FLOGI ACC from either port until the R_RDY is returned (the NL_Port must stay on the loop until its buffer clears and it can send an R_RDY);

b) Mandate that R_RDY be sent following the receipt of a CLS when receiving a FLOGI ACC, or maybe all of the time;

c) Bag Class 3 and use Class 2 instead;

d) Deviate from FC-PH and let the completion be indicated by the CLS instead of the R_RDY.

I'm guessing that (d) is preferred... Comments? Or has this been dealt with already somewhere else?}

6.2 Well-Known Addresses

This clause specifies the well-known addresses that shall be provided by the fabric, and the N_Port characteristics of those addresses. All services provided by well-known addresses are either directly performed by a fabric element, or by some other entity known to the fabric as a whole. This report does not make the distinction; rather, it mandates that NL_Ports and N_Ports have access to required addresses and services, and that this access be available in some manner through the FL_Port or F_Port.

Table 15 lists the well-known address requirements for fabrics with usage defined by this document.

Table 15 – Well-Known Address Support for Fabrics

Parameter	Fabric Support	Notes
Well-Known Address x'FFFFFF' (Broadcast)	R?	
Well-Known Address x'FFFFFE' (Link Service Facilitator) FLOGI RRQ FDISC? ADISC? All other Extended Link Services?	R R R R? R? ?	1
Well-Known Address x'FFFFFD' (Fabric Controller) Loop Management ELS (see annex A) ?	R? A ?	
Well-Known Address x'FFFFFC' (Directory Server) Simple Name Server Directory Server ARP Server	R R A A	 2
Well-Known Address x'FFFFFB' (Time Server)	A?	
Well-Known Address x'FFFFFA' (Management Server)	P?	
Well-Known Address x'FFFFF9' (Quality of Service Facillitator)	P?	
Well-Known Address x'FFFFF8' (Alias Server)	P?	
Well-Known Address x'FFFFF7' (Security Key Distribution Server)	P?	
Well-Known Address x'FFFFF6' (Clock Synchronization Server)	P?	

Notes:

1 In FC-PH, this port is called the “Fabric F_Port”; in essence, the address to which an N_Port sends a FLOGI.

2 The Simple Name Server is defined in FC-GS-2.

6.2.1 Features for N_Ports at Well-Known Addresses

This clause specifies the supported features for an N_Port residing at a well-known address. Table 16 lists the features with usage defined by this document.

Table 16 – Features for N_Ports at Well-Known Addresses

Parameter	N_Port at well-known address	Notes
Class of service for PLOGI and PLOGI ACC	3	1
Supported Classes of Service:		1
Class 1	P	
Class 2	P	
Class 3	R	
Class 4 and up	P	
Implicit Login with N_Port at any Well-Known Address	R	2

Notes:

1 If an N_Port at a well-known address receives a request for service (other than FLOGI to x'FFFFFE') in any Class of Service other than Class 3, it shall issue a P_RJT (not F_RJT). The P_RJT shall begin with SOFnX and end with EOFt. The Action Code shall be “Rejected frame is not retryable”, and the Reason Code shall be “Class not supported”.

2 An NL_Port or N_Port is allowed to assume the login parameters defined in this clause if it has knowledge that the fabric is compliant with this report {?!}. The N_Port at the well-known address is required to support this behavior.

6.2.2 PLOGI Common Service Parameters for N_Ports at Well-Known Addresses

Table 17 lists PLOGI Common Service Parameters for N_Ports at well-known addresses with usage defined by this document.

Table 17 – PLOGI Common Service Parameters for Well-Known Addresses

Common Service Parameter	NL_Port	Notes
FC-PH Version Highest Version Lowest Version	 X x'20'	
BB_Credit (min)	0	
Continously Increasing Relative Offset	1	
Random Relative Offset	0	
Valid Vendor Version Level	0	1
N_Port/F_Port	0	
Alternate BB_Credit Management	0	2
BB Receive Data Field Size (min)	256	
Total Concurrent Sequences (min)	1	
Relative Offset by Information Category	R	
E_D_TOV	-	3

Notes:

1 Profile versions are communicated via the Report Node Capabilities (RNC) ELS. The Valid Vendor Version Level shall not be set to 1 for the purposes of communicating support of this Technical Report.

2 Since the well-known address is accessed via the fabric, Alternate BB_Credit Management does not apply.

3 Fabric attach, whether directly or via loop, is not a point-point topology, therefore E_D_TOV in PLOGI does not apply. See clause 10 for rules regarding use of various timers and timeouts.

6.2.3 PLOGI Class 3 Service Parameters for N_Ports at Well-Known Addresses

Table 18 lists Class 3 Service Parameters for N_Ports at well-known addresses with usage defined by this document.

Table 18 – PLOGI Class 3 Service Parameters for Well-Known Addresses

Class 3 Service Parameter	NL_Port	Notes
Initial Process Associator	00	
Initiator Data compression capable	0	
Initiator Data compression History buffer size	-	
Recipient Error Policy Supported	00	
Recipient Categories per Sequence=00 (one)	R	
Recipient Data compression capable	0	
Maximum Receive data field size (min)	256	
Concurrent Sequences (min)	1	
Open Sequences per Exchange (min)	1	
Notes: 1		

6.2.4 Other FC-2 Features for N_Ports at Well-Known Addresses

Table 19 lists other FC-2 features for N_Ports at well-known addresses with usage defined by this document.

Table 19 – Other FC-2 Features for Well-Known Addresses

Feature	N_Port Originator	N_Port Recipient	Notes
X_ID reassignment	P	P	
X_ID invalidation	P	P	
R_CTL Routing Bits			
FC-4 Device_Data frame	R	R	
Extended Link_Data frame	R	R	
Basic Link_ Data frame	R	R	
Link_Control frame	P	P	
Nonzero Continue Sequence Condition values	P	-	
Ignore nonzero Continue Sequence values	-	R	
Sequence Retransmission	?	?	
Abort Sequence Condition	?	?	
Optional Headers (all)	P	P	
Node Name Format			1
IEEE	A	A	
IEEE Extended	A	A	
N_Port Name Format			1
IEEE	A	A	
IEEE Extended	A	A	

Notes:

1 All 64-bits of the Node Name and N_Port Name shall be significant.

6.2.5 Basic Link Services for N_Ports at Well-Known Addresses

All N_Ports at well-known addresses are prohibited from transmitting NOP and RMC. N_Ports at well-known addresses must to be able to transmit ABTS and to accept ABTS. {?}

6.2.6 Extended Link Services for N_Ports at Well-Known Addresses

Table 20 lists Extended Link Services with usage defined by this document. Devices which receive requests for Extended Link Services which are unsupported shall return LS_RJT with a reason code “Command not supported.”

Table 20 – Extended Link Services for Well-Known Addresses

Feature	N_Port Originator	N_Port Recipient	Notes
ABTX	P	P	
ADISC	P	P	
ADVC	P	P	
ECHO	P	P	
ESTC	P	P	
ESTS	P	P	
FDISC	P	P	2
FLOGI	P	P	2
LOGO	P	R	
PDISC	R	R	
PLOGI	R	R	
PRLI	R	R	
PRLI Common Service Parameters	P	P	
Single Service Parameter page per PRLI request	R	R	
Multiple Service Parameter pages per PRLI request	P	A	
ACC contains only those pages specified in PRLI	R	R	
Accept Response code=“Command executed”	R	R	
PRLO	R	R	
RCS	P	P	
RES	P	P	
RLS	A	A	3
RNC (previously called RVU)	R	R	
RRQ	R	R	
RSI	P	P	
RSS	P	P	
RTV	P	P	
TEST	P	P	
TPRLO	R	R	

Notes:

1 ADISC is the preferred method to authenticate addressing following loop initialization.

2 Since the enitities at the well-known addresses are “initimate” with the fabric, these services are not needed. Implicit login is assumed.

3 Not all fields in the LESB may be supported (see vendor-specific documentation). RLS can be a diagnostic tool for isolating link degradation in loop hub or backplane topologies. Using RLS, diagnostic software can use a combination of self-directed test frames, LPB, and LESB polling to determine which device or link segment is causing link errors.

6.3 FC-AL Features for FL_Ports

Table 21 lists Fibre Channel Arbitrated Loop features for FL_Ports with usage defined by this document.

Table 21 – FC-AL Features for FL_Ports

Feature	FL_Port Originator	FL_Port Recipient	Notes
Open Full Duplex Open L_Port on local loop	 R	 R	
Open Half Duplex Open L_Port on local loop	 P	 R	
Unfairness	I	R	1
Transfer mode (use of TRANSFER loop state)	I	R	1
LILP/LIRP	R	R	2
Multicast via Selective Replicate	P	P	3
Broadcast via Broadcast Replicate	P	P	3
Alternate BB_Credit model	R	R	4
Login_BB_Credit=0	R	R	5
Login_BB_Credit>0	I	A	5
LPEyx/LPByx/LPEfx	I	P	6
LIP (F7, F7) and (F7, AL_PS) (F8, F7) and (F8, AL_PS) (AL_PD, AL_PS)	 I I I	 R R P	 7 8 9
MRK	P	R	10

Notes:

1. FL_Ports shall tolerate the use of Unfairness and Transfer Mode by other L_Ports. They don't have to like it, but they shall tolerate it!
2. Some legacy devices that do not implement LILP and LIRP would otherwise be compliant with this profile.
3. An FL_Port shall allow OPNfr and OPNyr activity on the Local Loop without interference. The FL_Port shall not originate or participate in this activity.
4. Alternate BB_Credit management is mandatory in FC-AL
5. FL_Ports shall interoperate with NL_Ports which advertise any Login_BB_Credit. A recipient of frames may login with Login_BB_Credit>0, but the originator of frames is not required to take advantage of it.
6. LPEfx is useful for resetting the bypass circuits of NL_Ports which have been bypassed and have lost their AL_PAs due to a LIP (and therefore cannot be enabled using an addressed LPEyx).
7. This LIP may be issued by an FL_Port to request an AL_PA if it has none, or if it is unable to win arbitration within A_W_TOV. FL_Port response to a LIP is described in clause 6.6 on page 37.
8. This LIP may be issued by an FL_Port which detects Loss of Signal or Loss of Synchronization. {NOTE: This LIP would also be used in error recovery cases where OLS would be used by an F_Port. No?}
9. An FL_Port may invoke this LIP to reset a Private NL_Port. An FL_Port shall not receive this LIP as a Target Reset function. The FL_Port responds to this LIP in the same manner as a LIP(F7, F7).
10. An FL_Port that receives MRK shall attempt to forward the MRK. FC-AL states that an L_Port may remove a MRK if necessary for clock skew management.

6.4 Fabric Reject (F_RJT) Frames

The clause describes the generation of F_RJT frames. {this draws substantially from the N_Port-to-F_Port Interoperability profile v1.0}

6.4.1 F_RJT Frame Delimiters

An F_RJT in response to an SOFc1 frame shall begin with SOFn1 and shall end with EOFdt. An F_RJT in response to a Class 2 frame shall begin with SOFn2 and shall end with EOFn. An F_RJT shall not be sent in Class 3. An F_RJT to any other Class of Service recognized by the FL_Port shall begin with SOFnX (where X is the Class) and shall end with EOFn.

6.4.2 Generation of F_RJT Frames

Table 23 summarizes the F_RJT codes that may be generated by an FL_Port compliant with this report, the situations in which these codes may be generated, and the recovery action that shall be performed by the NL_Port. This table contains all Reason Codes applicable to Class 2 and Class 3; Reason Codes that are applicable to other Classes of Service are not listed here and shall not be generated by compliant FL_Ports.

Table 22 – F_RJT Reason Codes

Reason Code	Generated by Fabric in response to		NL_Port Recovery Action	Notes
	before or during FLOGI	after FLOGI		
Invalid D_ID	not generated	The D_ID specified has not been assigned or is otherwise not valid	Abnormally terminate Sequence	1, 9
Invalid S_ID	a) Could not grant S_ID requested by NL_Port; OR b) NL_Port not logged into fabric	S_ID not equal to identifier assigned during FLOGI	Perform FLOGI	10
N_Port or NL_Port not available, temporarily	not generated	Link recovery or loop initialization in progress at destination N_Port	Abnormally terminate Sequence	1, 2
N_Port or NL_Port not available, permanently	not generated	N_Port/NL_Port or F_Port/FL_Port off line, powered down, or non-existant at destination	Abnormally terminate Sequence	1, 2
Class not supported	not generated	a) Class of Service not supported by FL_Port b) Broadcast frame (D_ID=x'FFFFFF') received in Class other than Class 3	Whack NL_Port with a large hammer	3, 4
Delimiter usage error	not generated	not generated	none	5
TYPE not supported	not generated	not generated	none	5

Table 22 – F_RJT Reason Codes

Reason Code	Generated by Fabric in response to		NL_Port Recovery Action	Notes
	before or during FLOGI	after FLOGI		
Incorrect Length	Class 3 FLOGI frame Data_Field size is larger than 256 bytes	Frame in a supported Class Data_Field size is larger than Maximum BB Receive Data Field Size	Whack NL_Port with a large hammer	4
Login required	NL_Port not logged into fabric	NL_Port no longer logged into fabric	Perform FLOGI	
Fabric path not available	not generated	Class 2 path to destination N_Port or NL_Port does not exist	Perform PLOGI for Class 3 only	6
Class of Service not supported by entity at x'FFFFFE'	FLOGI not sent in Class 3	BLS or ELS not sent in Class 3	Perform FLOGI, BLS, or ELS in Class 3	7
Invalid CS_CTL field	not generated	not generated	none	5, 8
Invalid Class of Service	not generated?	not generated?	none?	11

Notes:

1 The NL_Port shall release all local resources associated with the Sequence.

2 The NL_Port shall perform the appropriate ULP recovery, such as ABTS {TBD}.

3 See “Class of Service not supported by entity at x’FFFFFE’ for the before or during FLOGI case.

4 This is a severe error that indicates a major inconsistency between the NL_Port and the FL_Port.

5 This Reason Code shall not be generated by an FL_Port.

6 This can occur when a fabric element that supports only Class 3 is used to join other fabric elements that support Classes 2 and 3. See {?}.

7 This may also indicate an FL_Port that does not support this Technical Report.

8 Non-zero CS_CTL bits are ignored, as required by FC-PH.

9 What likely happened is that the intended N_Port or NL_Port has had its ID changed. The Directory Server may be used to find the N_Port or NL_Port and re-login with it.

10 Since the NL_Port requested S_ID is of the form x’0000’||AL_PA, this may be an indication that two NL_Ports have somehow acquired the same AL_PA. The FL_Port may wish to initiate a LIP if it detects this situation.

11 {Can anyone elaborate on the use of this code (it’s in FC-PH-2)?}

6.5 Other Fabric Features

Table 23 lists other fabric features for FL_Ports communicating with attached NL_Ports, with usage defined by this document.

Table 23 – Other Fabric Features

Feature	Fabric	Notes
Sequential Delivery to NL_Ports: Class 2 Class 3	 R R	 1,2 1
Support Frame Piling	R	3
Support Frame Bundling	A	3
Support Sequence Collation	A	4
Support Tenancy Collation	?	5

Notes:

1 Actually, the fabric shall forward deliverable frames to an NL_Port in the same order that they were received from the source port. If the source port sends frames out of Sequence order, the fabric shall not re-order them. Also, gaps occur if the fabric returns F_BSY or F_RJT to a frame, or if the destination returns P_BSY or P_RJT to a frame, or if a frame is invalid. This is the behavior described FC-PH, 23.6.7.2, bit 27.

2 This does not imply that the fabric shall guarantee sequential delivery to N_Ports attached to F_Ports. {This is an interesting statement to make, since the Sequential Delivery bits in the FLOGI ACC make no distinction between loop and direct fabric-attach ports. comments?}

3 Frames are delivered to the destination when the specified collection of frames has been accumulated, or sooner if the fabric is congested or if the loop is idle.

4 If Sequence Collation is required by the destination, the fabric shall accumulate the entire Sequence before delivery, even if the fabric is congested or the Sequence is in danger of being discarded.

5 If Tenancy Collation is required by the destination, the fabric shall accumulate all of the Sequences in the Tenancy before delivery, even if the fabric is congested or one or more of the Sequences are in danger of being discarded. If supported by the fabric, Tenancy Collation shall only be supported between NL_Ports.

6

6.5.1 EOF Modification

The fabric may change an EOFn or EOFt to EOFni in the following situations:

– CRC error detected;

– FC-1 coding error or elastic store error in data field;

– Incorrect delimiter detected (e.g., EOFdt on a Class 2 frame).

{reaction to F_BSY and P_BSY in Class 2?}

6.6 FL_Port Initialization Features

FL_Ports shall initialize as follows:

a) Whenever an FL_Port receives any LIP, it shall begin initialization.

b) If a LIP(x'F8', X) is received, the FL_Port shall be implicitly logged out from all attached NL_Ports.

c) An FL_Port shall compete to become Loop Master as defined in FC-AL. (The FL_Port is expected to win. If it doesn't then the loop does not conform to this report.)

d) The FL_Port shall perform the LIFA, LIPA, LIHA, and LISA Loop Initialization Sequences.

e) {under what conditions would an FL_Port set the L_bit?}

f) After completion of Loop Initialization, the Fabric shall assume that all acquired AL_PAs are Private NL_Ports, until those ports complete FLOGI.

g) The Directory Server at x'FFFFFC' shall initialize all acquired NL_Ports {how?}.

{more?}

7 N_Port Feature Sets

The tables in this clause list features described in the Fibre Channel Physical and Signaling Interface standard (FC-PH) specific to the operation of N_Ports that are communicating across a fabric with Public NL_Ports. These tables indicate whether the feature is required, prohibited, or optional for compliance with this specification. Features which are not listed do not affect interoperability of N_Ports with Public NL_Ports.

The legend for table entries used in this clause may be found in clause 3.4 on page 5.

7.1 N_Port Login Request Parameters for N_Ports

This clause specifies the parameters sent by an N_Port to an NL_Port to request a login. Table 24 lists the login request parameters requested by N_Ports with usage defined by this document.

Table 24 – N_Port Login Request Parameters for N_Ports

Parameter	N_Port	Notes
Recommended Class of service for PLOGI and PLOGI ACC	3	1
Supported Classes of Service: Class 1 Class 2 Class 3 Class 4 and up	 P A R P	
Validation of Classes of service other than Class 3 via ADISC when communicating outside the Local Loop	R	2

Notes:

1. If an N_Port follows the convention of attempting PLOGI first in Class 2 (Class 1 presumably being skipped because the fabric does not support it), it shall tolerate the rejection by the receiving NL_Port and proceed to attempt PLOGI in Class 3. Likewise, the NL_Port shall also be tolerant of this behavior. The N_Port should not assume that the NL_Port does not support Class 2 because the Class 2 PLOGI was rejected {?!}.
2. See related note in clause 5.2 on page 15. Note that there are certainly many other ways to do the validation, but the only method that ensures interoperability with NL_Ports is this one.

7.1.1 N_Port Common Service Parameters for N_Ports

Table 25 lists N_Port Common Service Parameters for N_Ports logging in with NL_Ports with usage defined by this document.

Table 25 – N_Port Common Service Parameters for N_Ports

Common Service Parameter	N_Port	Notes
FC-PH Version Highest Version Lowest Version	 X 20 hex	
BB_Credit (min)	-	2
Continously Increasing Relative Offset	1	
Random Relative Offset	0	
Valid Vendor Version Level	0	1
N_Port/F_Port	0	
Alternate BB_Credit Management	-	2
BB Receive Data Field Size (min)	256	
Total Concurrent Sequences (min)	1	
Relative Offset by Information Category	R	
E_D_TOV	-	2

Notes:

1 Profile versions are communicated via the Report Node Capabilities (RNC) ELS. The Valid Vendor Version Level shall not be set to 1 for the purposes of communicating support of this Technical Report.

2 This field does not apply to a direct fabric-attached device.

7.1.2 N_Port Class 2 Service Parameters

Table 26 lists Class 2 Service Parameters with usage defined by this document.

Table 26 – Class 2 N_Port Service Parameters for N_Ports

Class 2 Service Parameter	N_Port	Notes
X_ID Reassignment	00	
Initial Process Associator Required (value = 11) All other values	 P A	
ACK Initiator Capability ACK_0 ACK_N	 X 0	
ACK Generation Assistance	X	1
ACK Recipient Capability ACK_0 ACK_N	 X 0	
Initiator Data compression capable	0	
Initiator Data compression History buffer size	-	
X_ID Interlock	X	
Recipient Error Policy Supported	00	
Recipient Categories per Sequence	00	
Recipient Data compression capable	0	
Maximum Receive data field size (min)	256	
Concurrent Sequences (min)	1	
End to End Credit (min)	1	
Open Sequences per Exchange (min)	1	

Notes:

1 ACK generation assistance is possibly useful for ACK_0 generation.

7.1.3 N_Port Class 3 Service Parameters

Table 27 lists Class 3 Service Parameters with usage defined by this document.

Table 27 – Class 3 N_Port Service Parameters for N_Ports

Class 3 Service Parameter	N_Port	Notes
Initial Process Associator Required (value = 11) All other values	 P A	
Initiator Data compression capable	0	
Initiator Data compression History buffer size	X	
Recipient Error Policy Supported	00	
Recipient Categories per Sequence=00 (one)	R	
Recipient Data compression capable	0	
Maximum Receive data field size (min)	256	
Concurrent Sequences (min)	1	
Open Sequences per Exchange (min)	1	
Initial Process Associator Required Supported	 P A	
Notes: 1		

7.2 Other FC-2 Features

Table 28 lists other FC-2 features for N_Ports with usage defined by this document.

Table 28 – Other FC-2 Features for N_Ports

Feature	N_Port Originator	N_Port Recipient	Notes
X_ID reassignment	P	A	
X_ID invalidation	P	A	
R_CTL Routing Bits FC-4 Device_Data frame Extended Link_Data frame Basic Link_ Data frame Link_Control frame	 R R R I	 R R R A	 1
Nonzero Continue Sequence Condition values	P	-	
Ignore nonzero Continue Sequence values	-	R	
X_ID Interlock in Class 2	A	A	
ACK Assist in Class 2	A	A	2
Sequence Retransmission	?	?	
Abort Sequence Condition	?	?	
Optional Headers (all)	P	A	
Payload size when frame is not the last frame of a sequence	(size MOD 4) == 0	(size MOD 4) == 0	3
Node Name Format IEEE IEEE Extended	 A A	 A A	 4
N_Port Name Format IEEE IEEE Extended	 A A	 A A	 4

Notes:

1 All Link Control frames are allowed for Class 2, except for ACK_N which is prohibited by this profile.

2 Values of 01 and 11 are allowed based on the result of login. A value of 10 (for ACK_N) is prohibited by this profile.

3 As per FC-PH, the size of a frame that is not the last frame of an Information Category must be a multiple of four. A frame length of zero is allowed.

4 All 64-bits of the Node Name and N_Port Name shall be significant.

7.3 Basic Link Services

All N_Ports are prohibited from transmitting NOP {?} and RMC. N_Ports shall be able to transmit ABTS and to accept ABTS.

7.4 Extended Link Services

Table 29 lists Extended Link Services with usage defined by this document. Devices which receive requests for Extended Link Services which are unsupported shall return LS_RJT with a reason code "Command not supported."

Table 29 – Extended Link Services for N_Ports

Feature	N_Port Originator	N_Port Recipient	Notes
ABTX	P	P	
ADISC	R	R	1
ADVC	P	P	
ECHO	P	P	
ESTC	P	P	
ESTS	P	P	
FDISC	I	P	2
FLOGI	R	P	2
LOGO	R	R	
PDISC	I	R	1
PLOGI	R	R	
PRLI PRLI Common Service Parameters Single Service Parameter page per PRLI request Multiple Service Parameter pages per PRLI request ACC contains only those pages specified in PRLI Accept Response code="Command executed"	R P R P R R	R P R A R R	
PRLO	R	R	
RCS	P	P	
RES	P	P	
RLS	I	A	3
RNC (previously called RVU)	R	R	
RRQ	R	R	
RSI	P	P	
RSS	P	P	
RTV	P	P	
TEST	P	P	
TPRLO	R	R	

Notes:

1 ADISC is the preferred method to authenticate addressing following initialization.

2 N_Ports do not receive FDISC or FLOGI requests.

3 Not all fields in the LESB may be supported (see vendor-specific documentation). RLS can be a diagnostic tool for isolating link degradation in loop hub or backplane topologies. Using RLS, diagnostic software can use a combination of self-directed test frames, LPB, and LESB polling to determine which device or link segment is causing link errors.

7.5 N_Port Initialization Features

N_Ports shall initialize as follows:

a) Whenever an N_Port receives any Primitive Sequence, it shall begin initialization. The N_Port shall suspend all open Exchanges with all other N_Ports and NL_Ports.

b) If an OLS or NOS is received, the N_Port shall be implicitly logged out from the fabric.

c) After completion of Link Initialization, the N_Port shall perform a FLOGI if required by a previous step. If the N_Port_ID received in the FLOGI ACC is different than the N_Port_ID of the N_Port prior to Link Initialization, the N_Port shall explicitly LOGO and clear all open Exchanges with all N_Ports and NL_Ports.

d) The N_Port shall perform PLOGI implicitly or explicitly with the Directory Server at x'FFFFFC' and register at least the following with the Simple Name Server: Native Address Identifier; Port_Name; Node_Name; supported Classes of Service; supported FC-4 types; and its IP address if it supports IP.

e) The N_Port shall perform {some sort of discovery of other N_Ports and NL_Ports using the Simple Name Server} and perform PLOGI with those N_Ports and/or NL_Ports. If the N_Port wishes to use Class 2 to communicate with an N_Port or NL_Port, and FLOGI and PLOGI with that port establish that it is possible, the N_Port shall originate an ADISC ELS with that port in Class 2. If the ADISC ACC is received successfully, the N_Port may use Class 2 with the other port; otherwise, Class 2 may be assumed to not be available on that path.

f) If the N_Port still has open Exchanges from before the start of initialization, it shall authenticate the Exchanges. For each N_Port or NL_Port for which the N_Port has an open Exchange, the N_Port shall originate an ADISC ELS. If the N_Port_ID||Port_Name||Node_Name do not match a logged-in N_Port or NL_Port, the N_Port shall originate LOGO to that port. If the N_Port_ID||Port_Name||Node_Name does match a logged-in N_Port or NL_Port, the N_Port may resume the associated Exchange(s).

7.5.1 Responses to Link Service Frames Before PLOGI

Table 11 summarizes the frame that shall be generated as a result of receiving different Link Service frames when the recipient port is not currently logged in with the sending port. The entries in this table only apply to Link Services and Sequences associated with Link Services, not FCP-level frames.

Table 30 – Responses to Link Service Frames from Logged Out Ports

Frame Received	N_Port or NL_Port Not Logged In (PLOGI)	N_Port or NL_Port Logged In (PLOGI)
ABTS	Discard and send LOGO	BA_ACC, BA_RJT (note 2)
ADISC	Discard and send LOGO	ACC (note 1), LS_RJT
LOGO	ACC	ACC
PDISC	Discard and send LOGO	ACC (note 1), LS_RJT
PLOGI	ACC	ACC
PRLI	Discard and send LOGO	ACC
PRLO	Discard and send LOGO	ACC (note 3)
RLS	Discard and send LOGO	ACC
RRQ	Discard and send LOGO	ACC

Notes:

1 all three identifiers: N_Port ID, N_Port Name, and Node Name must match a logged-in port for ACC to be returned.

2 BA_ACC or BA_RJT as specified in clause 8.12.2

3 if no PRLI, set the reason code to “image pair does not exist”

7.5.2 Responses to ULP-Level Frames

Table 12 summarizes the frame that shall be generated as a result of receiving different frames associated with a ULP-level Sequence when the recipient port is either currently not logged in with the sending port, or is logged in but has not received a Process Login from the sending port.

Table 31 – Responses to ULP-Level Frames from Logged Out Ports

Frame Received	No PLOGI	PLOGI but no PRLI
ABTS	Discard and send LOGO	Discard and send PRLO
ULP frame	Discard and send LOGO	Discard and send PRLO
RRQ	Discard and send LOGO	Discard and send PRLO

8 SCSI Features

This clause contains features for NL_Ports that implement the SCSI-FCP protocol and SCSI command sets. These features also apply to N_Ports that communicate with SCSI-FCP NL_Ports.

8.1 Auto Contingent Allegiance (ACA)

During the discovery process, FCP SCSI Initiators shall use INQUIRY to discover whether or not Targets support ACA behavior.

ACA shall be used (NACA=1 in CDB) only when both the sending SCSI Initiator and receiving Target support ACA functionality.

ACA shall not be used (NACA=0 in CDB) when either the sending SCSI Initiator or the receiving Target do not support ACA functionality. If a Target receives a command with NACA=0 and that command fails, autosense data and the contingent allegiance condition are cleared upon:

– Class 3: transmission of the FCP_RSP Sequence to the faulted SCSI Initiator.

– Class 2: receipt by the Target of the ACK sent in response to the FCP_RSP Sequence.

8.2 Command Linking

Command Linking is Prohibited by all devices conforming to this profile. The Link and Flag bits of the CDB shall be set to zeroes.

8.3 SCSI Disk Devices

8.3.1 Asynchronous Event Notification (AEN)

The use of Asynchronous Event Notification by disk devices is prohibited.

8.3.2 Disk Device Commands

Table 32 lists the SCSI Disk Device commands with usage defined by this report. Commands, and features within commands, which are not listed are optional. Interoperability between SCSI Initiators and Targets is not guaranteed if optional commands are used. If support of a field is listed as Required without specifying a value for that field, it is assumed that all possible values of the bit or field must be supported per the appropriate ANSI standard. Any unlisted commands/features/settings are implicitly Invokeable by the SCSI Initiator, and Allowed by the SCSI Target.

Table 32 – SCSI Disk Device Commands

Feature	Initiator	Target	Notes
FORMAT UNIT FmtData CmpLst Defect list format 0 0 N/A	 I	 R	
INQUIRY EVPD=1 Standard INQUIRY data (bytes 0-35)	I I I	R R R	
MODE SELECT(10) PF=1	I R	R R	
MODE SENSE(10) DBD=0	I I	R R	
READ(10)	I	R	
READ CAPACITY	I	R	
READ DEFECT DATA	I	R	

Table 32 – SCSI Disk Device Commands

Feature	Initiator	Target	Notes
RESERVE(10)/RESERVE(10) Extent=1 Extent=0 3rdPty=1	I P R P	R A R A	
REQUEST SENSE	I	R	
SEND DIAGNOSTIC SelfTest=1	I I	R R	
START/STOP UNIT	I	R	
TEST UNIT READY	I	R	
WRITE(10)	I	R	
WRITE BUFFER Mode=100b Mode=101b All other modes	I I I P	R R R A	

8.3.3 Mode Select/Sense (10) Parameters (Direct Access)

Table 33 lists Mode Select (10) and Mode Sense (10) command parameters with usage defined by this document. If a parameter is Required, a SCSI Target must be able to return a value for the parameter in a MODE SENSE(10) command. If the Required parameter is settable, a Target must be able to accept this parameter in a MODE SELECT(10) command. If a parameter is Allowed, a SCSI Target may be able to accept this parameter in a MODE SELECT(10) command, and may be able to return a value for the parameter in a MODE SENSE(10) command, but SCSI Initiators are prohibited from relying on SCSI Target support of these features. All unlisted features implicitly fall into this category. For example, in the Control Mode Page no value for RAC, (new for SCSI-3) is specified, which means both RAC=0 and RAC=1 are allowed.

Table 33 – Disk Mode Select/Sense Parameters

Parameter	Target	Notes
Mode Parameter Header (10) Medium Type per SBC WP=1 DPOFUA=1	R R A A	
Block Descriptors Number returned by target=1 Block length an integer multiple of 2 bytes	 R R	
Control Mode Page RLEC=0 Queue Algorithm Modifier=1 Qerr=0 Dque=0 RAC=? ByprtM=? BybthS=? EECA (from SCSI-2)=0 RAERP/UAAERP/EAERP=000 Ready AER holdoff period=0 Busy timeout period=?	R R R] R] R ? ? ? R R R ?	 7 7 7 7 7 7 7 7

Table 33 – Disk Mode Select/Sense Parameters

Parameter	Target	Notes
Disconnect/Reconnect Page	R	
Buffer Full/Empty Ratio	R	1
Maximum Burst Size	R	2
Connect Time Limit	A	3
Bus Inactivity Limit	A	4
Enable Modify Data Pointer (EMDP) bit=1	R	5

Notes:

1 Example: consider a target with ten 512-byte buffers and a buffer full ratio of 3Fh (63 decimal). The formula is:

$$\mathrm{INT}\left[\frac{\textit{buffer full ratio}}{256} \times \textit{number of buffers}\right] = \mathrm{INT}\left[\frac{63}{256} \times 10\right] = 2$$

Which means that on read operations, the target should attempt to arbitrate for the loop and transmit data whenever 2 or more buffers are full.

2 Regardless of whether or not FCP_XFER_RDY is used, the Maximum Burst Size corresponds to the maximum Sequence length (in bytes) of the SCSI Initiator or SCSI Target. The Maximum burst size must be a multiple of 512 bytes. More than one Sequence may be sent in a single loop tenancy. More than one loop tenancy may be used to transmit a single Sequence.

3 Connect Time Limit corresponds to the maximum time that the Target's NL_Port may keep the loop open. The Target's NL_Port shall send CLS no later than this time limit after transmitting each OPNy. If the Target has provided a non-zero Login_BB_Credit, the Target shall not send or re-transmit the CLS independent of the Connect Time Limit setting.

4 Bus Inactivity Limit indicates the maximum time that the Target's NL_Port may defer closing the loop without sending frames for FCP Information Units, or Link Control frames associated with those IUs. This is the maximum time from the end of one IU frame sent by the Target to the beginning of the next IU frame sent by the same NL_Port within the same Loop Tenancy. The Target shall close the loop immediately if it has exhausted its Available BB_Credit or it cannot send the next IU frame within this time limit. {what?}

5 EMDP=1 indicates that Sequences may be sent in any order. EMDP=0 indicates Sequences shall be sent in ascending offset order. This bit does not affect the order of frames within a Sequence. Targets which can return FCP_XFER_RDY DATA_RO parameters which are not continuously increasing within an Exchange implicitly support EMDP=1.

6 These features are still undecided in PLDA. We can either wait for them to figure it out, or make our own suggestions. The "R" values in some of these are suggestions of the PLDA editor.

8.4 SCSI Tape Devices

8.4.1 Asynchronous Event Notification (AEN)

The use of AEN by tape devices is (tbd).

8.4.2 Tape Device Commands

Table 34 lists the SCSI Disk Device commands with usage defined by this report. Commands and features within commands which are not listed are optional. Interoperability between SCSI Initiators and Targets is not guaranteed if optional commands are used. If support of a field is listed as Required without specifying a value for that field, it is assumed that all possible values of the bit or field must be supported per the appropriate ANSI standard. Any unlisted commands/features/settings are implicitly Invokeable by the SCSI Initiator, and Allowed by the SCSI Target.

Table 34 – SCSI Tape Device Commands

Feature	Initiator	Target	Doc	Notes
ERASE			SSC	
INITIALIZE ELEMENT STATUS	I	A	SMC	
INQUIRY	R	R	SPC	
LOAD UNLOAD			SSC	
LOCATE	I	A	SSC	
LOG SELECT	R	R	SPC	
LOG SENSE	R	R	SPC	
MODE SENSE/MODE SELECT	R	R	SPC	
Generic Pages	R	R	SPC	
Tape specific pages	R	R	SSC	
Media changer specific pages	I	A	SMC	
MOVE MEDIUM	I	A	SMC	
PREVENT/ALLOW MEDIUM REMOVAL	I	A	SSC	
READ	R	R	SPC	
READ BLOCK LIMITS	I	A	SSC	
READ BUFFER	I	A	SPC	
READ ELEMENT STATUS	I	A	SMC	
READ POSITION	I	A	SSC	
RECEIVE DIAGNOSTIC RESULTS	I	A	SPC	
REQUEST SENSE	R	R	SPC	1
RESERVE	I	A	SPC	
RELEASE	I	A	SPC	
REWIND	R	R	SSC	
SEND DIAGNOSTIC	I	A	SPC	
SPACE	R	R	SSC	
TEST UNIT READY	R	R	SPC	
VERIFY	I	A	SSC	
WRITE	R	R	SSC	
WRITE BUFFER	I	A	SPC	
WRITE FILEMARK	R	R	SSC	
1 FCP requires the use of Autosense				

8.5 SCSI Status

SCSI Targets shall return the following status values for the appropriate conditions. Status values not listed are optional (not required to be returned by Targets, not relied upon by SCSI Initiators).

- ACA ACTIVE (if ACA is supported by both the SCSI Initiator and Target)
- RESERVATION CONFLICT
- BUSY
- GOOD

– CHECK CONDITION

– TASK SET FULL

8.6 Target Discovery

Whenever a possibility of a topology change exists, a SCSI Initiator may want to rediscover the new topology. The Target Discovery procedure for a SCSI Initiator is:

{a procedure that involves going to the Simple Name Server and asking, and/or a procedure that involves poking at all loops in a manner similar to the one shown below for each Loop Base Address. Once a potential device is found, then proceed as indicated below with the ADISC step.}

{by the way, I'm tempted to redo the following as a list of steps to make it clearer (nested IFs can be a challenge to follow). I'm sure everyone will have their own (different) opinion...}

For all valid AL_PAs:

```
OPN(AL_PA)
  IF OPN successful, THEN
    ADISC(AL_PA) or PDISC(AL_PA)
    IF (LOGO returned or Node or Port Name changed) THEN
      PLOGI(AL_PA)
      IF PLOGI successful THEN
        IF (no hard address conflicts or application tolerant of hard address conflicts)
        PRLI(AL_PA)
        IF PRLI successful, THEN
        MODE SENSE(AL_PA,LUN_0)
        INQUIRY(AL_PA,LUN_0)
    etc...
        ENDIF
        ENDIF
      ENDIF
    ENDIF
  ENDIF
NEXT AL_PA
```

In order to determine whether or not an OPN was "successful" in the above procedure, an NL_Port must be able to either:

a) detect when an OPN has not been intercepted by the designated AL_PA

b) detect that R_RDYs have not been received from the AL_PA specified in an OPN within PL_TIME of sending that OPN.

Through the use of the Target Discovery procedure, the SCSI Initiator has the ability to assemble a database of [Hard Address:Node Name:Port Name:N_Port ID] tuples for all responding AL_PAs. There are several confirmations a SCSI Initiator can perform on that database to determine which Targets it can continue to communicate with which are not defined by this Technical Report. See 8.7 and 8.8 for rules regarding completion of open Exchanges.

Not all intiiators may perform the exact steps described in the above algorithm, although a SCSI Initiator is required to issue ADISC or PDISC to all Targets it is logged in with within RR_TOV of receiving LIP if it wants to remain logged in with those Targets.

The ADISC/PDISC procedure is designed to avoid the abnormal termination of all open Exchanges whenever a new device is inserted or removed from the loop, or whenever a device powers on or off.

Since devices are not required to respond to class 3 frames which have a D_ID which does not match the full 24-bit N_Port identifier of the receiving port, this may result in timeouts during the target discovery process. Therefore, for performance reasons ports shall originate PDISC Exchanges by transmitting the PDISC ELS and closing the loop without waiting for the ACC. SCSI Initiators shall have the capability of originating multiple concurrent Exchanges in order to hide multiple timeouts from the user.

{this following note from PLDA is fun...}

NOTE – Future public loop Technical Reports may wish to recommend a "two-pass" discovery process. The first pass would use the nonzero Domain+Area of the originating public loop in the D_ID for all ADISC/PDISC frames, and the second pass would attempt to discover all public loop ports which are sharing the loop by using zero for the Domain+Area. Timeouts on public loop ADISC/PDISC Exchanges would be handled in the same fashion as the public loop Technical Report - closing the loop without waiting for the ACC to the ADISC/PDISC and managing multiple concurrent ADISC/PDISC Exchanges.Practically speaking, public loop ports would not be able to use this two-pass process since they have no method of determing the nonzero Domain+Area of the attached FL_Port.

8.7 Exchange Authentication following LIP

If the login parameters or address identifier of any device (SCSI Initiator or Target) change, the device shall perform explicit logout with all logged-in ports before allowing the new parameters to take effect. If there are open and active Sequences with the NL_Ports to be logged out, the device shall respond to frames from those NL_Ports with LOGO.

8.7.1 SCSI Initiator Exchange Authentication

Following completion of loop initialization, SCSI Initiators shall authenticate the Targets with which they have open tasks before continuing those tasks.

a) A SCSI Initiator shall suspend execution of all open tasks with a Target following loop initialization until authentication is complete.

b) For each Target with which it has open tasks, a SCSI Initiator shall originate ADISC to that Target.

c) For every Target which returns an ACC, if the [N_Port ID:Port Name:Node Name] triplet for that Target:

   1) does not match the triplet of a logged-in Target, the SCSI Initiator shall transmit LOGO to that Target.

   2) does match the triplet of a logged-in Target, the SCSI Initiator may resume all tasks with that Target.

8.7.2 Target Exchange Authentication

Following completion of loop initialization, SCSI Targets are required to wait for exchange authenticate by the Initiator's with which they have open tasks before continuing those tasks.

a) For each Initiator with which it has open tasks, a SCSI Target shall suspend tasks associated with that Initiator until an ADISC is received from that Initiator. (see 8.7.3 for timeout information related to waiting for the ADISC.)

b) For every Initiator which sends an ADISC, if the [N_Port ID:Port Name:Node Name] triplet for that Initiator:

   1) does not match the triplet of a logged-in Initiator, the SCSI Target shall transmit LOGO to that Initiator.

   2) does match the triplet of a logged-in Initiator, the SCSI Target shall send the ACC to the Extended Link Service. The Target may then resume all tasks with that Initiator.

8.7.3 Target Resource Recovery

If a Target does not have an ACA condition in effect for a SCSI Initiator, and does not receive:

a) acknowledgement of Sequence Initiative transfer (that is, the next sequence in the logical operation following the Target's transmission of a FCP_XFER_RDY or FCP_DATA IU which transfers Sequence Initiative) within RR_TOV of transferring that initiative, or,

b) a ADISC from a logged-in SCSI Initiator within RR_TOV of receiving a LIP,

the Target shall implicitly logout the SCSI Initiator.

8.8 Post-LIP Frame Transmission and Reception

The following sections describe the behavior of SCSI Initiators and Targets following loop initialization.

{this needs rewriting in light of N_Ports and NL_Ports not being privy to LIPs on a Remote Loop.}

8.8.1 SCSI Target Behavior

Following loop initialization, any frames received (other than ADISC [or PDISC?]) from a SCSI Initiator which has not been authenticated shall be discarded. Those frames shall not be considered part of any valid Exchange, and only BB_Credit shall be balanced in response to those frames. SCSI Initiator authentication is complete when the ACC to an ADISC has been transmitted.

Following transmission or forwarding of an initialization LIP, only frames related to SCSI Initiator authentication (i.e., ACC to an ADISC or PDISC) shall be transmitted to that SCSI Initiator before frames for other Exchanges to that SCSI Initiator are transmitted.

8.8.2 SCSI Initiator Behavior

Following loop initialization, any frames received (other than ACC to an ADISC or PDISC) from a Target which has not been authenticated shall be discarded. Those frames shall not be considered part of any valid Exchange, and only BB_Credit shall be balanced in response to those frames. Target authentication is complete when the ACC to an ADISC has been received.

Following transmission or forwarding of an initialization LIP, only frames related to target authentication (i.e., ADISC or PDISC) shall be be transmitted to that Target before frames for other Exchanges to that Target are transmitted. SCSI Initiator timeout waiting for ACC to an ADISC or PDISC shall result in implicit logout of the Target.

8.9 Exchange Origination Capability

A Target shall have Exchange origination capability to originate a LOGO or PRLO request. All other Exchanges are originated by the SCSI Initiator.

8.10 Clearing Effects of ULP, FCP, FC-PH, and FC-AL Actions

Table 35 summarizes which FCP Target objects are cleared as a result in the listed FCP SCSI Initiator actions. A ‘Y’ in the corresponding column indicates the object is cleared to its default or power-on value within the device upon completion of specified action.

Table 35 – Clearing Effects of SCSI Initiator ULP, FCP, FC-PH, and FC-AL Actions

	FCP SCSI Initiator Action								
FCP Target Object	**Power Cycle**	**Reset LIP(y,x)**[3]	**LOGO, PLOGI**	**ABTS w/L_S**	**PRLI, PRLO**	**TPRLO**	**Target Reset**	**Clear Task Set**[4]	**Abort Task Set**[4]
PLOGI parms									
For all logged-in SCSI Initiator ports	Y	Y	N	N	N	N	N	N	N
Only for port initiating action	N	N	N	N	N	N	N	N	N
Open Sequences									
For all SCSI Initiator ports with open Sequences	Y	Y	N	N	N	Y	Y[2]	N	N
Only for port initiating action	N	N	Y	Y	Y	N	N	Y[2]	Y[2]
BB_Credit_CNT	Y	Y	Y	N	N	N	N	N	N
Hard Address Restored	Y[1]	Y[1]	N	N	N	N	N	N	N
PRLI parms cleared									
For all logged-in SCSI Initiator ports	Y	Y	N	N	N	N	N	N	N
Only for ports of specified TYPE	N	N	N	N	Y	Y	N	N	N
Only for port initiating action	N	N	Y	N	Y	N	N	N	N
Open Tasks (Exchanges)									
All tasks for all SCSI Initiator ports with open tasks	Y	Y	N	N	N	Y	Y	Y	N
All tasks, only for port initiating action	N	N	Y	N	Y	N	N	N	Y
Only specified task for port initiating action	N	N	N	Y	N	N	N	N	N
Target mode page parms = saved									
For all SCSI Initiators	Y	Y	N	N	N	Y	Y	N	N
Only for port initiating action	N	N	Y	N	Y	N	N	N	N
Pre-existing UA condition									
For all SCSI Initiators	Y	Y	N	N	N	Y	Y	N	N
Only for port initiating action	N	N	Y	N	Y	N	N	N	N
Pre-existing ACA condition									
For all SCSI Initiators	Y	Y	N	N	N	Y	Y	N	N
Only for port initiating action	N	N	Y	N	Y	N	N	N	N

Notes:

1 If the NL_Port has a hard address and the port experiences a power-cycle occurs or receives LIP(AL_PD,AL_PS), the hard address shall become the preferred address of the NL_Port, even if the previously acquired address is still available. {this needs repair for PLP}

2 Tasks are cleared internally within the Target, but open FC-PH Sequences must be individually aborted by the SCSI Initiator via ABTS-LS which also has the effect of aborting the associated Exchange.

3 This is also known as LIP(AL_PD,AL_PS). If the destination of this selective hard reset LIP matches the address of the receiving device, the receiving device shall observe the behavior described in this column. {this needs repair for PLP}

4 For multi-LUN Targets, Clear Task Set, and Abort Task set affect only the addressed LUN, not the entire Target.

8.10.1 Multiple Ports

In a dual-port implementation of a single-LUN disk Target with common BB_Credit management hardware shared between the two ports, the above objects are cleared regardless of which port the action occurs on. Tasks can only complete normally (i.e., status returned in FCP_RSP and the Exchange closed) on the same port that they were issued on.

8.11 SCSI-FCP Feature Set

The following Feature Sets are described in the Fibre Channel Protocol for SCSI (FCP).

8.11.1 Process Login Parameters

Table 36 lists Process Login paramaters with usage defined by this document.

Table 36 – PRLI Parameters

Feature	SCSI Initiator	SCSI Target	Notes
Originator Process Associator Valid	P	P	
Responder Process Associator Valid	P	P	
Originator Process Associator	P	P	
Responder Process Associator	P	P	
Establish Image Pair (bit 13)	R	R	2
Initiator Function	R	A	1
Target Function	A	R	1
Data Overlay Allowed	P	P	
Data + Response in same Sequence (Read)	P	P	
Command + Data in same Sequence (Write)	P	P	
Read XFER_RDY Disabled	R	R	
Write XFER_RDY Disabled	P	P	

Notes:

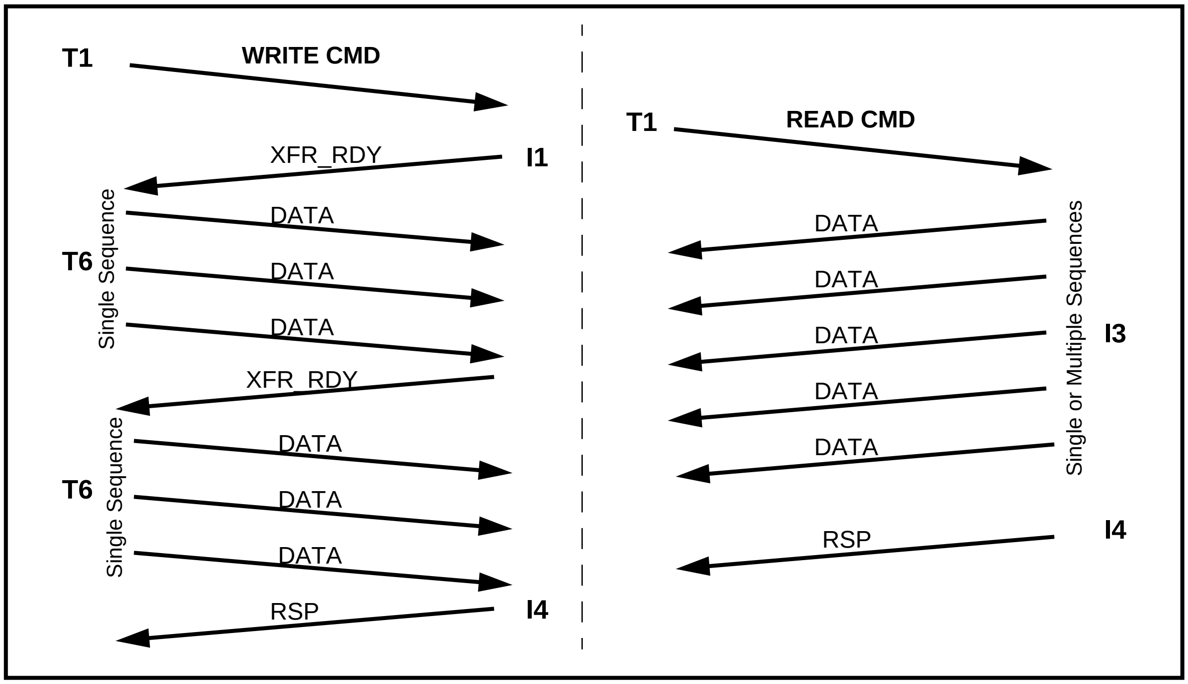
1 Allowed for XOR functionality

2 This bit must be set in the PRLI payload (requesting the image pair be established) and must be set to a one in the PRLI ACC payload when the image pair has been successfully established.

8.11.2 FCP Information Units

All T1 (FCP_CMND), I1 (FCP_XFER_RDY), and I4 (FCP_RSP) Sequences shall be single-frame Sequences. T6 (FCP_DATA for write data etc.) and I3 (FCP_DATA for read data etc.) Sequences may be multiple-frame Sequences. Figure 2 shows examples of FCP IU usage.

Figure 2 – FCP Read/Write IU Examples



8.11.2.1 SCSI Initiator to Target IUs

T1 (CMD/Task Management with SI transferred), and T6 (Write Data with SI transferred) are required. All others are prohibited.

8.11.2.1.1 FCP_CMND

FCP_DL in the FCP_CMND payload shall always be equal to the number of bytes expected to be transferred for the command. Transfer Length in the FCP_CDB may be expressed in in either bytes or logical blocks as appropriate for the specific device type and command. For device types which specify the Transfer Length in the FCP_CDB in logical blocks, the FCP_DL shall be equal to (Transfer_Length x Block_Size). Targets are not required to check for this consistency.

On read commands, the SCSI Initiator is responsible for ensuring the amount of data returned is equal to the amount specified by FCP_DL, even if the Status returned is GOOD. If the amount does not match FCP_DL, the appropriate ULP recovery action shall be invoked. Since there are no transfers of Sequence Initiative during read operations once the Target is transmitting data, status may be returned as GOOD when operating in Class 3 even though some of the data was not received by the SCSI initiator. This may occur as the result of lost or corrupted frames in the read data.

8.11.2.1.2 Write Data

The Parameter field of the first frame of a Write data Sequence shall be set to the Relative Offset specified by the corresponding FCP_XFER_RDY (FCP IU I1). On subsequent frames of the Sequence, continuously increasing Relative Offset shall be used.

8.11.2.2 Target to Initiator IUs

I1 (FCP_XFER_RDY on Write), I3 (Read Data), I4 (FCP_RSP) are required. All others are prohibited. Targets shall respond to Task Management functions using I4.

8.11.2.2.1 FCP_XFER_RDY

On transfers consisting of multiple Write Data Sequences, the DATA_RO parameter contained in consecutive FCP_XFER_RDY Sequences does not have to be continuously increasing. This provides the logical equivalent of SCSI Modify Data Pointers. This is not the same as the Random Relative Offset PLOGI parameter, which refers to Relative Offset on consecutive frames within a Sequence. Within an Exchange, only the FCP_DATA frame containing the largest Relative Offset may contain fill bytes.

The DATA_RO parameter in the FCP_XFER_RDY payload on Write operations shall be a multiple of 8 bytes. The initial Relative Offset of each Read Data Sequence shall also be a multiple of 8 bytes.

The BURST_LEN parameter in the FCP_XFER_RDY IU shall be less than or equal to the amount of data remaining to be transferred for the write task.

8.11.2.2.2 Read Data

Within Read Data Sequences, the Relative Offset must be continuously increasing. Between Read Data Sequences in a Read Exchange with multiple data Sequences, the starting Relative Offsets (the Relative Offsets in the first frames of each Sequence) need not be continuously increasing.

8.11.2.2.3 FCP_RSP

According to FC-PH, the first two bits [31:30] of the first word of a command status frame payload fall into one of the following categories: 00 (successful and complete), 01 (successful but incomplete), 10 (unsuccessful but complete), 11 (unsuccessful and incomplete).

Since the first word of FCP_RSP frames are reserved in FCP, these bits are set to zero, regardless of the content of the SCSI Status portion of the payload (which may indicate the command was unsuccessful and/or incomplete). SCSI Initiators shall not rely on bits 31:30 of word 0 in FCP_RSP to determine success or completion status of a command.

8.11.3 Task Management Flags and Information Units

All SCSI Initiators shall send Task Management functions using T1. All Targets shall return FCP_RSP to Task Management functions using I4. The RSP_CODE in the FCP_RSP_INFO field shall indicate the result of the Task Management function. The SCSI Status byte and FCP_SNS_INFO shall be ignored for I4 information units sent in response to a Task Management function.

The result of the Task Management function is indicated in the RSP_CODE of the FCP_RSP_INFO field and shall be as shown in table 37.

Table 37 – Task Management Function RSP_CODES

RSP_CODE	SAM Mapping	Meaning
00	Function Complete	Task management function was performed
04	Function Rejected	Task management function was not performed because the function is unsupported
05	Service Delivery or Target Failure	Task management function was not performed, but the function is supported.

Table 38 lists FCP Task Management functions with usage defined by this document.

Table 38 – FCP Task Management Flags

Feature	SCSI Initiator	SCSI Target	Notes
Terminate Task	P	P	
Clear ACA	I	R	1

Table 38 – FCP Task Management Flags

Feature	SCSI Initiator	SCSI Target	Notes
Target Reset	I	R	2
Clear Task Set	I	R	2
Abort Task Set	I	R	2

Notes:

1 Clear ACA is only required by the SCSI Target if it advertises support of ACA via the NormACA bit in the INQUIRY data. Clear ACA is required by SCSI Initiators which use ACA (NACA bit in CDB=1).

2 See 8.12.2 for SCSI Initiator requirements following transmission of these Task Attributes.

8.11.4 FCP Task Attributes

Table 39 lists FCP Task Attributes with usage defined by this document.

Table 39 – FCP Task Attributes

Feature	SCSI Initiator	SCSI Target	Notes
Untagged	P	P	1
Simple Queue Type	R	R	
Ordered Queue Type	I	R	2
Head of Queue Type	I	R	3
Auto Contingent Allegiance Type	R	R	4

Notes:

1 All FCP operations are implicitly tagged with an Exchange ID. Therefore, SCSI Initiators using untagged queueing rules shall use the Simple Queue task attribute, and not issue overlapped commands.

2 Ordered Queue Types are allowed to be invoked by SCSI Initiators, since delivery order is preserved.

3 Head of Queue Type is useful over public loops which can ensure delivery order, but may be less useful across fabrics/topologies which cannot. {for notes 2 & 3, do we have enough delivery ordering to make this usable?}

4 ACA Type is only required by the SCSI Target if it advertises support of ACA via the NormACA bit in the INQUIRY data. ACA Type is required by SCSI Initiators which use ACA (NACA bit in CDB=1).

8.11.5 FCP_RSP Payload

Table 40 lists FCP_RSP fields with usage defined by this document.

Table 40 – FCP_RSP Payload

Feature	SCSI Initiator	SCSI Target	Notes
FCP_SNS_LEN	-	<=96	1
FCP_SNS_INFO	R	I	

Table 40 – FCP_RSP Payload

Feature	SCSI Initiator	SCSI Target	Notes
Length of Additional Sense Bytes in FCP_SNS_INFO	N/A	<=78	1
FCP_RSP_LEN	0 or 8	0 or 8	
FCP_RSP_INFO	R	I	2

Notes:

1 SCSI single-LUN direct access Targets have this restriction to keep FCP_RSP payload to less than or equal to 128 bytes. SCSI Initiators have no such restrictions. Other device types (e.g., controller devices) may return sense information (FCP_SNS_INFO) up to the SCSI limit of 255 bytes, and SCSI Initiators which support multiple device types may have to support the full value afforded by the SCSI-3 Primary Command set. The FCP_RSP Payload therefore consists of a maximum of 96 bytes of Sense Data, and 32 bytes of other information.

2 If present, this field shall contain the FCP_RSP code in an 8 byte field in the format specified by FCP.

8.11.6 Other FCP Features

Table 41 lists Other FCP features with usage defined by this document.

Table 41 – Other FCP Features

Feature	SCSI Initiator	SCSI Target	Notes
FCP_LUN	R	R	2
FCP_LUN (0)	I	R	2
INQUIRY of FCP_LUN (0)	I	R	2
Process Login	R	R	
Process Logout	P	P	1

Notes:

1 The single exception to this rule is that PRLO may be issued following receipt of a frame from a device which has not performed PRLI as a notification. A PRLI has the effect of implicit PRLO.

2 Some direct access Targets will require multiple LUNs (e.g., disks over 8Gbyte with 512 byte block size on some OS’s necessitates more than 1 LUN). Multiple-LUN Targets shall address the LUNs using the “first level addressing” field of SCC (reference [9]) with “peripheral device addressing method” (00b) reported per Table 5 in SCC.

8.12 Error Detection and Recovery

This clause describes the error actions to be taken by the SCSI Initiator and Target upon detection of an error condition.

8.12.1 Error Detection

The ABTS protocol (described in 8.12.2) shall be invoked by a SCSI Initiator when the ULP timer has expired and FCP_RSP has not been received.

Targets are not required to time Sequences as Sequence Recipients or as Sequence Initiators. Targets are not required to either assign RX_IDs or detect overlapped commands (duplicate OX_ID/RX_ID combinations). SCSI Initiators shall guarantee X_ID uniqueness by only reusing OX_IDs to a Target N_Port or NL_Port after the Exchange is closed.

8.12.2 Error Recovery Using ABTS Protocol

The ABTS-L_S protocol is required, which uses ABTS to abort entire Exchanges. The unit of error recovery for this Technical Report is an Exchange, not a Sequence. The protocol defining retransmission of Exchanges following ABTS is beyond the scope of this document.

Only a SCSI Initiator may initiate FCP Recovery Abort. This Technical Report does not define the protocol by which multiple SCSI Initiators communicate or synchronize shared peripherals.

The ABTS protocol shall be invoked on each Open Exchange:

a) following receipt of “Function Complete” to Abort Task Set

b) following receipt of “Function Complete” to Clear Task Set

c) following receipt of “Function Complete” to Target Reset

d) upon receipt of CHECK CONDITION status with Sense Key=Unit Attention and ASC/ASCQ of either:

   1) “Command cleared by another Initiator” (2F 00), or

   2) “Power on, reset, or bus device reset occurred” (29 00)

e) upon ULP_TOV expiration; i.e., no STATUS returned for a Command within an application-specified amount of time (see clause 10.7 on page 67)

8.12.2.1 Abort Sequence (ABTS) Frame

Only the SCSI Initiator shall transmit an ABTS frame. When it does so, the specified fields shall be set as follows:.

Table 42 – ABTS Contents

	Field	Sub-field	Content
Frame Header	F_CTL	Sequence Context	Initiator (even though the ABTS initiator may not have Sequence Initiative for the Sequence being aborted).
	SEQ_ID		SEQ_ID of the open Sequence in the Exchange being aborted. If no Sequence is open for this Exchange, the SEQ_ID is any SEQ_ID not currently open (for any Exchange) between the SCSI Initiator and Target.
	SEQ_CNT		SEQ_CNT of last frame transmitted in an Open Sequence + 1. If no Sequence is open, then SEQ_CNT=x’0000’.
	OX_ID		OX_ID=same as that assigned by SCSI Initiator (Exchange Originator) for the command being aborted.
	RX_ID		Set to x’FFFF’ or the same as that assigned by Target for the task being aborted.

8.12.2.2 Basic Accept (BA_ACC) Frame

Only a SCSI Target is required to accept ABTS with BA_ACC.

Table 43 – BA_ACC Contents

	Field	Sub-field	Content
Frame Header	OX_ID		OX_ID from ABTS frame
	RX_ID		RX_ID from ABTS frame
	F_CTL	L_S bit	1
		Sequence Context	Recipient

Table 43 – BA_ACC Contents

	Field	Sub-field	Content
Payload	SEQ_ID validity		x'00'
	SEQ_ID byte		invalid (don't-care)
	OX_ID aborted		OX_ID from ABTS frame
	RX_ID aborted		RX_ID from ABTS frame
	Lowest SEQ_CNT		x'0000'
	Highest SEQ_CNT		x'FFFF'

8.12.2.3 Basic Reject (BA_RJT) Frame

BA_RJT is transmitted by a Target in response to ABTS only if:

a) The Target has assigned an RX_ID to an OX_ID in a previous frame, and

b) The SCSI Initiator has issued ABTS which contains an unknown OX_ID/RX_ID combination. An RX_ID of x'FFFF' is not considered "unknown" since the SCSI Initiator may issue ABTS without knowing the RX_ID assigned by the Target, due to lost frames.

The reason code shall be "Logical Error" with a reason code explanation of "Invalid OX_ID/RX_ID combination".

8.12.2.4 Reinstate Recovery Qualifier (RRQ) Frame

Following reciept of the ACC in response to an ABTS, the SCSI Initiator shall transmit RRQ. The RRQ shall be transmitted after R_A_TOV expires (see clause 10.6). The format of the RRQ is,

Table 44 – RRQ Request Contents

	Field	Content
Frame Header	OX_ID	Identifier of a new exchange
	RX_ID	x'FFFF'
Payload	Originator S_ID	Source_ID of the SCSI Initiator
	OX_ID aborted	OX_ID of XCHG that was previously aborted with ABTS
	RX_ID aborted	RX_ID of XCHG that was previously aborted with ABTS

Following successful completion of the RRQ, the Target shall respond with ACC.

8.12.3 SCSI Initiator Behavior

ABTS may be transmitted even if Sequence Initiative is not held. Following the transmission of ABTS, any Device_Data frames received on this Exchange shall be discarded until the BA_ACC with L_S=1 bit is received from the Target. If a proper BA_ACC or BA_RJT is not received from the Target within E_D_TOV, second level error recovery shall be initiated.

ABTS may not take effect immediately. For example, if ABTS is sent following transmission of a Read command, the Target may send back all or some of the requested read data before replying with the ACC to the ABTS (or the data may already be in flight at the time the ABTS was sent). The SCSI Initiator shall be able to receive this data and balance BB_Credit in order for the Target to send the ACC.

There are circumstances where a configuration change event could occur very quickly, which results in link errors and a ULP timeout, but no LIP is generated to reconfigure the loop. A SCSI Initiator shall attempt the ABTS protocol following ULP_TOV timeout. If the ABTS protocol fails, second level error recovery as described in clause 8.12.6. shall be performed.

8.12.4 SCSI Target Behavior

When ABTS is received at the Target, it shall return either BA_ACC or BA_RJT.

The Target shall qualify ABTS based on the OX_ID and RX_ID. If RX_ID has not been assigned, and an unknown OX_ID is detected in the ABTS frame, then BA_RJT shall be transmitted. If RX_ID has been assigned, and an unknown OX_ID or a known OX_ID with an RX_ID which is both unknown and not equal to x'FFFF' is detected in the ABTS frame, then BA_RJT shall be transmitted. (The RX_ID is not guaranteed to be known by a SCSI Initiator if FCP_CMND, XFR_RDY, or the first Read Data frame of an Exchange is lost.) Otherwise, all Exchange resources shall be reclaimed and BA_ACC shall be transmitted with L_S=1 set in F_CTL.

If an error is detected by a Target while it has Sequence Initiative, the only permissible recovery action is the transmission of FCP_RSP with CHECK CONDITION status and an appropriate Sense Key/ASC/ASCQ.

If an error is detected by a Target while it does not have Sequence Initiative, it must wait until it has been given Sequence Initiative before it can return FCP_RSP.

If ULP resources in the Target are unavailable when a command is issued, the Target shall return TASK SET FULL status in response to the command. A ULP resource at the Target is a buffer associated with storing the SCSI Command prior to execution, or any buffer required to maintain Command context during execution.

If a Target receives a PLOGI request which it cannot accept due to a limitation on login resources,

a) if there are no open tasks with a logged-in SCSI Initiator, that SCSI Initiator may be implicitly logged out so that the PLOGI request can be processed (no algorithm is specified for how a Target selects the SCSI Initiator to be implicitly logged out)

b) if all logged-in SCSI Initiators have open tasks, the Target may respond to the PLOGI with LS_RJT and reason code “unable to perform command request” and reason code explanation “insufficient resources to support Login”.

8.12.5 SEQ_ID Reuse

For single-frame Sequences in Class 3, a device may reuse a SEQ_ID immediately following transmission of the single-frame Sequence. For single-frame Sequences in Class 2, the SEQ_ID a device may reuse a SEQ_ID immediately following receipt of the ACK for the Sequence, or after an R_A_TOV timeout.

For multi-frame Sequences in Class 3,

a) A device may reuse a SEQ_ID for which it has Sequence Initiative following confirmation of transfer of Sequence Initiative. For example, the SEQ_ID for the FCP_DATA portion of a write command may be reused by an SCSI Initiator following receipt of the FCP_RSP for that command.

b) On Read operations, the maximum number of data Sequences per Exchange is 256. SCSI Initiators shall size their Read operations appropriately so that Targets, using minimum-sized frames, can return all read data within 255 Sequences. A Target may not reuse a SEQ_ID for a multi-frame Sequence within a Read Exchange.

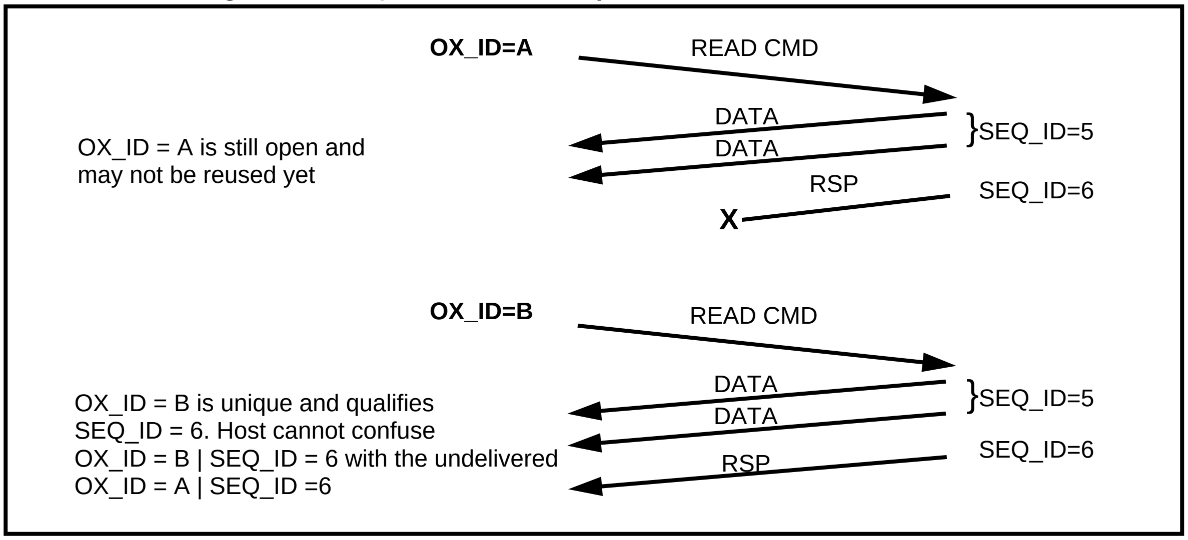
c) A Target may reuse a SEQ_ID for a Read Data Sequence immediately following transmission of FCP_RSP.

Behavior (c) is justifed through the following argument (see figure 3):

– All Sequences in FCP are qualified by an X_ID (per FCP) and X_ IDs must always be checked before a Sequence within the Exchange is delivered to the ULP.

- All open X_IDs between N_Port/NL_Port pairs must be unique (per FC-PH).
- Targets are prohibited from retransmitting just FCP_RSP (per Exchange-based error recovery in this Technical Report and FCSI profiles).
- Therefore, SEQ_IDs are not required to be unique across all Exchanges, and the previous Exchange is guaranteed to never complete.

Figure 3 – SEQ_ID Reuse Example for Class 3, In-order Fabrics:



8.12.6 Second Level Error Recovery

If the ABTS protocol fails, the SCSI Initiator shall:

a) Retry the ABTS protocol once.

b) If the retried ABTS protocol fails, the SCSI Initiator shall implicitly logout (FC-PH LOGO) the Target. All outstanding Exchanges with that Target are implicitly terminated at the SCSI Initiator.

8.12.7 Task Management and Multiple-Initiator Targets

If a Target Reset or Clear Task Set management function is received by a Target which has multiple SCSI Initiators logged in with it, then:

a) Task Management functions are sent using the I1 Information Unit (FCP_CMND with Sequence Initiative transferred).

b) The Target shall create a Unit Attention Condition for all other SCSI Initiators, regardless of whether or not the Target believes a command is oustanding to those SCSI Initiators. This is because FCP_RSP may have been transmitted but not received by the SCSI Initiator, or the SCSI Initiator may have transmitted a command which has not yet been received by the Target.

c) The Target may clear all resources associated with the cleared Exchanges, per SAM.

d) The Target shall return FCP_RSP upon completion of (a) and (b). The payload shall be zeroes with the exception of the FCP_RSP_LEN_VALID bit, FCP_RSP_LEN (which shall be set equal to 8), and the FCP_RSP_INFO. If the SCSI Initiator does not receive FCP_RSP, it may retry the Task Management. If the retry fails, it shall initiate second level error recovery.

e) Upon discovery of a power-on Unit Attention Condition, SCSI Initiators shall issue ABTS for all commands which are outstanding to that Target. From a SCSI Initiator perspective, this refers to all commands for which FCP_RSP has not been received.

f) If the Target receives an ABTS frame for which the referenced S_ID | OX_ID | RX_ID does not exist, it shall respond with BA_ACC with the L_S=1.

9 IP Features

9.1 Something...

....

10 Timers on Public Loop

FC-PH requirements for timers are summarized in {?}. This clause defines the times and timeout values used for public loop implementations. Since a loop is neither a switched fabric not a point to point topology, most FC-PH timer rules do not apply.

Table 45 – Timer Summary

Timer	Description	Value
PL_TIME	Maximum physical delay of Local Loop, including elasticity buffers in each L_Port	1 ms
LIS_HOLD_TIME	Maximum time for each L_Port to forward a Loop Initialization Sequence	1 ms
LM_TOV	Minimum time Loop Master waits for a Loop Initialization Sequence to return	128 ms
AW_TOV	Minimum time to wait for access to loop before LIP allowed	1 sec
E_D_TOV	Minimum time to wait for Sequence completion before initiating recovery (more for fabric...)	2 sec
R_A_TOV	Minimum time to wait before allowing Target to reinstate recovery qualifier	10 sec
ULP_TOV	Minimum time to wait for ULP-specific result before ULP initiates recovery	>E_D_TOV
RR_TOV	Minimum time NL_Port waits for ADISC/PDISC after LIP or Sequence completion before implicitly logging out another N_Port or NL_Port	2 sec

10.1 Physical Loop Delay (PL_TIME)

AL_TIME is defined by FC-AL to be 15ms, which represents approximately twice the worst case round trip latency for a loop with 134 nodes, each with a latency of 6 transmission words, and 10km optical segments between each node. Since this time is dominated by the physical propagation latency, it is the same regardless of whether the loop is operating at 133Mb or 1062Mb. Since a 1,340km loop (838 miles) loop is deemed unrealistic, a Local Loop shall be limited to a total cable length of 67km (42 miles). Therefore, PL_TIME can be made equal to 1ms.

10.2 Loop Initialization Sequence Hold Time (LIS_HOLD_TIME)

LIS_HOLD_TIME is the maximum amount of time between when an L_Port receives a Loop Initialization Sequence (LIFA, LIPA, LIHA, LISA, LIRP, LILP), and when the L_Port shall forward it to the next L_Port.

10.3 Loop Master Timeout Value (LM_TOV)

LM_TOV is the minimum time a Loop Master shall wait for a Loop Initialization frame to be returned before it may assume the loop is not functioning properly. It is equal to PL_TIME + (LIS_HOLD_TIME x #participating nodes). For a Local Loop, LM_TOV = 1ms + (1ms x 127) = 128ms.

10.4 Abitration Wait Timeout Value (AW_TOV)

AW_TOV is equal to the minimum period of time an NL_Port shall wait to win arbitration before it may assume a malfunction or excessive unfairness and transmit an initialization LIP. This is based on 125 ports all transmitting 64k bytes per tenancy, followed by each of the 125 ports also gaining unfair access to the loop and sending another 64k bytes. PL_TIME includes the time it takes to transmit (64k x 250) bytes, including 250 PL_TIMEs.

10.5 Sequence Timeout (E_D_TOV)

E_D_TOV is the minimum time an NL_Port shall wait between the time it transmits a frame requesting transfer of Sequence Initiative and the time it receives a frame within the same Exchange which acknowledges the transfer of Sequence Initiative before invoking error recovery. It includes at least one AW_TIME, two PL_TIMEs, and the time it takes within the Sequence Recipient for it to formulate a response and any internal queueing delays . The value is rounded up to 2 seconds.

SCSI Initiators may choose not to define this specifically, but group it into a ULP timer. E_D_TOV may also be used as a SCSI Initiator Resource Recovery timer to recover internal resources associated with a task to an unresponsive Target. For example, if an information unit is placed into an outbound hardware queue in the SCSI Initiator port and still exists there E_D_TOV later with other tasks unable to be serviced, the SCSI Initiator may immediately reclaim the outbound queue resources upon expiration of E_D_TOV, and initiate ABTS error recovery after the resources have been reclaimed.

10.6 Resource Allocation Timeout (R_A_TOV)

{text TBD}

Since ?, R_A_TOV=10 seconds for the purposes of reinstating Recovery Qualifiers....

NOTE – RRQ timing is only one of many uses that FC-PH has for R_A_TOV (see {?}). One disadvantage of this definition is that for at least two topologies (class 3 in-order fabrics and public loops) which cannot misorder or busy/retry frames, Recovery Qualifiers can be reused as soon the ABTS Initiator has received the ACC. In order to allow such topologies to optimize their recovery operations, X3T11 should review FC-PH (with the help of {?}) to determine which clauses refer to only fabrics which can misorder frames, and which clauses refer to all fabrics.

10.7 ULP Timer (ULP_TOV)

ULP_TOV is the minimum time that a ULP process waits for some ulp-specific result (like SCSI status or ...?) before initiating ULP recovery....?

10.8 Resource Recovery Timer (RR_TOV)

RR_TOV is the minimum time an NL_Port shall wait for a specific N_Port or NL_Port action following LIP or transfer of Sequence Initiative before the reclaiming resources of a logged-in N_Port or NL_Port.

Annex A
(normative)

New Extended Link Services for Public Loop Operation

New Extended Link Services are defined for the support of Arbitrated Loops, that form part of a Fabric:

- Loop Change Notification (LCN): Informs subscribers of Public Loop configuration changes;
- Loop Change Subscribe (LCS): Enables an entity to be informed when:
  - new Public Loops are attached to the Fabric;
  - Public Loops vanish from the Fabric;
  - changes occur on attached Public Loops.
- Loop Change Unsubscribe (LCU): Stops sending of Loop Change Notifications to the requestor.
- Loop Initialize (LINIT): Enables a requestor to initiate Loop Initialization.
- Loop Port Control (LPC): Enables a requestor to bypass or enable Loop Devices on a given Loop.
- Loop Status (LSTS): Enables a requestor to ascertain the state of a given Loop and its Loop Devices.

A.1 Loop Initialize (LINIT)

An entity may wish to initiate Loop Initialization on a given Loop. The Loop Initialize Extended Link Service request shall be used to achieve this.

Acceptance of this request shall not indicate that the requested LIP has been delivered.

Protocol:

Loop Initialize request Sequence
Reply Sequence

Format: FT–1

Addressing: The S_ID designates the entity requesting Loop Initialization of the identified Loop. The D_ID field designates the destination for the LINIT; this shall be the Loop Base Address of the Loop to initialize.

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

Table 1 – LINIT Payload

Item	Size Bytes
hex ‘xx 00 00 00’	4
Reserved	2
LIP Byte 3	1
LIP Byte 4	1

LIP Byte 3: The 3rd character of the LIP Primitive Sequence to be employed.

LIP Byte 4: The 4th character of the LIP Primitive Sequence to be employed.

Reply Extended Link Service Sequence:

Service Reject (LS_RJT)
Signifies the rejection of the LINIT command
Accept (ACC)
Signifies acceptance of the LINIT request and initiation of the Loop Initialization process.

– Accept Payload
The format of the Accept Payload is shown in table 2.

Table 2 – LINIT Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4

A.2 Loop Port Control (LPC)

An entity may wish to bypass or enable Loop Devices in a given Loop. The Loop Port Control Extended Link Service request shall be used to achieve this.

Acceptance of this request shall not indicate that the requested Bypass and/or Enable has been acheived.

If the FL_Port addressed by the Loop Base Address is unable to successfully deliver a requested LPB or LPE within two PL_TIME delays, the FL_Port shall begin loop initialization.

Protocol:

Loop Port Control request Sequence
Reply Sequence

Format: FT–1

Addressing: The S_ID designates the entity requesting bypassing or enabling of Loop Devices on a specific Loop. The D_ID field designates the destination for the LPC; this shall be the Loop Base Address of the loop to initialize.

Payload: The format of the LPC request Payload is shown in table 3.

Table 3 – LPC Payload

Item	Size Bytes
hex ‘xx 00 00 00’	4
Loop Devices to Bypass (LPB) (AL_PA bit map)	16
Loop Devices to Enable (LPE) (AL_PA bit map)	16

Loop Devices to Bypass: The format of this field follows the AL_PA bit mapped format defined in FC–AL (see FC–AL table 15). If a bit is set = 1 in this field, then attempts at bypassing the Loop Device, sending the LPByx Primitive Sequence at the identified location shall be made.

Loop Devices to Enable: The format of this field follows the AL_PA bit mapped format defined in FC–AL (see FC–AL table 15). If a bit is set = 1 in this field, then attempts at enabling the Loop Device, sending the LPEyx Primitive Sequence at the identified location shall be made. If Word 0 bit 31 is set = 1 then the LPEfx Primitive Sequence shall be sent on the specified Loop.

Reply Extended Link Service Sequence:

Service Reject (LS_RJT)
Signifies the rejection of the LPC command
Accept (ACC)
Signifies acceptance of the LPC request.
– Accept Payload
The format of the Accept Payload is shown in table 4.

Table 4 – LPC Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4

A.3 Loop Change Notification (LCN)

A Loop Change Notification Extended Link Service request shall be sent to subscribing destinations when an event occurs on an Arbitrated Loop, which may have affected current PLOGI states, or makes new Arbitrated Loop resources available.

The Fabric Controller may delay the sending of LCNs for newly added Loop Devices, to allow sufficient time to determine whether the new Loop Devices are Public or Private.

The LCN shall be sent in the Class of Service initially used to subscribe to the LCN service.

Protocol:

Loop Change Notification request Sequence
No reply Sequence
Format: FT–1

Addressing: The S_ID designates the N_Port reporting the LCN, which shall be the Fabric Controller (hex ‘FF FF FD‘) {why not the Loop Base Addr?}. The D_ID field designates the destination for the LCN.

Payload: The format of the LCN Payload is shown in table 5.

Table 5 – LCN Payload

Item	Size Bytes
hex ‘xx 00 00 00’	4
Reserved	1
Loop Base address	3
Notification Count	4
New Loop Devices (AL_PA bit map)	16
Retired Loop Devices (AL_PA bit map)	16
Current Public Loop Devices (AL_PA bit map)	16
Current Private Loop Devices (AL_PA bit map)	16
Current Bypassed Loop Devices (AL_PA bit map)	16
Initial AL_PA position list (AL_PA position list from LILP)	128

Loop Base Address: As defined in Clause 3.

Notification Count: The Notification Count is incremented every time a new LCN is generated for a specific Loop. It can be used by the recipient to check that it has received all LCNs. The Notification Count of the first LCN for a given Loop Base address shall be hex ‘00 00 00 00’.

{“first LCN”: relative to what? what resets the count?}

New Loop Devices: The format of this field follows the AL_PA bit mapped format defined in FC–AL (see FC–AL table 15). If a bit is set = 1 in this field, then a Loop Device that was not present on the Loop when the last LCN was issued, or when the LCS was received, is present at the identified location.

Retired Loop Devices: The format of this field follows the AL_PA bit mapped format defined in FC–AL (see FC–AL table 15). If a bit is set = 1 in this field, then a Loop Device that was present on the Loop when the last LCN was issued, or when the LCS was received, is no longer present at the identified location.

Current Public Loop Devices: The format of this field follows the AL_PA bit mapped format defined in FC–AL (see FC–AL table 15). If a bit is set = 1 in this field, then a Public Loop Device is present at the identified location.

Current Private Loop Devices: The format of this field follows the AL_PA bit mapped format defined in FC–AL (see FC–AL table 15). If a bit is set = 1 in this field, then a Private Loop Device is present at the identified location.

Current Bypassed Loop Devices: The format of this field follows the AL_PA bit mapped format defined in FC–AL (see FC–AL table 15). If a bit is set = 1 in this field, then the Loop Device at the identified location has been bypassed. If Word 0 bit 31 in this field is set = 1, then information on bypassed Loop Devices is unavailable from the specific Loop.

Initial AL_PA position list: The format of this field follows the AL_PA position list format defined in FC–AL for the LILP Sequence.

A.4 Loop Change Subscribe (LCS)

The Loop Change Subscribe Extended Link Service request shall be used by an entity to subscribe to the Loop Change Notification (LCN) service.

Protocol:

Loop Change Subscribe request Sequence
Accept reply Sequence

Format: FT–1

Addressing: The S_ID designates the entity requesting subscription to the LCN service. The D_ID field designates the destination N_Port for the LCS, which shall be the Fabric Controller (hex ‘FF FF FD‘).

Payload: The format of the LCS Payload is shown in table 6. If the requesting entity must change the value of any of the fields of the LCS payload, it shall issue a new LCS which (if accepted) shall supercede the prior LCS.

Table 6 – LCS Payload

Item	Size Bytes
hex ‘xx 00 00 00’	4
Reserved	1
Loop Base Address	3
OX_ID for proxy ADISC	2
OX_ID for proxy PDISC	2
Port Name of requestor	8
Node Name of requestor	8
PLOGI payload	116 or 256

Loop Base Address: As defined in Clause 3.

OX_ID for proxy ADISC: If this field is non-zero, the Fabric Controller masquerading as the requestor entity, using the provided information, shall send an ADISC to all surviving Loop Devices on the identified Loop, every time Loop initialization completes on the identified Loop. If this field is zero, the Fabric Controller does not send an ADISC.

OX_ID for proxy PDISC: If this field is non-zero, the Fabric Controller masquerading as the requestor entity, using the provided information, shall send an PDISC to all surviving Loop Devices on the identified Loop, every time Loop initialization completes on the identified Loop. If this field is zero, the Fabric Controller does not send an PDISC.

Port Name of requestor: The 8 byte Port Name for the requestor as defined in FC–PH.

Node Name of requestor: The 8 byte Node Name for the requestor as defined in FC–PH.

PLOGI payload: The complete PLOGI payload, including the Extended Link Services command code, used by the requestor. The length of this field is either 116 or 256 bytes (see FC–PH–3).

Reply Extended Link Service Sequence:

Service Reject (LS_RJT)
Signifies the rejection of the LCS command
Accept (ACC)
Signifies the acceptance of the requestor as a subscriber to the LCN service.
– Accept Payload
The format of the Accept Payload is shown in table 7.

Table 7 – LCS Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4

A.5 Loop Change Unsubscribe (LCU)

An entity may wish to stop being notified of changes to Fabric attached Arbitrated Loop Configurations. The Loop Change Unsubscribe Extended Link Service request shall be used to achieves this.

Protocol:

Loop Change Unsubscribe request Sequence
Reply Sequence

Format: FT–1

Addressing: The S_ID designates the entity requesting termination of the subscription to the LCN service. The D_ID field designates the destination N_Port for the LCU, which shall be the Fabric Controller (hex ‘FF FF FD‘).

Payload: The format of the LCU request Payload is shown in table 8.

Table 8 – LCU Payload

Item	Size Bytes
hex ‘xx 00 00 00’	4

Reply Extended Link Service Sequence:

Service Reject (LS_RJT)
Signifies the rejection of the LCU command
Accept (ACC)
Signifies a termination of the LCN service to the requesting entity. An Accept is returned even if the requestor was not a LCN subscriber.
– Accept Payload

The format of the Accept Payload is shown in table 9.

Table 9 – LCU Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4

A.6 Loop Status (LSTS)

A Loop Status Extended Link Service request shall be sent to ascertain the state of a specific Loop.

Protocol:

Loop Status request Sequence
Reply Sequence

Format: FT–1

Addressing: The S_ID designates the entity requesting status for a specific Loop. The D_ID field designates the destination for the LSTS, which shall be the Loop Base address, the lowest 24 bit address identifier assigned to the specified Arbitrated Loop.

Payload: The format of the LSTS Payload is shown in table 10.

Table 10 – LSTS Payload

Item	Size Bytes
hex ‘xx 00 00 00’	4

Reply Extended Link Service Sequence:

Service Reject (LS_RJT)
Signifies the rejection of the LSTS command
Accept (ACC)
Supplies the requested status for the identified Loop.
– Accept Payload

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

Table 11 – LSTS Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4
Reserved	3
Loop State	1
Current Public Loop Devices (AL_PA bit map)	16
Current Private Loop Devices (AL_PA bit map)	16
Current Bypassed Loop Devices (AL_PA bit map)	16
Initial AL_PA position list (AL_PA position list from LILP)	128

Loop State: The format of the Loop State field is shown in table 12.

Table 12 – Loop State

State	value
Reserved	0
Link Failure {detected at FL_Port?}	1
Offline {meaning what?}	2
Reserved	3
Initializing {assume this means somewhere between first LIP sent or received and LISM or LIFA sent or received?}	4
Address assignment {break this into LISM, LIFA, ...?}	5
Reserved	6 - 254
Online {define by loop state?}	255

Current Public Loop Devices: The format of this field follows the AL_PA bit mapped format defined in FC–AL (see FC–AL table 15). If a bit is set = 1 in this field, then a Public Loop Device is present at the identified location.

Current Private Loop Devices: The format of this field follows the AL_PA bit mapped format defined in FC–AL (see FC–AL table 15). If a bit is set = 1 in this field, then a Private Loop Device is present at the identified location.

Current Bypassed Loop Devices: The format of this field follows the AL_PA bit mapped format defined in FC–AL (see FC–AL table 15). If a bit is set = 1 in this field, then the Loop Device at the identified location has been bypassed. If Word 0 bit 31 in this field is set = 1, then information on bypassing of Loop Devices is unavailable for the specific Loop.

Initial AL_PA position list: The format of this field follows the AL_PA position list format defined in FC–AL for the LILP Sequence.

{PROPOSED VERSIONS DISCUSSED AT DALLAS LATE ON MAY 2. NOT A DONE DEAL!}

A.7 Loop Change Notification (LCN)

A Loop Change Notification Extended Link Service request shall be sent to subscribing destinations when an event occurs on an Arbitrated Loop, which may have affected current PLOGI states, or makes new Arbitrated Loop resources available.

The Fabric Controller may delay the sending of LCNs for newly added Loop Devices, to allow sufficient time to determine whether the new Loop Devices are Public or Private.

The LCN shall be sent in the Class of Service initially used to subscribe to the LCN service.

Protocol:

Loop Change Notification request Sequence
No reply Sequence

Format: FT–1

Addressing: The S_ID designates the N_Port reporting the LCN, which shall be the Fabric Controller (hex ‘FF FF FD‘) {why not the Loop Base Addr?}. The D_ID field designates the destination for the LCN.

Payload: The format of the LCN Payload is shown in table 5.

Table 13 – LCN Payload

Item	Size Bytes
hex ‘xx 00 00 00’	4
Reserved	1
Loop Base Address	3
Participating AL_PAs (AL_PA bit map from LISA)	16
Initial AL_PA position list (AL_PA position list from LILP)	128

Loop Base Address: As defined in Clause 3.

Participating AL_PAs: The format of this field follows the AL_PA bit mapped format defined in FC–AL (see FC–AL table 15) that was the payload of the LISA Sequence.

Initial AL_PA position list: The format of this field follows the AL_PA position list format defined in FC–AL for the LILP Sequence.

A.8 Loop Change Subscribe (LCS)

The Loop Change Subscribe Extended Link Service request shall be used by an entity to subscribe to the Loop Change Notification (LCN) service. Once subscribed, the entity will receive an LCN anytime a loop within the Fabric changes state.

Protocol:

Loop Change Subscribe request Sequence
Accept reply Sequence

Format: FT–1

Addressing: The S_ID designates the entity requesting subscription to the LCN service. The D_ID field designates the destination N_Port for the LCS, which shall be the Fabric Controller (hex ‘FF FF FD‘).

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

Table 14 – LCS Payload

Item	Size Bytes
hex ‘xx 00 00 00’	4

Reply Extended Link Service Sequence:

Service Reject (LS_RJT)
Signifies the rejection of the LCS command
Accept (ACC)
Signifies the acceptance of the requestor as a subscriber to the LCN service.
– Accept Payload
The format of the Accept Payload is shown in table 7.

Table 15 – LCS Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4

A.9 Loop Change Unsubscribe (LCU)

An entity may wish to stop being notified of changes to Fabric attached Arbitrated Loop Configurations. The Loop Change Unsubscribe Extended Link Service request shall be used to achieves this.

Protocol:

Loop Change Unsubscribe request Sequence
Reply Sequence

Format: FT–1

Addressing: The S_ID designates the entity requesting termination of the subscription to the LCN service. The D_ID field designates the destination N_Port for the LCU, which shall be the Fabric Controller (hex ‘FF FF FD‘).

Payload: The format of the LCU request Payload is shown in table 8.

Table 16 – LCU Payload

Item	Size Bytes
hex ‘xx 00 00 00’	4

Reply Extended Link Service Sequence:

Service Reject (LS_RJT)
Signifies the rejection of the LCU command

Accept (ACC)

Signifies a termination of the LCN service to the requesting entity. An Accept is returned even if the requestor was not a LCN subscriber.

– Accept Payload

The format of the Accept Payload is shown in table 9.

Table 17 – LCU Accept Payload

Item	Size Bytes
hex ‘02 00 00 00’	4

Annex B
(informative)

BB_Credit Management for Public Loop Devices

B.1 Login_BB_Credit

OPN Originators may open full or half duplex, regardless of the value of Login_BB_Credit.

B.2 Available_BB_Credit

Available_BB_Credit is a variable used by a transmitter to determine permission to transmit frames, and if so how many. The transmitter may transmit a frame when Available_BB_Credit is greater than 0. The rules for modifying Available_BB_Credit are:

a) In an OPN Originator, Available_BB_Credit may be initialized to a value less than or equal to the Login_BB_Credit of the OPN Recipient upon transmission of an OPN by the OPN Originator.

b) In an OPN Recipient (acting as a Sequence Initiator), Available_BB_Credit may be initialized to a value less than or equal to Login_BB_Credit of the OPN Originator (acting as a Sequence Recipient) upon receipt of the OPN.

c) Available_BB_Credit is decremented upon transmission of a frame.

d) Available_BB_Credit is incremented upon receipt of R_RDY, except that following an OPN to or from a Sequence Recipient with Login_BB_Credit>0, a number of R_RDYs is discarded equal to the value to which Available_BB_Credit was initialized.

B.3 Login_BB_Credit

The Login_BB_Credit of a Loop Device is the maximum number of frames that the device commits to receive immediately following the receipt of an OPN. At any time, a device may have more than Login_BB_Credit receive buffers available, and may make those buffers available to a Sequence Initiator with which the device is currently OPN. The process by which a device decides that buffers are available is specific to the device.

B.3.1 Login_BB_Credit=0

Interoperability with devices which advertise Login_BB_Credit=0 is required by this report. When Login_BB_Credit=0 at the Sequence Recipient, the following rules apply:

a) For each frame transmitted, Sequence Initiators must receive one or more R_RDYs from the Sequence Recipient before transmitting the frame.

b) If the Sequence Recipient has one or more available receive buffers, it shall transmit the number of R_RDYs equal to the number of available receive buffers upon receipt of OPN or transmission of OPN.

c) If the Sequence Recipient is an OPN Recipient which has no available receive buffers, it shall respond with an immediate CLS rather than keep the circuit open until a buffer becomes available.

B.3.2 Login_BB_Credit>0

a) When a device with Login_BB_Credit>0 becomes open (either as OPN Originator or as OPN Recipient), it shall transmit no fewer than Login_BB_Credit R_RDYs back to the other device. This provides compatibility with OPN Originators which wait for R_RDYs before transmitting, and provides the full available credit to devices that initialize Available_BB_Credit to a value less than Login_BB_Credit.

b) If a Sequence Initiator (either as OPN Originator or OPN Recipient) chooses to take advantage of nonzero Login_BB_Credit advertised by a Sequence Recipient, the following rules apply:

   1) The Sequence Initiator must be aware of the value of Login_BB_Credit at the Sequence Recipient in order to know how many frames it may transmit without receiving any R_RDYs.

   2) The Sequence Initiator may initialize its value for Available_BB_Credit to any value between one and the Login_BB_Credit of the Sequence Recipient immediately following the transmission or receipt of the OPN.

   3) The Sequence Initiator may transmit Available_BB_Credit frames immediately following the transmission or receipt of the OPN without waiting for R_RDYs.

   4) The Sequence Initiator shall discard a number of initial R_RDYs equal to the initial value of Available_BB_Credit.

c) When Login_BB_Credit>0 at an NL_Port, that NL_Port shall neither originate nor forward a CLS until it can guarantee that the number of available receive buffers upon receipt of an immediate, subsequent OPN will be greater than or equal to its Login_BB_Credit. A number of strategies may be employed to ensure this:

   1) The NL_Port never transitions from RECEIVED CLOSE to MONITORING, or from XMITTED CLOSE to MONITORING, or from TRANSFER to MONITORING, unless it has Login_BB_Credit buffers available.

   2) The NL_Port never makes available more than Login_BB_Credit buffers, but always has two times Login_BB_Credit buffers actually available before arbitrating and originating an OPN.

B.3.3 OPN and CLS latencies

When Login_BB_Credit=0, CLS latency is minimized but a latency exists on every OPN while waiting for an R_RDY before frames can be transmitted.

When Login_BB_Credit>0, OPN latency is minimized but CLS latency may increase while waiting for Login_BB_Credit buffers to become available. It is possible to reduce CLS latency when Login_BB_Credit>0 by increasing the available receive buffers within a Sequence Recipient. For example, a Sequence Recipient may choose to have:

Available_Rx_Buffers=Available BB_Credit (of current Sequence Initiator) + Login_BB_Credit

at all times during a Loop Tenancy with a Sequence Initiator. This would allow the Sequence Recipient as a CLS Recipient to forward a CLS immediately without having to wait for Login_BB_Credit to become available. A Sequence Recipient controls the Available BB_Credit variable in a Sequence Initiator through the return of R_RDYs.

Some ports may be tempted to reduce CLS latency in another fashion. To prevent buffer overruns, a CLS Recipient is only required to have Login_BB_Credit buffers available before the next OPN is received, not necessarily before the CLS is forwarded. Ports which initiate or forward a CLS before Login_BB_Credit is available and rely on a (topology-specific) delay between the time a CLS is transmitted or forwarded and the time an OPN is received do so at their own risk and with methods not defined in this Technical Report.

Annex C
(informative)

Getting the Most From Your Fabric

{NOTE: This section under construction! TBD by Bent S.}

C.1 Introduction

...and stuff...

A

B

C

D

E

F

H

I

L

M

N

O

P

Q

R

S

T

U