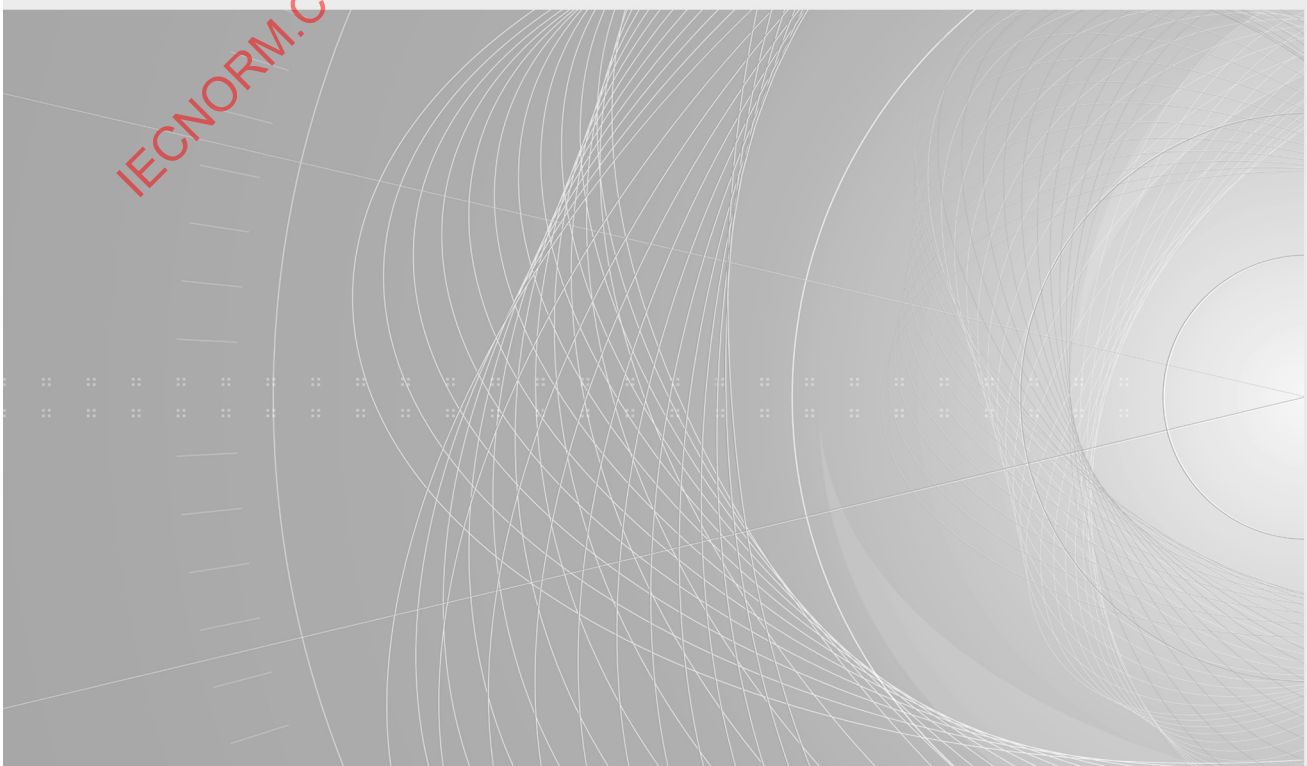
ISO/IEC 14165-243

Edition 1.0 2012-12

INTERNATIONAL STANDARD

Information technology – Fibre channel – Part 243: Fibre channel backbone-3 (FC-BB-3)



ISO/IEC 14165-243:2012(E)

IEC Central Office
3, rue de Varembé
CH-1211 Geneva 20
Switzerland

Tel.: +41 22 919 02 11
Fax: +41 22 919 03 00
info@iec.ch
www.iec.ch

About the IEC

The International Electrotechnical Commission (IEC) is the leading global organization that prepares and publishes International Standards for all electrical, electronic and related technologies.

About IEC publications

The technical content of IEC publications is kept under constant review by the IEC. Please make sure that you have the latest edition, a corrigenda or an amendment might have been published.

Useful links:

IEC publications search - www.iec.ch/searchpub

The advanced search enables you to find IEC publications by a variety of criteria (reference number, text, technical committee,…).
It also gives information on projects, replaced and withdrawn publications.

IEC Just Published - webstore.iec.ch/justpublished

Stay up to date on all new IEC publications. Just Published details all new publications released. Available on-line and also once a month by email.

Electropedia - www.electropedia.org

The world's leading online dictionary of electronic and electrical terms containing more than 30 000 terms and definitions in English and French, with equivalent terms in additional languages. Also known as the International Electrotechnical Vocabulary (IEV) on-line.

Customer Service Centre - webstore.iec.ch/csc

If you wish to give us your feedback on this publication or need further assistance, please contact the Customer Service Centre: csc@iec.ch.

ISO/IEC 14165-243

Edition 1.0 2012-12

INTERNATIONAL STANDARD

Information technology – Fibre channel – Part 243: Fibre channel backbone-3 (FC-BB-3)

INTERNATIONAL ELECTROTECHNICAL COMMISSION

PRICE CODE XB

ICS 35.200

ISBN 978-2-83220-542-6

CONTENTS

INFORMATION TECHNOLOGY – FIBRE CHANNEL – Part 243: Fibre channel backbone-3 (FC-BB-3)

FOREWORD

1) ISO (International Organization for Standardization) and IEC (International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards. Their preparation is entrusted to technical committees; any ISO and IEC member body interested in the subject dealt with may participate in this preparatory work. International governmental and non-governmental organizations liaising with ISO and IEC also participate in this preparation.

2) In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

3) The formal decisions or agreements of IEC and ISO on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC and ISO member bodies.

4) IEC, ISO and ISO/IEC publications have the form of recommendations for international use and are accepted by IEC and ISO member bodies in that sense. While all reasonable efforts are made to ensure that the technical content of IEC, ISO and ISO/IEC publications is accurate, IEC or ISO cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.

5) In order to promote international uniformity, IEC and ISO member bodies undertake to apply IEC, ISO and ISO/IEC publications transparently to the maximum extent possible in their national and regional publications. Any divergence between any ISO/IEC publication and the corresponding national or regional publication should be clearly indicated in the latter.

6) ISO and IEC provide no marking procedure to indicate their approval and cannot be rendered responsible for any equipment declared to be in conformity with an ISO/IEC publication.

7) All users should ensure that they have the latest edition of this publication.

8) No liability shall attach to IEC or ISO or its directors, employees, servants or agents including individual experts and members of their technical committees and IEC or ISO member bodies for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication of, use of, or reliance upon, this ISO/IEC publication or any other IEC, ISO or ISO/IEC publications.

9) Attention is drawn to the normative references cited in this publication. Use of the referenced publications is indispensable for the correct application of this publication.

10) Attention is drawn to the possibility that some of the elements of this International Standard may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

International Standard ISO/IEC 14165-243 was prepared by subcommittee 25: Interconnection of information technology equipment, of ISO/IEC joint technical committee 1: Information technology.

A list of all currently available parts of the ISO/IEC 14165 series, under the general title *Information technology – Fibre channel,* can be found on the IEC web site.

This International Standard has been approved by vote of the member bodies and the voting results may be obtained from the address given on the second title page.

This publication has been drafted in accordance with the ISO/IEC Directives, Part 2.

INTRODUCTION

This International Standard specifies mechanisms that allow extension of Fibre Channel links and/or switched networks across Wide Area Networks. FC-BB-3 defines four distinct Fibre Channel backbone mappings: FC over ATM, FC over SONET, FC over TCP/IP, and FC over GFPT.

INFORMATION TECHNOLOGY – FIBRE CHANNEL – Part 243: Fibre channel backbone-3 (FC-BB-3)

1 Scope

This part of ISO/IEC 14165-243 consists of four distinct Fibre Channel mappings resulting in the following four models:

- FC-BB-3_ATM (FC over ATM backbone network)
- FC-BB-3_SONET (FC over SONET backbone network)
- FC-BB-3_IP (FC over TCP/IP backbone network)
- FC-BB-3_GFPT (FC over SONET/SDH/OTN/PDH backbone network using GFPT adaptation)

Figure 1, figure 2, and figure 3 illustrate the scope and the major components of the FC-BB-3 models and its relationship to the FCIP standard and the ATM Forum/ITU-T standards. Table 1 shows the organization of this standard. FC-BB-3_IP, FC-BB-3_ATM, FC-BB-3_SONET, and FC-BB-3_GFPT do not interoperate in any way and are independent models.

Table 1 – FC-BB-3 Organization

Model type	Applicable Clauses and Annexes
FC-BB-3_ATM, FC-BB-3_SONET, FC-BB-3_IP, FC-BB-3_GFPT	1-4
FC-BB-3_ATM, FC-BB-3_SONET	5, 6
FC-BB-3_ATM	7, Annexes A, B, C
FC-BB-3_SONET	8, Annexes A, C
FC-BB-3_IP	9, Annex A
FC-BB-3_GFPT	10

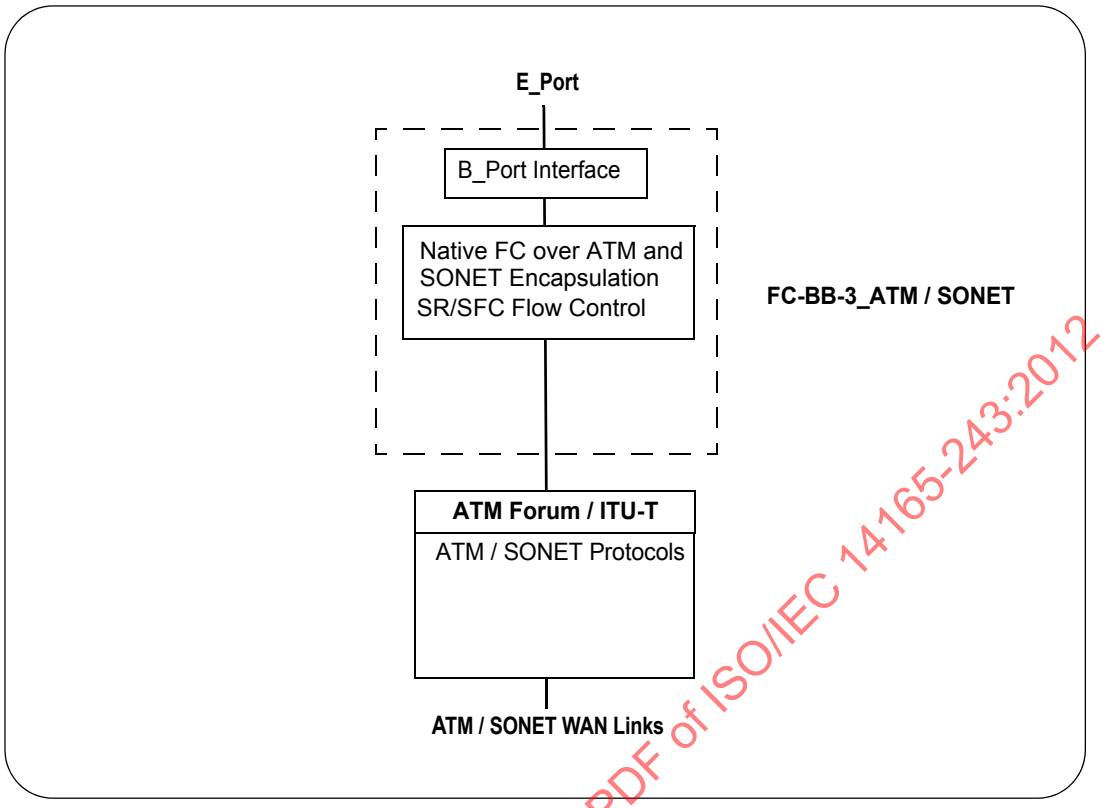


Figure 1 – Scope and components of FC-BB-3_ATM/SONET models

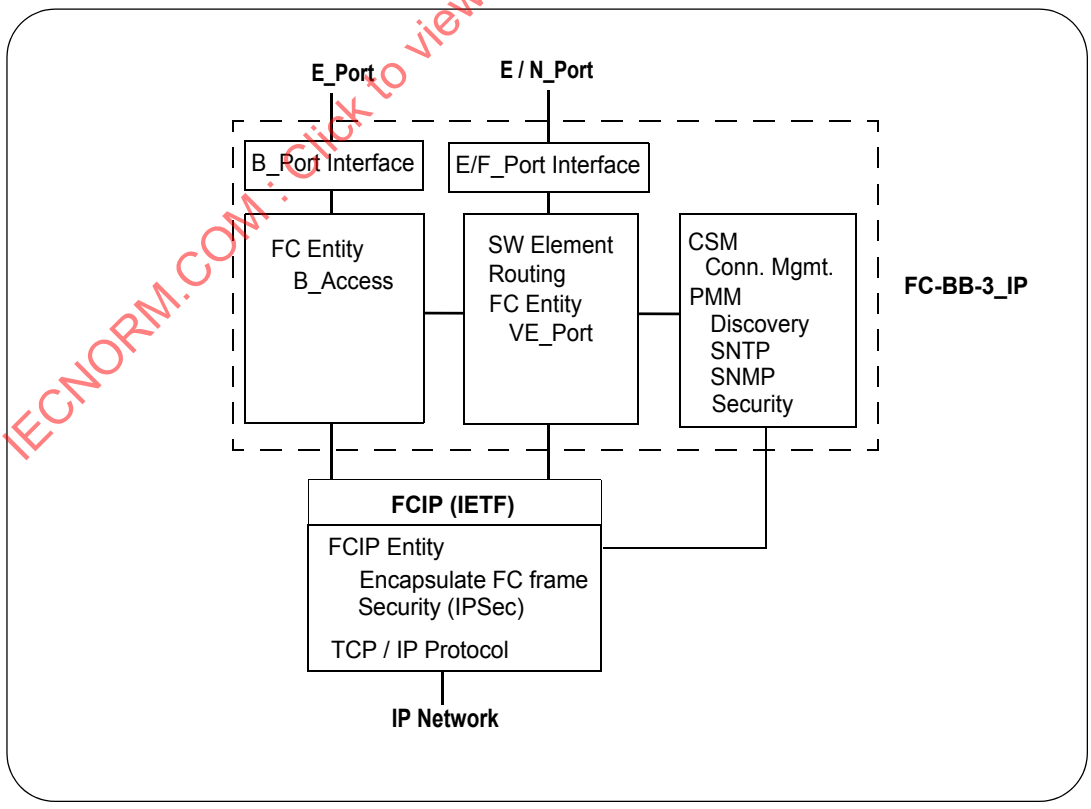


Figure 2 – Scope and components of FC-BB-3_IP model

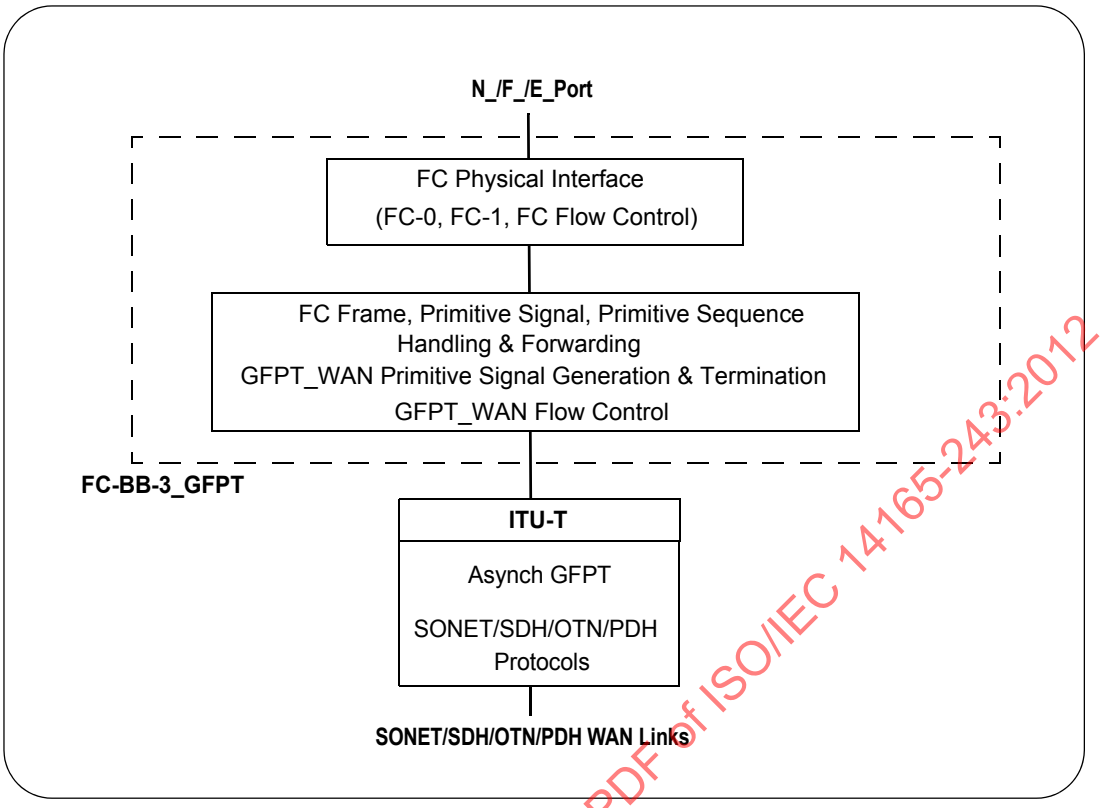


Figure 3 – Scope and components of FC-BB-3_GFPT model

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document, including any amendments, applies.

ISO/IEC 3309:1993, *Information technology - Telecommunications and information exchange between systems - High-level Data Link Control (HDLC) Procedures - Frame structure* (withdrawn)

ISO/IEC 13239:1997(E), *Information technology - Telecommunications and information exchange between systems - High-level data link control (HDLC) procedures*

INCITS 426-2007, *Information technology - Fibre Channel Security Protocols (FC-SP)* [1]
(planned as ISO/IEC 14165-431)

ANSI T1.105-2001, *Synchronous Optical Network (SONET) - Basic Description Including Multiplex Structures, Rates, and Formats*

For electronic copies of references under development by INCITS T11, see www.t11.org

INCITS 418-2006, *Fibre Channel - Switch Fabric - 4 (FC-SW-4)*

INCITS 424-2007, *Fibre Channel - Framing and Signaling -2 (FC-FS-2)*

INCITS 433-2007, *Fibre Channel - Link Services (FC-LS)*

Copies of the following approved ITU-T standards may be obtained through the ITU-T Publications department at http://www.itu.int

ITU-T Rec. I.356 (2000), *B-ISDN ATM layer cell transfer performance*

ITU-T Rec. I.363.5 (1996), *B-ISDN ATM Adaptation Layer specification: Type 5 AAL*

ITU-T X.25-1997, *Interface between Data terminal Equipment (DTE) and Data Circuit-terminating Equipment (DCE) for terminals operating in the packet mode and connected to public data networks by dedicated circuit, X.25-1997*

ITU-T Q.2931 (1995), *Broadband Integrated Services Digital Network (B-ISDN) Digital Subscriber Signaling System No. 2 (DSS2); User-Network Interface (UNI) Layer 3 Specification for Basic Call/Connection Control*

ITU-T Q.2971 (1995), *Broadband Integrated Services Digital Network (B-ISDN) Digital Subscriber Signaling System No. 2 (DSS2); User-Network Interface (UNI) Layer 3 Specification for Point-to-Multipoint Call/Connection Control*

ITU-T Rec. G.707/Y.1322, *Network/Node Interface for the Synchronous Digital Hierarchy (SDH),* 2003

ITU-T Rec. G.709/Y.1331, *Interfaces for the Optical Transport Network (OTN),* 2004

ITU-T Rec. G.7041/Y.1303, (2003), *Generic Framing Procedure (GFP)*

ITU-T Rec. G.7042/Y.1305 (2004), *Link capacity adjustment scheme (LCAS) for virtual concatenated signals*

ITU-T Rec. G.7043/Y.1343 (2004) *Virtual concatenation of PDH signals*

ITU-T Rec. G.783, (2000), *Characteristics of synchronous digital hierarchy (SDH) equipment functional blocks*

ITU-T Rec. G.798, (2002), *Characteristics of optical transport network hierarchy equipment functional blocks*

ITU-T Rec. G.8040/Y.1340, (2003), *GFP Frame Mapping into Plesiochronous Digital Hierarchy (PDH)*

ITU-T Rec. G.806, *Characteristics of Transport Equipment - Description Methodology and Generic Functionality*

Copies of the following approved IETF standards may be obtained through the Internet Engineering Task Force (IETF) at www.ietf.org.

RFC 1619, *PPP over SONET/SDH*, May 1994

1. T11/Project 1570D/Rev. 1.6

RFC 1661, *The Point-to-Point Protocol (PPP),* July 1994

RFC 1662, *PPP in HDLC-like Framing*, July 1994

RFC 2030, *Simple Network Time Protocol (SNTP) Version 4 for IPv4, IPv6 and OSI,* October 1996

RFC 3643, *Fibre Channel (FC) Frame Encapsulation,* December 2004

RFC 3821, *Fibre Channel Over TCP/IP (FCIP)*, July 2004

RFC 3822, *Finding Fibre Channel over TCP/IP (FCIP) Entities Using Service Location Protocol version 2 (SLPv2),* July 2004

Copies of the following approved ATM Forum standards may be obtained through the MFA Forum at http://www.mfaforum.org/tech/atm_specs.shtml

ATM Forum (1996), *Traffic Management Specification 4.0*. AF-TM-0056.000

3 Terms, definitions and conventions

3.1 Terms and definitions

For the purposes of this document the following terms and definitions apply.

3.1.1
BBW
refers to either an FC-BB-3_ATM or an FC-BB-3_SONET device

3.1.2
B_Port
Bridge Port on a device that implements FC-BB-3_ATM, FC-BB-3_SONET or FC-BB-3_IP, and connects to an E_Port on an FC switch

3.1.3
B_Port_Name
name_Identifier (see 3.1.20) that identifies a B_Port (see 3.1.2) for identification purposes

Note 1 to entry: The format of the name is specified in FC-SW-4.

3.1.4
BSW
generic term for a backbone switch

Note 1 to entry: See FC-SW-4.

3.1.5
codeword
sequence of bits of a code corresponding to a symbol

3.1.6
E_Por
Fabric expansion port that attaches to another E_Port to create an Inter-Switch Link

Note 1 to entry: See FC-SW-4.

3.1.7
E_Port_Name
Name_Identifier (see 3.1.20) that identifies an E_Port (see 3.1.6)

3.1.8
Fabric_Name
Name_Identifier (see 3.1.20) associated with a Fabric

Note 1 to entry: See FC-SW-4.

3.1.9
F_Port
port by which non-loop N_Ports are attached to a Fabric

Note 1 to entry: This port does not include FL_Ports.

Note 2 to entry: See FC-SW-4 and FC-FS-2.

3.1.10
F_Port_Name
Name_Identifier (see 3.1.20) that identifies an F_Port (see 3.1.9)

3.1.11
Fabric Initialization
process for configuring and building a Fabric

Note 1 to entry: See FC-SW-4.

3.1.12
FC-BB-3_ATM
model defining equipment that interfaces with a Fibre Channel switched network on one side and an ATM network on the other side

3.1.13
FC-BB-3_GFPT
equipment model defining gateway functionality for the interconnection of two non-Arbitrated Loop FC physical ports across a GFPT WAN infrastructure

EXAMPLE SONET, SDH, OTN, PDH

Note 1 to entry: Supports both arbitrary-rate WAN transport and distance extension of buffer-to-buffer flow control.

3.1.14
FC-BB-3_IP
model defining equipment that interfaces with a Fibre Channel switched network on one side and an IP network on the other side

3.1.15
FC-BB-3_SONET
model defining equipment that interfaces with a Fibre Channel switched network on one side and a SONET/SDH network on the other side

3.1.16
FC_Port
port generating/terminating and/or forwarding FC frames, and generating/terminating FC Primitive Signals and Primitive Sequences

Note 1 to entry: FC_Ports include N_Ports, F_Ports, E_Ports, B_Ports, VE_Ports, and B_Access.

3.1.17
Fibre Channel Backbone link
Transport Trail or equivalent network channel connection used for communications between two FC-BB-3 devices

Note 1 to entry: This encompasses FC-BBW_ATM, FC-BBW_SONET, FC-BB-3_IP, and GFPT_WAN links.

Note 2 to entry: Note that a Fibre Channel Backbone link may, in some cases, be comprised of more than one physical or logical connection.

3.1.18
Generic Framing Procedure
GFP
procedure for adaptation of data (i.e., PDUs or 8B/10B encoded characters) to octet-synchronous (i.e., SONET, SDH, OTN) and bit-synchronous (i.e., PDH) Wide Area Network transport infrastructures, specified by ITU-T

Note 1 to entry: Wide Area Network transport infrastructures is specified by ITU-T.

Note 2 to entry: See ITU-T Rec. G.7041/Y.1303.

3.1.19
keep alive timeout value
K_A_TOV
timer that is used by the Link Keep Alive (LKA) ELS as a trigger for issuing LKA

Note 1 to entry: For LKA ELS see FC-LS.

3.1.20
Name_Identifier
64-bit identifier, with a 60-bit value preceded with a 4-bit Network_Address_Authority Identifier, used to identify entities in Fibre Channel (e.g., N_Port, node, F_Port, or Fabric) (see FC-FS-2)

3.1.21
Node_Name
Name_Identifier (see 3.1.20) associated with a node (see FC-FS-2)

3.1.22
N_Port
device port that generates/terminates FC-4 channel traffic

3.1.23
N_Port_Name
name_Identifier (see 3.1.20) that identifies an N_Port (see 3.1.22)

3.1.24
Ordered Set
See FC-FS-2.

3.1.25
outstanding poll condition
condition where the BBW has sent a command message with the P bit (see 5.5.4.7) set to one and has not yet received a response message with the F bit (see 5.5.4.7) set to one (see 6.3.5)

3.1.26
Simple Flow Control
SFC
flow control protocol that may be applied between two FC-BB-3_ATM or FC-BB-3_SONET devices across a ATM/SONET WAN

3.1.27
Selective Retransmission flow control
SR flow control
sliding window flow control protocol that may be applied between two FC-BB-3_ATM or FC-BB-3_SONET devices across a ATM/SONET WAN

Note 1 to entry: SR flow control provides for both flow control and error recovery.

3.1.28
Switch_Name
Name_Identifier (see 3.1.20) that identifies a Switch or a Bridge device

Note 1 to entry: The format of the name is specified in FC-FS-2. Each Switch and Bridge device shall provide a unique Switch_Name within the Fabric.

3.1.29
WAN interface
interface that connects to a Wide Area Network

Note 1 to entry: A WAN interface may be physical (e.g., ATM, SONET) or logical (e.g., GFPT_WAN)

3.2 FC-BB-3_ATM definitions

3.2.1
AAL
ATM Adaptation Layer
collection of standardized protocols that adapt user traffic to 48-octet payloads that may be placed in a cell-formatted stream

Note 1 to entry: The AAL is subdivided into the Convergence Sublayer (CS) and the Segmentation and Reassembly (SAR) sublayer. There are currently four types of AALs (i.e., AAL1, AAL2, AAL3/4, and AAL5) to support the various service categories.

3.2.2
AAL Service Categories
the ATM Forum has defined five Traffic Service Categories supported by the AALs: Constant Bit Rate (CBR), Variable Bit Rate-Real Time (VBR-RT), Variable Bit Rate-Non Real Time (VBR-NRT), Available Bit Rate (ABR), and Unspecified Bit Rate (UBR)

3.2.3
AAL Type 5
AAL5
protocol standard that is used in FC-BB-3_ATM

Note 1 to entry: AAL5 was originally intended for variable bit rate traffic not requiring source-destination timing relation.

Note 2 to entry: AAL5 is now also used with applications that have constant bit rate traffic where source-destination timing relation is important.

3.2.4
ATM
Asynchronous Transfer Mode

Note 1 to entry: Broadband-ISDN standards defined by ITU-T and the ATM Forum.

3.2.5
ATM QoS parameters
set of performance characteristics of the contracted ATM connection

Note 1 to entry: Six ATM QoS parameters are currently defined: Peak-to-peak Cell Delay Variation (CDV), Maximum Cell Transfer Delay (maxCTD), Cell Loss Ratio (CLR), Cell Error Ratio (CER), Severely Errored Cell Block Ratio (SECBR), and Cell Misinsertion Rate (CMR).

3.2.6
ATM Traffic Descriptor
ttraffic characteristics of an ATM connection

Note 1 to entry: A Connection Traffic Descriptor includes a Source Traffic Descriptor, CDV Tolerance (CDVT), and a Conformance definition. A Source Traffic Descriptor is described by the following parameters: Peak Cell Rate (PCR), Sustainable Cell Rate (SCR), Maximum Burst Size (MBS), and a Minimum Cell Rate (MCR).

3.2.7
Cell Loss Priority
CLP
one-bit field in the ATM cell header specifying whether the cell is more (i.e., CLP=1) or less (i.e., CLP=0) likely to be discarded by an ATM network experiencing congestion

3.2.8
Cell Loss Ratio
CLR
QoS parameter that gives the ratio of lost cells to the total number of transmitted cells

3.2.9
Connection Admission Control
CAC
actions taken by the ATM network to accept or reject a connection request based on its QoS and traffic parameter requirements and then route this connection across the network

3.2.10
Convergence Sublayer Protocol Data Unit
PDU used at the Convergence Sublayer (CS) for passing information between the higher layers and the Segmentation and Reassembly (SAR) sublayer that is located below the CS where the cell conversion takes place (see 3.2.1)

3.2.11
GCRA
method applied at the network side of the UNI (see 3.2.19) to test the conformance of an ATM cell to its traffic contract

3.2.12
Operations, Administration, And Maintenance
OAM
management framework defined by the ITU

Note 1 to entry: OAM cells are special purpose ATM cells exchanged between an ATM end-system and an ATM switch and between ATM switches. OAM cells are used for network fault and performance management and analysis.

3.2.13
Permanent Virtual Circuit
PVC
preconfigured logical connection between two ATM devices

3.2.14
Permanent Virtual Connection
PVC
ATM term for a Permanent Virtual Circuit (see 3.2.13) between ATM devices

Note 1 to entry: The terms may be used interchangeably.

3.2.15
Switched Virtual Call
generic term that refers to switched virtual circuits (see 3.2.16) and switched virtual connections (see 3.2.17)

3.2.16
Switched Virtual Circuit
logical ATM connection established via signaling

Note 1 to entry: End systems transmit a UNI 3.1/4.0 signaling request via the Q.2931 Signaling Protocol.

3.2.17
Switched Virtual Connection
SVC
ATM term for Switched Virtual Circuit (see 3.2.16)

3.2.18
Usage Parameter Control
UPC
set of policing mechanisms implemented by the network at the UNI (see 3.2.19) to monitor and control traffic submitted by each end user

3.2.19
User-network Interface
UNI
interface, defined as a set of protocols and traffic characteristics, between the customer premises equipment and the ATM networks

3.2.20
Virtual Channel
one of several logical connections defined within one Virtual Path (see 3.2.23) between two ATM devices

3.2.21
Virtual Channel Connection
VCC
concatenation of Virtual Channel Links

Note 1 to entry: Switching cells within an ATM switch for a given VCC is based on the VPI/VCI (see 3.2.25) value indicated on the cell header.

3.2.22
Virtual Circuit
VC
connection that is set up across the network between a source and a destination where a fixed route is chosen for the entire session and bandwidth and ID dynamically allocated to the user

3.2.23
Virtual Path
VP
logical connection between two ATM devices (e.g., CPEs, switches)

Note 1 to entry: A Virtual Path consists of a set of Virtual Channels (see 3.2.20).

3.2.24
Virtual Path Connection
VPC
concatenation of Virtual Path Links (VPLs)

Note 1 to entry: Switching cells within an ATM switch for a given VPC is based on the VPI value indicated on the cell header.

3.2.25
Virtual Path Identifier/Virtual Channel Identifier
VPI/VCI
combination of two numbers, namely the VPI and the VCI, used to identify a Virtual Connection (VC) and switch cells in an ATM network

3.3 FC-BB-3_SONET definitions

3.3.1
Administrative Unit
AU
SDH-specific information structure, consisting of an STS SPE (see 3.3.14) and its associated set of STS pointer/pointer action bytes

3.3.2
Concatenated Synchronous Transport Signal Level *N*
STS-*N*c
STS-*N* Line layer signal in which the STS Envelope Capacities from the *N* STS-1s have been combined to carry an STS-*N*c Synchronous Payload Envelope (SPE) (see 3.3.14) that is transported not as several separate signals but as a single entity

Note 1 to entry: The equivalent SDH term for an STS-3c SPE is a VC-4.

3.3.3
container
SDH term that is equivalent to the payload capacity of a Synchronous Payload Envelope (SPE) (see 3.3.14)

3.3.4
Operations, Administration, and Maintenance
OAM
management framework defined by the ITU

Note 1 to entry: OAM cells are special purpose ATM cells exchanged between an ATM end-system and an ATM switch and between ATM switches.

Note 2 to entry: OAM cells are used for network fault and performance management and analysis.

3.3.5
Optical Carrier Level N
OC-N
optical signal that results from an optical conversion of an STS-*N* signal

Note 1 to entry: SDH does not make the distinction between a logical signal (e.g., STS-1 in SONET) and a physical signal (e.g., OC-1 in SONET). The equivalent SDH term for both logical and physical signals is Synchronous Transport Module Level M (STM-M), where M = (*N*/3). There are equivalent STM-M signals only for values of *N* = 3, 12, 48, and 192.

3.3.6
Optical Transport Network
OTN
framework of ITU standards for optical signal multiplexing and transport

3.3.7
Path
logical connection between the point at which a standard frame format for the signal at the given rate is assembled, and the point at which the standard frame format for the signal is disassembled

Note 1 to entry: The equivalent SDH term is also Path.

3.3.8
Payload Pointer
pointer that indicates the location of the beginning of the Synchronous Payload Envelope (see 3.3.14)

Note 1 to entry: The equivalent SDH term is pointer.

3.3.9
Plesiochronous Digital Hierarchy
PDH
bit-oriented telecommunications multiplexing and transport protocols (e.g., DS-3, E3)

3.3.10
SONET
acronym for Synchronous Optical NETwork

Note 1 to entry: SONET is a term in general usage, that refers to the rates and formats specified in ANSI T1.105.

3.3.11
STS Path Terminating Equipment
STS PTE
network elements that multiplex/demultiplex the STS payload

Note 1 to entry: STS PTEs may originate, access, modify, or terminate the STS Path Overhead necessary to transport the STS payload, or may perform any combination of these actions.

3.3.12
super-rate signals
signal that has to be carried by a Concatenated Synchronous Transport Signal level Nc (STS-Nc)

Note 1 to entry: There is no equivalent SDH term.

3.3.13
Synchronous Digital Hierarchy
SDH
family of ITU-T standards whose technical contents closely resemble that found for the SONET family of ANSI standards

3.3.14
STS Synchronous Payload Envelope
STS SPE
125-microsecond frame structure composed of STS Path Overhead and bandwidth for payload

Note 1 to entry: The term generically refers to STS-1 SPEs and STS-Nc SPEs. The equivalent SDH term for STS-1 SPE is Virtual Container level 3 (VC-3). The equivalent SDH term for STS-3c SPE is Virtual Container level 4 (VC-4). The equivalent SDH term for STS-Nc SPE (N>3) is Virtual Container level 4-Xc (VC-4-Xc), where X=(N/3).

3.3.15
Synchronous Transport Module Level M
STM-M
transport signals for the Synchronous Digital Hierarchy (SDH) (see 3.3.13)

Note 1 to entry: Defined signals exist at rates of M times 155.52 Mbit/s, where M = 1, 4, 16, or 64 .

Note 2 to entry: These are equivalent to SONET OC-N signals, where N = 3M.

3.3.16
Synchronous Transport Signal Level N
STS-N
signal obtained by byte interleaving N STS-1 signals together

Note 1 to entry: The rate of the STS-N is N times 51.840 Mbit/s.

Note 2 to entry: SDH does not make the distinction between a logical signal (e.g. STS-N in SONET) and a physical signal (e.g. OC-N in SONET).

Note 3 to entry: The equivalent SDH term for both logical and physical signals is Synchronous Transport Module Level M (STM-M), where M = (N/3). There are equivalent STM-M signals only for values of N = 3, 12, 48, and 192.

3.3.17
Tributary Unit
TU
SDH term for SONET Virtual Tributary (see 3.3.19)

3.3.18
Virtual Container
VC
SDH term for either an STS or VT SPE

3.3.19
Virtual Tributary
VT
structure designed for transport and switching of sub-STS-1 payloads

Note 1 to entry: There are currently four sizes of VT. The equivalent SDH term is Tributary Unit (see 3.3.17).

3.4 FC-BB-3_IP definitions

3.4.1
B_Access
component of the FC Entity (see 3.4.8) that interfaces with the FCIP_LEP (see 3.4.20) component of the FCIP Entity (see 3.4.15) on one side and the B_Port on the other side

3.4.2
B_Access_Name
Name_Identifier (see 3.1.20) of B_Access portal

3.4.3
B_Access Virtual ISL
Virtual ISL (see 3.4.27) that connects two B_Access portals

3.4.4
control and service module
CSM
control component of the FC-BB-3_IP interface that mainly handles connection management

Note 1 to entry: CSM interfaces with the PMM (see 3.4.22).

3.4.5
encapsulated FC frame
SOF/EOF delimited FC frame prefixed with a 28-byte FC frame Encapsulation Header (see RFC 3643)

3.4.6
Encapsulated Frame Receiver Portal
TCP access point through which an encapsulated FC frame (see 3.4.5) is received from the IP network by an FCIP_DE (see 3.4.14)

3.4.7
Encapsulated Frame Transmitter Portal
TCP access point through which an encapsulated FC frame (see 3.4.5) is transmitted to the IP network by the FCIP_DE (see 3.4.14)

3.4.8
FC Entity
principal interface point to the FC switched network on one side and in combination with the FCIP Entity to the IP network on the other side.

Note 1 to entry: It is the data forwarding component of the FC-BB-3_IP interface consisting of VE_Port(s) (see 3.4.24) and/or B_Access (see 3.4.1) portals.

3.4.9
FC Entity Protocol Layer
protocol layer that lies between the Fibre Channel level FC-2 and the FCIP Entity Protocol Layer (see 3.4.16)

Note 1 to entry: Its primary function is to support one or more Virtual E_Ports (see 3.4.24) or B_Access (see 3.4.1) portals and to communicate with the FCIP Entity (see 3.4.8).

3.4.10
FC Receiver Portal
access point through which an FC frame and timestamp enters an FCIP_DE (see 3.4.14) from the VE_Port/ B_Access (see 3.4.24/3.4.1)

3.4.11
FC Transmitter Portal
access point through which an FC frame and timestamp leaves an FCIP_DE (see 3.4.14) to the VE_Port/ B_Access (see 3.4.24/3.4.1)

3.4.12
FC-BB-3_IP device
device defined by the FC-BB-3_IP model

3.4.13
FC-BB-3_IP interface
the point that has interfaces to the FC switched network on one side and the IP network on the other side

Note 1 to entry: FC-BB-3_IP interface consists of a Switching Element, FC/FCIP Entity pair(s), the CSM, and the PMM.

3.4.14
FCIP Data Engine
FCIP_DE
the data forwarding component of the FCIP Entity's (see 3.4.15) FCIP_LEP (see 3.4.20) that handles FC frame encapsulation, de-encapsulation, and transmission of encapsulated frames through a single TCP connection

3.4.15
FCIP Entity
data forwarding component of the FC-BB-3_IP interface consisting of the FCIP_LEP (see 3.4.20) and it is the principal interface point to the IP network on one side and in combination with the FC Entity (see 3.4.8) to the FC switched network on the other side

Note 1 to entry: Its primary function is formatting, encapsulating, and forwarding encapsulated FC frames (see 3.4.5) across the IP network interface.

3.4.16
FCIP Entity Protocol Layer
protocol layer that lies between the FC Entity (see 3.4.8) layer and the TCP layer

3.4.17
FCIP frame
FCIP term for an encapsulated FC frame (see 3.4.5)

3.4.18
FCIP Link
virtual link that connects an FCIP_LEP (see 3.4.20) in one FC-BB-3_IP device (see 3.4.12) with another

Note 1 to entry: It consists of one or more TCP connections.

3.4.19
FCIP Link Originator and Acceptor
the FC-BB-3_IP FCIP_LEP (see 3.4.20) that originates an FCIP Link is defined as the FCIP Link Originator

Note 1 to entry: The corresponding FCIP_LEP that accepts this link is defined as the FCIP Link Acceptor.

3.4.20
FCIP Link Endpoint
FCIP_LEP
component of an FCIP Entity (see 3.4.15) that contains one or more FCIP_DEs (see 3.4.14)

3.4.21
FCIP Transit Time
FTT
total transit time of an encapsulated Fibre Channel frame in the IP network

3.4.22
Platform Management Module
PMM
management component of the FC-BB-3_IP interface that handles time synchronization, discovery, and security

Note 1 to entry: It interfaces with the CSM (see 3.4.4).

3.4.23
Request For Comment
RFC
documents put out by the IETF

3.4.24
Virtual E_Port
VE_Port
data forwarding component of the FC Entity (see 3.4.8) that emulates an E_Port (see 3.1.6)

Note 1 to entry: The term virtual indicates the use of a non Fibre Channel link connecting the VE_Ports. In the case of the FC-BB-3_IP model, a VE_Port interfaces with the FCIP_LEP component (see 3.4.20) of the FCIP Entity (see 3.4.15) on one side and a Fibre Channel Switching Element on the other side.

3.4.25
VE_Port_Name
Name_Identifier (see 3.1.20) of the VE_Port (see 3.4.24)

3.4.26
VE_Port Virtual ISL
Virtual ISL (see 3.4.27) that connects two VE_Ports (see 3.4.24)

3.4.27
Virtual ISL
ISL that connects two VE_Ports (see 3.4.24) or two B_Access portals (see 3.4.1) across a non-FC link

3.5 FC-BB-3_GFPT definitions

3.5.1
ASFC_PAUSE
GFPT_WAN Primitive Signal used to pause flow on a GFPT_WAN link (see 10.3.4)

Note 1 to entry: ASFC_PAUSE is never transmitted to, or expected from, FC_Ports.

3.5.2
ASFC_RESUME
GFPT_WAN Primitive Signal used to resume flow on a GFPT_WAN link (see 10.3.4)

Note 1 to entry: ASFC_RESUME is never transmitted to, or expected from, FC_Ports.

3.5.3
ELP
Exchange Link Parameters SW_ILS (see FC-SW-4)

3.5.4
FLOGI
Fabric Login ELS (see FC-LS)

3.5.5
F_BSY
Fabric Busy (see FC-FS-2)

3.5.6
GFP Server
Generic Framing Procedure (see 3.1.18) adaptation/de-adaptation engine

3.5.7
GFPT
(Asynchronous) Transparent Generic Framing Procedure (see 3.1.18)

3.5.8
GFPT_WAN interface
Transport network-side interface, on an FC-BB-3_GFPT device, corresponding to one GFPT_WAN facility (see 3.5.10), and to one Transport Trail (see 3.5.24)

Note 1 to entry: May or may not correspond to the full SONET/SDH/OTN/PDH access facility/bandwidth.

3.5.9
GFPT_WAN link
Transport Trail (see 3.5.24) assigned to one GFPT_WAN facility (see 3.5.10)

3.5.10
GFPT_WAN facility
Transport Trail (see 3.5.24), GFP Server (see 3.5.6), FC-BB-3_GFPT devices, and their respective GFPT_WAN interfaces (see 3.5.8), corresponding to one interconnected FC_Port pair

3.5.11
inbound
sent from the FC-BB-3_GFPT device to the attached FC_Port

3.5.12
ISL
Inter-Switch Link (see FC-SW-4)

3.5.13
LEM
Login Exchange Monitor (see 10.3.3)

3.5.14
LS_ACC
Link Service Accept frame (see FC-LS)

3.5.15
LS_RJT
Link Service Reject frame (see FC-LS)

3.5.16
outbound
sent from the attached FC_Port to the FC-BB-3_GFPT device

3.5.17
PING
GFPT_WAN Primitive Signal used to initiate latency measurement on a GFPT_WAN link

Note 1 to entry: PING is never transmitted to, or expected from, FC_Ports.

3.5.18
PING_ACK
GFPT_WAN Primitive Signal used to reply to a PING and complete round-trip latency measurement on a GFPT_WAN link

Note 1 to entry: PING_ACK is never transmitted to, or expected from, FC_Ports.

3.5.19
PLOGI
Port Login (see FC-LS)

3.5.20
P_BSY
N_Port Busy (see FC-FS-2)

3.5.21
RPSC ELS
Report Port Speed Capabilities ELS (see FC-LS)

3.5.22
SW_ACC
SW Accept Reply frame (see FC-SW-4)

3.5.23
SW_RJT
SW Reject Reply frame (see FC-SW-4)

3.5.24
Transport Trail
contiguously or virtually-concatenated signal group (see T1.105-2001) comprised of one or more standardized SONET/SDH/OTN/PDH synchronous transport signals

3.5.25
WAN Primitive Signal
ASFC_PAUSE (see 3.5.1), ASFC_RESUME (see 3.5.2), PING (see 3.5.17), or PING_ACK (see 3.5.18) Primitive Signal

Note 1 to entry: These Primitive Signals are always generated and terminated by FC-BB-3_GFPT devices and transmitted only between FC-BB-3_GFPT devices.

Note 2 to entry: They are never transmitted to nor received from FC_Ports.

3.5.26
WAN_HOLDOFF_TOV
time-out value, specific to FC-BB-3_GFPT devices, which defines the period that elapses, following detection/indication of a GFPT_WAN link failure, before a GFPT_WAN Down condition is declared for the purposes of the state machine described in 10.3.2

Note 1 to entry: The criteria for such detection are WAN-specific and outside the scope of this standard.

3.6 Editorial conventions

In FC-BB-3, a number of conditions, mechanisms, sequences, parameters, events, states, or similar terms are printed with the first letter of each word in uppercase and the rest lowercase (e.g., Exchange, Sequence). Any lowercase uses of these words have the normal technical English meaning.

Lists sequenced by letters (e.g., a-red, b-blue, c-green) show no ordering relationship between the listed items. Numbered lists (e.g., 1-red, 2-blue, 3-green) show an ordering relationship between the listed items.

The ISO convention of numbering is used (i.e., the thousands and higher multiples are separated by a space and a comma is used as the decimal point). A comparison of the American and ISO conventions are shown in table 2.

Table 2 – ISO and American Conventions

ISO	American
0,6	0.6
1 000	1,000
1 323 462,9	1,323,462.9

In case of any conflict between figure, table, and text, the text, then tables, and finally figures take precedence. Exceptions to this convention are indicated in the appropriate sections.

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

When the value of the bit or field is not relevant, x or xx appears in place of a specific value. If a field or a control bit in a frame is specified as not meaningful, the entity that receives the frame shall not check that field or control bit.

Numbers that are not immediately followed by lower-case b or h are decimal values.

Numbers immediately followed by lower-case b (xxb) are binary values.

Numbers or upper case letters immediately followed by lower-case h (xxh) are hexadecimal values.

3.7 List of commonly used acronyms and abbreviations

Abbreviations and acronyms applicable to this standard are listed. Definitions of several of these items are included in clause 3.

3.7.1 General

BB Backbone
BB-2 Backbone -2
BBW Backbone (ATM or SONET) WAN
BSW Border Switch
EBP Exchange B_Access Parameters
ELP Exchange Link Parameters
EOF End of Frame
ESC Exchange Switch Capabilities
FCIP Fibre Channel Over TCP/IP
FCS Frame Check Sequence
FC-SP Fibre Channel - Security Protocol
FC-SW-3 Fibre Channel - Switched Fabric - 3
ISL Inter-switch Link
ITU-T International Telecomm. Union - Telecommunication Standardization Section
K_A_TOV Keep Alive Timeout value
LKA Link Keep Alive
MTU Maximum Transfer Unit
PDU Protocol Data Unit
SFC Simple Flow Control
SOF Start of Frame
SR Selective Retransmission
SW_ACC Switch Fabric Internal Link Service Accept
SW_CS Switch Fabric Common Services
SW_ILS Switch Fabric Internal Link Services
SW_RJT Switch Fabric Internal Link Service Reject
WAN Wide Area Network

3.7.2 FC-BB-3_ATM

AAL5 ATM Adaptation Layer 5
ATM Asynchronous Transfer Mode
BER Bit Error Rate
CLR Cell Loss Ratio (ATM)
CPCS Common Part Convergence Sublayer
PVC Permanent Virtual Circuit, Permanent Virtual Connection
QoS Quality of Service
SAAL Signaling ATM Adaptation Layer
SVC Switched Virtual Circuit, Switched Virtual Connection
UBR Unspecified Bit Rate (ATM)
UNI User Network Interface (ATM)
VBR-NRT Variable Bit Rate - Non Real Time (ATM)
VC Virtual Circuit

3.7.3 FC-BB-3_SONET

HDLC High-level Data Link Control
nm Nanometer
OC-N Optical Carrier Level ***N***
ppm Parts per Million
PPP Point-to-Point Protocol
PTE Path Terminating Equipment
RFC Request for Comment
SDH Synchronous Digital Hierarchy
SMT Station Management (FDDI)
SONET Synchronous Optical Network
SPE Synchronous Payload Envelope
STM-M Synchronous Transport Module level ***M***
STS Synchronous Transport Signal
STS-N Synchronous Transport Module level ***N***
STS-Nc Synchronous Transport Module level ***Nc***
TU Tributary Unit
ULA 48-bit Universal LAN MAC Address
ULP Upper Level Protocol
ULP_TOV Upper_Level_Protocol_Timeout value
VC Virtual Container
VP Virtual Path
VT Virtual Tributary

3.7.4 FC-BB-3_IP

B_Access B_Access Portals
CSM Control and Service Module
FCIP FC over TCP/IP
FCIP_DE FCIP Data Engine
FCIP_LEP FCIP Link Endpoint
IETF IETF Internet Engineering Task Force (www.ietf.org)
PMM Platform Management Module
RFC Request For Comment
VE_Port Virtual E_Port

3.7.5 FC-BB-3_GFPT

ASFC Alternate Simple Flow Control

ELP Exchange Link Parameters
F_BSY Fabric Busy
FLOGI Fabric Login
GFP Generic Framing Procedure
GFPT (Asynchronous) Transparent Generic Framing Procedure
GFPT_WAN GFPT Wide Area Network
ISL Inter-Switch Link
LEM Login Exchange Monitor
LS_ACC Link Service Accept Reply Frame
LS_RJT Link Service Reject Reply Frame
P_BSY N_Port Busy
PLOGI Port Login
SW_ACC SW Accept Reply Frame
SW_RJT SW Reject Reply Frame

3.8 Symbols

Unless indicated otherwise, the following symbol has the listed meaning.
|| concatenation

3.9 Keywords

3.9.1
expected
keyword used to describe the behavior of the hardware or software in the design models assumed by this standard. Other hardware and software design models may also be implemented

3.9.2
ignored
keyword used to describe an unused bit, byte, word, field or code value. The content or value of an ignored bit, byte, word, field or code value shall not be examined by the receiving device and may be set to any value by the transmitting device

3.9.3
invalid
keyword used to describe an illegal or unsupported bit, byte, word, field or code value. Receipt of an invalid bit, byte, word, field or code value shall be reported as an error

3.9.4
mandatory
keyword indicating an item that is required to be implemented as defined in this standard

3.9.5
may
keyword that indicates flexibility of choice with no implied preference (equivalent to “may or may not”)

3.9.6
may not
keyword that indicates flexibility of choice with no implied preference (equivalent to “may or may not”)

3.9.7
optional
keyword that describes features that are not required to be implemented by this standard. However, if any optional feature defined by this standards is implemented, then it shall be implemented as defined in this standard

3.9.8
reserved
keyword referring to bits, bytes, words, fields and code values that are set aside for future standardization. A reserved bit, byte, word or field shall be set to zero, or in accordance with a future extension to this standard.

Recipients are not required to check reserved bits, bytes, words or fields for zero values. Receipt of reserved code values in defined fields shall be reported as an error

3.9.9
shall
keyword indicating a mandatory requirement. Designers are required to implement all such mandatory requirements to ensure interoperability with other products that conform to this standard

3.9.10
should
keyword indicating flexibility of choice with a strongly preferred alternative; equivalent to the phrase “it is strongly recommended”

3.9.11
x or **xx**
the value of the bit or field is not relevant

4 FC-BB-3 structure and concepts

4.1 FC-BB-3 backbone mappings

FC-BB-3 models (i.e., FC-BB-3_ATM, FC-BB-3_SONET, FC-BB-3_IP, and FC-BB-3_GFPT), specified in this standard, define equipment capable of extending Fibre Channel switched networks and/or links over Wide Area Network infrastructures and over distance (i.e., over non-negligible link latencies). One important distinction among the models discussed in this standard is the emphasis placed on the backbone (i.e., WAN) type. The FC-BB-3_ATM and FC-BB-3_SONET models assume ATM and SONET transport network technologies. The FC-BB-3_IP models assume the use of TCP connections over IP networks. The FC-BB-3_GFPT model makes use of the Asynchronous Transparent Generic Framing Procedure (GFPT) (see ITU-T Rec. G.7041/ Y.1303). GFPT may be used for adaptation to different transport facilities including SONET, SDH, OTN and PDH. Details regarding the mapping of GFPT-adapted traffic to such transport facilities are elaborated in various ITU-T standards (see clause 2).

A second important distinction among the mappings discussed in this standard relates to supported architectures (i.e., network and/or link topologies) and the place of the defined devices within them. FC-BB-3_ATM and FC-BB-3_SONET define a fabric bridge device that interconnects the E_Ports of external FC switches. The switch-facing interfaces resident on these bridges are called B_Ports. B_Ports have selected fabric functions (see FC-SW-4). B_Ports are fabric ports, and FC-BB-3_ATM and FC-BB-3_SONET devices are components of an FC fabric. FC-BB-3_IP also defines support for FC bridge devices with fabric-facing B_Ports. However, FC-BB-3_IP also supports functional integration within an FC Switch. Thus FC-BB-3_IP devices may also have E_Ports and F_Ports. FC-BB-3_GFPT defines a device that is not a component of a fabric, and supports no fabric functionality. Instead, it interconnects two Fibre Channel physical ports (i.e., attached FC_Ports), appearing architecturally as a wire to those ports.

NOTE 1 - Future extension of FC-BB-3_ATM and FC-BB-3_SONET models to include support for resident E_Ports and F_Ports is not precluded.

4.2 FC-BB-3 reference models

FC-BB-3 defines four reference models, corresponding to the FC-BB-3_ATM, FC-BB-3_SONET, FC-BB-3_IP, and FC-BB-3_GFPT models. These reference models are shown in figure 4, figure 5, figure 6, and figure 7, respectively. In figure 4 and figure 5 (i.e., FC-BB-3_ATM and FC-BB-3_SONET, respectively), frames destined for a remote FC network enter a B_Port and are forwarded on the backbone network to their destination. In figure 6 (i.e., FC-BB-3_IP), frames destined for a remote FC network enter a B_Port, an E_Port, or an F_Port, and are forwarded on the backbone network to their destination. In figure 7 (i.e., FC-BB-3_GFPT), FC physical signals (i.e., relevant 8B/10B codewords) enter an FC physical port on an FC-BB-3_GFPT device, and are forwarded on the backbone network to their destination.

The FC-BB-3_ATM and FC-BB-3_SONET models support the attachment of FC switches (i.e., E_Ports) via one or more B_Ports. The FC-BB-3_IP model supports the attachment of FC switches (i.e., E_Ports) via one or more B_Ports or E_Ports and the attachment of N_Ports via one or more F_Ports. The FC-BB-3_GFPT model supports the attachment of N_Ports, F_Ports, and E_Ports, and the following Fibre Channel port interconnections:

a) N_Port to N_Port;

b) N_Port to F_Port; and

c) E_Port to E_Port.

Table 3 summarizes the resident FC_Port types for the four different FC-BB-3 models.

Table 3 – Models and resident FC_Port types

	Reference Model			
	FC-BB-3_ATM	**FC-BB-3_SONET**	**FC-BB-3_IP**	**FC-BB-3_GFPT**
Resident FC_Port Type(s)	B_Port	B_Port	B_Port, E_Port, F_Port	None (FC Physical Interface)

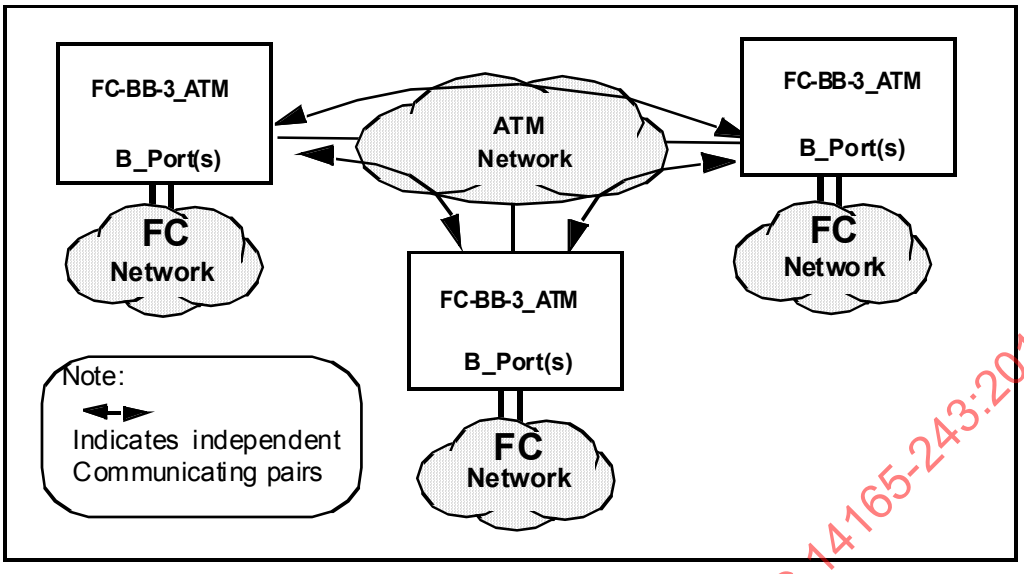


Figure 4 – FC-BB-3_ATM reference model

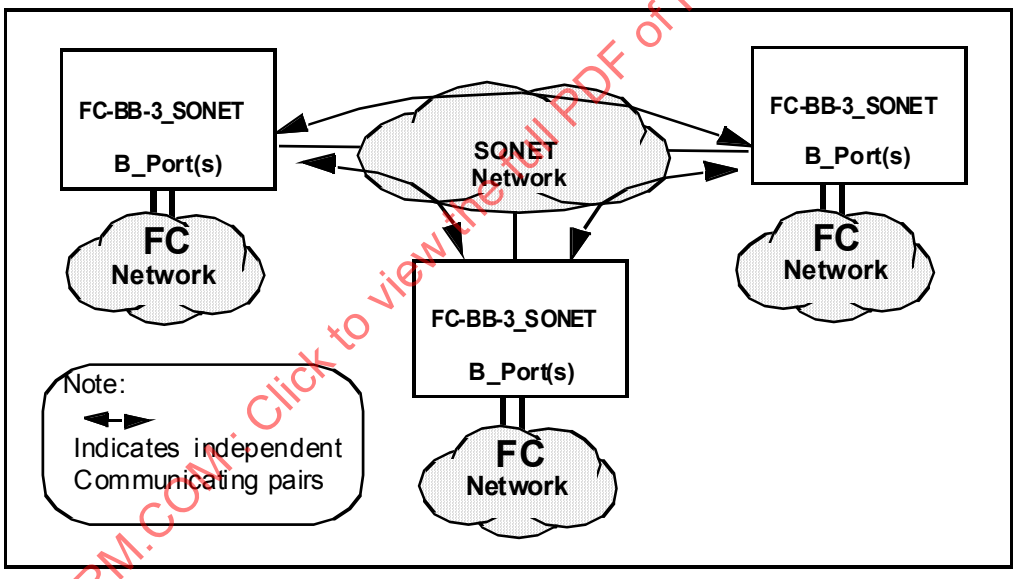


Figure 5 – FC-BB-3_SONET reference model

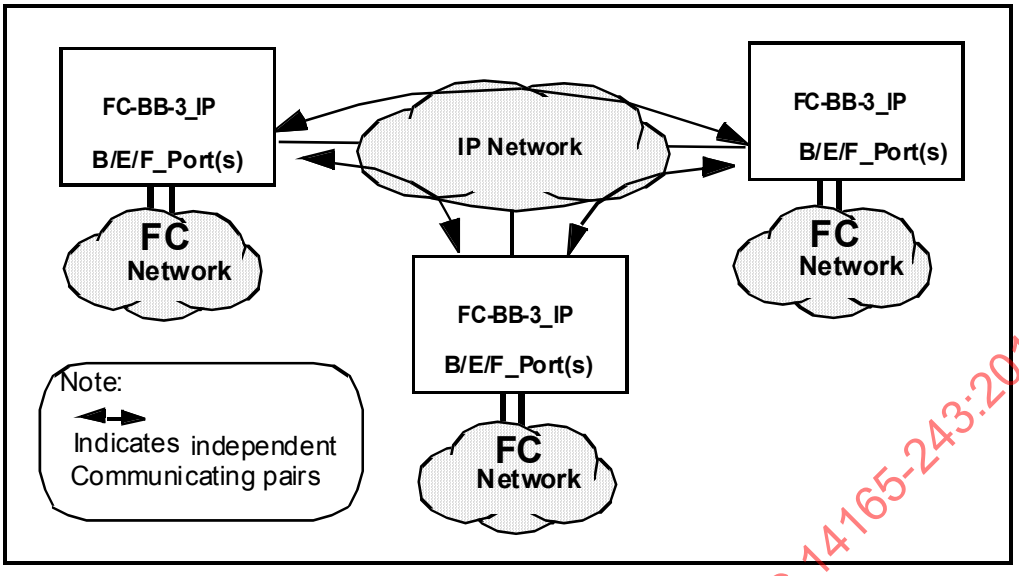


Figure 6 – FC-BB-3_IP reference model

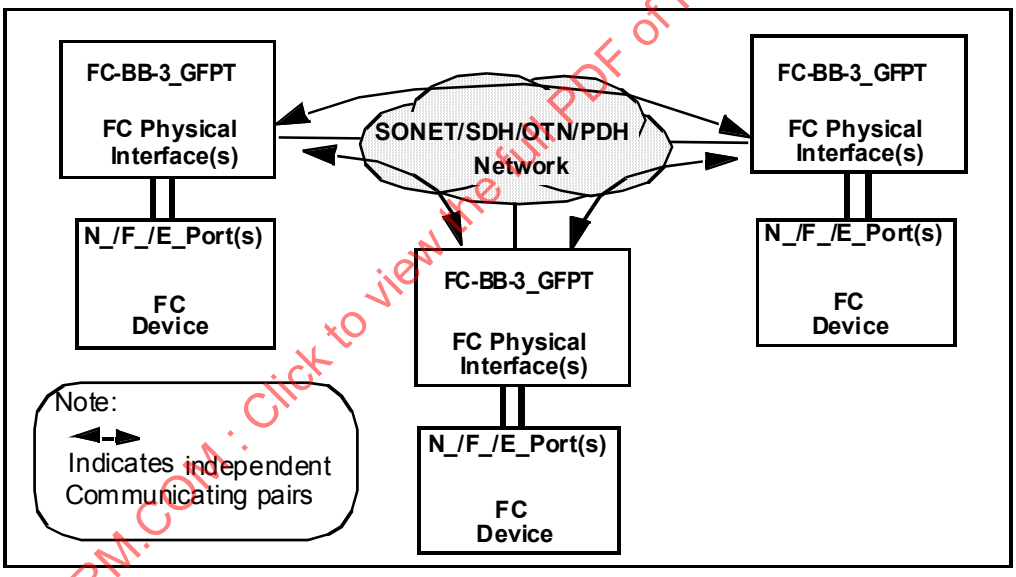


Figure 7 – FC-BB-3_GFPT reference model

4.3 FC-BB-3 models overview

4.3.1 FC-BB-3_ATM

The FC-BB-3_ATM model defines the means by which Fibre Channel networks interface with and connect across a wide-area ATM network. FC-BB-3_ATM defines the frame mapping, encapsulation, and any signaling required by the ATM protocols. FC-BB-3_ATM also defines the frame handling, call handling, addressing, flow control protocol, and error recovery required to support the Fibre Channel mapping over ATM. FC-BB-3_ATM makes use of the ATM Adaptation Layer 5 for payload transport.

FC-BB-3_ATM messages are formed by encapsulating byte-encoded Class 2, 3, 4, or F Fibre Channel frames into a suitable format for carriage over the WAN. Clause 5 describes the FC-BB-3_ATM message in detail.

SR and SFC are two flow control protocols that may be used over ATM networks. The SR protocol provides a reliable transport of frames between two FC-BB-3_ATM devices. Use of the SR protocol is optional. The SR protocol is an efficient sliding window link-layer full-duplex protocol that supports data transport with flow control

and error recovery functions. SR has been adopted from ITU's Link Access Protocol B (LAPB), that is derived from ISO/IEC's High-level Data Link Control (HDLC) (Balanced Classes). Use of LAPB in SR is limited to a subset of the synchronous modulo 32768 super sequence numbering service option. Clause 6 describes the SR protocol in detail. The SFC protocol (see 6.4) provides a mechanism to temporarily pause the transmission of frames from a remote FC-BB-3_ATM device. Use of the SFC protocol is optional.

4.3.2 FC-BB-3_SONET

The FC-BB-3_SONET model defines the means by which Fibre Channel networks interface with and connect across a wide-area SONET/SDH network. FC-BB-3_SONET defines the frame mapping, encapsulation, and any signaling required by the SONET protocols. FC-BB-3_SONET also defines the frame handling, call handling, addressing, flow control protocol, and error recovery required to support the Fibre Channel mapping over SONET/SDH. FC-BB-3_SONET makes use of either the High-level Data Link Control (HDLC) for payload transport (described in detail in this standard), or the relevant Frame Mapped Generic Framing Procedure (GFPF) mapping described in ITU-T Rec. G.7041/Y.1303.

NOTE 2 - The FC over SONET/GFP mappings defined by FC-BB-3_SONET and FC-BB-3_GFPT, are fundamentally different from the native FC transport supported by the Synchronous, Full-Rate Transparent Mapped Generic Framing Procedure defined in ITU-T Rec. G.7041/Y.1303. The latter mapping constitutes a full-rate, fully transparent FC wire extension between two FC_Ports and is not discussed in this standard.

The SR and SFC flow control protocols may be used over SONET networks.

4.3.3 FC-BB-3_IP

The FC-BB-3_IP model defines the means by which Fibre Channel networks interface with and connect across an IP network. FC-BB-3_IP makes use of the FCIP standard (see RFC 3821) to define the mapping and control required by the TCP/IP protocol and the FC frame encapsulation standard (see RFC 3643) to define the encapsulation. FC-BB-3_IP also defines the connection management, addressing, time synchronization, discovery, security, switching, routing, and error recovery required to support Fibre Channel over TCP/IP. FC-BB-3_IP is agnostic about the underlying physical technology that exists beneath the IP layer. In this sense, the IP network could use ATM, SONET, Gigabit Ethernet, or any other link-level technology below it.

FC-BB-3_IP encapsulates byte-encoded Class 2, 3, 4, or F Fibre Channel frames into a suitable format (i.e., encapsulated FC frames) for carriage over the IP network. Subclause 9.6 describes encapsulated FC frames in detail. The TCP/IP protocol suite provides a reliable transport of frames over the IP network. TCP provides flow control and error recovery.

The FC-BB-3_IP protocol provides mechanisms to create VE_Port or B_Access connectivity as the case may be, over the IP network (see 9).

4.3.4 FC-BB-3_GFPT

The FC-BB-3_GFPT model defines the means by which FC physical links may be extended over any WAN transport infrastructure for which GFP mapping is defined. FC-BB-3_GFPT supports the interconnection of arbitrary, legal, non-Arbitrated Loop FC_Port combinations, imposing no requirements, and making no suppositions, regarding the topology, or even the presence of an FC fabric. FC-BB-3_GFPT supports efficient transport of FC data over transport facilities of arbitrary bandwidths and potentially large distances. FC-BB-3_GFPT supports Class 2, 3, and F traffic.

FC-BB-3_GFPT devices do not generate FC frames and do not directly participate in port initialization or other Exchanges. FC-BB-3_GFPT devices are exempt from any requirements regarding FC_Port authentication (see FC-SP), and they do not impede or interfere with any such processes that may occur between the attached and interconnected FC_Ports. FC-BB-3_GFPT devices have no FC identity or visibility, and administratively they may be kept strictly separated and distinct from FC fabrics and ports.

FC_Ports are interconnected pair-wise over SONET/SDH/OTN/PDH networks, via FC-BB-3_GFPT devices, in a point-to-point fashion. Although multiple FC_Ports may interface with a single FC-BB-3_GFPT device, each opposing FC_Port pair is connected via a dedicated Transport Trail (e.g., a contiguously or virtually-concatenated group). Since trail and access section configurations may differ, FC-BB-3_GFPT devices have both physical interfaces to the transport network, and individual FC-BB-3_GFPT devices may have more than one

such physical interface, as well as logical interfaces associated with individual circuits. Logical interfaces are referred to as GFPT_WAN interfaces. A GFPT_WAN interface corresponds to a specific Transport Trail, and always to a single attached FC_Port pair. Governance of the relationship of GFPT_WAN interfaces to physical SONET/SDH/OTN/PDH interfaces, and of any changes of such relationships (e.g., as may occur during network protection events), is specified in the appropriate ITU-T and ANSI-T1 standards (see clause 2), and is therefore outside the scope of this standard. Multiple GFPT_WAN links originating on one FC-BB-3_GFPT device may be terminated on different, and geographically disparate, FC-BB-3_GFPT devices. The routing and provisioning of network facilities underlying GFPT_WAN links is outside the scope of this standard.

The FC-BB-3_GFPT device interfaces to attached FC_Ports are FC physical interfaces operating at standard rates. The FC physical interfaces on FC-BB-3_GFPT devices may support link speed negotiation with the attached FC_Ports.

4.4 FC-BB-3 requirements

4.4.1 Fibre Channel Class support

Class F shall be supported between FC-BB-3 devices. Class 1 is not supported, and Class 2, 3 or 4 may be supported between FC-BB-3 devices.

4.4.2 Payload transparency

4.4.2.1 FC-BB-3_ATM, FC-BB-3_SONET, FC-BB-3_IP

Arriving Class 2, 3, 4, and F Fibre Channel frames from an FC network and destined to a remote FC network shall be encapsulated using the FC-BB-3_ATM/SONET/IP defined mechanisms and transmitted to the appropriate FC-BB-3_ATM/SONET/IP device.

Arriving encapsulated frames received from remote FC-BB-3_ATM/SONET/IP device shall be de-encapsulated and sent to an FC network.

Primitive Signals and Primitive Sequences shall not be transported between FC-BB-3_ATM/SONET/IP devices.

4.4.2.2 FC-BB-3_GFPT

FC frames inbound from one attached FC_Port shall be delivered to the remote FC_Port in native form (i.e., without further encapsulation) across the transport network according to the adaptation processes described in 10.3.8. Frames received from remote FC-BB-3_GFPT devices shall be forwarded to the attached FC_Port. Selected ELP, SW_ACC, FLOGI, PLOGI, and LS_ACC frames may be subject to inspection and/or minor modifications, in transiting one or the other FC-BB-3_GFPT device, as described in 10.3.4.

Primitive Signals transmitted by an attached FC_Port may be forwarded across the transport network for delivery to the remote FC_Port according to the rules described in 10.3.5. When they are forwarded, they are forwarded in native form (i.e., without further encapsulation) according to the adaptation processes described in 10.3.8.

Primitive Sequences transmitted by an attached FC_Port are forwarded across the transport network for delivery to the remote FC_Port according to the rules described in clause 10. Primitive Sequences are forwarded in native form (i.e., without further encapsulation) according to the adaptation (including rate adaptation) processes described in 10.3.8.

4.4.3 Latency delay and timeout value

FC-BB-3_IP shall ensure that the incoming encapsulated FC frames whose FCIP Transit Time (FTT) exceeds 1/2 E_D_TOV shall be discarded and not admitted into the FC network. Fibre Channel timeout values shall be administratively set to accommodate the FTT.

FC-BB-3_IP shall allow Class F encapsulated FC frames to be transmitted with a zero timestamp value.

FC-BB-3_GFPT requires that the inherent latency between two FC-BB-3_GFPT devices be bounded and included within one-half of the E_D_TOV and R_T_TOV values of the attached devices, and within the R_A_TOV values of the attached fabric(s), if any.

4.4.4 QoS and bandwidth

FC-BB-3_ATM shall use the VBR-NRT ATM service. VBR-NRT ATM service provides cell loss and bandwidth guarantees. It is recommended that FC-BB-3_ATM make use of a single Virtual Circuit (VC). Use of additional VCs to address special traffic QoS requirements is allowed but not recommended. FC-BB-3_ATM recommends allocating a minimum bandwidth for each FC-BB-3_ATM VC that is used in order to avoid starvation; however, the service discipline prioritization for the VCs is implementation-specific and beyond the scope of this standard.

FC-BB-3_SONET has no specific SONET service requirements.

FC-BB-3_IP recommends that some form of preferential QoS be used for the FCIP traffic in the IP network to minimize latency and packet drops although no particular form of QoS is recommended. See RFC 3821.

FC-BB-3_GFPT has no specific transport service requirements.

4.4.5 In-order delivery

FC-BB-3_ATM shall guarantee in-order delivery of frames within each ATM VC. No other ordering relationship among ATM VCs need be preserved.

FC-BB-3_SONET shall guarantee in-order delivery of frames within each SONET/SDH provisioned path. No other ordering relationship among SONET/PDH provisioned paths need be preserved.

FC-BB-3_IP shall guarantee in-order delivery of frames within the scope of any TCP connection.

FC-BB-3_GFPT shall provide in-order delivery within each provisioned Transport Trail for all transmitted data (i.e., frames, Primitive Signals and Primitive Sequences), as discussed and with the exceptions detailed in clause 10.

4.4.6 Flow control

FC-BB-3_ATM or FC-BB-3_SONET devices may use the Selective Retransmission (SR) protocol to provide for reliable delivery of frames over the WAN between two devices. In the case of FC-BB-3_ATM, if the SR protocol is used, then the flow control is separately applied to each ATM VC.

FC-BB-3_ATM or FC-BB-3_SONET devices may use the Simple Flow Control (SFC) protocol to temporarily pause the transmission of frames from a remote device. In the case of FC-BB-3_ATM, if the SFC protocol is used, then the flow control is separately applied to each ATM VC.

FC-BB-3_IP devices shall ensure that TCP flow control and error recovery acts in proper concert with the Fibre Channel BB_Credit flow control mechanism.

Flow control at E_Ports, F_Ports, VE_Ports, and B_Ports shall operate as defined in FC-SW-4.

The Alternate Simple Flow Control (ASFC) mechanism (see 10.3.4) shall be used between FC-BB-3_GFPT devices. Flow control on FC_Port-facing links is specified in clause 10.

Flow control at FC-BB-3_GFPT physical interfaces to attached FC devices shall operate as defined in FC-SW-4 or FC-FS-2, as appropriate.

4.5 FC-BB-3 SW_ILS codes

Table 4 shows the SW_ILS codes allocated for FC-BB-3 use.

Table 4 – FC-BB-3 SW_ILS codes

Encoded Value (hex)	Description	Abbr.	Reference
28 03 00 00	Authentication Special Frame Request	ASF	9.7.2.3.2
28 01 00 00	Exchange B_Access Parameter	EBP	7.3.5

Table 5 shows the ELS codes allocated for FC-BB-3 use.

Table 5 – FC-BB-3 ELS codes

Encoded Value (hex)	Description	Abbr.	Reference
80 00 00 00	Link Keep Alive Request	LKA	FC-LS

5 FC-BB-3_ATM and FC-BB-3_SONET Messages and Formats

5.1 General

In all text to follow, the term BBW applies to both FC-BB-3_ATM and FC-BB-3_SONET. The structure of a BBW message is given in table 6. It consists of 3 fields: LLC/SNAP Header, BBW_Header, and the BBW message payload. The structures of the fields are given in table 7, table 8, table 10, and table 11.

Table 6 – BBW message structure

	Field	Size (Bytes)
BBW message	LLC/SNAP Header	8
	BBW_Header	4
	BBW message payload	Max: 2148

5.2 LLC/SNAP header format

LLC/SNAP Header (8 bytes): The Logical Link Control (LLC)/Sub Network Access Protocol (SNAP) Header consists of a 3-byte LLC field and a 5-byte SNAP sub-field.

Table 7 – LLC/SNAP header

Field		Word	Byte	Encoded Value (hex)
LLC	DSAP	0	0	AA
	SSAP		1	AA
	CTRL		2	03
SNAP	OUI	0	3	00
		1	0	00
		1	1	00
	PID	1	2	88
		1	3	8D

LLC (3 bytes): The LLC field consists of three 1-byte sub-fields: Destination Service Access Point (DSAP), Source Service Access Point (SSAP), and Control (CTRL). The encoding for LLC given in table 7 indicates that an IEEE 802.2 SNAP follows.

SNAP (5 bytes): The SNAP field consists of two sub-fields: A 3-byte Organizationally Unique Identifier (OUI) sub-field and a 2-byte Protocol Identification (PID) sub-field. The encoding for OUI given in table 7 indicates the presence of an IEEE 802.2 Routed protocol in the payload. The encoding for PID given in table 7 indicates the payload protocol type is Fibre Channel.

5.3 BBW_Header format

The 4-byte BBW_Header (see table 8) consists of three fields: A 1-byte Flow Control Type field, a 2-byte PAUSE field and a 1-bit Address Bit field. The structure of the Flow Control Type field is given in table 9.

Table 8 – BBW_Header

Word	Byte	Field	Size (Bytes)	Remarks
2	0	Flow Control Type	1	
	1-2	PAUSE	2	Applicable only when SFC protocol is specified as the Flow Control Type.
	3	<bit 0>: Address Bit 1= Command; 0 = Response <bits 1-7>: **Reserved**	1	Applicable only when SR protocol is specified as the Flow Control Type.

Flow Control Type (1 byte): This field defines encodings for Simple Flow Control (SFC) and SR flow control.

Table 9 – Flow control protocol type encodings

Encodings (hex)	Flow Control Type
00	SFC
01	SR
Others	Reserved

PAUSE (2 bytes): The 2-byte PAUSE field is applicable only when the Flow Control Type is SFC. The PAUSE field defines the number of 512-bit time units to pause transmission. A value of zero indicates zero pause transmission time units. This field is also set to zero value when no flow control is desired with the Flow Control Type specified as SFC.

Address Bit (bit 0: Byte 3): This field is applicable only when the Flow Control Type is SR flow control. This bit identifies the SR message as either a command or a response. Messages containing commands shall set this bit to 1. Messages containing responses shall set this bit to 0. This field is used in conjunction with the Poll Bit of the SR protocol.

5.4 BBW message payload format for SFC

The general structure of the BBW message payload when SFC is specified as the Flow Control Type is given in table 10. It consists of the following fields: 4-byte SOF, 24-byte FC-Header, FC frame payload, 4-byte CRC, and 4-byte EOF. The SOF and EOF byte encodings are defined in annex A. The FC Header, FC frame payload, and the CRC are standard Fibre Channel frame fields that arrive at the B_Port or the E_Port interface.

NOTE 3 - The format also applies when the Flow Control Type is specified as SFC and the PAUSE field carries a zero value (i.e., when no flow control is desired).

Table 10 – BBW message payload structure for SFC

Field	Size (Bytes)
SOF	4
FC Header	24
FC frame payload (includes optional header)	Min: 0 Max:2 112
CRC	4
EOF	4

5.5 BBW message payload format for SR

5.5.1 General

The general structure of the BBW message payload when SR is specified as the Flow Control Type is given in table 11. It consists of a 4-byte SR_Header and the SR_BBW message payload. SR_Header encodings specify the SR_BBW message type. SR_Header is described in 5.5.2.

Table 11 – BBW message payload structure for SR

Field	Field	Size (Bytes)
SR_BBW message	SR_Header	4
	SR_BBW message payload	Max: 2 148

The SR_BBW message payload format depends on the type of the SR_BBW message and is described in 5.5.5.1, 5.5.5.4, and 5.5.5.9.

5.5.2 SR_Header formats

5.5.2.1 Field formats

The SR_Header defines three types of field formats that are used to perform numbered information transfer (i.e., I-format), numbered supervisory functions (i.e., S-format), and unnumbered control functions (i.e., U-format) (see table 12). SR makes use of nine different types of messages:

a) I-format (1): SR_I;

b) S-format (3): SR_RR, SR_RNR, SR_REJ; and

c) U-format (5): SR_SM, SR_DISC, SR_FRMR, SR_UA, SR_DM.

Bits 0 and 1 of the SR_Header provide the encodings for the three field formats. Bits 2 and 3 provide the encodings for the different S-format messages and are shown in table 13. Bits 2, 3, 5, 6, and 7 provide the encodings for the different U-format messages and are shown in table 14.

Table 12 – SR_Header format

Format	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
I	N(R)															P	N(S)															0
S	N(R)															P/F	Reserved												S	S	0	1
U	Reserved																Reserved							M	M	M	P/F	M	M	1	1

N(S): Transmitter send sequence number; N(R): Transmitter receive sequence number.
SS: Supervisory function bits defined in table 13.
MMMMM: Modifier function bits defined in table 14.
P/F: Poll/Final bit; P Poll bit (1 = Poll).

Table 13 – SS bits encoding

SS Bits		Supervisory message
3	2	
0	0	SR_RR
0	1	Reserved
1	0	SR_RNR
1	1	SR_SREJ

Table 14 – MMMMM bit encoding

MMMMM Bits					Unnumbered message
7	6	5	3	2	
1	1	0	0	0	SR_SM
0	1	0	0	0	SR_DISC
1	0	0	0	1	SR_FRMR
0	1	1	0	0	SR_UA
0	0	0	1	1	SR_DM

5.5.2.2 Information transfer I-format

The I-format is used to perform an information transfer. The functions of the N(S), N(R), and P fields are independent (i.e., each SR_I message has a N(S), a N(R), that may or may not acknowledge additional SR_I messages received by the BBWs, and a P bit that may be set to a 0 or 1).

5.5.2.3 Supervisory S-format

The S-format is used to perform data link supervisory control functions such as acknowledge SR_I messages, request retransmission of SR_I messages, and to request a temporary suspension of transmission of SR_I messages. The functions of the N(R) and P/F fields are independent (i.e., each supervisory message has a N(R) that may or may not acknowledge additional SR_I messages received by the BBW, and a P/F bit that may be set to 0 or 1).

5.5.2.4 Unnumbered U-format

The U-format is used to provide additional data link control functions. This format contains no sequence numbers, but does include a P/F bit that may be set to a 0 or a 1.

5.5.3 SR_BBW messages

A description of the nine different SR_BBW messages appears in table 15. Only the SR_I, SR_SREJ, and SR_FRMR messages carry a payload, all other messages carry a null payload.

.

Table 15 – SR_BBW messages

Purpose	Message	Command/ Response[a]	Description
Information transfer	SR_I	Command	Carries encapsulated Class 2, 3, 4, or F frames as payload.
Control (Supervisory messages)	SR_RR	Command or Response	Indicates Ready to Receive SR_I messages (negates busy condition) and acknowledges previous SR_I messages; carries no payload.
	SR_RNR	Command or Response	Indicates Receiver Not Ready to accept more SR_I messages (busy condition) and acknowledges previous SR_I messages; carries no payload.
	SR_SREJ	Command or Response	Indicates Selective Retransmission of errored SR_I messages; carries a payload.
Control (Unnumbered messages)	SR_SM	Command	Mode setting command to set up link and resets all messages counters to 0; carries no payload.
	SR_UA	Response	Unnumbered response to the SR_SM command and indicates an acceptance and information transfer phase; carries no payload
	SR_DM	Response	Unnumbered response to the SR_SM command and indicates a disconnected phase; carries no payload
	SR_FRMR	Response	Unnumbered response to the SR_SM command and indicates message reject for the SR_SM message; carries a payload
	SR_DISC	Command	Command indicates the sender is suspending operation and enters the disconnected mode after receiving an SR_UA response; carries no payload
a) Command/Response indicated by the Address Bit in byte 3 of the BBW_Header.			

5.5.4 Format field parameters

5.5.4.1 General

The following subclauses describe the different format fields and other related aspects of the SR protocol.

5.5.4.2 Modulus of SR

Each SR_I message is sequentially numbered and may have the value 0 through modulus minus 1, where “modulus” is equal to 32 768 (i.e., the maximum value of the sequence numbers). The sequence numbers cycle through the entire range.

5.5.4.3 Send state variable V(S)

The send state variable V(S) denotes the sequence number of the next-in-sequence SR_I message to be transmitted. V(S) may take on the values 0 through modulus minus 1. The value of V(S) is incremented by 1 with each successive SR_I message transmission, but cannot exceed the N(R) of the last received SR_I or supervisory message by more than the maximum number of outstanding SR_I messages ***k***. The value of ***k*** is defined in 6.3.8.4.

5.5.4.4 Send sequence number N(S)

Only SR_I messages contain N(S), the send sequence number of the transmitted SR_I message. At the time that an in-sequence SR_I message is designated for transmission, the value of N(S) is set to the value of the send state variable V(S).

5.5.4.5 Receive state variable V(R)

The receive state variable V(R) denotes the sequence number of the next-in-sequence SR_I message expected to be received. V(R) may take on the values 0 through modulus minus 1. The value of V(R) is incremented by 1 by the receipt of an error-free, in-sequence SR_I message whose send sequence number N(S) equals the receive state variable V(R).

5.5.4.6 Receive sequence number N(R)

All SR_I messages and supervisory messages, except SR_SREJ messages with the F bit set to 0, shall contain N(R), the expected send sequence number of the next received SR_I message. At the time that a message of the above types is designated for transmission, the value of N(R) is set to the current value of the receive state variable V(R). N(R) indicates that the BBW transmitting the N(R) has correctly received all SR_I messages numbered up to and including N(R)-1.

5.5.4.7 Functions of the Poll/Final bit (P/F)

All messages contain P/F, the Poll/Final bit. In command messages, the P/F bit is referred to as the P bit. In response messages it is referred to as the F bit.

The Poll bit set to 1 is used by the BBW to solicit (i.e., poll) a response from the remote BBW.

The Final bit set to 1 is used by the BBW to indicate the response message transmitted by the remote BBW, as a result of the soliciting (i.e., poll) command.

The use of the P/F bit is described in 6.3.3.

5.5.5 SR commands and responses

5.5.5.1 Information (SR_I) command

The function of the information (SR_I) command is to transfer a sequentially numbered message containing an information field across a data link.

The SR_I message command carries the mapped byte-encoded Class 2, 3, 4, or F frames. The following steps are involved in generating the SR_I message:

a) constructing the SR_I message payload by prefixing the proper 32-bit SOF delimiter to the incoming FC-Header, FC frame payload, and the CRC; and appending the corresponding 32-bit EOF delimiter to the CRC; and

NOTE 4 - The original Fibre Channel frame CRC field remains and the sender does not have to send a valid CRC and the receiver does not have to validate the CRC.

b) constructing the 4-byte SR_Header and prefixing it to the SR_I message payload.

Table 16 illustrates the format of the SR_I message information field (i.e., payload). The maximum size of the BBW message is 2152 bytes corresponding to a maximum size FC frame payload of 2 112 bytes.

Table 16 – SR_I message format

Field	Description	Size (Bytes)
SR_Header		4
SR_I message payload	SOF	4
	FC-Header	24
	FC frame payload (includes optional header)	Min: 0 Max:2 112
	CRC	4
	EOF	4

The FC frame payload uses the SOF and EOF codes defined in annex A.

NOTE 5 - SR protocol-generated control messages do not carry the SOF and the EOF fields nor the FC Header in the payload.

5.5.5.2 Receive ready (SR_RR) command and response

The Receive Ready (SR_RR) supervisory message is used by the BBW to:

a) indicate it is ready to receive an SR_I message; and

b) acknowledge previously received SR_I messages numbered up to and including N(R)-1.

An SR_RR message may be used to indicate the clearance of a busy condition that was reported by the earlier transmission of an SR_RNR message by the same device. In addition to indicating the BBW status, the SR_RR message with the P bit set to 1 may be used to ask for the status of the remote BBW.

5.5.5.3 Receive not ready (SR_RNR) command and response

The Receive Not Ready (SR_RNR) supervisory message is used to indicate a busy condition (i.e., temporary inability to accept additional incoming SR_I messages). SR_I messages numbered up to and including N(R)-1 are acknowledged. SR_I message N(R) and any subsequent SR_I messages received, if any, are not acknowledged (i.e., the acceptance status of these SR_I messages shall be indicated in subsequent exchanges). In addition to indicating the status, the SR_RNR command with the P bit set to 1 may be used by a BBW to ask for the status of the remote BBW.

5.5.5.4 Selective reject (SR_SREJ) response

The SR_REJ supervisory message shall be used by a BBW to request retransmission of one or more, not necessarily contiguous, SR_I messages. The N(R) field shall contain the sequence number of the earliest SR_I message to be retransmitted and the information field (see table 17) shall contain, in ascending order (i.e., 32 767 is higher than 32 766 and 0 is higher than 32 767 for modulo 32 768), the sequence numbers of additional SR_I message(s), if any, that needs to be retransmitted.

The payload field shall be encoded such that there is a 2-byte field for each standalone SR_I message in need of retransmission, and a 4-byte span list for each sequence of two or more contiguously numbered SR_I messages in need of retransmission, as depicted in table 17. Standalone SR_I messages are identified in the payload field by the appropriate N(R) value preceded by a 0 bit in the 2-byte field used. Span lists are identified in the payload field by the N(R) value of the first SR_I message in the span list preceded by a 1 bit in the 2-byte field used, followed by the N(R) value of the last message in the span list preceded by a 1 bit in the 2-byte field used.

Table 17 – SR_SREJ payload format example

Bit 1 Value	Information Field	Size (Bytes)
<Bit 1> = 0	<Bits 2-16> = 1-N(R) of standalone SR_I message	2
<Bit 1> = 1	<Bits 2-16> = N(R) of first SR_I message in span list	2
<Bit 1> = 1	<Bits 2-16> = N(R) of last SR_I message in span list	2
<Bit 1> = 0	<Bits 2-16> = N(R) of standalone SR_I message	2
<Bit 1> = 1	<Bits 2-16> = N(R) of first SR_I message in span list	2
<Bit 1> = 1	<Bits 2-16> = N(R) of last SR_I message in span list	2
. . .		

NOTE 6 - The maximum size of the BBW message payload carrying the SR_SREJ message is 2 148 bytes corresponding to a maximum possible encoding of 1 074 standalone SR_I messages or a maximum possible encoding of 537 span list sets.

If the P/F bit in an SREJ message is set to 1, then SR_I messages numbered up to N(R)-1 inclusive, N(R) being the value in the SR_Header field, shall be considered as acknowledged. If the P/F bit in an SREJ message is set to 0, then the N(R) in the SR_Header field of the SREJ message does not indicate acknowledgement of SR_I messages.

5.5.5.5 Set Mode (SR_SM) command

The procedures to be followed on receipt of the Set Mode (SR_SM) command are specified in 6.3.4.

The SR_SM unnumbered command is used to initialize the BBW device.

No information field shall be permitted with the SR_SM command. The transmission of an SR_SM command indicates the clearance of a busy condition that was reported by the earlier transmission of an SR_RNR message by the same BBW device. The BBW device shall confirm the acceptance of the SR_SM command by the transmission, at the first opportunity, of an SR_UA response. Upon acceptance of this command, the BBW device send state variable V(S) and receive state variable V(R) shall be set to 0.

Previously transmitted SR_I messages that are unacknowledged when this message is processed remain unacknowledged. It is the responsibility of a higher layer to recover from the possible loss of the contents of SR_I messages that are not acknowledged.

5.5.5.6 Disconnect (SR_DISC) command

The SR_DISC unnumbered command is used to terminate the link that had been previously set. It is used to inform the BBW receiving the SR_DISC command that the remote BBW is suspending operation. No information field shall be provided with the SR_DISC command. Prior to actioning the SR_DISC command, the BBW receiving the SR_DISC command shall confirm the acceptance of the SR_DISC command by the transmission of an SR_UA response. The BBW sending the SR_DISC command shall enter the disconnected state when it receives the acknowledging SR_UA response.

Previously transmitted SR_I messages that are unacknowledged when this command is processed shall remains unacknowledged. It is the responsibility of a higher layer to recover from the possible loss of the contents of such SR_I messages.

5.5.5.7 Unnumbered acknowledgement (SR_UA) response

The SR_UA unnumbered response is used by the BBW device to acknowledge the receipt and acceptance of the SR_SM command. The received SR_SM command shall not be processed until the SR_UA response is transmitted. The transmission of an SR_UA response indicates the clearance of a busy condition that was re-

ported by the earlier transmission of an SR_RNR message by that same BBW device. No information field shall be provided with the SR_UA response.

5.5.5.8 Disconnected mode (SR_DM) response

The SR_DM unnumbered response is used to indicate that a BBW device is logically disconnected from the data link and is in the disconnected state. The SR_DM response may be sent to indicate that the BBW has entered the disconnected state without having received an SR_DISC command. Alternatively, if the SR_DM response was transmitted in response to the SR_SM command, the SR_DM command informs the remote BBW that the BBW is still in the disconnected state and cannot execute the SR_SM command. No information field shall be provided with the SR_DM response.

A BBW in a disconnected state shall monitor received commands and shall process an SR_SM command as specified in 6.3.4 and shall respond with an SR_DM response with the F bit set to 1 to any other command received with the P bit set to 1.

5.5.5.9 Message reject (SR_FRMR) response

An SR_FRMR unnumbered response shall be used by the BBW device to report an error condition that is not recoverable by retransmission of the identical message for the following conditions resulting from the receipt of a valid message:

a) the receipt of a command or response SR_Header sub-field that is undefined;

b) the receipt of an invalid N(R); or

c) the receipt of a message with an information field that is not permitted or the receipt of a supervisory or unnumbered message with incorrect length.

A valid N(R) shall be within the range from the lowest send sequence number N(S) of the still unacknowledged message(s) to the current BBW send state variable inclusive or to the current internal variable x if the BBW is in the timer recovery condition as described in 6.3.5.10.

An information field that immediately follows the SR_Header consists of nine bytes, is returned with this response, and provides the reason for the SR_FRMR response. Table 18 specifies the SR_FRMR payload format.

Table 18 – SR_FRMR payload format

Word	Bit Number	Field	Field Information	Size (bits)
0	0 -31		Rejected SR_Header Field	32
1	32		Set to 0	1
	33-47	V(S)	V(S) is the current send state variable value at the BBW reporting the rejection condition (bit 33= low-order bit)	15
	48	C/R	C/R set to 1 indicates the rejected message was a response; C/R set to 0 indicates the rejected message was a command	1
	49-63	V(R)	V(R) is the current receive state variable value at the BBW reporting the rejection condition (bit 49= low-order bit)	15
2	65	W	W set to 1 indicates that the SR_Header field received and returned in bits 1 through 32 was undefined.	1
	66	X	X set to 1 indicates that the SR_Header field received and returned in bits 1 through 32 was considered invalid because the message contained a payload that was not permitted with this type of message or is a supervisory or unnumbered message with incorrect length. Bit W shall be set to 1 in conjunction with this bit.	1
	67	Y	Reserved	1
	68	Z	Z set to 1 indicates that the SR_Header field received and returned in bits 1 through 32 contained an invalid N(R).	1
	69-72		Set to 0.	4
	73-96		Reserved	24

5.5.6 Exception condition reporting and recovery

5.5.6.1 Exception conditions

The error recovery procedures that are available to effect recovery following the detection/occurrence of an exception condition are described in this subclause. Exception conditions described are those situations that may occur as the result of transmission errors, BBW device malfunction, or operational situations.

5.5.6.2 Busy condition

The busy condition results when the BBW is temporarily unable to continue to receive SR_I messages due to internal constraints (e.g., receive buffering limitations). Upon entering the busy condition, a BBW transmits an SR_RNR message. SR_I messages pending transmission may be transmitted from the busy BBW prior to or following the SR_RNR message.

An indication that the busy condition has cleared is communicated by the transmission of an SR_UA (i.e., only in response to an SR_SM command), SR_RR, SR_SREJ, or SR_SM message.

5.5.6.3 N(S) sequence error condition

5.5.6.3.1 General

The information field of all received SR_I messages whose N(S) is not in the range V(R) and V(R)+k-1 inclusive, shall be discarded. The information field of all SR_I messages received by the BBW whose N(S) is in the range V(R) and V(R) + k -1 inclusive, shall be saved in the receive buffer.

An N(S) sequence error exception condition occurs in the receiver when a received SR_I message contains an N(S) that is not equal to the receive state variable V(R) at the receiver. The receiver shall not acknowledge (i.e., increment its receive state variable) the SR_I message causing the sequence error, or any SR_I message that may follow, until an SR_I message with the correct N(S) is received.

A BBW device that receives one or more valid SR_I messages having sequence errors or subsequent supervisory messages (i.e., SR_RR, SR_RNR, or SR_SREJ) shall accept the N(R) field and the P or F bit to perform data link control functions (e.g., to receive acknowledgement of previously transmitted SR_I messages and to cause the BBW to respond with the P bit set to 1).

The means specified in 5.5.6.3.2 and 5.5.6.3.3 shall be available for initiating the retransmission of lost or errored SR_I messages following the occurrence of an N(S) sequence error condition.

5.5.6.3.2 SR_SREJ recovery

The SR_SREJ message shall be used to initiate more efficient error recovery by selectively requesting the retransmission of one or more, not necessarily contiguous, lost or errored SR_I message(s) following the detection of sequence errors, rather than requesting the retransmission of all SR_I messages. When a BBW receives an out-of-sequence message, the SR_I message shall be saved in a receive buffer. The SR_I message shall be delivered to the upper layer only when all SR_I messages numbered below N(S) are correctly received. If message number N(S)-1 has not been received previously, then an SR_SREJ response message with the F bit set to 0 shall be transmitted containing the sequence numbers of the block of consecutive missing SR_I messages ending at N(S)-1. On receiving such an SR_SREJ message the BBW device shall retransmit all requested SR_I messages. After retransmitting these SR_I messages, the BBW may transmit new SR_I messages, if they become available.

When a BBW receives a command message with the P bit set to 1, if there are out-of-sequence SR_I messages saved in the receive buffer, it shall transmit an SR_SREJ message, with the F bit set to 1, containing a complete list of missing sequence numbers. The BBW that receives the SR_SREJ message shall retransmit all requested SR_I messages, except those that were transmitted subsequent to the last command message with the P bit set to 1.

5.5.6.3.3 Time-out recovery

If a BBW, due to a transmission error, does not receive, or receives and discards, a single SR_I message or the last SR_I message in a sequence of SR_I messages, it shall not detect a N(S) sequence error condition and, therefore, shall not transmit an SR_SREJ message.

The BBW that transmitted the unacknowledged SR_I message(s) shall, following the completion of a system specified time-out period (see 6.3.5.2 and 6.3.5.10), send a supervisory command message (i.e., SR_RR or SR_RNR) with the P bit set to 1. SR_I messages shall be retransmitted on the receipt of an SR_RR response message with the F bit set to 1 or an SR_SREJ message.

5.5.6.3.4 Invalid message condition

Any message that is invalid shall be discarded, and no action is taken as the result of that message. An invalid message is defined as one that contains:

a) the BBW_Header defined (see table 8) with an invalid encoding; or

b) the SR_Header defined (see table 12) with an invalid encoding.

5.5.6.3.5 Message rejection condition

A message rejection condition is established upon the receipt of an error-free message with one of the conditions listed in 5.5.5.9. At the BBW, this message rejection exception condition is reported by an SR_FRMR response for an appropriate BBW action. Once a BBW has established an exception condition, no additional SR_I messages shall be accepted until the condition is reset by the remote BBW, except for examination of the P bit. The SR_FRMR response may be repeated at each opportunity as described in 6.3.7.3, until recovery is effected by the remote BBW, or until the BBW initiates its own recovery in case the remote BBW does not respond.

6 SR and SFC Protocol Procedures

6.1 Applicability

This clause only applies to FC-BB-3_ATM and FC-BB-3_SONET.

The SR protocol is described in 6.2 and 6.3, and the SFC protocol in 6.4.

6.2 SR protocol overview

The Selective Retransmission (SR) protocol is an efficient sliding window link-layer full-duplex protocol that supports both the flow control and error recovery functions. SR has been adopted from ITU's Link Access Protocol B (LAPB), that was derived from ISO/IEC's High-level Data Link Control (HDLC) balanced classes. Use of LAPB in SR is limited to a subset of the synchronous modulo 32 768 super sequence numbering service option.

SR works between two BBW devices (see figure 8). SR flow control works by streaming multiple messages within an allowed window, bounded by the system parameter ***k***, and awaits acknowledgements before sending more messages. Acknowledgements indicate which messages were correctly received and there is a provision for requesting retransmission of selected messages in the current window. Fibre Channel Sequences and Exchanges are not visible to the SR flow control protocol which sees the BBW messages constructed from the FC frames.

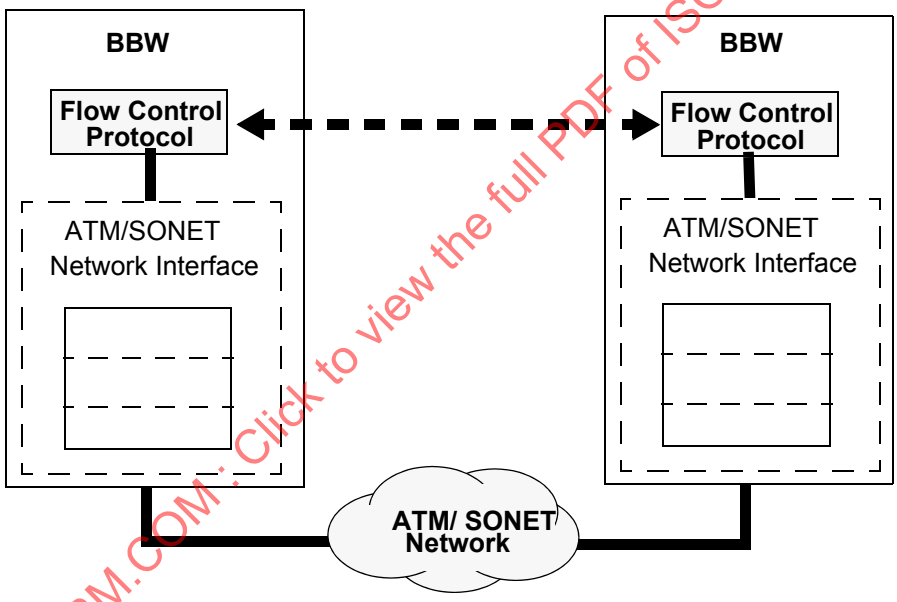


Figure 8 – SR flow control protocol between two BBWs

Some benefits of the SR protocol are summarized below:

a) it is used for reliable transport of all Class 2, 3, 4, and F frames between two BBW devices;

b) it synchronizes the BBW Sender and the BBW Receiver at the BBW message level;

c) it optimizes buffer management at the BBW devices;

d) it acts as a congestion avoidance technique to match the capacity of the sender to the capacity of the network that carries the payload;

e) it ensures correct delivery of messages (i.e., an error control and recovery function); and

f) it provides a continuous stream of traffic across the WAN thus leading to a higher throughput (i.e., optimizes bandwidth utilization at each BBW device).

The nine different SR messages listed in table 15 have a correspondence to the LAPB frame types. Note that only the information transfer SR_I message is flow-controlled while all other messages are control messages of the protocol.

The SR protocol specifies the maximum number (***k***) of outstanding messages at any given time. ***k*** is a system parameter that is not negotiated and is fixed in a given implementation. The value of this system parameter depends on the WAN delay characteristics and the number of buffers available. Typically, the value of ***k*** is expected to be far below the maximum number of 32 767.

The following subclauses describe the SR protocol procedures and reference the different message fields discussed in clause 5.

6.3 Description of the SR procedure

6.3.1 SR mode of operation

The SR protocol shall be limited to a subset of the synchronous modulo 32768 super sequence numbering service option operation of the LAPB protocol. See 5.5 for a description of the SR BBW message formats. The mode-setting command employed to initialize (i.e., set-up) or reset the protocol is the SR_SM command.

6.3.2 SR procedure for addressing

The Address Bit field (see table 8) identifies a message as either a command or a response.

This field is used in conjunction with the Poll/Final bit.

6.3.3 SR procedure for the use of the P/F bit

The BBW receiving an SR_SM, SR_DISC, supervisory command (i.e., SR_RR, SR_RNR, SR_SREJ), or SR_I message with the P bit set to 1 shall set the F bit to 1 in the next response message it transmits.

The response message returned by the BBW to an SR_SM or SR_DISC command with the P bit set to 1 shall be an SR_UA or SR_DM response with the F bit set to 1.

The response message returned by the BBW to an SR_I message with the P bit set to 1, received in the information transfer state, shall be an SR_RR, SR_SREJ, SR_RNR, or SR_FRMR response with the F bit set to 1.

The response message returned by the BBW to a supervisory command with the P bit set to 1, received in the information transfer state, shall be an SR_RR, SR_RNR, SR_SREJ or SR_FRMR response with the F bit set to 1.

The response message returned by the BBW to an SR_I message or supervisory message with the P bit set to 1, received in the disconnected state, shall be an SR_DM response with the F bit set to 1.

The P bit may be used by the BBW in conjunction with the timer recovery condition (see 6.3.5.10).

6.3.4 SR procedure for data link set-up and disconnection

6.3.4.1 Data link set-up

The BBW shall indicate to the SR protocol layer that it is able to set up the data link after it has provisioned an underlying ATM VC or a SONET Path.

Either BBW may initiate data link set-up. Prior to initiation of data link set-up, either BBW may initiate data link disconnection (see 6.3.4.3) for the purpose of ensuring that both BBW devices are in the same state. A BBW may also transmit an unsolicited SR_DM response to request the remote BBW to initiate data link set-up.

The BBW shall initiate data link set-up by transmitting an SR_SM command. If, upon correct receipt of the SR_SM command, the BBW device determines that it may enter the information transfer state, it shall return an SR_UA response to the sender, reset its send and receive state variables V(S) and V(R) to zero and shall consider that the link is set up.

If, upon receipt of the SR_SM command, the BBW device determines that it cannot enter the information transfer state, it shall return an SR_DM response as a denial to the link set up initialization and shall consider that the data link is **not** set up. In order to avoid misinterpretation of the SR_DM response received, it is suggested that the BBW always send its SR_SM command with the P bit set to 1. Otherwise, it is not possible to differentiate an SR_DM response intended as a denial to data link set up from an SR_DM response that is issued in a separate unsolicited sense as a request for a mode-setting command as described in 6.3.4.4.2.

The BBW device shall initiate link set up by transmitting an SR_SM command and starting its timer T1 in order to determine when too much time has elapsed waiting for a reply (see 6.3.8.1). Upon reception of an SR_UA response, the BBW shall reset its send and receive state variables V(S) and V(R) to zero, shall stop its timer T1, and shall consider that the link is set up. Upon reception of an SR_DM response as a denial to the link set-up initialization, the BBW shall stop its timer T1 and shall consider that the link is not set up.

The BBW having sent the SR_SM command, shall ignore and discard any messages except an SR_SM, or SR_DISC command, or an SR_UA or SR_DM response received from the remote BBW. The receipt of an SR_SM or SR_DISC command results in a collision situation that is resolved per 6.3.4.5. Messages other than the SR_UA and the SR_DM responses sent in response to a received SR_SM or SM_DISC command shall be sent only after the link is set up and if no outstanding SR_SM command exists.

After the BBW sends the SR_SM command, if an SR_UA or SR_DM response is not received correctly, timer T1 shall run out in the BBW. The BBW shall then resend the SR_SM command and shall restart timer T1. After transmission of the SR_SM command N2 times by the BBW, appropriate higher layer recovery action shall be initiated. The value of N2 is defined in 6.3.8.3.

6.3.4.2 Information transfer state

After having transmitted the SR_UA response to the SR_SM command or having received the SR_UA response to a transmitted SR_SM command, the BBW shall accept and transmit SR_I messages and supervisory messages (i.e., SR_RR, SR_RNR, and SR_SREJ) according to the procedures defined in 6.3.5.

When receiving the SR_SM command while in the information transfer phase, the BBW shall conform to the data link resetting procedure described in 6.3.7.

6.3.4.3 Data link disconnection

The BBW shall initiate a disconnect of the link by transmitting an SR_DISC command. On correctly receiving an SR_DISC command while in the information transfer state, the BBW shall send an SR_UA response and enter the disconnected state. On correctly receiving an SR_DISC command in the disconnected state, the remote BBW shall send an SR_DM response and remain in the disconnected state. In order to avoid misinterpretation of the SR_DM response received, it is suggested that the BBW always send its SR_DISC command with the P bit set to 1. Otherwise, it is not possible to differentiate an SR_DM response intended as an indication that the device is already in the disconnected state from an SR_DM response that is issued in a separate unsolicited sense as a request for a mode setting command (i.e., SR_SM) as described in 6.3.4.4.2.

The BBW shall initiate a disconnect of the data link by transmitting an SR_DISC command and starting its timer T1 (see 6.3.8.1). Upon reception of an SR_UA response from the remote BBW, the BBW shall stop its timer T1 and shall enter the disconnected state. Upon reception of an SR_DM response from the remote BBW as an indication that the remote BBW was already in the disconnected state, the BBW shall stop its timer T1 and shall enter the disconnected state.

The BBW having sent the SR_DISC command shall ignore and discard any messages except an SR_SM or SR_DISC command, or an SR_UA or SR_DM response received from the remote BBW. The receipt of an SR_SM or SR_DISC command from the remote BBW shall result in a collision situation that is resolved per 6.3.4.5.

After the BBW sends the SR_DISC command, if an SR_UA or SR_DM response is not received correctly, timer T1 shall run out in the BBW. The BBW shall then resend the SR_DISC command and shall restart timer T1. After transmission of the SR_DISC command N2 times by the BBW, appropriate higher layer recovery action shall be initiated. The value of N2 is defined in 6.3.8.3.

6.3.4.4 Disconnected state

6.3.4.4.1 Procedure 1

After having received an SR_DISC command from the remote BBW and returned an SR_UA response to the remote BBW, or having received the SR_UA response to a transmitted SR_DISC command, the BBW shall enter the disconnected state.

In the disconnected state, the BBW may initiate data link set-up. In the disconnected state, the BBW shall react to the receipt of an SR_SM command as described in 6.3.4.1, and shall transmit an SR_DM response in answer to a received SR_DISC command. When receiving any other command, defined or undefined, with the P

bit set to 1, the BBW shall transmit an SR_DM response with the F bit set to 1. Other messages received while in the disconnected state shall be ignored by the BBW.

6.3.4.4.2 Procedure 2

When the BBW enters the disconnected state after detecting error conditions as listed in 6.3.6, or after an internal malfunction, it may indicate this by sending an SR_DM response rather than an SR_DISC command. In these cases, the BBW shall transmit an SR_DM response and start its timer T1 (see 6.3.8.1).

If timer T1 runs out before the reception of an SR_SM or SR_DISC command from the remote BBW, the BBW shall retransmit the SR_DM response and restart timer T1. After retransmission of the SR_DM response N2 times, the BBW shall remain in the disconnected state and appropriate recovery actions shall be initiated. The value of N2 is defined in 6.3.8.3.

Alternatively, after an internal malfunction, the BBW may either initiate a data link resetting procedure (see 6.3.7) or disconnect the data link (see 6.3.4.3) prior to initiating a data link set-up procedure (see 6.3.4.1).

6.3.4.5 Collision of unnumbered commands

6.3.4.5.1 Procedure 1

If the sent and received unnumbered commands are the same, the BBWs shall each send the SR_UA response at the earliest possible opportunity. The BBW shall enter the indicated state either:

a) after receiving the SR_UA response;

b) after sending the SR_UA response; or

c) after timing out waiting for the SR_UA response having sent an SR_UA response.

In the case of item b) above, the BBW shall accept a subsequent SR_UA response to the mode-setting command it issued without causing an exception condition if received within the time-out interval.

6.3.4.5.2 Procedure 2

If the sent and received unnumbered commands are different, the BBWs shall each enter the disconnected state and issue an SR_DM response at the earliest possible opportunity.

6.3.4.6 Collision of SR_DM response with SR_SM or SR_DISC command

When an SR_DM response is issued by the BBW as an unsolicited response to request the remote BBW to issue a mode-setting command as described in 6.3.4.4, a collision between an SR_SM or SR_DISC command and the unsolicited SR_DM response may occur. In order to avoid misinterpretation of the SR_DM response received, the remote BBW always sends its SR_SM or SR_DISC command with the P bit set to 1.

6.3.4.7 Collision of SR_DM responses

A contention situation may occur when both the BBWs issue an SR_DM response. In this case, either BBW may issue an SR_SM command to resolve the contention situation.

6.3.5 Procedures for information transfer using multi-selective reject

6.3.5.1 Procedures for SR_I messages

The procedures that apply to the transmission of SR_I messages in each direction during the information transfer phase using multi-selective reject are described below.

6.3.5.2 Sending new SR_I messages

When the BBW has a new SR_I message to transmit (i.e., an SR_I message not already transmitted), it shall transmit it with a N(S) equal to its current send state variable V(S), and a N(R) equal to its current receive state variable V(R). At the end of the transmission of the SR_I message, it shall increment its send state variable V(S) by 1.

If the BBW timer T1 is not running at the time of transmission of the SR_I message, it shall be started.

If the BBW send state variable V(S) is equal to the last value N(R) received plus ***k***, where ***k*** is the maximum number of outstanding SR_I frames (see 6.3.8.4), the BBW shall not transmit any new SR_I frames.

If the remote BBW is busy, the BBW shall not transmit any new SR_I messages.

When the BBW is in the busy condition, it may still transmit SR_I messages, provided that the remote BBW is not busy.

6.3.5.3 Receiving an in-sequence SR_I message

When the BBW is not in a busy condition and receives a valid SR_I message whose send sequence number N(S) is equal to its receive state variable V(R), the BBW shall accept the information field of this message and increment by one the receive state variable V(R). If the SR_I message, whose N(S) is equal to the incremented value of V(R), is present in the receive buffer, then the BBW shall remove it from the receive buffer, deliver it to the upper layer and increment V(R) by one. The BBW shall repeat this procedure until V(R) reaches a value such that the SR_I message whose N(S) is equal to V(R) is not present in the receive buffer. The BBW shall then take one of the following actions:

a) if the BBW is now in the busy condition, it shall transmit an SR_RNR message with N(R) equal to the value of the BBW receive variable V(R) (see 6.3.5.9); or

b) if the BBW is still not in a busy condition:

   1) if the P bit is set to 1, then the BBW shall transmit a response message with the F bit set to 1, as specified in 6.3.5.12;

   2) if an SR_I message is available for transmission as specified in 6.3.8.4, the BBW shall act as described in 6.3.5.2, sending new SR_I messages and acknowledging the received SR_I message by setting N(R) in the SR_Header field of the next transmitted SR_I message to the value of the BBW receive state variable V(R), or the BBW shall acknowledge the received SR_I message by transmitting an SR_RR message with the N(R) equal to the value of the BBW receive state variable V(R); or

   3) the BBW shall transmit an SR_RR message with N(R) equal to the value of the BBW receive state variable V(R).

   When the BBW is in a busy condition, it may ignore the information field contained in any received SR_I message.

6.3.5.4 Reception of invalid messages

When the BBW receives an invalid message (see 5.5.6.3.4), it shall discard the message.

6.3.5.5 Reception of out-of-sequence SR_I messages

When the BBW is not in a busy condition and it receives a valid SR_I message whose send sequence number N(S) is out-of-sequence, (i.e., not equal to the receive state variable V(R)), then it shall perform one of the following actions:

a) if N(S) is less than V(R) or greater than or equal to V(R) + ***k***, then it shall discard the information field of the SR_I message. If the P bit of the SR_I message is set to 1, then the BBW shall transmit a response message with the F bit set to 1, as specified in 6.3.5.12; or

b) if N(S) is greater than V(R) and less than V(R) + ***k***, then it shall save the SR_I message in the receive buffer. It shall then perform one of the following actions:

   1) if the P bit of the SR_I message is set to 1, then the BBW shall transmit a response message with the F bit set to 1, as specified in 6.3.5.12;

   2) if the BBW is now in a busy condition, it shall transmit an SR_RNR message with N(R) equal to the value of the receive variable V(R), as specified in 6.3.5.9; or

   3) if the SR_I message numbered N(S)-1 has not yet been received, then the BBW shall transmit an SR_SREJ response message with the F bit set to 0. The BBW shall create a list of contiguous sequence numbers N(X), N(X)+1, N(X)+2,..., N(S)-1, where N(X) is greater than or equal to V(R) and none of the SR_I messages N(X) to N(S)-1 have been received. The N(R) field of the SR_SREJ message shall be set to N(X) and the information field set to the list N(X)+1,...,N(S)-1. If the list of sequence numbers is too large to fit into the information field of the SR_SREJ message, then the list shall be truncated to fit in one SR_SREJ message, by including only the earliest sequence numbers.

When the BBW is in the busy condition, it may ignore the information field contained in any received SR_I message.

6.3.5.6 Receiving acknowledgement

When correctly receiving an SR_I message or a supervisory message (i.e., SR_RR, SR_RNR, or SR_SREJ with the F bit set to 1), even in the busy condition, the BBW shall consider the N(R) contained in this message as an acknowledgement for all the SR_I messages it has transmitted with a N(S) up to and including the received N(R)-1. The BBW shall stop the timer T1 if the received supervisory message has the F bit set to 1 or if there is no outstanding poll condition and the N(R) is higher than the last received N(R), actually acknowledging some SR_I messages.

If timer T1 has been stopped by the receipt of an SR_I message, an SR_RR command message, an SR_RR response message with the F bit set to 0, or an SR_RNR message, and if there are outstanding SR_I messages still unacknowledged, the BBW shall restart timer T1. If timer T1 has been stopped by the receipt of an SR_SREJ message with the F bit set to 1, the BBW shall follow the retransmission procedure specified in 6.3.5.7.2. If timer T1 has been stopped by the receipt of an SR_RR message with the F bit set to 1, the BBW shall follow the retransmission procedure specified in 6.3.5.11.

6.3.5.7 Receiving an SR_SREJ response message

6.3.5.7.1 Receiving an SR_SREJ response message with the F bit set to 0

When receiving an SR_SREJ response message with the F bit set to 0, the BBW shall retransmit all SR_I messages, whose sequence numbers are indicated in the N(R) field and the information field of the SR_SREJ message, in the order specified in the SR_SREJ message. Retransmission shall conform to the following:

a) if the BBW is transmitting a supervisory or SR_I message when it receives the SR_SREJ message, it shall complete that transmission before commencing transmission of the requested SR_I messages;

b) if the BBW is transmitting an unnumbered command or response message when it receives the SR_SREJ message, it shall ignore the request for retransmission; or

c) if the BBW is not transmitting any message when it receives the SR_SREJ message, it shall commence transmission of the requested SR_I messages immediately.

If there is no outstanding poll condition, then a poll shall be sent, either by transmitting an SR_RR command, or SR_RNR command if the BBW is in the busy condition, with the P bit set to 1 or by setting the P bit in the last retransmitted SR_I message and timer T1 shall be restarted.

If there is an outstanding poll condition, then timer T1 shall not be restarted.

6.3.5.7.2 Receiving an SR_SREJ response message with the F bit set to 1

When receiving an SR_SREJ response message with the F bit set to 1, the BBW shall retransmit all SR_I messages, whose sequence numbers are indicated in the N(R) field and the information field of the SR_SREJ message, in the order specified in the SR_SREJ message, except those messages that were sent after the message with the P bit set to 1 was sent. Retransmission shall conform to the following:

a) if the BBW is transmitting a supervisory message or SR_I message when it receives the SR_SREJ message, it shall complete that transmission before commencing transmission of the requested SR_I messages;

b) if the BBW is transmitting an unnumbered command or response when it receives the SR_SREJ message, it shall ignore the request for retransmission; or

c) if the BBW is not transmitting any message when it receives the SR_SREJ message, it shall commence transmission of the requested SR_I messages immediately.

If any messages are retransmitted, then a poll shall be sent, either by transmitting an SR_RR command, or SR_RNR command if the BBW is in the busy condition, with the P bit set to 1 or by setting the P bit in the last retransmitted SR_I message.

timer T1 shall be restarted.

6.3.5.8 Receiving an SR_RNR message

After receiving an SR_RNR message, the BBW shall stop transmission of SR_I messages until an SR_RR or SR_SREJ message is received.

The BBW shall start timer T1, if necessary, as specified in 6.3.8.1.

When timer T1 runs out before receipt of a busy clearance indication, the BBW shall transmit a supervisory message (i.e., SR_RR, SR_RNR), with the P bit set to 1 and shall restart timer T1, in order to determine if there is any change in the receive status of the remote BBW. The remote BBW shall respond to the P bit set to 1 with a supervisory response message (i.e., SR_RR, SR_RNR, SR_SREJ) with the F bit set to 1 indicating continuation of the busy condition (i.e., SR_RNR message) or clearance of the busy condition (i.e., SR_RR, SR_SREJ). Upon receipt of the remote BBW response, timer T1 shall be stopped. The BBW shall process the supervisory response message as follows:

a) if the response is an SR_RR message, the busy condition shall be assumed to be cleared and the BBW may retransmit messages as specified in 6.3.5.11. New SR_I messages may be transmitted as specified in 6.3.5.2;

b) if the response is an SR_SREJ message, the busy condition shall be assumed to be cleared and the BBW may retransmit messages as specified in 6.3.5.7.2. New SR_I messages may be transmitted as specified in 6.3.5.2; or

c) if the response is an SR_RNR message, the busy condition shall be assumed to still exist and the BBW, after a period of time (e.g., the duration of timer T1), shall repeat the enquiry of the remote BBW receive status.

If timer T1 runs out before a status response is received, the enquiry process above shall be repeated. If N2 attempts to get a status response fail, the BBW shall initiate the link resetting procedure as described in 6.3.7.

If, at any time during the enquiry process, an unsolicited SR_RR or SR_SREJ message is received from the remote BBW, it shall be considered to be an indication of clearance of the busy condition. Should the unsolicited SR_RR message be a command message with the P bit set to 1, the appropriate response message with the F bit set to 1 shall be transmitted (see 6.3.5.12) before the BBW may resume transmission of SR_I messages. The BBW shall not clear the outstanding poll condition. The BBW shall not stop timer T1. If an unsolicited SR_SREJ message is received, then the BBW shall perform retransmissions as specified in 6.3.5.7.1.

6.3.5.9 BBW busy condition

When the BBW enters a busy condition, it shall transmit an SR_RNR message at the earliest opportunity. The SR_RNR message shall be a command frame with the P bit set to 1 if an acknowledged transfer of the busy condition indication is required, otherwise the SR_RNR message may be a command or response message. While in the busy condition, the BBW shall accept and process supervisory messages, accept and process the N(R) field of SR_I, SR_RR, and SR_SREJ messages with the F bit set to 1, and return an SR_RNR response with the F bit set to 1 if it receives a supervisory command or SR_I command message with the P bit set to 1. Received SR_I messages may be discarded or saved as specified in 6.3.5.3 and 6.3.5.5, however, SR_RR or SR_SREJ messages shall not be transmitted. To clear the busy condition, the BBW shall transmit an SR_RR message, with the N(R) field set to the current receive state variable V(R). The SR_RR message shall be a command message with the P bit set to 1 if an acknowledged transfer of the busy-to-non-busy transition is required, otherwise the SR_RR message may be either a command or response message.

6.3.5.10 Awaiting acknowledgement

If the timer T1 runs out while waiting for the acknowledgement of an SR_I message from the remote BBW, the BBW shall restart timer T1 and transmit an appropriate supervisory command message (i.e., SR_RR, SR_RNR) with the P bit set to 1. The BBW may transmit new SR_I messages after sending this enquiry message.

If the BBW receives an SR_SREJ response message with the F bit set to 1, the BBW shall restart timer T1 and retransmit SR_I messages as specified in 6.3.5.7.2.

If the BBW receives an SR_SREJ response message with the F bit set to 0, the BBW shall retransmit SR_I messages as specified in 6.3.5.7.2.

If the BBW receives an SR_RR response message with the F bit set to 1, the BBW shall restart timer T1 and retransmit SR_I messages as specified in 6.3.5.11.

If the BBW receives an SR_RR response message with the F bit set to 0, or an SR_RR command message or SR_I message with the P bit set to 0 or 1, the BBW shall not restart timer T1, but shall use the received N(R) as an indication of acknowledgement of transmitted SR_I messages up to and including SR_I message numbered N(R)-1.

If timer T1 runs out before a supervisory response message with the F bit set to 1 is received, the BBW shall retransmit an appropriate supervisory command message (i.e., SR_RR, SR_RNR) with the P bit set to 1. After N2 such attempts, the BBW shall initiate the link resetting procedure as described in 6.3.7.

6.3.5.11 Receiving an SR_RR response messages with the F bit set to 1

When receiving an SR_RR response message with the F bit set to 1, the BBW shall process the N(R) field as specified in 6.3.5.6. If there are outstanding SR_I messages that are unacknowledged and no new SR_I messages have been transmitted subsequent to the last message with the P bit set to 1, then the BBW shall retransmit all outstanding SR_I messages except those that were sent after the message with the P bit set to 1 was sent. Retransmission shall conform to the following:

a) if the BBW is transmitting a supervisory or SR_I message when it receives the SR_RR message, it shall complete that transmission before commencing transmission of the requested SR_I messages;

b) if the BBW is transmitting an unnumbered command or response when it receives the SR_RR message, it shall ignore the request for retransmission; or

c) if the BBW is not transmitting any message when it receives the SR_RR message, it shall commence transmission of the requested SR_I messages immediately.

If any messages are retransmitted, then a poll shall be sent, either by transmitting an SR_RR command, or SR_RNR command if the BBW is in the busy condition, with the P bit set to 1 or by setting the P bit in the last retransmitted SR_I message.

The timer T1 shall be stopped. If any SR_I messages are outstanding, then timer T1 shall be started.

6.3.5.12 Responding to command messages with the P bit set to 1

When receiving an SR_RR, SR_RNR, or SR_I command message with the P bit set to 1, the BBW shall generate an appropriate response message as follows:

a) if the BBW is in the busy condition, it shall transmit an SR_RNR response message with the F bit set to 1;

b) if there are some out-of-sequence messages in the receive buffer, then it shall transmit an SR_SREJ message with the F bit set to 1; N(R) shall be set to the receive state variable V(R) and the information field set to the sequence numbers of all missing SR_I messages, except V(R). If the list of sequence numbers is too large to fit in the information field of the SR_SREJ message, then the list shall be truncated by including only the earliest sequence numbers; or

c) if there are no out-of-sequence messages in the receive buffer, then an SR_RR response message with the F bit set to 1 shall be sent.

6.3.6 SR conditions for data link resetting or data link re-initialization

6.3.6.1 Condition 1

When a BBW receives, during the information transfer state, a message that is not valid (see 5.5.6.3.4) with one of the conditions listed in 5.5.5.9, the BBW shall request the remote BBW to initiate a data link resetting procedure by transmitting an SR_FRMR response to the remote BBW as described in 6.3.7.3.

6.3.6.2 Condition 2

When the BBW receives, during the information transfer state, an SR_FRMR response from the remote BBW, the BBW shall either initiate the data link resetting procedures itself as described in 6.3.7.2 or return an SR_DM response to ask the remote BBW to initiate the data link set-up (i.e., initialization) procedure as described in 6.3.4.1. After transmitting an SR_DM response, the BBW shall enter the disconnected state as described in 6.3.4.4.2.

6.3.6.3 Condition 3

When the BBW receives, during the information transfer state, an SR_UA response, or an unsolicited response with the F bit set to 1, the BBW may either initiate the data link resetting procedures itself as described in 6.3.7.2, or return an SR_DM response to ask the remote BBW to initiate the data link set-up (i.e., initialization)

procedure as described in 6.3.4.1. After transmitting an SR_DM response, the BBW shall enter the disconnected state as described in 6.3.4.4.2.

6.3.6.4 Condition 4

When the BBW receives, during the information transfer state, an SR_DM response from the remote BBW, the BBW shall either initiate the data link set-up (i.e., initialization) procedure as described in 6.3.4.1, or return an SR_DM response to ask the remote BBW to initiate the data link set-up (i.e., initialization) procedures as described in 6.3.4.1. After transmitting an SR_DM response, the BBW shall enter the disconnected state as described in 6.3.4.4.2.

6.3.7 SR procedures for data link resetting

6.3.7.1 Overview

The data link resetting procedures are used to initialize both directions of information transfer. The data link resetting procedures only apply during the information transfer state.

6.3.7.2 Procedure 1

Either BBW may initiate a data link reset procedure. The data link reset procedure indicates a clearance of a BBW and/or remote BBW busy condition, if present.

The remote BBW shall initiate a data link resetting by transmitting an SR_SM command to the BBW. If, upon correct receipt of the SR_SM command, the BBW determines that it is able to continue in the information transfer state, it shall return an SR_UA response to the remote BBW, shall reset its send and receive state variables V(S) and V(R) to zero, and shall remain in the information transfer state. If, upon the receipt of the SR_SM command, the BBW determines that it cannot remain in the information transfer state, it shall return an SR_DM response as a denial to the resetting request and shall enter the disconnected state.

The BBW shall initiate a data link resetting by transmitting an SR_SM command to the remote BBW and starting its timer T1 (see 6.3.8.1). Upon reception of an SR_UA response from the remote BBW, the BBW shall reset its send and receive state variables V(S) and V(R) to zero, shall stop its timer T1, and shall remain in the information transfer state. Upon reception of an SR_DM response from the remote BBW as a denial to the data link resetting request, the BBW shall stop its timer T1 and shall enter the disconnected state.

The BBW, having sent an SR_SM command shall ignore and discard any messages received from the remote BBW except an SR_SM or SR_DISC command, or an SR_UA or SR_DM response. The receipt of an SR_SM or SR_DISC command from the remote BBW shall result in a collision situation that is resolved per 6.3.4.5. Messages other than the SR_UA or SR_DM response sent in response to a received SR_SM or SR_DISC command shall be sent only after the data link is reset and if no outstanding SR_SM command exists.

After the BBW sends the SR_SM command, if an SR_UA or SR_DM response is not received correctly, timer T1 shall run out in the BBW. The BBW shall then resend the SR_SM command and shall restart timer T1. After N2 attempts to reset the data link, the BBW shall initiate appropriate higher layer recovery action and shall enter the disconnected state. The value of N2 is defined in 6.3.8.3.

6.3.7.3 Procedure 2

The BBW may ask the remote BBW to reset the data link by transmitting an SR_FRMR response (see 6.3.6.1). After transmitting an SR_FRMR response, the BBW shall enter the message rejection condition.

The message rejection condition is cleared when the BBW receives an SR_SM command, an SR_DISC command, an SR_FRMR response, or an SR_DM response; or if the BBW transmits an SR_SM command, an SR_DISC command, or an SR_DM response. Other commands received while in the message rejection condition shall cause the BBW to retransmit the SR_FRMR response with the same information field as originally transmitted.

The BBW may start timer T1 on transmission of the SR_FRMR response. If timer T1 runs out before the message rejection condition is cleared, the BBW may retransmit the SR_FRMR response, and restart T1. After N2 attempts (i.e., time outs) to get the remote BBW to reset the data link, the BBW may reset the data link itself as described in 6.3.7.2. The value of N2 is defined in 6.3.8.3.

In the message rejection condition, SR_I messages and supervisory messages shall not be transmitted by the BBW. Also, the BBW shall ignore and discard the N(S) and information fields of any received SR_I messages and the N(R) fields of any received SR_I messages and supervisory messages. When an additional SR_FRMR response is transmitted by the BBW as a result of the receipt of a command message while timer T1 is running, timer T1 shall continue to run. Upon reception of an SR_FRMR response, even during a message rejection

condition, the BBW shall initiate a resetting procedure by transmitting an SR_SM command as described in 6.3.7.2, or shall transmit an SR_DM response to ask the remote BBW to initiate the data link set-up procedure as described in 6.3.4.1, and enter the disconnected state.

6.3.8 List of SR system parameters

6.3.8.1 Timer T1

The same value of the timer T1 shall be made known and agreed to by all BBWs.

The period of timer T1, at the end of which retransmission of a message may be initiated (see 6.3.4 and 6.3.5), shall take into account whether T1 is started at the beginning or the end of the transmission of a message.

The proper operation of the procedure requires that the transmitter's timer T1 be greater than the maximum time between transmission of a message (i.e., SR_SM, SR_DISC, SR_I, or supervisory command, or SR_DM or SR_FRMR response) and the reception of the corresponding message returned as an answer to that message (i.e., SR_UA, SR_DM, or acknowledging message). Therefore, the receiver should not delay the response or acknowledging message returned to one of the above messages by more than a value T2, where T2 is a system parameter (see 6.3.8.2).

The BBW shall not delay the response or acknowledging message returned to one of the above remote BBW messages by more than a period T2.

6.3.8.2 Parameter T2

The same value of the parameter T2 shall be made known and agreed to by all BBWs.

The period of parameter T2 shall indicate the amount of time available at the BBW before the acknowledging message shall be initiated in order to ensure its receipt by the remote BBW, prior to timer T1 running out at the BBWs (parameter T2 < timer T1).

The period of parameter T2 shall take into account the following timing factors:

a) the transmission time of the acknowledging message;

b) the propagation time over the access link;

c) the stated processing times at the BBWs; and

d) the time to complete the transmission of the message(s) in the BBW transmit queue that are neither displaceable nor modifiable in an orderly manner.

Given a value for timer T1 for the BBWs, the value of parameter T2 shall be no larger than T1 minus 2 times the propagation time over the access data link, minus the message processing time at the BBW, minus the message processing time at the remote BBW, and minus the transmission time of the acknowledging message by the BBW.

Annex C provides guidelines for tuning the SR protocol parameters.

6.3.8.3 Maximum number of attempts to complete a transmission N2

The same value of the N2 system parameter shall be made known and agreed to by the BBWs.

The value of N2 shall indicate the maximum number of attempts made by the BBW to complete the successful transmission of a message to the remote BBW.

6.3.8.4 Maximum number of outstanding SR_I messages *k*

The same value of the ***k*** system parameter shall be made known and agreed to by the BBWs.

The value of ***k*** shall indicate the maximum number of sequentially numbered SR_I messages that the BBWs may have outstanding (i.e., unacknowledged) at any given time. The value of ***k*** shall never exceed 32767 for modulo 32768 operation.

NOTE 7 - Annex C provides guidelines for selecting appropriate values of ***k*** and message size to maximize the efficiency of links with long propagation delays.

6.4 Simple Flow Control (SFC)

Simple Flow Control (SFC) is a mechanism that requests the remote BBW from pausing transmission for a time period defined by the number of time units in the PAUSE bytes of the BBW_Header. Each time unit corresponds to the transmission time for 512 bits (i.e., 64 bytes). A zero value in the PAUSE bytes indicates that the remote BBW does not need to pause transmission, the effect is the same as non-use of flow control. If a subsequent message is received with the PAUSE field set to a different value, then the pause time is reset to this new value.

Use of SFC is optional and may result in the remote BBW in simply ignoring the PAUSE bytes. In this case, pausing is not accomplished.

7 FC-BB-3_ATM Structure and Concepts

7.1 Applicability

Clause 4 discusses the FC-BB-3_ATM reference model. Clause 5 describes the required messages and clause 6 describes the flow control mechanisms applicable to FC-BB-3_ATM. This clause discusses the FC-BB-3_ATM functional model.

7.2 FC-BB-3_ATM overview

Figure 9 shows a network configuration consisting of three FC-BB-3_ATM devices. FC-BB-3_ATM is a Fibre Channel backbone transport protocol that tunnels AAL5 encapsulated FC frames across the ATM network. An FC-BB-3_ATM device has interfaces to both the ATM and the FC network. The FC network interface supports multiple B_Ports (see figure 11). The model applies equally to both private or public ATM networks.

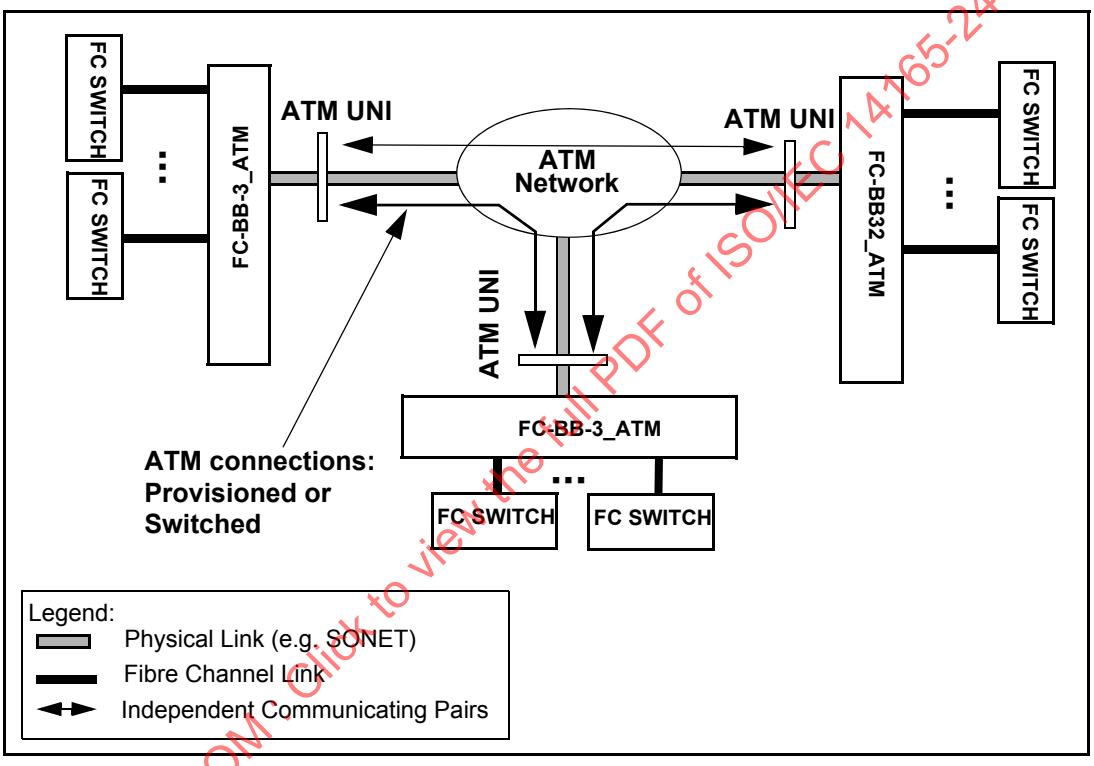


Figure 9 – FC-BB-3_ATM network configuration

FC-BB-3_ATM devices that support B_Port do not require FC switching. The FC-BB-3_ATM protocol provides mechanisms to create B_Access connectivity over the ATM network.

FC-BB-3_ATM protocol communication occurs between pairs of FC-BB-3_ATM devices over virtual constructs (i.e., FCATM links) that are described in 7.3.4.6. Although FC-BB-3_ATM protocol communication occurs between pairs of FC-BB-3_ATM devices, a single FC-BB-3_ATM device may communicate with more than one device at the same time.

NOTE 8 - The current scheme allows a single FC-BB-3_ATM device to independently connect to more than one FC-BB-3_ATM device, but does not specify a point-to-multipoint connection.

The FC-BB-3_ATM protocol creates BBW messages that consist of an 8-byte LLC/SNAP Header and an 4-byte BBW_Header followed by the BBW message payload. The specific format and content of the BBW message payload depends on the type of flow control protocol used. The BBW message payload carries byte-encoded SOF/EOF delimited Class 2, 3, 4, or F FC frames.

The BBW messages are encapsulated in ATM Adaptation Layer 5 (AAL5) format for carriage over the ATM network. The AAL5-encapsulated BBW messages are segmented into ATM cells and routed to the proper destination ATM address. FC-BB-3_ATM does not interpret the data content of the FC frames other than capturing and retaining their SOF/EOF identities in the encapsulated FC frame. As such, FC Sequences and Exchanges

are not visible to the FC-BB-3_ATM protocol. All AAL5 encapsulated FC frames are transparently transported over the ATM network.

The LLC/SNAP Header indicates the payload type as Fibre Channel (see 5.2). The BBW_Header indicates the type of flow control used (i.e., Selective Retransmission (SR), Simple Flow Control (SFC), or none). The SR protocol makes the transport of FC frames between two FC-BB-3_ATMs reliable. The SR protocol supports both flow control and error recovery functions. Use of the SR protocol is optional. When SR flow control is used, the 4-byte BBW_Header is followed by a 4-byte SR_Header that is inserted at the beginning of the BBW message payload. The SFC protocol provides a mechanism to temporarily pause the transmission of frames from a remote BBW device. Use of the SFC protocol is optional. When SFC is used, the 4-byte BBW_Header is directly followed by the BBW message payload. No SFC header is prefixed or used. See 5.4 .

In-order delivery is guaranteed within the scope of an ATM Virtual Connection (VC) and frames are transmitted from the FC-BB-3_ATM in the same order as they are received. The FC Entity is expected to specify and handle all other FC frame delivery ordering requirements.

FC-BB-3_ATM devices also exchange SW_ILS control information using Class F FC frames. These FC frames are encapsulated and tunneled in the same way as the incoming FC frames.

7.3 FC-BB-3_ATM B_Access functional model

7.3.1 Protocol layers

Figure 11 shows the B_Access functional model of an FC-BB-3_ATM device that consists of the B_Port FC interface, the ATM network interface, and the FC-BB-3_ATM interface. The protocol layers at these interfaces are listed below:

a) B_Port FC interface: FC-0, FC-1, and FC-2 Levels;

b) ATM network interface: PHY, ATM, and Adaptation Layers (i.e., AAL5 and SAAL); and

c) FC-BB-3_ATM interface: FC Entity and FCATM Entity protocol layers.

Figure 10 illustrates the protocol layers across these interfaces.

7.3.2 B_Port FC interface

The FC-BB-3_ATM FC interface supports one or more B_Ports thus requires the support of the FC-0, FC-1, and FC-2 levels. The B_Ports in general connect to different external FC switches, but connectivity to the same FC switch is also allowed.

B_Ports are uniquely identified by an 8-byte B_Port_Name.

7.3.3 ATM network interface

The ATM network interface includes the PHY, ATM, and Adaptation Layers. The basic FC-BB-3_ATM reference model supports one ATM port using different media types and/or different rates. The ATM Adaptation Layer-5 (AAL5) is used for BBW message transport while SAAL Adaptation is used for ATM signaling. FC-BB-3_ATM may use either provisioned Permanent Virtual Circuit (PVC) or Switched Virtual Connection (SVC) to transport messages. SVC requires the use of the User Network Interface (UNI) Signaling Protocol specifying the desired Service Category, QoS, and Traffic Parameters. Both Public UNI and Private UNI shall be supported.

FC-BB-3_ATM uses Variable Bit Rate Non Real Time traffic (VBR-NRT) service. AAL5 is particularly well suited for carriage of VBR-NRT. VBR-NRT ATM service provides cell loss and bandwidth guarantees.

FC-BB-3_ATM recommends use of a single Virtual Circuit (VC). Use of additional VCs to address special traffic QoS requirements is allowed but not recommended. If SR or SFC flow control is used, then flow control is separately applied to each VC.

FC-BB-3_ATM recommends allocating a minimum bandwidth for each VC that is used in order to avoid starvation. However, the service discipline prioritization for the VCs is implementation specific and beyond the scope of this standard.

7.3.4 FC-BB-3_ATM protocol interface

7.3.4.1 Major components

The FC-BB-3_ATM protocol interface is a point that has interfaces to the FC network on one side and the ATM network on the other. In addition to the two network interfaces, it consists of the following major components:

a) FC and FCATM Entities;

b) Control and Service Module (CSM); and

c) Platform Management Module (PMM).

7.3.4.2 FC Entity and FCATM Entity

The FC Entity is the principal interface point to the FC network on one side and operates in combination with the FCATM Entity to the ATM network on the other side. The primary functions of the FC Entity is to support one or more B_Access portals and to communicate with the FCATM Entity. The FC Entity layer lies between the FC-2 level and the FCATM Entity layer as shown in figure 10.

The FCATM Entity is the principal interface point to the ATM network on one side and operates in combination with the FC Entity to the FC network on the other side. The primary function of the FCATM Entity is formatting, encapsulating, and forwarding AAL5 encapsulated FC frames across the ATM network interface.

The FC/FCATM Entity pair interfaces with the CSM and the PMM through an implementation defined interface.

7.3.4.3 FC Entity

The FC-BB-3_ATM interface may support multiple instances of the FC/FCATM Entity pairs. Each instance of the FC/FCATM Entity pair consists of one or more B_Access/FCATM_LEP pairs. A B_Access portal is a component of the FC Entity that interfaces with the FCATM_LEP component of the FCATM Entity. The B_Access portal receives FC frames from the B_Port and sends them to the FCATM_LEP for encapsulation and transmission on the ATM network. The B_Access portal may also exchange Class F control frames with the remote B_Access portal via the FCATM_LEPs. There is a one-to-one relationship between a B_Access portal and an FCATM_LEP. B_Access portals communicate via B_Access Virtual ISLs (see 7.3.5).

There is no switching and routing required in the case of the B_Port functional model. However, the forwarding of FC frames across the B_Access/FCATM_LEP pair is still required. When multiple FCATM_DEs within an FCATM_LEP are in use, the selection of which FCATM_DE to use is decided by policies in the FCATM Entity.

A B_Access portal is uniquely identified by an 8-byte B_Access_Name.

Within an FC-BB-3_ATM device, each FC/FCATM Entity pair instance is uniquely identified by an 8-byte identifier called the FC/FCATM identifier. The FC/FCATM identifier uses the Name_Identifier format.

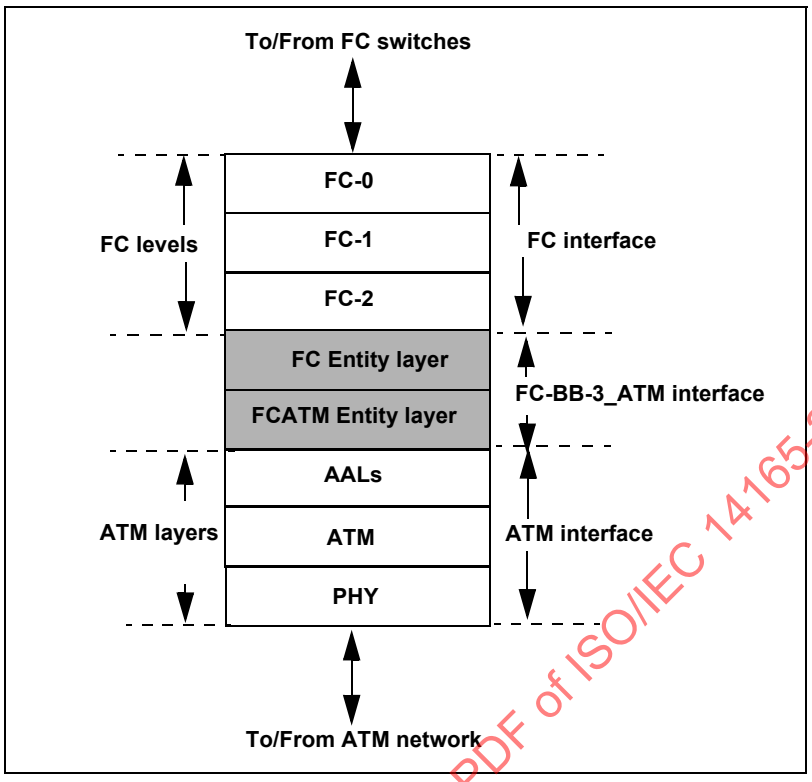


Figure 10 – FC-BB-3_ATM protocol layers

7.3.4.4 FCATM Entity – Components

The FCATM_LEP is a component of the FCATM Entity that formats, encapsulates, flow controls, and forwards AAL5 encapsulated FC frames. AAL5 encapsulated FC frames are sent as ATM cells over the ATM network. The FCATM_LEP receives byte-encoded SOF/EOF delimited FC frames from its B_Access portal.

The FCATM Data Engine (FCATM_DE) is the data-forwarding component of the FCATM_LEP. The FCATM_DE handles all encapsulation, de-encapsulation, and transmission and reception of the AAL5 encapsulated FC frames on the FCATM link. The FCATM_LEP contains one or more FCATM_DEs, each corresponding to an ATM VC. The FCATM_DE has four interface points:

a) **FC Receiver Portal:** The access point through which an byte-encoded SOF/EOF delimited FC frame enters an FCATM_DE from the B_Access portal;

b) **FC Transmitter Portal:** The access point through which a reconstituted byte-encoded SOF/EOF delimited FC frame leaves an FCATM_DE to the B_Access portal;

c) **AAL5 Encapsulated Frame Receiver Portal:** The ATM access point through which an AAL5 encapsulated FC frame is received from the ATM Network by the FCATM_DE; and

d) **AAL5 Encapsulated Frame Transmitter Portal:** The ATM access point through which an AAL5 encapsulated FC frame is transmitted to the ATM network by the FCATM_DE.

7.3.4.5 FCATM_DE

7.3.4.5.1 BBW messages

The FCATM_DE engine creates the 8-byte LLC/SNAP Header and the 4-byte BBW_Header that are prefixed to the BBW message payload. The specific format and content of the BBW message payload depends on the type of flow control protocol used. The BBW message payload carries the byte-encoded SOF/EOF delimited Class 2, 3, 4, or F FC frames.

FCATM_DE does not interpret the data content of the FC frames other than capturing and retaining their SOF/EOF identities in the encapsulated FC frame. As such, FC Sequences and Exchanges are not visible to the

FC-BB-3_ATM protocol. When flow control is not used, the FCATM_DE sets the PAUSE field in the BBW_Header to a zero value and the Flow Control Type to SFC.

NOTE 9 - The setting of Flow Control Type to SFC in combination with a zero value in the PAUSE field amounts to non-use of any flow control protocol and avoids specifying another Flow Control Type encoding.

When SFC is used, the FCATM_DE sets the PAUSE field to an appropriate value indicating the number of 512-bit-time units to pause transmission.

When SR protocol is used, the FCATM_DE prefixes a 4-byte SR_Header at the beginning of the encapsulated frame that is mapped into the payload of the SR_I message. The SR_Header indicates the type of SR message along with other control information.

NOTE 10 - All SR protocol-generated control messages carry neither the SOF and the EOF fields nor the FC frame headers in the payload.

7.3.4.5.2 Encapsulation using AAL5

The FCATM_DE encapsulates the BBW messages using ATM Adaptation Layer 5 (AAL5). The BBW messages form the payload of the AAL5 CPCS-PDU. The AAL5 layer pads the CPCS-PDU if necessary, and appends a trailer at the end. The AAL5 CPCS-PDU along with the trailer is then segmented into 48 bytes to form the SAR PDUs. A 5-byte ATM cell header is prefixed to each SAR PDU forming an ATM cell. See 7.5 .

7.3.4.5.3 ATM QoS

FC-BB-3_ATM recommends use of a single Virtual Circuit (VC). Use of additional VCs to address special traffic QoS requirements is allowed but not required. The SR protocol is separately applied to each VC that is used. In order to avoid starvation, a minimum bandwidth allocation **is recommended** for each FC-BB-3_ATM VC that is used. If more than one VC is used, then the service discipline prioritization for the VCs is implementation-specific.

In-order delivery is guaranteed within the scope of the ATM Virtual Connection (VC). Frames shall be shipped from the FC-BB-3_ATM in the same order as they are received.

7.3.4.5.4 Forwarding

The ATM cell is forwarded to the proper destination using the ATM destination address from a mapping table that corresponds to the D_ID address.

7.3.4.6 B_Access Virtual ISL and FCATM Links

The FC/FCATM Entity pair provides a data forwarding path between itself and a remote FC/FCATM Entity pair via virtual constructs. Two types of virtual constructs are defined:

a) a B_Access Virtual ISL is a logical construct that is created between two FC Entity B_Access portals for the explicit purpose of sending and receiving byte-encoded SOF/EOF delimited FC frames via the FCATM Entity. Conceptually, communication between two B_Access portals is similar to communication between two E_Ports; and

b) an FCATM Link is a logical construct that is created between two FCATM Entity LEPs for the explicit purpose of sending and receiving AAL5 encapsulated FC frames and AAL5 encapsulated FCATM control information. Conceptually, communication between two LEPs is similar to the communication between two instances of an ATM application.

There is a one-to-one mapping between a B_Access Virtual ISL and an FCATM Link.

Each FCATM Link consists of one or more ATM VCs, all between the same two FC-BB-3_ATM devices. Although more than one FCATM Link may be formed between a pair of FC-BB-3_ATM devices, a typical configuration may only consist of a single FCATM Link. See figure 15 for some examples of allowed network topologies. The FC-BB-3_ATM FCATM_LEP that originates an FCATM Link is defined as the FCATM Link Originator. The corresponding FCATM_LEP that accepts this link is defined as the FCATM Link Acceptor. An FCATM Link is fully characterized by its FCATM Link Originator and FCATM Link Acceptor identities. An FCATM Link Originator or FCATM Link Acceptor is fully identified by all of the following:

a) an 8-byte Fabric_Name;

b) an 8-byte B_Access_Name; and

c) an 8-byte FC/FCATM Entity identifier.

To uniquely identify an FCATM link, the following items are required:

a) the 8-byte Fabric_Name of the FCATM Link Originator;

b) the 8-byte B_Access_Name of the FCATM Link Originator;

c) the 8-byte FC/FCATM Entity identifier of the FCATM Link Originator; and

d) the 8-byte Fabric_Name of the FCATM Link Acceptor.

NOTE 11 - The FCATM Link Acceptor's 8-byte FC/FCATM Entity identifier and the B_Access_Name of the acceptor provide additional information about an FCATM Link but are not required to uniquely identify it.

7.3.4.7 CSM – Signaling

FC-BB-3_ATM supports the use of both Provisioned Permanent Virtual Connections (PVCs) or Switched Virtual Connections (SVCs) to transport messages. If PVCs are used, then no ATM signaling is required and connections are provisioned (i.e., preconfigured).

If SVC is used, then the Call Handling Function initiates the ATM User Network Interface (UNI) Signaling Protocol to set up a Virtual Connection (VC) (see ITU-T Q.2971) and the VC is torn down after use. The FC-BB-3_ATM shall use the ATM UNI signaling connection request messages to establish a connection and a traffic contract. The traffic contract establishes the FC-BB-3_ATM defined QoS and traffic parameters. If the requested connection is acceptable to the network, then a connection is set up between the FC-BB-3_ATMs. FC-BB-3_ATM shall support UNI 3.1 and higher. A dedicated channel (i.e., Virtual Path Identifier (VPI) = 0 and Virtual Channel Identifier (VCI) = 5) is reserved for signaling between the end user and the interfacing ATM device (switch). ATM connections allow traffic to flow in one or both directions (i.e., unidirectional or bi-directional) with the bandwidth the same or different in each direction. FC-BB-3_ATM requires bidirectional connectivity.

7.3.4.8 PMM

PMM functionality is reserved for future definition.

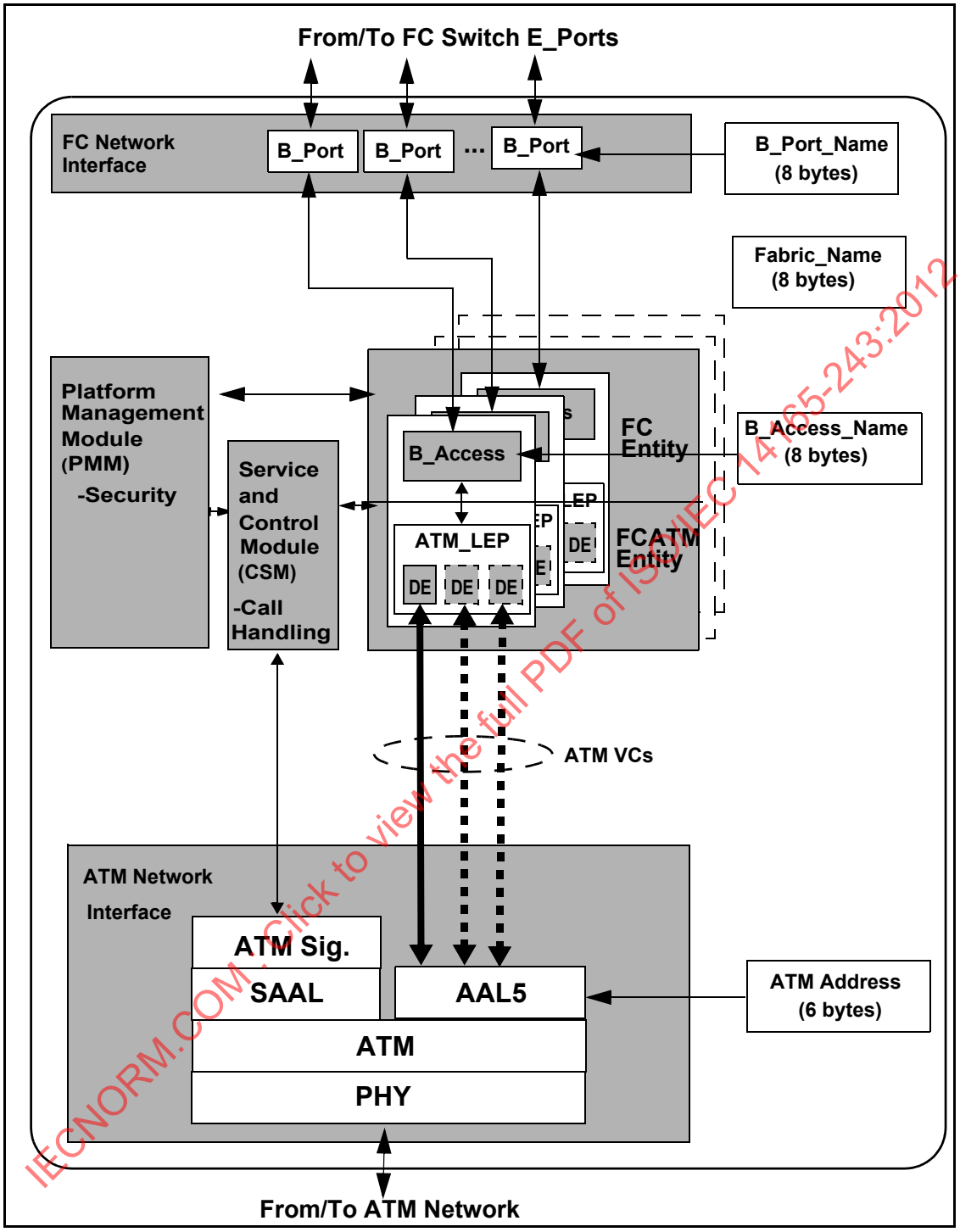


Figure 11 – FC-BB-3_ATM B_Access functional model

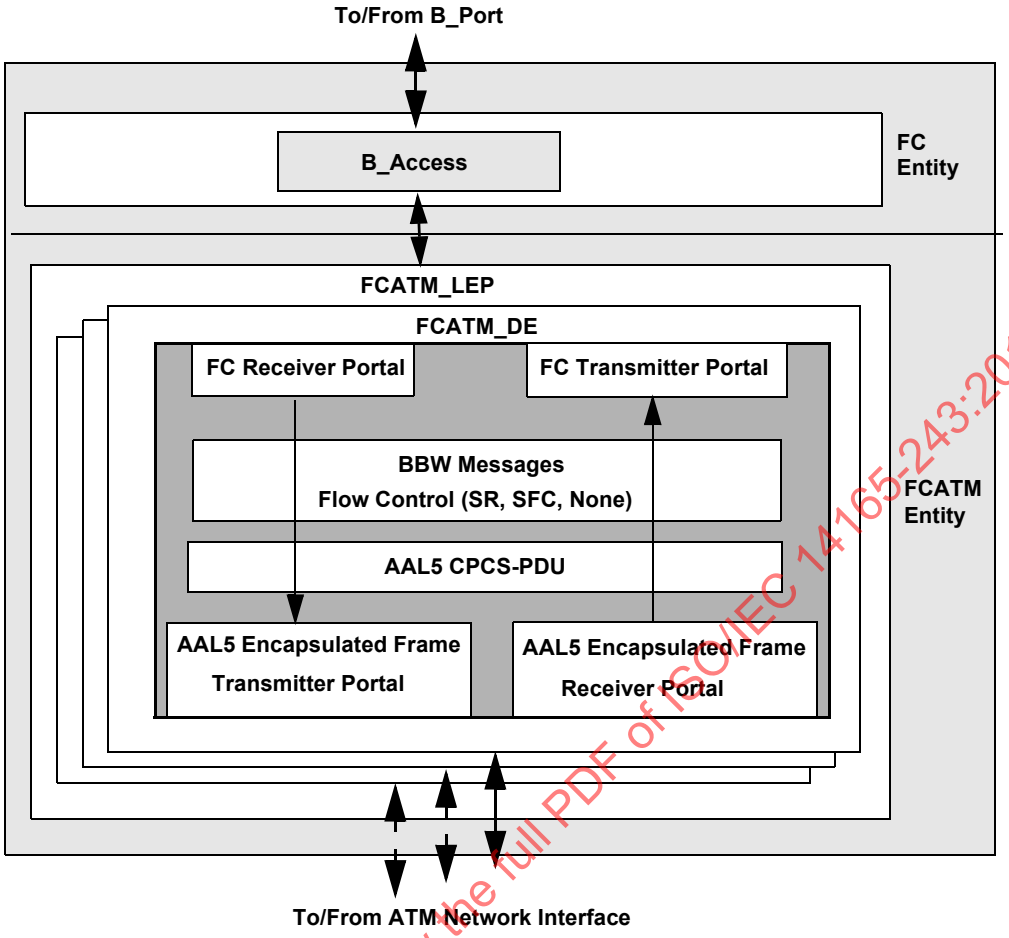


Figure 12 – FCATM_LEP and FCATM_DE

7.3.5 B_Access Virtual ISL exchanges – Exchange B_Access Parameters (EBP) SW_ILS

B_Access portals exchange SW_ILSs on the B_Access Virtual ISL. The Link Services that occur on the B_Access Virtual ISL are the EBP SW_ILS and the LKA ELS (see FC-LS). Figure 13 shows the scope of the FC end node frames, B_Access frames, B_Access Virtual ISL, FCIP Link and TCP connections.

.

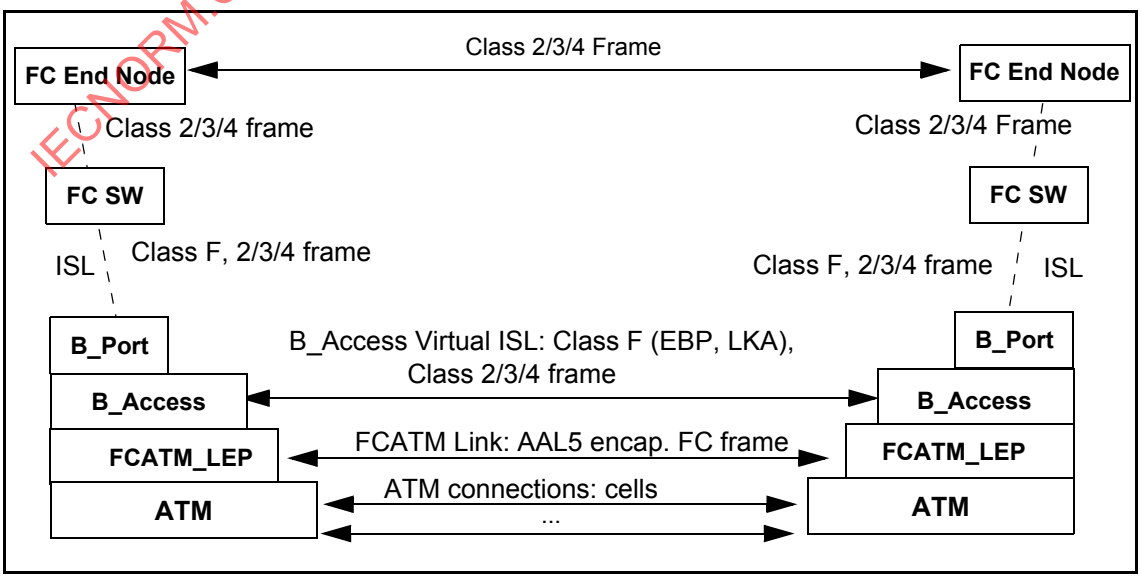


Figure 13 – Scope of B_Access Virtual ISL

The Exchange B_Access Parameters (EBP) Switch Fabric Internal Link Service (SW_ILS) is sent by a B_Access portal to a remote B_Access portal in order to establish operating link parameters and port capabilities for the B_Access Virtual ISL formed by the two B_Access portal peers. Successful acceptance of EBP SW_ILS shall be completed before the B_Ports begin switch port mode initialization.

Protocol: Exchange B_Access Parameters (EBP) request Sequence

Accept (SW_ACC) reply Sequence

Addressing: For use in switch port configuration, the S_ID field shall be set to FFFFFDh, indicating the Fabric Controller of the originating B_Access portal. The D_ID field shall be set to FFFFFDh, indicating the Fabric Controller of the destination B_Access portal.

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

Table 19 – EBP request payload

Item	Size Bytes
28 01 00 00h	4
R_A_TOV	4
E_D_TOV	4
K_A_TOV	4
Requester B_Access_Name	8
Class F Service Parameters	16

Requester B_Access_Name: This field shall contain the name of the device that originated the EBP request.

R_A_TOV: This field shall be set to the value of R_A_TOV required by the FC-BB-3_ATM device.

E_D_TOV: This field shall be set to the value of E_D_TOV required by the FC-BB-3_ATM device.

K_A_TOV: This field shall be set to the value of K_A_TOV required by the FC-BB-3_ATM device.

Class F Service Parameters: This field shall contain the B_Access Class F Service Parameters and its format is identical with its use in the ELP SW_ILS (see FC-SW-4).

Reply Switch Fabric Internal Link Service Sequence:

Service Reject (SW_RJT):

Signifies the rejection of the EBP command.

Accept (SW_ACC):

Signifies acceptance of the EBP command.

Accept Payload:

Addressing: For use in switch port configuration, the S_ID field shall be set to FFFFFDh indicating the Fabric Controller of the originating B_Access. The D_ID field shall be set to FFFFFDh, indicating the Fabric Controller of the destination B_Access.

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

Table 20 – EBP accept payload

Item	Size Bytes
02 00 00 00h	4
R_A_TOV	4
E_D_TOV	4
K_A_TOV	4
Responder B_Access_Name	8
Class F Service Parameters	16

The fields in table 20 are the same as defined for table 19 except for the Responder B_Access_Name field.

Responder B_Access_Name: This field shall contain the B_Access_Name of the remote device that responds to the EBP request.

The SW_RJT reply payload format is given in FC-SW-4. The EBP reject reason code explanation is shown in table 21.

Table 21 – EBP reject reason code explanation

Encoded Value (Bits 23-16)	Description
0000 0000	No additional explanation
0000 0001	Class F Service Parameter error
0000 0010	Invalid B_Access_Name
0000 0011	K_A_TOV mismatch
0000 0100	E_D_TOV mismatch
0000 0101	R_A_TOV mismatch
others	Reserved

7.3.6 B_Access initialization state machine

The B_Access initialization state machine is shown in figure 14.

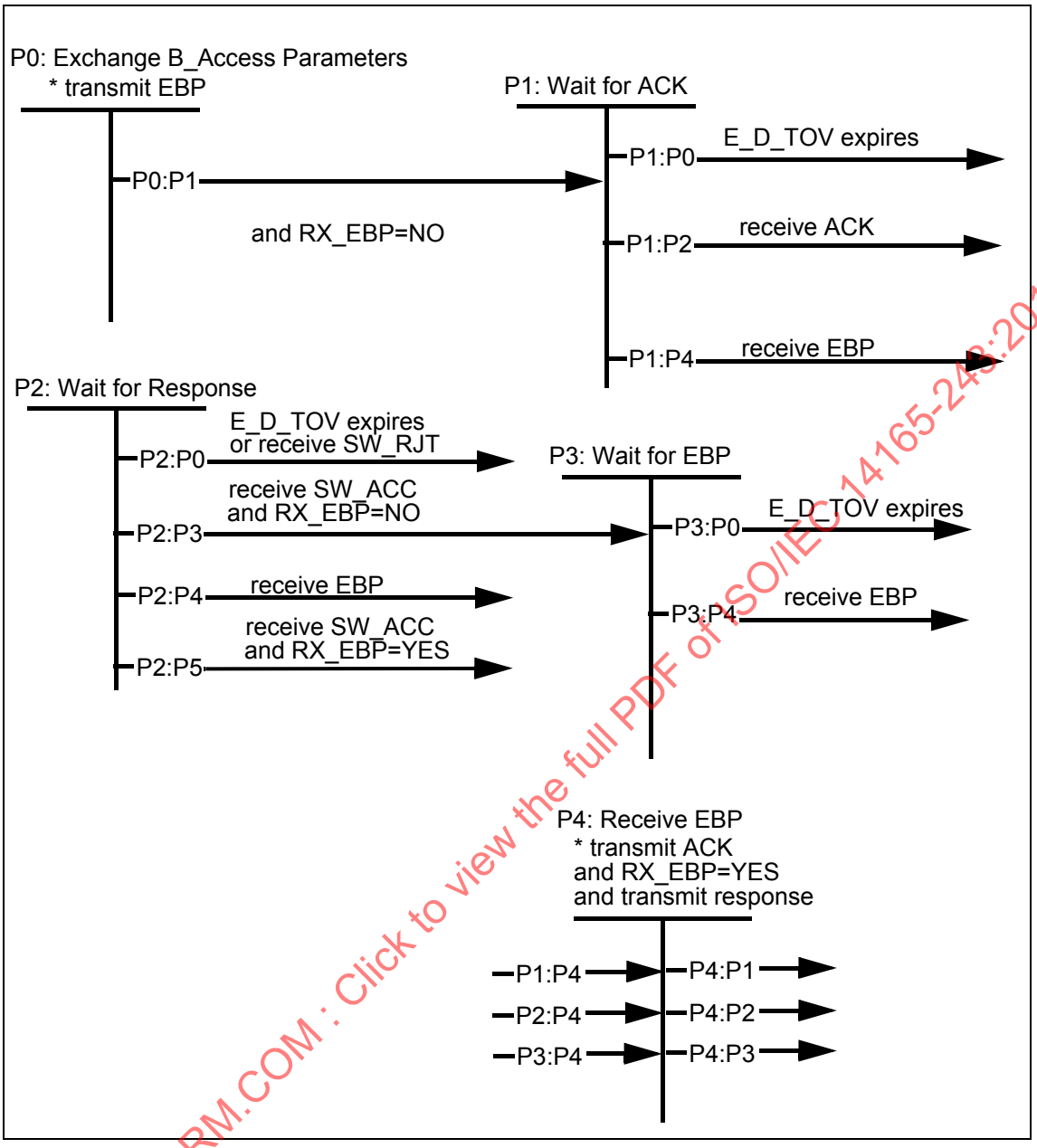


Figure 14 – B_Access initialization state machine

State P0: Exchange B_Access Parameters. This state marks the beginning of the B_Access initialization. Activity other than that described within the state machine is suspended until initialization is complete.

Transition P0:P1. The B_Access resets the RX_EBP flag.

State P1: Wait for ACK. In this state the B_Access waits until an ACK for the B_Access's transmitted EBP is received.

Transition P1:P0. This transition occurs when the B_Access has not received an ACK within E_D_TOV after the transmission of an EBP.

Transition P1:P2. This transition occurs when the B_Access receives an ACK before E_D_TOV expires.

Transition P1:P4. This transition occurs when the B_Access receives an EBP while waiting for an ACK.

State P2: Wait for Response. In this state the B_Access has received an ACK for its EBP and is waiting for a response.

Transition P2:P0. This transition occurs when the B_Access has not received a response within E_D_TOV after the transmission of an EBP or receives an SW_RJT.

Transition P2:P3. This transition occurs when the B_Access receives an SW_ACC and has not received an EBP.

Transition P2:P4. This transition occurs when the B_Access receives an EBP while waiting for a response.

Transition P2:P5. This transition occurs when the B_Access receives an SW_ACC and has received an EBP.

State P3: Wait for EBP. In this state the B_Access has received an ACK for its EBP and is waiting for an EBP.

Transition P3:P0. This transition occurs when the B_Access has not received an EBP within E_D_TOV of the transmission of an EBP.

Transition P3:P4. This transition occurs when a B_Access receives an EBP while waiting for a response.

State P4: Receive EBP. In this state the B_Access has received an EBP. The B_Access responds with an ACK and transmits an SW_ACC or SW_RJT depending upon whether or not the received configuration parameters contained within the EBP are acceptable. The B_Access sets RX_EBP to indicate an EBP has been received and is accepted.

Transition P4:P1. This transition occurs when a B_Access receives an EBP from its peer yet hasn't received an ACK for a previously transmitted EBP.

Transition P4:P2. This transition occurs when a B_Access receives an EBP from its peer yet hasn't received a response for a previously transmitted EBP.

Transition P4:P3. This transition should be removed from the diagram as it is termination point of the machine.

7.4 FC-BB-3_ATM network topologies

Figure 15 shows some example FC-BB-3_ATM network topologies that exist between three FC-BB-3_ATM sites.

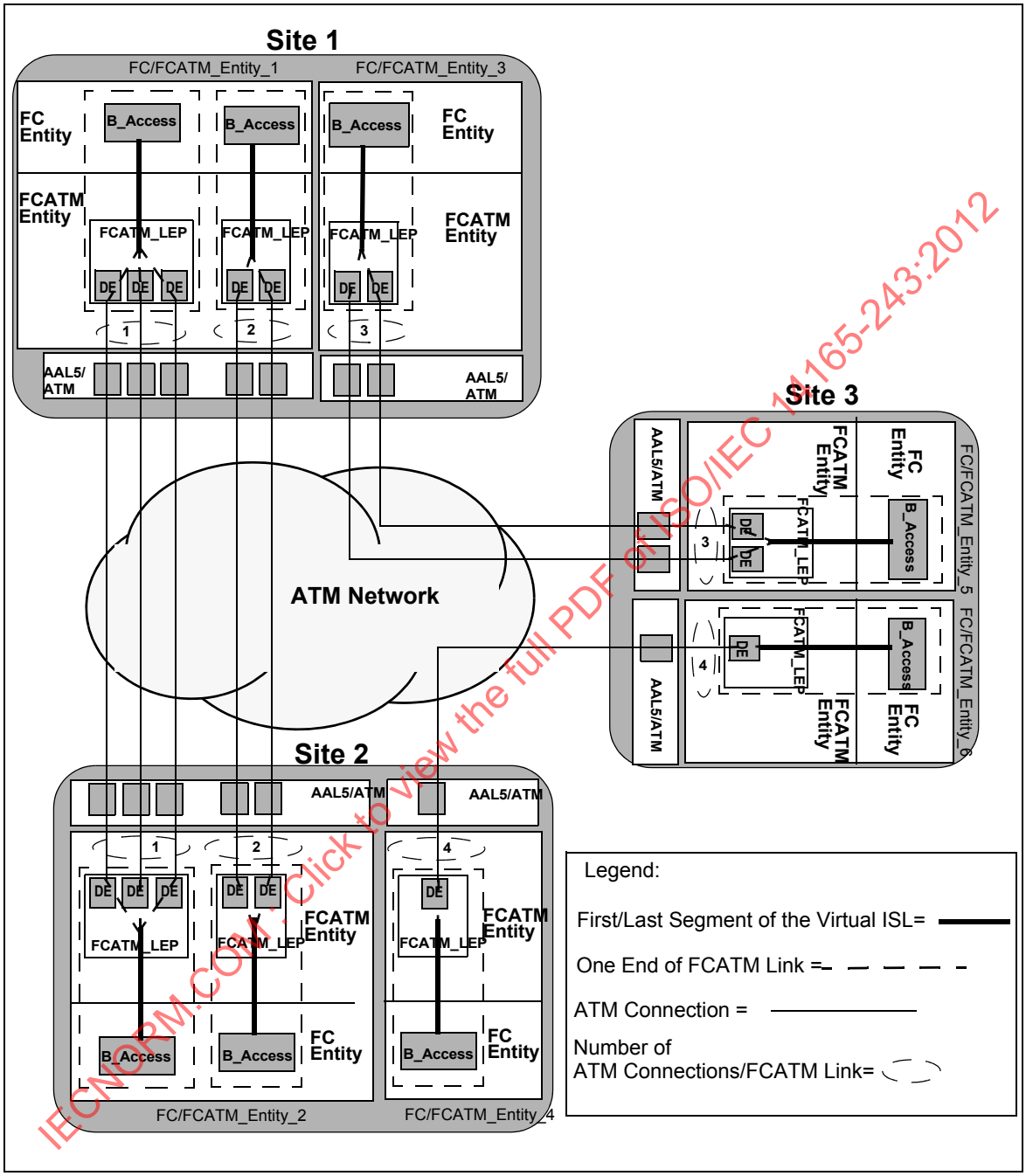


Figure 15 – FC-BB-3_ATM network topologies

As depicted in figure 15:

a) FCATM Link 1 connects Sites 1 and 2 and consists of three ATM connections;

b) FCATM Link 2 connects Sites 1 and 2 and consists of two ATM connections. FCATM Link 2, however, is distinct from Link 1 although it exists between the same two FC/FCATM Entity pairs (i.e., FC/FCATM_Entity_1 and FC/FCATM_Entity_2);

c) FCATM Link 3 connects Sites 1 and 3 and consists of two ATM connections. FCATM Link 3 exists between FC/FCATM_Entity_3 and FC/FCATM_Entity_5; and

d) FCATM Link 4 connects Sites 2 and 3 and consists of one ATM connection. FCATM Link 4 exists between FC/FCATM_Entity_4 and FC/FCATM_Entity_6.

7.5 Mapping and message encapsulation using AAL5

7.5.1 Overview

BBW messages are transparently transported over the ATM WAN. However, before it may be transported, it first has to be adapted. This adaptation is done using the ATM Adaptation Layer (AAL5). The AAL5 encapsulated BBW message is then segmented into ATM cells and routed to the proper destination ATM address.

7.5.2 Mapping BBW messages to AAL5

The BBW message is first mapped to a null AAL5 Service Specific Convergence Sublayer (SSCS) and then to a Common Part Convergence Sublayer (CPCS) to form the AAL5 CPCS-PDU that has a maximum size of 2 160/2 164 bytes. See note 12. The AAL5 CPCS-PDU is padded, if necessary up to 47 bytes, and then appended with an 8-byte CPCS-Trailer. The CPCS-PDU, CPCS-Pad, and CPCS-Trailer is then segmented into 48 bytes to form the Segmentation and Reassembly PDU (SAR-PDU). A 5-byte ATM cell header is attached to each SAR PDU to form an ATM cell.

CPCS-PDU: The BBW message maps into this field that consists of the LLC/SNAP Header, BBW_Header, and the BBW message payload.

CPCS-Pad: A CPCS-Pad ensures an exact mapping of the CPCS-PDU into SAR 48-byte payloads. A CPCS-Pad may range from 0-47 bytes. The maximum pad value of 47 bytes never occurs when the CPCS-PDU carries the BBW message payload because the payload is always a multiple of 4 bytes and aligned on a 4-byte boundary

CPCS-Trailer: A CPCS-Trailer is 8 bytes long and consists of a 1-byte User-to-User (UU) field, a 1-byte Common Part Indicator (CPI) field, a 2-byte length field, and a 4-byte CRC check sum field.

The UU and CPI fields are currently not used. The CPCS-PDU length field indicates the length in bytes of the CPCS-PDU payload. The length indicates the useful payload size. Therefore, the CPCS-PDU size may vary with byte increments. The AAL5 CRC field is set as defined in the relevant ATM standards (see table 22).

Table 22 – Mapping of BBW messages to AAL5 CPCS

Field	Item	Size Bytes
CPCS-PDU	LLC/SNAP Header	8
	BBW_Header	4
	BBW message payload (See Note 12)	Max: 2 148/2 152 (See note 13)
CPCS-Pad		0-47
CPCS-Trailer	Reserved (CPCS-UU, not used)	1
	Reserved (CPI, not used)	1
	CPCS-PDU Length (in bytes)	2
	CPCS-PDU CRC	4

NOTE 12 - If SR is used, then only the SR_I, SR_SREJ, and SR_FRMR carry a non-zero payload.

NOTE 13 - The maximum CPCS-PDU value indicated in the table is based on the maximum Fibre Channel frame size. CPCS-PDU for other non Fibre Channel transport may be much larger and up to 65 535 bytes. The maximum of 2 148 bytes of BBW message payload is due to a maximum of 2 112 bytes of FC frame payload, 4 bytes of SOF, 24 bytes of FC Header, 4 bytes of EOF, 4 bytes of CRC. If SR is used then 4 bytes of SR_Header yields a total maximum of 2 152 bytes.

Figure 16 illustrates the AAL5 Mapping for an FC frame when SFC is used.

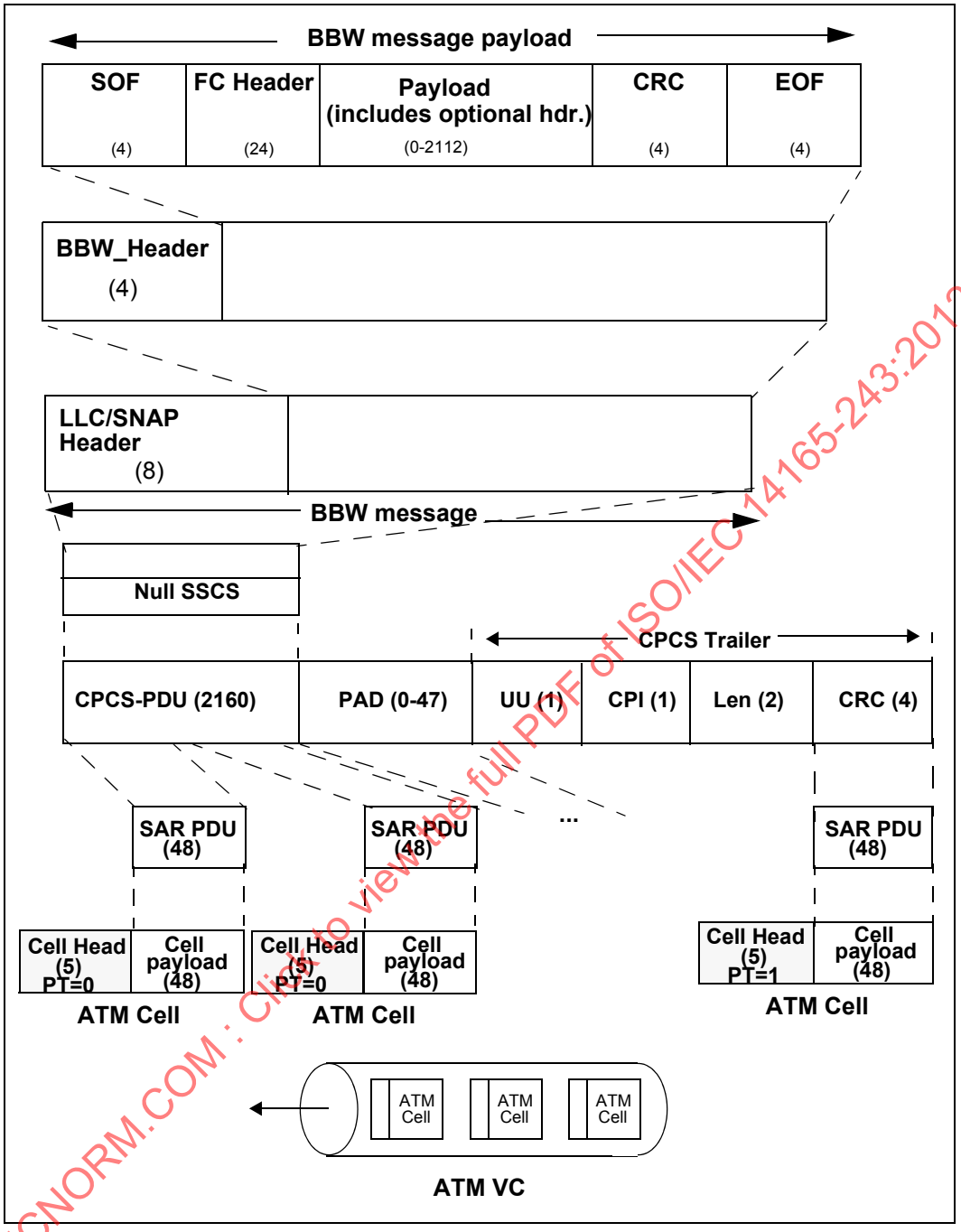


Figure 16 – AAL5 Mapping of a BBW message with SFC

Figure 17 illustrates the AAL5 mapping for an FC frame when SR is used.

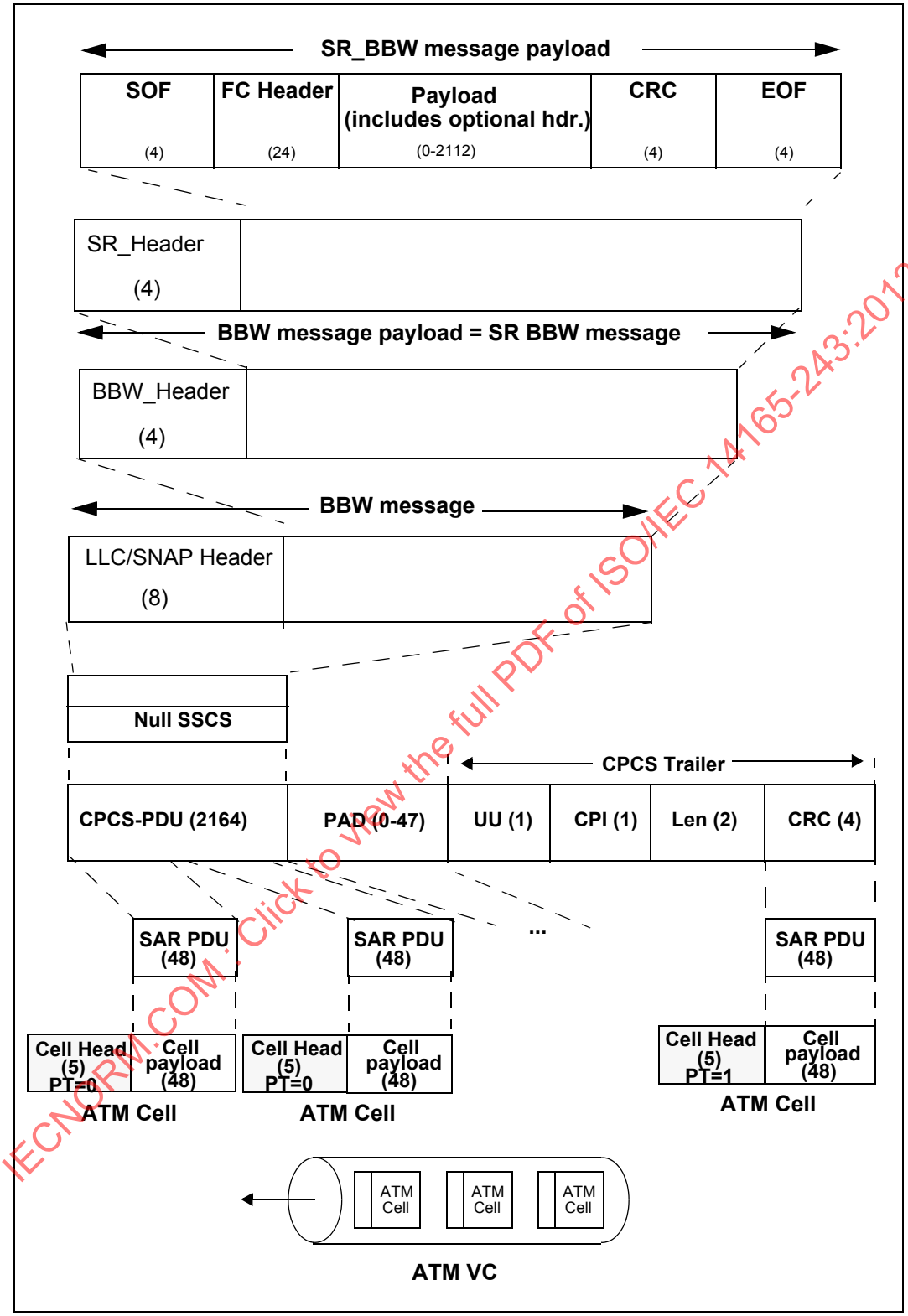


Figure 17 – AAL5 Mapping of a BBW message with SR

7.6 FC-BB-3_ATM service considerations

7.6.1 ATM service type

Different types of ATM service provide different levels of service features. FC-BB-3_ATM recommends use of the VBR-NRT ATM service or better. See annex B.

7.6.2 Latency delay and timeout value

FC-BB-3_ATM and the ATM network introduce latency delays that warrant special considerations with respect to Fibre Channel E_D_TOV and R_A_TOV values. The total path delay between a source FC-BB-3_ATM and a destination FC-BB-3_ATM consists of the latency delay components due to the queuing time at the FC-BB-3_ATM devices and all intermediate ATM switches, cell transmission time, propagation time and SVC setup time if applicable. It is recommended that this total path delay be less than 1/2 E_D_TOV to conform to normal Fibre Channel time out values.

NOTE 14 - VBR-NRT does not provide delay guarantees. Delay guarantees are practically realized by Service Level Agreements (SLAs) with the ATM Service Provider.

7.6.3 Bandwidth sharing and allocation

In ATM, bandwidth sharing is accomplished by multiplexing different upper layer traffic (e.g., Fibre Channel, IP). Multiplexing different upper layer protocol traffic may occur in two ways, multiplexing within a single VC and multiplexing using different VCs. The latter method is not recommended.

Multiplexing within a single VC, also referred to as **VC multiplexing,** is applicable for all traffic intended for the same destination and when using the same ATM service category. Upper layer multiplexed protocol data is distinguished based on the BBW_Header. The biggest reason to use VC multiplexing is to minimize the number of VCCs established especially in a PVC environment. FC-BB-3_ATM recommends using a single VC to multiplex all traffic.

Use of more than one VC to the same destination to address special traffic QoS requirements is allowed but introduces an increased level of complexity and is therefore not recommended. The SR flow control protocol is separately applied to each VC in such a case. A minimum bandwidth allocation **is recommended** for each VC that is used in order to avoid starvation. FC-BB-3_ATM does not specify any particular service discipline when more than one VC is used, but recommends a minimum bandwidth for each VC to protect it from starvation. This is illustrated in figure 18. The service discipline prioritization for the VCs is implementation specific and outside the scope of FC-BB-3.

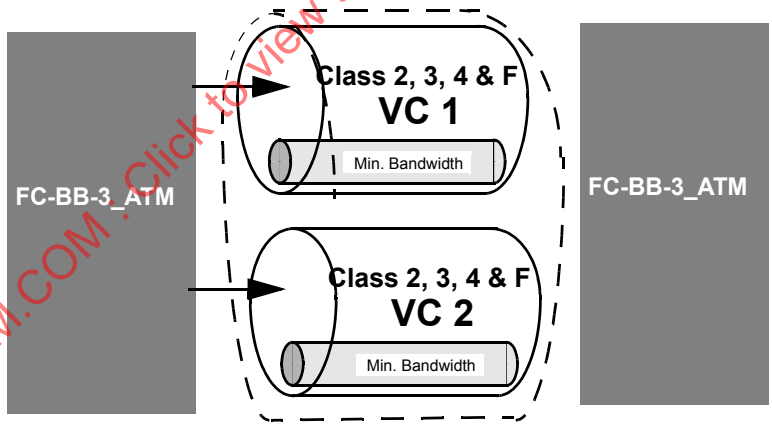


Figure 18 – Recommended ATM bandwidth allocation for multiple VCs

7.6.4 Quality of Service (QoS)

FC-BB-3_ATM specifies the use of VBR-NRT service for all VCs. Table 23 shows the QoS parameters and traffic descriptors specific to VBR-NRT and the guarantees provided by this service.

VBR-NRT service is best suited for non-time-based critical data, that require guarantees for loss and bandwidth but not delay. This service matches the requirements of FC Classes 2, 3, 4, and F.

QoS is a term used to refer to the set of performance characteristics of the contracted ATM connection. Although, a total of six QoS parameters are defined and available with other ATM Services, VBR-NRT only specifies a single QoS parameter (i.e., Cell Loss Ratio).

NOTE 15 - Some QoS parameters specified with other ATM Services include: Peak-to-peak Cell Delay Variation (CDV), Maximum Cell Transfer Delay (maxCTD). maxCTD provides delay guarantees. See annex B for details.

ATM Traffic Descriptor is a term used to describe the traffic characteristics of an ATM connection. A Connection Traffic Descriptor includes a Source Traffic Descriptor, CDV Tolerance (CDVT), and a Conformance definition. A Source Traffic Descriptor is described by four parameters: Peak Cell Rate (PCR), Sustainable Cell Rate (SCR), Maximum Burst Size (MBS), and a Minimum Cell Rate (MCR). See annex B for more details. Service guarantees are realized by these traffic descriptors.

Bandwidth guarantees are achieved by SCR, PCR, and MBR.

Table 23 – ATM VBR-NRT service specification

ATM Traffic Descriptors	VBR-NRT Service Category	Remark
QoS Parameters	CLR*	Cell Loss Ratio; guarantees Loss
Source Traffic Descriptors: (SCR, PCR, MBS guarantee bandwidth)	PCR*, CDVT*, SCR, MBS MCR	Peak Cell Rate, CDV Tolerance, Sustainable Cell Rate, Maximum Burst Size, Minimum Cell Rate;
Conformance Definition	GCRA*	Generic Cell Rate Algorithm (Leaky Bucket Algorithm)
NOTE 1 * Items are supplied by telco and are negotiable. NOTE 2 Cell Transfer Delay (CTD) (not in table) a QoS parameter associated with VBR-RT is also negotiable.		

7.6.5 Delivery Order

FC-BB-3_ATM shall guarantee in-order delivery of frames within a VC. No other ordering relationship among VCs is normally preserved or assumed. When the number of VCs is greater than 1, then the traffic management entity within the FC-BB-3_ATM device shall ensure that using separate VCs does not result in out-of-order delivery. In other words, once message transmission begins on a VC, then it shall continue using the same VC until completion of the message.

NOTE 16 - The out-of-sequence delivery problem associated with datagram networks is not present here. This benefit is a consequence of the strict requirement in ATM that requires all cells to always follow the same route during the call's duration. However, the possibility of missing or errored messages still remains and is addressed by the SR protocol.

7.6.6 Loss and Flow Control

ATM networks are lossy and they may drop cells, typically due to network congestion. When a cell loss occurs, the end applications are expected to recover from this loss. Recovery from such losses occurs at the FC-BB-3_ATM devices using the SR protocol that supports error recovery.

NOTE 17 - The SFC protocol has no error recovery support.

Use of a flow control protocol (i.e., SFC or SR) at the FC-BB-3_ATM device provides the ability to deal with speed mismatches between the FC and the ATM interface.

8 FC-BB-3_SONET Structure and Concepts

8.1 Applicability and related clauses

Clause 4 discusses the FC-BB-3_SONET reference model. Clause 5 describes the required messages and clause 6 describes the flow control mechanisms applicable to FC-BB-3_SONET. This clause discusses the FC-BB-3_SONET functional model.

8.2 FC-BB-3_SONET overview

FC-BB-3_SONET is a Fibre Channel backbone transport protocol that tunnels HDLC encapsulated FC frames across the SONET/SDH network. Figure 19 shows a network configuration consisting of three FC-BB-3_SONET devices. An FC-BB-3_SONET device has interfaces to both the SONET network and the FC fabric. The FC fabric interface supports multiple B_Ports. The model applies equally well to both private and public SONET/SDH networks.

FC-BB-3_SONET devices that support B_Ports do not require FC switching. FC-BB-3_SONET protocol communication occurs between pairs of FC-BB-3_SONET devices. Although communication occurs between pairs of FC-BB-3_SONET devices, a single FC-BB-3_SONET device may communicate with more than one device at the same time.

NOTE 18 - The current scheme allows an FC-BB-3_SONET device to independently connect to more than one FC-BB-3_SONET device, but does not specify a point-to-multipoint connection.

No distinction is made in this standard regarding the topology of the SONET/SDH network, be it point-to-point using path-terminationg equipment pairs, hub networks, or ring architectures. The current model supports a configuration of one or more point-to-point connections only.

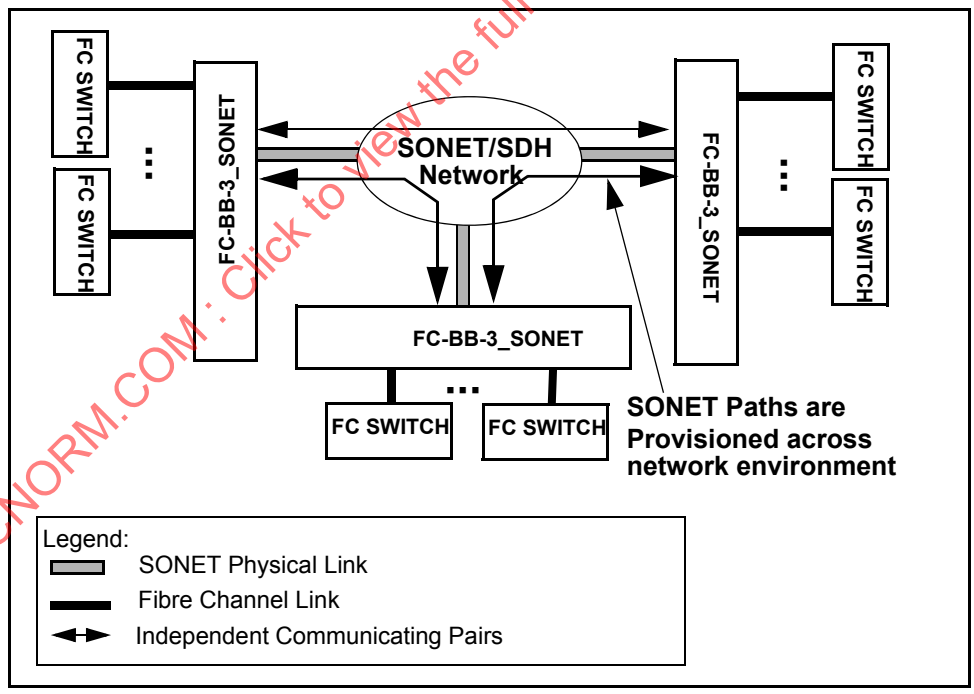


Figure 19 – FC-BB-3_SONET network configuration

The FC-BB-3_SONET protocol creates BBW messages that consist of an 8-byte LLC/SNAP Header and a 4-byte BBW_Header followed by the BBW message payload. The specific format and content of the BBW message payload depends on the type of flow control protocol used. The BBW message payloads carry byte-encoded SOF/EOF delimited Class 2, 3, 4, or F FC frames.

The BBW messages are encapsulated in HDLC-like format for carriage over the SONET/SDH network. The HDLC-encapsulated BBW messages are mapped into SPE/Virtual Containers and finally transmitted to the destination. HDLC encapsulation is the typical method of preparing frames for transmission over SONET/SDH

and is described in RFCs 1662 and 2615. FC-BB-3_SONET does not interpret the data content of the FC frames other than capturing and retaining their SOF/EOF identities in the encapsulated FC frame. As such, FC Sequences and Exchanges are not visible to the FC-BB-3_SONET protocol. All HDLC encapsulated FC frames are transparently transported over the SONET/SDH network.

Prior to an FC-BB-3_SONET transmitting data to a remote FC-BB-3_SONET, the required provisioning of the SONET/SDH path to the remote BBW needs to be completed. The details of this configuration are dependent upon the network topology and are beyond the scope of this standard.

All FC frames are encapsulated with the All-Stations address a binary sequence 11111111b (hexadecimal FFh) in the HDLC header therefore there is no requirement for an FC-BB-3_SONET to examine the destination address (D_ID) field in the Fibre Channel frame header. Frames are simply forwarded to the attached FC-BB-3_SONET egress device across the SONET network.

The LLC/SNAP Header indicates the payload type as Fibre Channel (see 5.2). The BBW_Header indicates the type of flow control used, Selective Retransmission (SR), Simple Flow Control (SFC), or none. The SR protocol makes the transport of FC frames between two FC-BB-3_SONETs reliable. The SR protocol supports both flow control and error recovery functions. Use of the SR protocol is optional. When SR flow control is used, the 4-byte BBW_Header is followed by a 4-byte SR_Header which is prefixed at the begin of the BBW message payload (see 5.5). The SFC protocol provides a mechanism to temporarily pause the transmission of frames from a remote BBW device. Use of the SFC protocol is optional. When SFC is used, the 4-byte BBW_Header is directly followed by the BBW message payload (see 5.4). No SFC header is prefixed or used.

In-order delivery is guaranteed for each BBW message and frames shall be transmitted from the FC-BB-3_SONET in the same order as they are received.

8.3 FC-BB-3_SONET functional model

8.3.1 Fibre Channel network interface

Figure 20 shows a functional model of the FC-BB-3_SONET. The Fibre Channel interface nominal port rate is assumed to full-rate, unless otherwise specified.

The FC-BB-3_SONET FC interface supports one or more B_Ports thus requiring the support of the FC-0, FC-1, and FC-2 levels. The B_Ports in general connect to different external FC switches, but connectivity to the same FC switch is also allowed. B_Ports are uniquely identified by an 8-byte B_Port_Name.

FC-BB-3 initialization occurs across the B_Port interface facing the FC network. The initialization of any generic B_Port is described in FC-SW-4. A B_Port indicates, using the ELP/ESC SW_ILS, that it is capable of parameter negotiation. Since FC-BB-3 does not support Class 1, the Class 1 Port Parameter VAL bit in the ELP shall be set to 0 (i.e., invalid). An ELP received at a B_Port may be rejected (i.e., SW_RJT) due to many reasons, including port-mismatch.

NOTE 19 - Initialization across the SONET WAN interface may use mechanisms similar to the one described in 9.4.3.3.2.

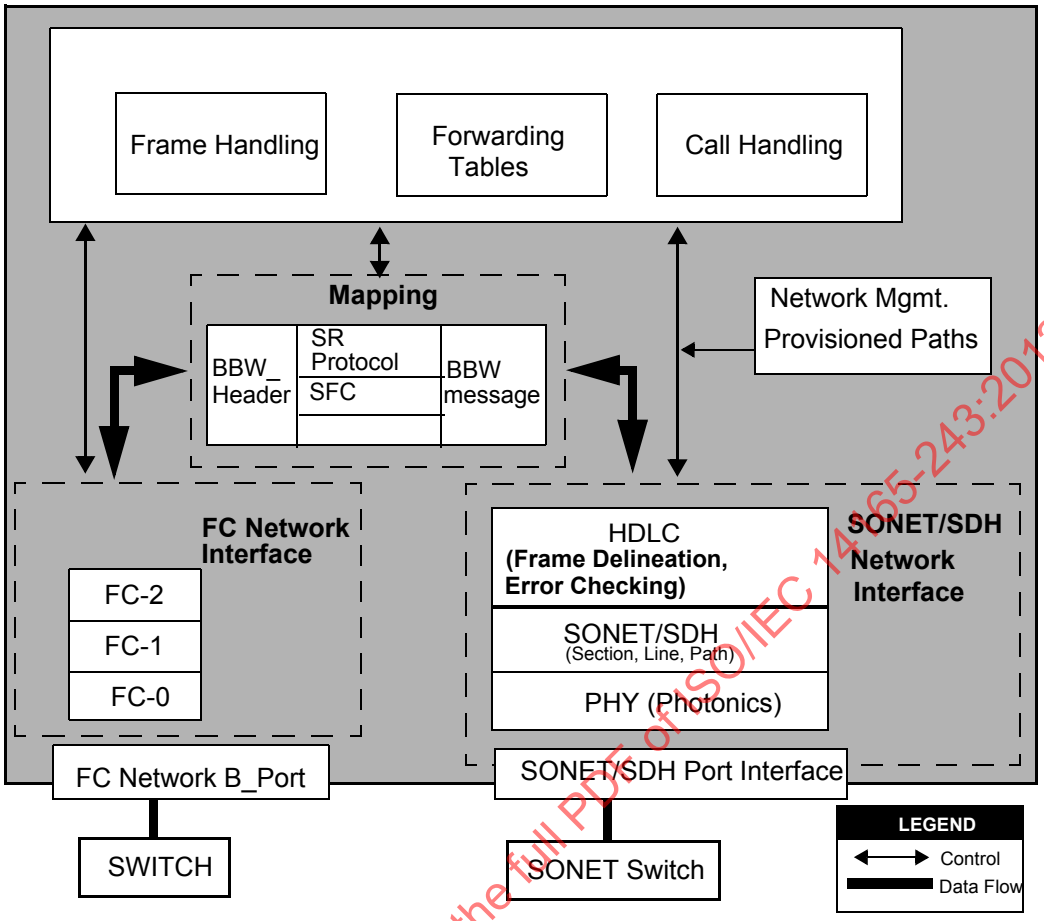


Figure 20 – FC-BB-3_SONET functional block diagram

8.3.2 SONET network interface

The SONET rate is nominally assumed to be STS-3c/STM-1 at 155.52 Mbit/s or higher. In the case of STS-3c/STM-1, the available information bandwidth is 149.760 Mbit/s, which is the STS-3c/STM-1 SPE with section, line, and path overhead removed. This is the same super-rate mapping that is used for ATM and FDDI. While the STS-3c/STM-1 rate is specified as the basic rate, the mapping specified within this standard is extended down to the STS-1 SONET rate (i.e., 51.84 Mbit/s).

Higher signal rates shall conform to the SDH STM series, rather than the SONET STS series. The STM series progresses in powers of 4, instead of 3, and employs fewer steps, which simplifies multiplexing and integration. For applications of Fibre Channel over SONET/SDH, it is envisioned that higher rates such as 622.08 Mbit/s and 2 488.32 Gbit/s may be developed and deployed as indicated in table 24.

Table 24 – SONET/SDH data rates

SONET	SDH Equivalent	Basic Rate
STS-3c-SPE	VC-4	155.52 Mbit/s
STS-12c-SPE	VC-4-4c	622.08 Mbit/s
STS-48c-SPE	VC-4-16c	2.4 Gbit/s
STS-192c-SPE	VC-4-64c	9.95 Gbit/s

Mappings for sub-STS-1 rates and rates of STS-192c/STM-48 or greater require further study and are beyond the scope of this standard.

The SONET/SDH interface includes the photonic, section/line/path, and HDLC encapsulation layers. The basic FC-BB-3_SONET reference model supports one SONET port using different rates.

The HDLC layer is used to prepare FC frame payloads for transport in SONET/SDH payload envelopes.

8.3.3 Mapping and encapsulation

FC-BB-3_SONET creates the 8-byte LLC/SNAP Header and the 4-byte BBW_Header that are prefixed to the BBW message payload. The specific format and content of the BBW message payload depends on the type of flow control protocol used. The BBW message payload carries the byte-encoded SOF/EOF delimited Class 2, 3, 4, or F FC frames.

When flow control is not used, FC-BB-3_SONET sets the PAUSE field in the BBW_Header to a zero value and the Flow Control Type to SFC.

NOTE 20 - This setting of Flow Control Type in combination with a zero value in the PAUSE field amounts to non use of any flow control protocol and avoids specifying another Flow Control Type encoding.

When SFC is used, FC-BB-3_SONET sets the PAUSE field to an appropriate value indicating the number of 512-time units to pause transmission. See 6.4 .

When SR Protocol is used, FC-BB-3_SONET prefixes a 4-byte SR Header at the beginning of an encapsulated frame that is mapped into the payload of the SR_I message. The SR Header indicates the type of SR message type along with other control information. See 6.2 and 6.3.

See 8.3.6 for details on encapsulation using HDLC-like framing.

8.3.4 FC-BB-3_SONET forwarding

FC-BB-3_SONET forwards FC frames that enter its B_Ports to a remote FC-BB-3_SONET using a mapping table that contains a list of FC-BB-3_SONET devices corresponding to a list of D_ID addresses.

8.3.5 Call handling

FC-BB-3_SONET provides a point-to-point service for all classes of FC frames transmitted between two switches.

8.3.6 Frame handling

Frame handling is mainly concerned with the following two tasks:

a) processing the incoming FC frames from the external switch that emerges from the FC-2 level that has to be transported across the WAN. Processing includes tasks such as BBW header and message generation, and mapping to HDLC and SONET SPE; and

b) processing the FC-BB-3_SONET message that has successfully made it across the WAN and that is to be sent to the external switch. Processing includes decoding the SONET SPE containing the HDLC frames, decoding the BBW message, and removing the message headers.

8.4 Mapping and Message encapsulation using HDLC-like framing

8.4.1 Overview

BBW messages are transparently transported over the SONET WAN. However, before BBW messages may be transported, they first have to be adapted. This adaptation is done using the HDLC layer. Similar to Packet-over-SONET and Frame Relay-over-SONET, the FC-BB-3_SONET model is based on the HDLC-like framing used in PPP-over-SONET/SDH, and described in RFC 1662. The BBW messages form the payload of the HDLC frame that is mapped into SPE/Virtual Containers.

An alternative method for adaptation/mapping of BBW messages to SONET/SDH, also applicable to OTN and PDH transport infrastructures, that uses the Frame Mapped Generic Framing Procedure, is also available and is defined in ITU-T Rec. G.7041/Y.1303. Where the Frame Mapped Generic Framing Procedure is used for BBW adaptation in FC-BB-3_SONET, the HDLC-specific descriptions of this standard shall not apply.

8.4.2 Mapping of BBW messages to HDLC format

Table 25 shows the mapping of the BBW message to HDLC format according to RFC 1662. The contents of the fields are transmitted from left to right. HDLC framing provides for the delineation of the SONET payloads using a technique called ‘stuffing/unstuffing.’ Each HDLC frame begins and ends with the flag sequence. During transmission, if the flag sequence occurs anywhere within the information field of the HDLC frame, it is

changed to an escape sequence. At the receiver, the escape sequences are removed and replaced with the original fields. A 32-bit FCS is calculated across the HDLC frame for error-checking purposes.

The HDLC frames are then mapped byte-synchronously into the SONET SPE / SDH Virtual Container including any necessary inter-frame byte stuffing. The STS-SPE/SDH Higher-Order VC is then scrambled using the self-synchronizing $x^{43}+1$ scrambler. Since the FC-BB-3_SONET interface is comprised of path terminating equipment, the SONET section, line, and path layers (i.e., regenerator, multiplex and path layers for SDH) are required. Any of the many physical interfaces specified for SONET and SDH may be accommodated depending on the distances required and the WAN service offering being utilized.

Flag sequence: The flag sequence is used to encapsulate and delineate the HDLC frame (i.e., frame synchronization). Each frame begins and ends with the flag sequence 7Eh. If a frame immediately follows another, one flag sequence may be treated as the end of the preceding frame and the beginning of the immediately following frame (i.e., there does not need to be two flags separating the frames). When there are no HDLC frames to be transmitted, the flag sequence is to be transmitted continuously in the SONET/SDH envelope/ VC. Back-to-back flags are considered empty frame indications.

Address: The Address field contains the destination HDLC address. The address FFh is an All Stations Address / Broadcast Address. Any station on the link connection shall accept this address. Frames with invalid addresses are silently ignored.

Control: The Control field identifies the HDLC frame type (i.e., information, supervisory, unnumbered). The Control field of 03h is the Unnumbered Information (UI) command. Unnumbered frames are used for transferring data when the location of the data in a sequence of frames is not to be checked (i.e., no send or receive counts are utilized).

Protocol: The Protocol field is as defined in RFC 1661. It is one or two octets, and its value identifies the payload encapsulated in the Information field. The field is transmitted and received most significant octet first.

Table 25 – Mapping of BBW messages to HDLC format

Field		Encoding (hex)	Size (Bytes)	Remarks
Begin Flag		7Eh	1	
Address		FFh	1	Set to FFh for Broadcast
Control		03h	1	Only Information Type used
Protocol			2	
BBW message	LLC/SNAP Header		8	
	BBW_Header		4	
	BBW message payload		Max: 2 148/2 152 (See Note 21)	Variable Length
FCS			4	
End Flag		7E	1	
Fill or Address			>=1	Inter-frame fill or next Address

NOTE 21 - The maximum of 2 148 bytes is due to a maximum of 2 112 bytes of FC frame payload, 4 bytes of SOF, 24 bytes of FC Header, 4 bytes of EOF, 4 bytes of CRC. If SR is used then 4 bytes of SR_Header yields a total maximum of 2 152 bytes.

Information: The Information field contains the BBW message.

Frame Check Sequence (FCS): By default, the 32-bit Frame Check Sequence (FCS) field is required as described in RFC 1662. The FCS is calculated most-significant byte to least-significant byte and from least-significant bit to most-significant bit within each such byte over all bits of the address, control, and information fields prior to escape conversions. The least significant byte of the result is transmitted first as it contains the coefficient of the highest term. The FCS is calculated based upon the following polynomial:

$x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$

Inter-frame fill: A sending FC-BB-3_SONET shall continuously transmit the flag sequence as inter-frame fill after the FCS field. The inter-frame flag sequences shall be silently discarded by the receiving station. When an underrun occurs during DMA in the sending FC-BB-3_SONET, it shall abort the frame transfer by continuously transmitting the flag sequence.

Framing and byte stuffing: The framing and byte stuffing for octet-oriented synchronous links are described in RFC 1662. HDLC frames (i.e., octet streams) are mapped into the SONET STS-SPE/SDH Higher Order VC with octet boundaries aligned using $x^{43}+1$ scrambling.

Escape sequences are defined to minimally escape the flag sequence and control escape octet. Prior to sending the frame, but after the FCS computation, every occurrence of the flag sequence, control escape octet, or Async-Control-Character-Map (ACCM) found within the octets of the payload are converted to a two-octet sequence that includes the control escape octet followed by the original octet exclusive-or'd with 20h. For example:

a) 7Eh is encoded as 7Dh, 5Eh (flag sequence);

b) 7Dh is encoded as 7Dh, 5Dh (control escape); and

c) 03h is encoded as 7Dh, 23h (ETX).

Upon receiving a frame, this conversion shall be reversed prior to FCS computation.

Abort sequence: A flag sequence inserted into the octet stream between the initial frame flag sequence and the FCS constitutes sequence abort. The receiver considers the frame invalid until a subsequent flag sequence is found in the octet stream.

For example, when an underrun condition occurs at the sending station (e.g., the sending station cannot complete the data transfer for one reason or another), the sending station transmits a control escape octet followed immediately by the flag sequence, the frame is ignored and not counted as an FCS error.

8.4.3 Mapping HDLC frames to SONET/SDH

The mapping of HDLC framed signals according to ISO/IEC 3309 is performed by aligning the byte structure of every HDLC frame with the byte structure of the SONET SPE / SDH Virtual Container. The HDLC frames are located by row within the SPE payload. Since the HDLC frames are of variable length (i.e., this mapping does not impose any restrictions on the maximum length), a frame may cross the SPE/Virtual Container frame boundary. See figure 21.

The HDLC flag sequence shall be used for inter-frame fill to buffer out the asynchronous nature of the arrival of the HDLC framed SONET PDUs according to the effective payload of the SPE/Virtual Container used, this excludes any fixed stuff bytes.

The HDLC framed signal plus the inter-frame fill shall be scrambled before they are inserted as payload of the SPE/Virtual Container used. In the reverse operation, following termination of the SPE/Virtual Container signal, the payload shall be descrambled before it is passed on to the HDLC mapping layer. A self-synchronizing scrambler with generator polynomial $x^{43} + 1$ (see RFC 1619) shall be used. Scrambling of the HDLC framed signal is required to provide security against emulation of the SONET/SDH set-reset scrambler pattern and replication of the STM-N frame alignment word.

The $x^{43} + 1$ scrambler shall operate continuously through the bytes of the SPE, bypassing bytes of SONET Path Overhead. The scrambling state at the beginning of an SPE shall be the state at the end of the previous SPE. Thus, the scrambler runs continuously and is not reset per frame. An initial seed of the scrambler is unspecified. Consequently, the first 43 transmitted bits following start-up or a SONET/SDH re-frame operation shall not be descrambled correctly.

The $x^{43} + 1$ scrambler operates on the input data stream with the most significant bit (MSB) first, consistent with the bit ordering and transmission ordering defined for SONET in ANSI T1.105.

The above mapping procedure shall be used for the mapping of HDLC framed signals in SONET STS-3c, STS-12c, and STS-48c SPEs and for equivalent SDH Virtual Containers.

.

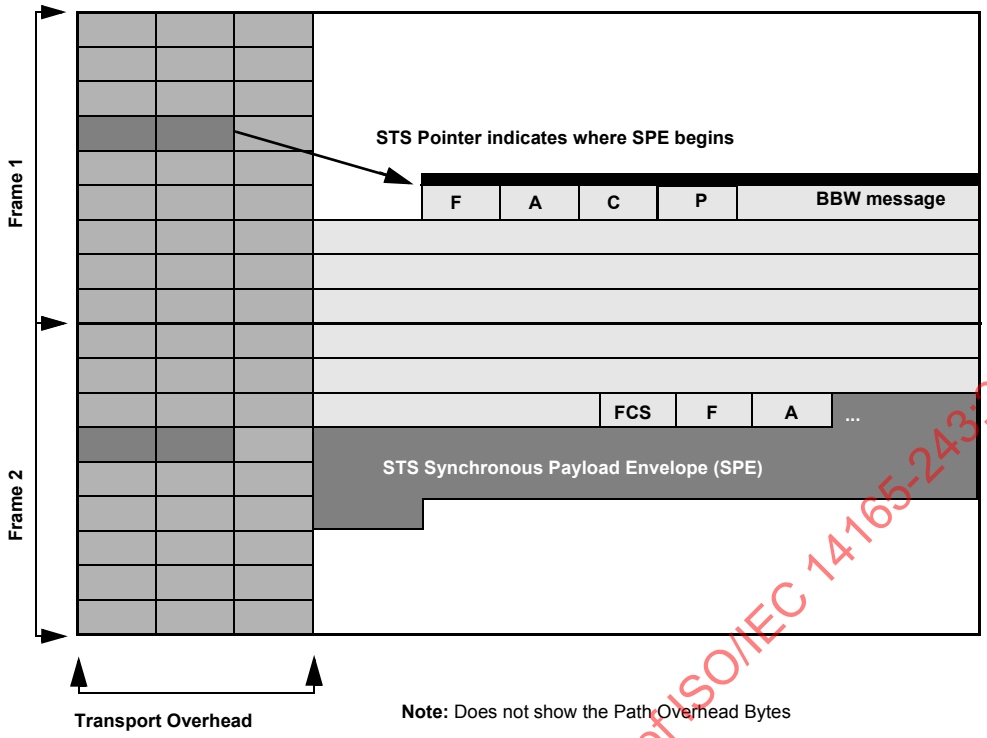


Figure 21 – SONET SPE HDLC mapping example

The Path Signal Label (C2) (see figure 22) indicates the contents of the SPE/Virtual Container. The value of 22d (16h) shall be used to indicate a variable-length HDLC frame with $x^{43} + 1$ scrambling enabled. Implementations shall not use a Path Signal Label (C2) value of 207 (CFh) that indicates a variable-length packet or frame without scrambling. The Multi-frame Indicator (H4) is unused, and shall be zero.

00010110b	16h	Mapping of HDLC framed signal

Figure 22 – Path signal label: C2

Table 26 shows the FC-BB-3_SONET protocol stack.

Table 26 – FC-BB-3_SONET protocol ptack

Interface Layer	Functionality
HDLC Mapping	– Frame Delineation – Link & Mapping Error Checking
SONET/SDH (Section, Line, path)	SONET/SDH – Section layer – Line layer – Path layer
Photonics	Optical layer

Figure 23 illustrates the encapsulation of BBW message into HDLC frame using SFC.

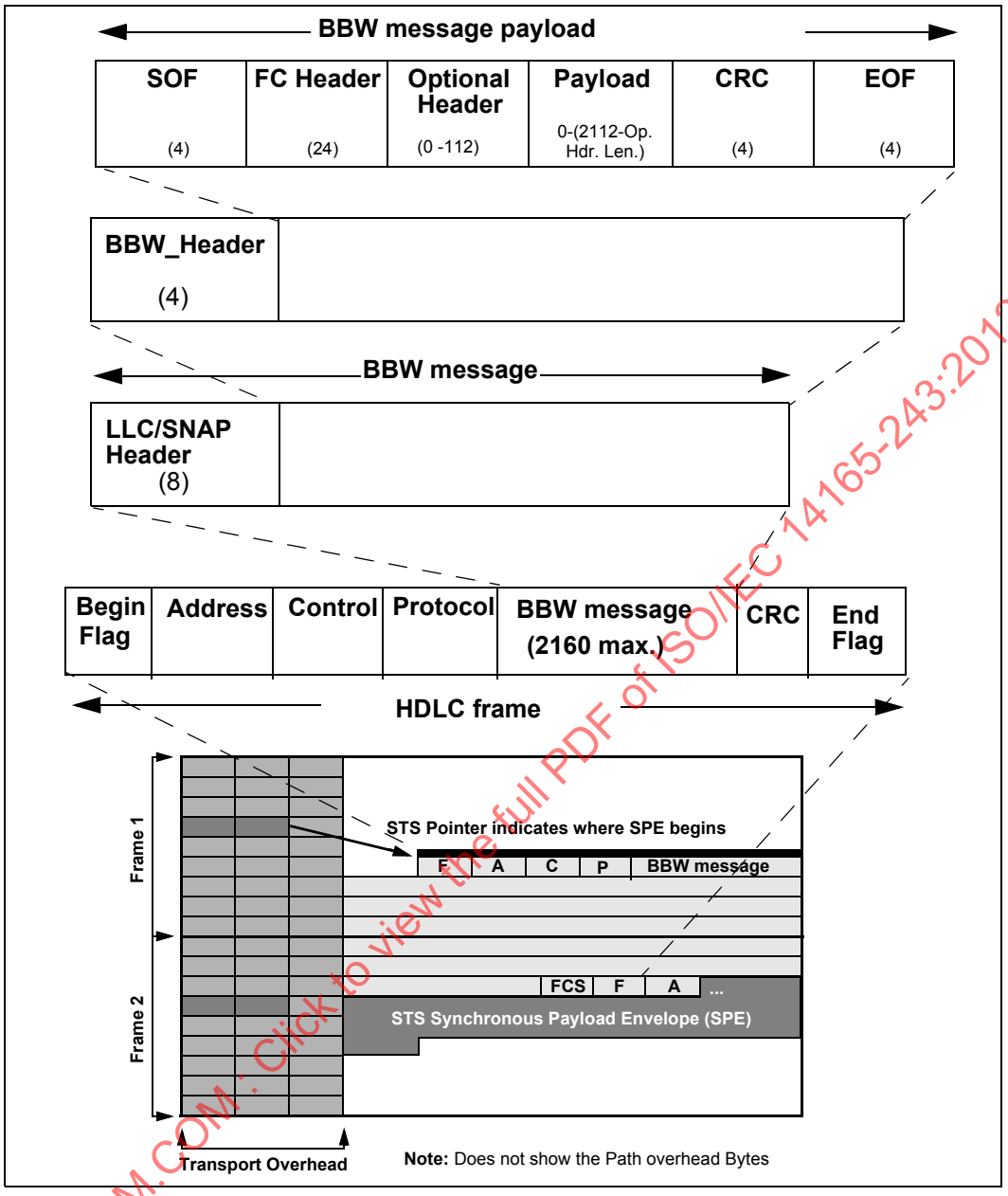


Figure 23 – Encapsulation of BBW message into HDLC frame using SFC

Figure 24 illustrates the encapsulation of the BBW message into HDLC frame using SR.

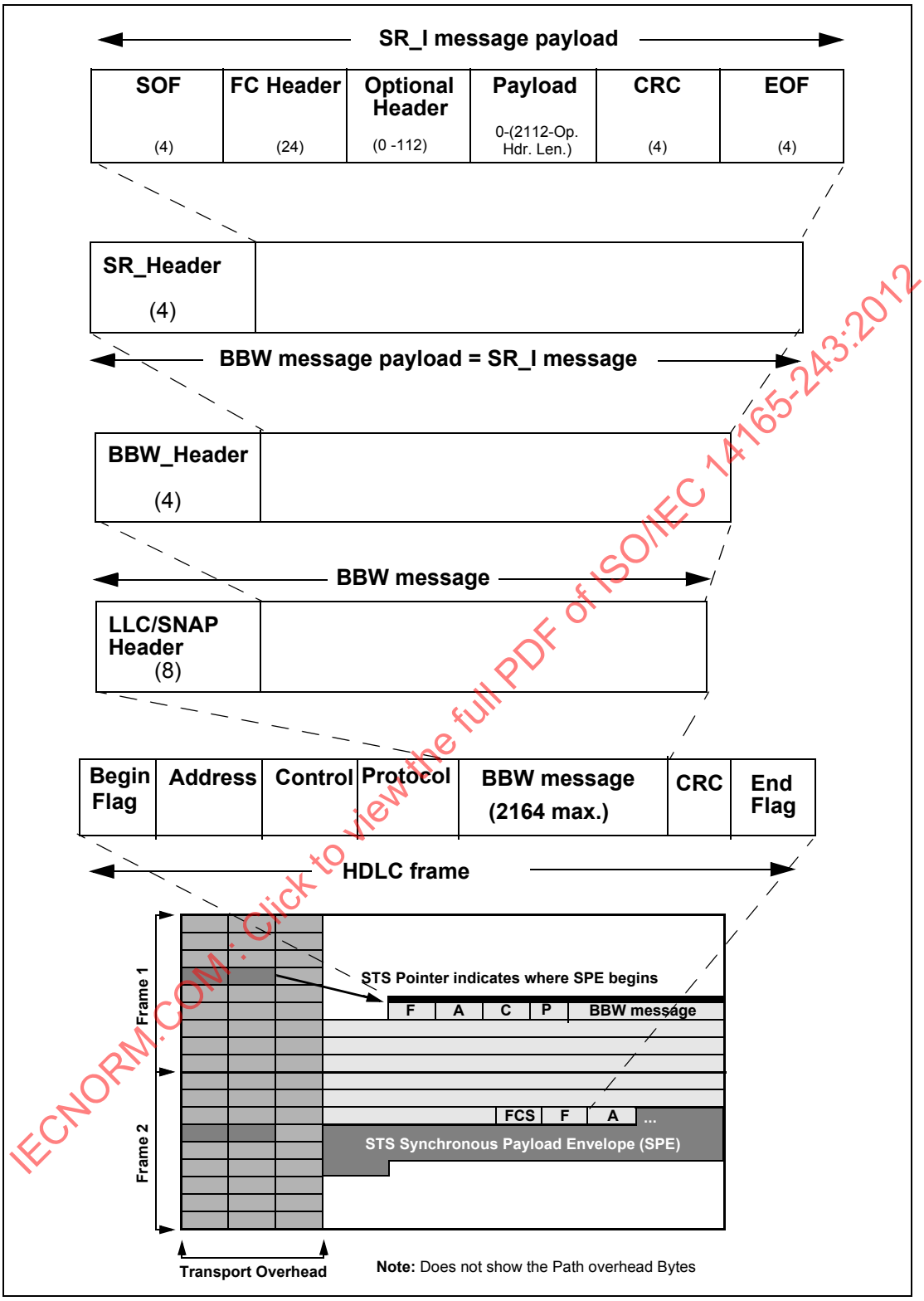


Figure 24 – Encapsulation of BBW message into HDLC frame using SR

8.5 FC-BB-3_SONET service considerations

8.5.1 Latency delay and timeout value

FC-BB-3_SONET and the SONET/SDH network introduce latency delays that warrant special considerations with respect to Fibre Channel E_D_TOV and R_A_TOV values. The total path delay between a source FC-BB-3_SONET and a destination FC-BB-3_SONET consists of the latency delay components due to queu-

ing time at the two FC-BB-3_SONET devices and all intermediate SONET switches, transmission time, and propagation time. It is recommended that this total path delay be less than 1/2 E_D_TOV to conform to normal Fibre Channel time out values.

8.5.2 Delivery order

FC-BB-3_SONET shall guarantee in -order delivery of frames.

NOTE 22 - The out-of-sequence delivery problem associated with datagram networks is not present here. This benefit is a consequence of the use of SONET technology. However, the possibility of missing or errored messages still remains and is addressed by the SR protocol.

8.5.3 Loss and flow control

SONET/SDH networks are not lossy but may suffer from occasional loss of frame due to imperfect bit error rate (BER). When such a loss occurs, the end application is expected to recover from this loss. Recovery from such losses occurs at the FC-BB-3_SONET devices using the SR protocol that supports error recovery.

NOTE 23 - The SFC protocol has no error recovery support.

Use of a flow control protocol (i.e., SFC or SR) at the FC-BB-3 device provides the abiltiy to deal with speed mismatches between the FC and the SONET/SDH interface.

A typical list of reliability specifications for SONET/SDH networks is as follows:

a) MTTF (Mean time to frame) = approximately 1.5 packets;

b) MTTS (Mean time to synchronization) = same as MTTF;

c) PFF (Probability of false frame) = 232.8E-12;

d) PFS (Probability of false synchronization) = same as PFF; and

e) PLF (Probability of loss of frame) = square of the BER multiplied by 500.

9 FC-BB-3_IP Structure and Concepts

9.1 Applicability

Clause 4 discussed the FC-BB-3_IP reference model. This clause discusses the FC-BB-3_IP functional model.

9.2 FC-BB-3_IP overview

Figure 25 shows a network configuration consisting of three FC-BB-3_IP devices. FC-BB-3_IP is a Fibre Channel backbone transport protocol that tunnels encapsulated FC frames across the IP network. An FC-BB-3_IP device has interfaces to both the IP and the FC network. The FC network interface supports multiple E_Ports/ F_Ports (see figure 26) or multiple B_Ports (see figure 30).

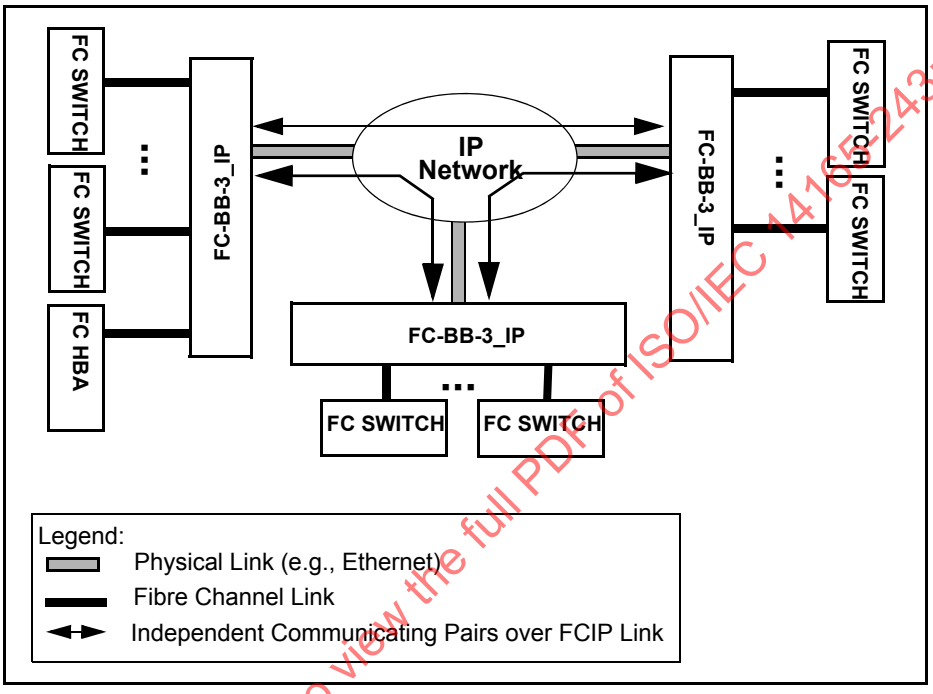


Figure 25 – FC-BB-3_IP network configuration

Only FC-BB-3_IP devices that support E_Ports or F_Ports require FC switching.

The FC-BB-3_IP protocol provides mechanisms to create VE_Port or B_Access connectivity over the IP network. The FC-BB-3_IP protocol communication occurs between pairs of FC-BB-3_IP devices over virtual constructs (i.e., FCIP Links) that are described in 9.3.4.2.4. Although the communication occurs between pairs of FC-BB-3_IP devices, a single FC-BB-3_IP device may communicate with more than one device at the same time (see figure 33).

NOTE 24 - Although the current scheme allows an FC-BB-3_IP device to independently connect to more than one FC-BB-3_IP device, it does not specify a point-to-multipoint connection.

The FC-BB-3_IP protocol uses encapsulated FC frames created by prefixing a 28-byte FC Encapsulation Header to the incoming SOF/EOF delimited FC frame (see RFC 3643). FC-BB-3_IP devices are not required to interpret the data content of the FC frames other than capturing and recording their SOF/EOF identities in the encapsulated FC frame. As such, FC Sequences and Exchanges are not visible to the FC-BB-3_IP protocol. All encapsulated FC frames are transparently transported over the IP network

FC-BB-3_IP devices also exchange SW_ILS control information using Class F FC frames (see figure 28 and figure 31). These FC frames are encapsulated and tunneled in the same way as the incoming FC frames.

Encapsulated FC frames join the TCP byte stream in order (see figure 34). TCP segments are created from TCP byte streams without any visibility or regard to encapsulated FC frame boundaries.

TCP flow control between two FC-BB-3_IP devices provides a reliable transport of encapsulated FC frames across the IP network. The only delivery order guarantee provided by TCP with respect to the FCIP protocol is

the correctly ordered delivery of encapsulated FC frames within a single TCP connection. The FC Entity is expected to specify and handle all other FC frame delivery ordering requirements.

9.3 VE_Port functional model

9.3.1 FC-BB-3_IP interface protocol layers

Figure 26 shows the VE_Port functional model of an FC-BB-3_IP device that consists of the E_Port/F_Port FC interface, the FC-BB-3_IP interface, and the IP network interface. The protocol layers at these interfaces are listed below:

a) E_Port/F_Port FC interface: FC-0, FC-1, and FC-2 levels;

b) An FC Switching Element (SE) with FC routing;

c) FC-BB-3_IP protocol interface: FC Entity and FCIP Entity protocol layers; and

d) IP network interface: TCP and IP layers.

Figure 27 illustrates the protocol layers across these interfaces.

9.3.2 E_Port/F_Port FC interface

The FC-BB-3_IP FC interface supports one or more E_Ports or F_Ports thus requiring the support of the FC-0, FC-1, and FC-2 levels. The E_Ports in general connect to different external FC switches, but connectivity to the same FC switch is also allowed. The data emerging from the FC-2 level is fed into an FC switching element.

The initialization of any generic E_Port or F_Port is described in FC-SW-4. An E_Port indicates its support for the ELP/ESC parameters using the ELP/ESC SW_ILS that is capable of parameter negotiation. Since FC-BB-3 does not support Class 1, the Class 1 Port Parameter VAL bit in the ELP shall be set to 0 (i.e., invalid). An ELP received at an E_Port may be rejected using SW_RJT for many reasons (see FC-SW-4).

An E_Port/F_Port is uniquely identified by an 8-byte E_Port_Name/F_Port_Name.

9.3.3 FC Switching Element (SE) with FC routing

The FC Switching Element (SE) switches and routes the incoming FC frames from the E_Port or F_Port to the proper VE_Port (see FC-SW-4). Routing is accomplished with the support of the FSPF routing protocol. Similarly, the FC SE switches and routes the data arriving from a VE_Port to the proper E_Port or F_Port.

The SE is uniquely identified by an 8-byte Switch_Name.

9.3.4 FC-BB-3_IP protocol interface

9.3.4.1 Major components

The FC-BB-3_IP protocol has interfaces to the FC network on one side and the IP network on the other. In addition to the two network interfaces, it consists of the following major components:

a) FC and FCIP Entities;

b) Control and Service Module (CSM); and

c) Platform Management Module (PMM).

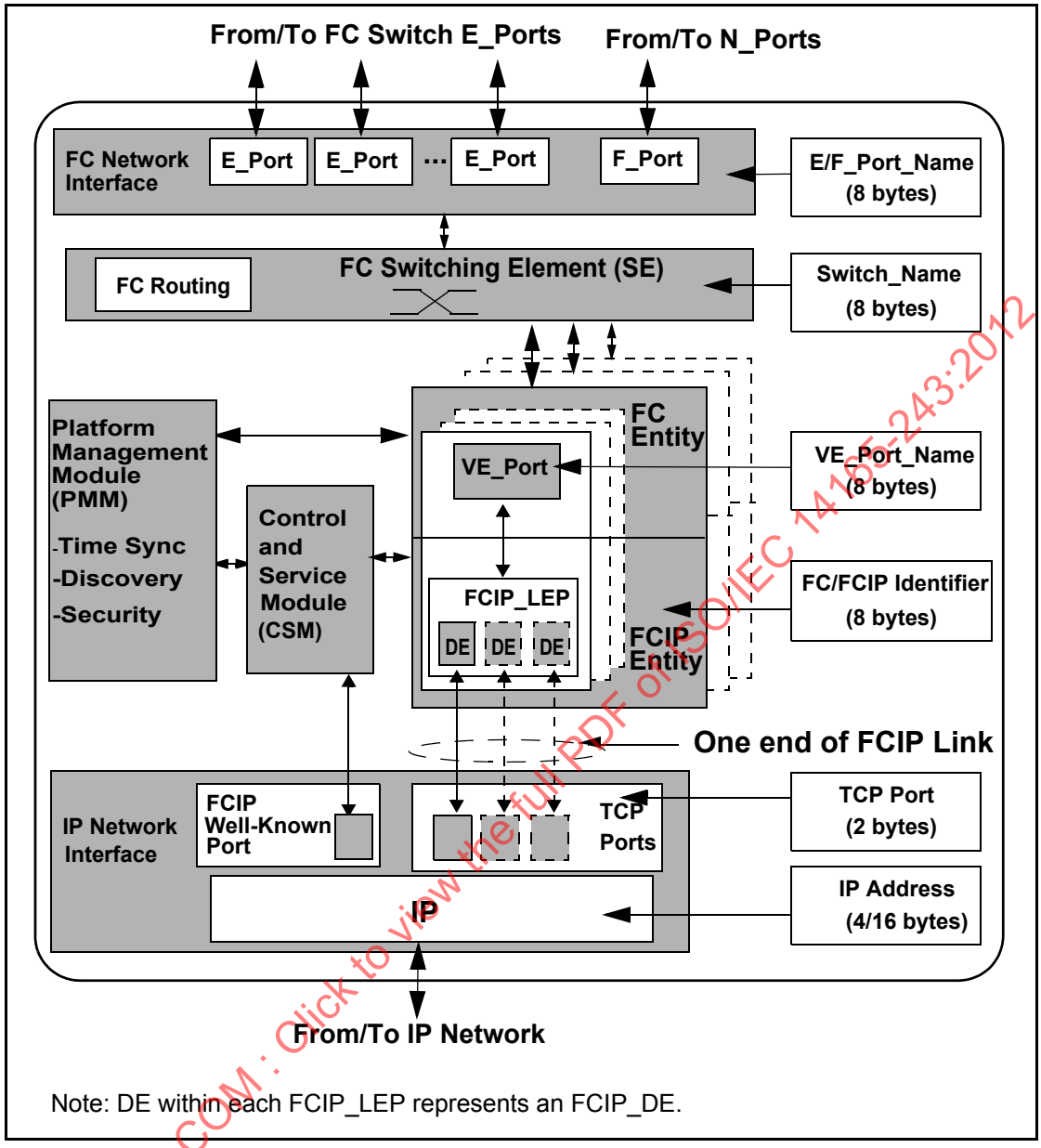


Figure 26 – FC-BB-3_IP VE_Port functional model

FC routing occurs at a higher level than IP routing. FC/FCIP Entities themselves do not actively participate in FC frame routing. FC routing uses the FSPF protocol described in FC-SW-4. FSPF routes are mapped onto the FCIP Links interconnecting FC-BB-3_IP devices. An FC frame's FSPF route decides the selection of the VE_Port/FCIP_LEP pair within a selected FC/FCIP Entity pair, when multiple pairs are in use. When multiple DEs within an FCIP_LEP are in use, the selection of which FCIP_DE to use is described in 9.7.3.5.

9.3.4.2 FC and FCIP Entities

9.3.4.2.1 Function

The FC Entity is the principal interface point to the FC network on one side and, in combination with the FCIP Entity, to the IP network on the other side. The primary functions of the FC Entity are to support one or more VE_Ports and to communicate with the FCIP Entity. The FC Entity layer lies between the FC-2 FC level and the FCIP Entity layer as shown in figure 27.

The FCIP Entity is the principal interface point to the IP network on one side, and in combination with the FC Entity, to the FC network on the other. The primary function of the FCIP Entity is formatting, encapsulating, and forwarding encapsulated FC frames across the IP network interface.

The FC/FCIP Entity pair interfaces with the CSM and the PMM through a vendor-specific mechanism.

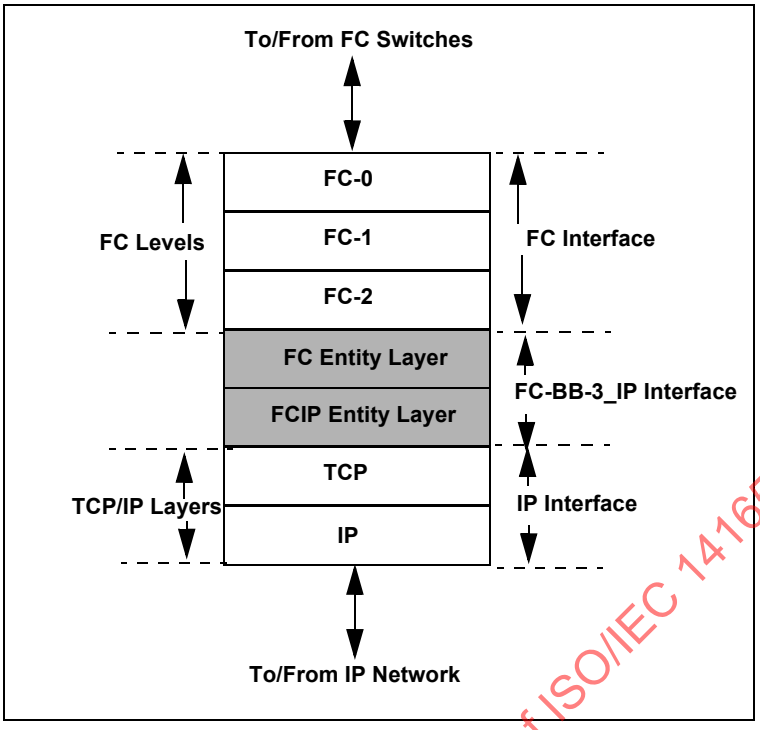


Figure 27 – FC-BB-3_IP Protocol Layers

9.3.4.2.2 FC Entity

The FC-BB-3_IP interface may support multiple instances of the FC/FCIP Entity pair. Each instance of the FC/FCIP Entity pair consists of one or more VE_Port/FCIP_LEP pairs. A VE_Port emulates an E_Port and interfaces with the FCIP_LEP component of the FCIP Entity. The term “Virtual” in VE_Port indicates the use of a non Fibre Channel link connecting the VE_Ports.

The VE_Port receives FC frames from the FC side and sends them to the FCIP_LEP for encapsulation and transmission on the IP network. The VE_Port may also exchange Class F control frames with the remote VE_Port via the LEPs. There is a one-to-one relationship between a VE_Port and an FCIP_LEP. VE_Ports communicate via VE_Port Virtual ISLs (see 9.3.4.2.4).

NOTE 25 - The term Virtual ISL when used unqualified refers to both a VE_Port Virtual ISL and a B_Access Virtual ISL.

A VE_Port is uniquely identified by an 8-byte VE_Port_Name.

Within an FC-BB-3_IP device, each FC/FCIP Entity pair instance is uniquely identified by an 8-byte identifier called the FC/FCIP identifier. The FC/FCIP identifier uses the Name_Identifier format.

Initialization at the FC-BB-3 protocol interface occurs between VE_Ports in a manner identical to standard E_Ports and is described in 9.3.4.3.

9.3.4.2.3 FCIP Entity

The FCIP_LEP is a component of the FCIP Entity that formats, encapsulates, and forwards encapsulated FC frames. Encapsulated FC frames are sent as TCP segments over the IP network.

The FCIP_LEP receives byte-encoded SOF/EOF delimited FC frames and a timestamp (see 9.3.4.5.2.2) from its VE_Port. The FCIP Data Engine (FCIP_DE) is the data forwarding component of the FCIP_LEP. The FCIP_DE handles all encapsulation and de-encapsulation, and transmission and reception of the encapsulated FC frames on the FCIP Link. The FCIP_LEP contains one or more FCIP_DEs, each corresponding to a TCP connection.

The FCIP_DE has four interface points (see RFC 3821):

a) **FC Receiver Portal:** The access point through which a byte-encoded SOF/EOF delimited FC frame and timestamp enters an FCIP_DE from the VE_Port;

b) **FC Transmitter Portal:** The access point through which a reconstituted byte-encoded SOF/EOF delimited FC frame and timestamp leaves an FCIP_DE to the VE_Port;

c) **Encapsulated Frame Receiver Portal:** The TCP access point through which an encapsulated FC frame is received from the IP network by the FCIP_DE; and

d) **Encapsulated Frame Transmitter Portal:** The TCP access point through which an encapsulated FC frame is transmitted to the IP network by the FCIP_DE.

9.3.4.2.4 VE_Port Virtual ISL and FCIP Link

The FC/FCIP Entity pair provides a data forwarding path between itself and a remote FC/FCIP Entity pair via virtual constructs. Two types of virtual constructs are defined:

a) a VE_Port Virtual ISL is a logical construct that is created between two FC Entity VE_Ports for the explicit purpose of sending and receiving byte-encoded SOF/EOF delimited FC frames via the FCIP Entity. Conceptually, communication between two VE_Ports is similar to communication between E_Ports; and

b) an FCIP Link is a logical construct that is created between two FCIP Entity LEPs for the explicit purpose of sending and receiving encapsulated FC frames and encapsulated FCIP control information. Conceptually, communication between two LEPs is similar to the communication between two instances of a TCP application.

There is a one-to-one mapping between a VE_Port Virtual ISL and an FCIP Link. Each FCIP Link consists of one or more TCP connections, all between the same two FC-BB-3 IP devices. Although more than one FCIP Link may be formed between a pair of FC-BB-3_IP devices, a typical configuration may only consist of a single FCIP Link. See figure 33 for some examples of allowed network topologies.

The FCIP_LEP that originates an FCIP Link is defined as the FCIP Link Originator. The corresponding FCIP_LEP that accepts this link is defined as the FCIP Link Acceptor. An FCIP Link is fully characterized by its FCIP Link Originator and FCIP Link Acceptor identities. An FCIP Link Originator or FCIP Link Acceptor is fully identified by all of the following:

a) an 8-byte Switch_Name;

b) an 8-byte VE_Port_Name; and

c) an 8-byte FC/FCIP Entity identifier.

To uniquely identify an FCIP Link, all of the following are required:

a) the 8-byte Switch_Name of the FCIP Link Originator;

b) the 8-byte VE_Port_Name of the FCIP Link Originator;

c) the 8-byte FC/FCIP Entity identifier of the FCIP Link Originator; and

d) the 8-byte Switch_Name of the FCIP Link Acceptor.

The FCIP Link Acceptor's 8-byte FC/FCIP Entity identifier and its VE_Port_Name provide additional information about an FCIP Link but are not required to uniquely identify it.

9.3.4.3 VE_Port Virtual ISL exchanges – SW_ILS exchanges

VE_Ports exchange SW_ILSs on the VE_Port Virtual ISL. The SW_ILSs that occur on the VE_Port Virtual ISL are the standard E_Port SW_ILSs (e.g., ELP, ESC, EFP, etc.), and in addition the LKA ELS (see FC-LS). Figure 28 shows the scope of the VE_Port Virtual ISLs.

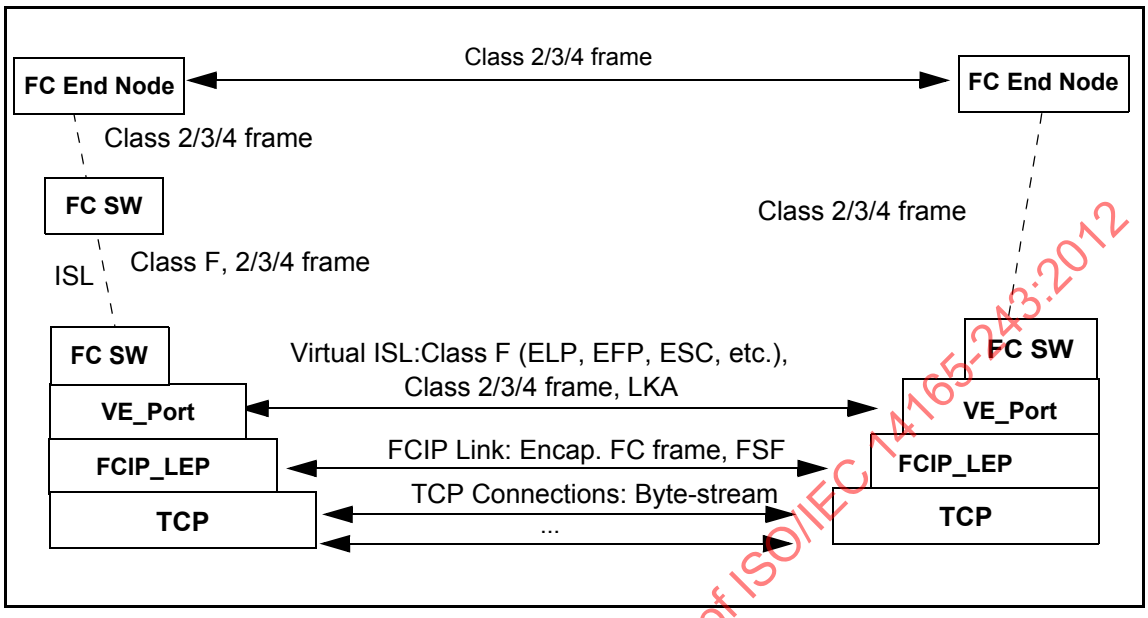


Figure 28 – Scope of VE_Port Virtual ISL

9.3.4.4 Control and Service Module (CSM)

The CSM is a control component of the FC-BB-3_IP interface that mainly deals with connection management. The CSM creates the FC/FCIP Entity pair during the Virtual ISL/FCIP Link setup. The CSM processes all requests for a link setup via the FCIP-registered TCP Port 3225 or optionally another TCP Port. The CSM processes requests to add additional TCP connections over the same FCIP Link. The CSM is also responsible for tearing down existing FCIP Links and TCP connections and deleting the FC/FCIP Entity pair.

NOTE 26 - Some aspects of the CSM functions are discussed only in RFC 3821.

9.3.4.5 Platform Management Module (PMM)

9.3.4.5.1 Function

The PMM is a management component of the FC-BB-3_IP interface that handles time synchronization, discovery, and security. The PMM is also the intended component for any miscellaneous housekeeping functions such as maintenance of event logs (see 9.8.3.5)

9.3.4.5.2 Time synchronization

9.3.4.5.2.1 FCIP Transit Time (FTT)

FCIP Transit Time (FTT) is defined as the total transit time of an encapsulated Fibre Channel frame in the IP network. The FCIP Transit Time is calculated by subtracting the timestamp value in the arriving encapsulated FC frame from the synchronized time in the FCIP Entity.

9.3.4.5.2.2 Building outgoing FC frame encapsulation headers

The FC Entity shall establish and maintain a synchronized time value in Simple Network Time Protocol (SNTP) Version 4 format (see RFC 2030) for use in computing the IP network transit times. The FC Entity shall use suitable internal clocks and one of the following mechanisms to establish and maintain the synchronized time value:

a) Fibre Channel time services; or

b) IP network SNTP server(s).

Each byte-encoded SOF/EOF delimited FC frame that the FC Entity delivers to the FCIP_DE through the FC receiver portal shall be accompanied by a timestamp value obtained from the synchronized time service. The FCIP_DE places the timestamp in the encapsulation header part of the encapsulated FC frame that carries the FC frame (see RFC 3643). If no synchronized timestamp value is available to accompany an entering Class 2, 3, or 4 FC frame, the frame should not be delivered to the FCIP_DE. However, FC-BB-3_IP shall allow any class F encapsulated FC frames to be transmitted with a zero timestamp value.

9.3.4.5.2.3 Checking IP network transit times in incoming FC frame encapsulation headers

Each byte-encoded SOF/EOF delimited FC frame delivered to the FC Entity through the FCIP_DE FC transmitter portal is to be accompanied by the timestamp value taken from the encapsulation header of the encapsulated FC frame. As noted in 9.3.4.5.2.2, the timestamp may be zero indicating that no valid timestamp was supplied by the sending FC Entity. Any frame other than a Class F frame whose timestamp is zero shall be discarded. A Class F frame whose timestamp is zero shall be processed as if it met all Fibre Channel timeout requirements.

When the timestamp is non-zero, the FTT of the arriving encapsulated Fibre Channel frame shall be compared to 1/2 E_D_TOV. If the FTT exceeds 1/2 E_D_TOV, then the frame shall be discarded. Otherwise the frame shall be processed normally. Fibre Channel timeout values shall be administratively set to accommodate the FTT.

9.3.4.5.3 Discovery

Discovery of FC-BB-3_IP devices is handled in accordance with the procedures outlined in 9.7.2.2 and in RFC 3821 and RFC 3822.

9.3.4.5.4 Security

Security in FC-BB-3_IP is defined at two levels, FC and FCIP. The FC level is secured through FC-SP mechanisms that are extended by FC-BB-3_IP. The FCIP level is secured through IPSec mechanisms (see RFC 3821). Figure 29 illustrates the scope of the two security mechanisms.

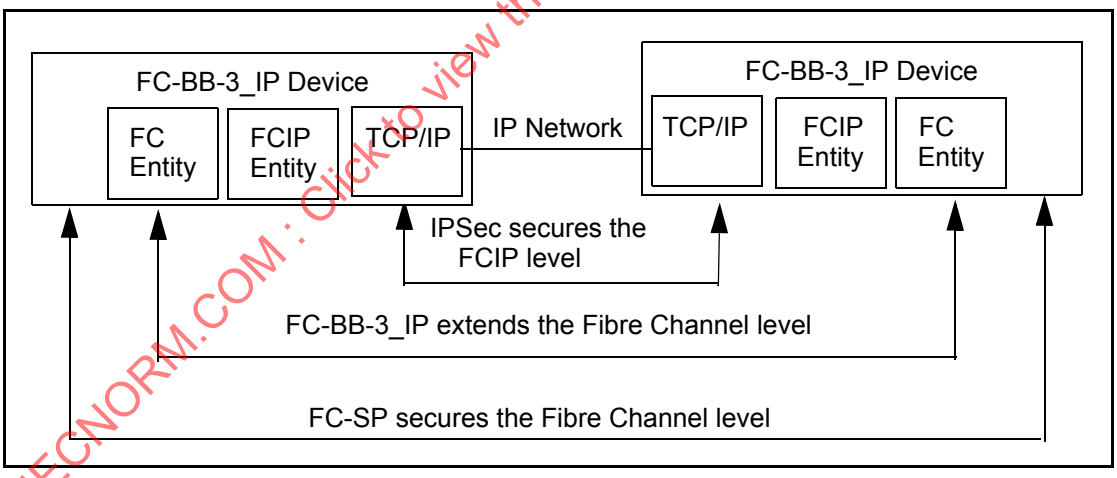


Figure 29 – Security layers

In most cases, the security requirements of an FC/FCIP Entity pair are satisfied outside the scope of this standard as follows:

a) security for the FC fabric is provided by the FC-SP capabilities (e.g., switch-to-switch authentication, frame authentication and confidentiality); and

b) security for the TCP connections used to transit the IP network is provided by the security features described in FCIP (e.g., IPSec packet authentication and confidentiality).

Depending on the security requirements of a given configuration, any or all of the security capabilities described in other standards may be enabled or disabled. However, it is important to note that the public IP network is subject to a large variety of security attacks, meaning that serious consideration should be given to enabling

the full suite of security features described in RFC 3821 whenever the public IP network is to be used to transmit FCIP frames.

An FC/FCIP entity pair has a potential security vulnerability where interactions may not be fully secured by either the FC-SP or FCIP security features. This vulnerability occurs when two or more TCP connections are aggregated in a single FCIP Link. The first TCP connection in an FCIP Link and its associated Virtual ISL may be authenticated using the FC-SP mechanisms. However, no such authentication is defined for subsequent TCP connections, since to FC they all appear to be part of an already authenticated Virtual ISL.

To prevent attacking entities in the IP network from forging additional invalid TCP connections, the FC-BB-3_IP mechanism described in 9.7.2.3 extends the protection of FC-SP authentication to subsequently-added TCP connections. The extension to FC-SP authentication described in 9.7.2.3.2 is based on the exchange of Class F requests and responses between FC Entities. This mechanism works in concert with the FC-SP Virtual ISL authentication mechanism, handling the Class F requests and responses over a previously authenticated TCP connection. In some configurations, this overhead may be unnecessary. However, in cases where fabric entities are capable of being authenticated without having their behavior fully trusted, the extension to FC-SP authentication should be used in combination with other FC-SP and FCIP security mechanisms to assure trustworthy formation of FCIP Links and Virtual ISLs.

9.3.5 IP network interface

The FC-BB-3_IP VE_Port reference model supports one logical IP interface and allows sharing a 4-byte IPv4 or 16-byte IPv6 address in the following ways:

a) a single IP address per FC-BB-3_IP device (i.e., a single IP address shared by all FC/FCIP Entity pairs);

b) multiple IP addresses per FC-BB-3_IP device (i.e., a single IP address per FC/FCIP Entity pair);

c) multiple IP addresses per FC/FCIP Entity pair (i.e., single IP address per VE_Port/FCIP_LEP pair); and

d) multiple IP Addresses per FCIP Link (i.e., a single IP address per TCP Port).

Use of different IP address schemes at the two ends of an FCIP Link is not expected to cause inter operability problems.

As shown in figure 27, the IP network interface consists of the TCP and IP layers. The encapsulated FC frame emerging from the FCIP_DE interfaces with the TCP layer. The IP layer interfaces with the TCP layer above it and the IP network below it. The TCP layer supports multiple TCP connections, each corresponding to an FCIP_DE. Each client-side TCP connection within an FCIP Link is assigned a unique TCP Port Number. Either the FCIP well-known TCP Port 3225 or optionally another TCP Port is used for accepting connection requests. These ports interface with the CSM through a vendor-specific mechanism.

IP routing occurs inside the IP network. Within the IP network, the route taken by an encapsulated FC frame is determined by the normal routing procedures of the IP network.

9.4 B_Access functional model

9.4.1 FC-BB-3_IP interface protocol layers

Figure 30 shows the functional model of an FC-BB-3_IP device that consists of the B_Port FC interface, the FC-BB-3_IP protocol interface, and the IP network interface. Figure 27 shows the details of the protocol layers across these interfaces.

NOTE 27 - Because of the similarity between the E_Port and B_Port functional models this subclause only describes any unique definitions for the B_Access. Other definitions and descriptions from 9.3 apply equally well and remain unchanged.

9.4.2 B_Port FC interface

The FC-BB-3_IP FC network interface supports one or more B_Ports thus requiring the support of the FC-0, FC-1, and FC-2 levels. B_Ports in general connect to different external FC switches, but connectivity to the same FC switch is allowed.

B_Ports are uniquely identified by an 8-byte B_Port_Name.

9.4.3 FC-BB-3_IP protocol interface

9.4.3.1 Major components

The B_Port FC-BB-3_IP interface consists of all the components of the VE_Port functional model (see 9.3.4.1) except the FC Switching Element with FC routing.

9.4.3.2 FC and FCIP Entities

9.4.3.2.1 Function

The primary function of the FC Entity is to support one or more B_Access portals and to communicate with the FCIP Entity.

The function of the FCIP Entity is identical to its function in the VE_Port functional model described in 9.3.4.2.3.

The FC/FCIP Entity pair interfaces with the CSM and the PMM through a vendor-specific mechanism.

9.4.3.2.2 FC Entity

The FC-BB-3_IP interface may support multiple instances of the FC/FCIP Entity pairs. Each instance of the FC/FCIP Entity pair consists of one or more B_Access/FCIP_LEP pairs. A B_Access portal is a component of the FC Entity that interfaces with the FCIP_LEP component of the FCIP Entity. The B_Access portal receives FC frames from the B_Port and sends them to the FCIP_LEP for encapsulation and transmission on the IP network. The B_Access portal may also exchange Class F control frames with the remote B_Access portal via the LEPs. There is a one-to-one relationship between a B_Access portal and an FCIP_LEP. B_Access portals communicate via B_Access Virtual ISLs (see 9.4.3.2.4).

There is no switching and routing required in the case of the B_Port functional model. However, the forwarding of FC frames across the B_Access/FCIP_LEP pair is still required. When multiple DEs within an FCIP_LEP are in use, the selection of which FCIP_DE to use is described in 9.7.3.5.

Initialization at the FC-BB-3 protocol interface occurs with EBP SW_ILS exchanges between B_Access portals in a manner identical to standard E_Ports and is described in 9.4.3.2.4. The B_Access initialization state machine is described in 9.4.3.3.2.

9.4.3.2.3 FCIP Entity

The FCIP_LEP receives byte-encoded SOF/EOF delimited FC frames and a timestamp from its B_Access portals. All other functions are identical to the functions of the FCIP Entity in the VE_Port functional model (see 9.3.4.2.3).

9.4.3.2.4 B_Access Virtual ISL and FCIP Links

A B_Access Virtual ISL is a logical construct that is created between two FC Entity B_Access portals for the explicit purpose of sending and receiving byte-encoded SOF/EOF delimited FC frames via the FCIP Entity. Conceptually, communication between two B_Access portals is similar to communication between two VE_Ports.

There is a one-to-one mapping between a B_Access Virtual ISL and an FCIP Link.

An FCIP Link Originator or FCIP Link Acceptor is fully identified by all of the following:

a) an 8-byte Fabric_Name;

b) an 8-byte B_Access_Name; and

c) an 8-byte FC/FCIP Entity identifier.

To uniquely identify an FCIP Link, the following items are required:

a) the 8-byte Fabric_Name of the FCIP Link Originator;

b) the 8-byte B_Access_Name of the FCIP Link Originator;

c) the 8-byte FC/FCIP Entity identifier of the FCIP Link Originator; and

d) the 8-byte Fabric_Name of the FCIP Link Acceptor.

The FCIP Link Acceptor's 8-byte FC/FCIP Entity identifier and its B_Access_Name provide additional information about an FCIP Link but are not required to uniquely identify it.

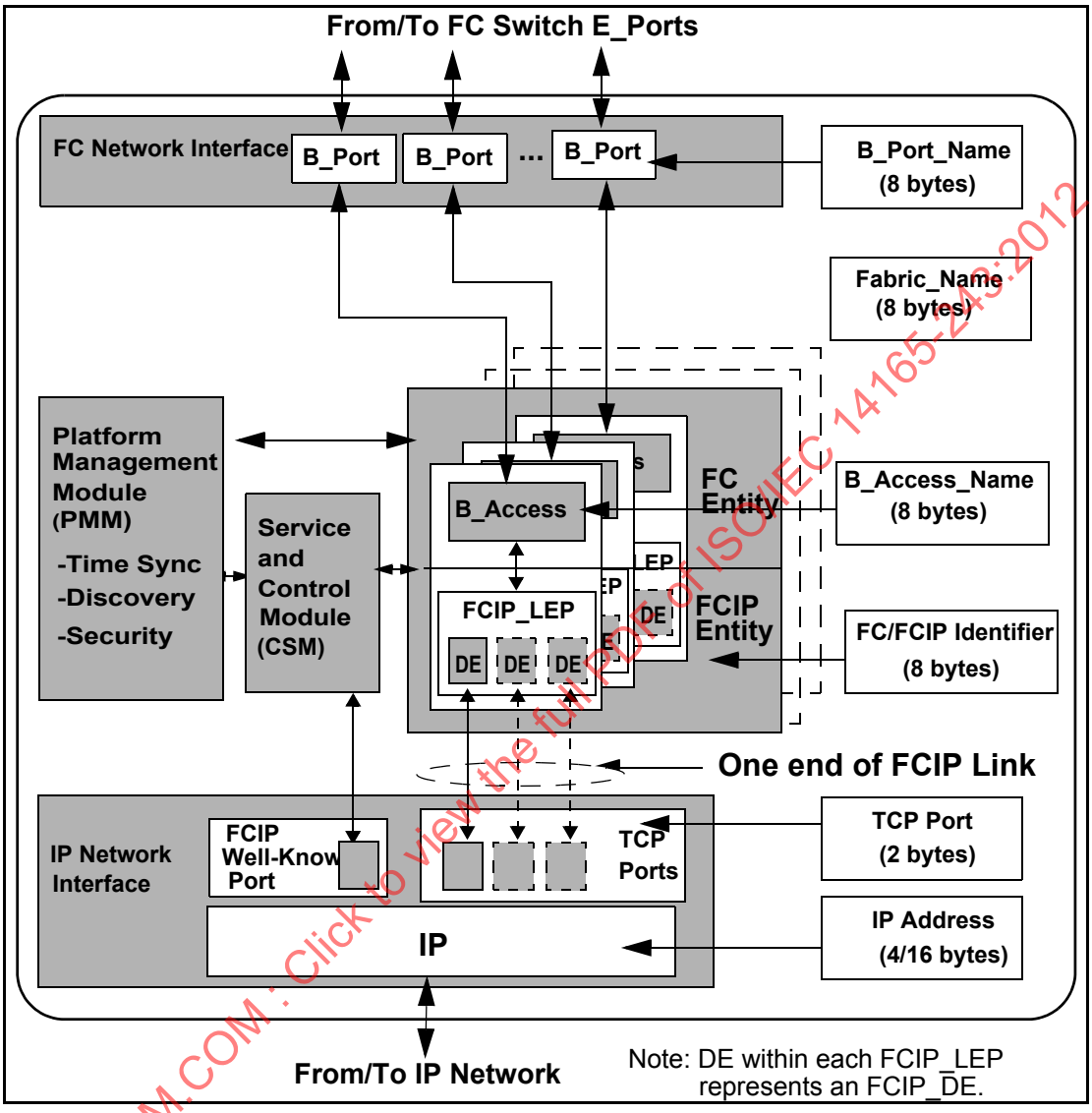


Figure 30 – FC-BB-3_IP B_Access functional model

9.4.3.3 B_Access Virtual ISL exchanges

9.4.3.3.1 Exchange B_Access Parameters (EBP) SW_ILS

B_Access portals exchange SW_ILSs on the B_Access Virtual ISL. The SW_ILSs that occur on the B_Access Virtual ISL are the EBP SW_ILS and the LKA ELS (see FC-LS). Figure 31 shows the scope of the B_Access Virtual ISLs.

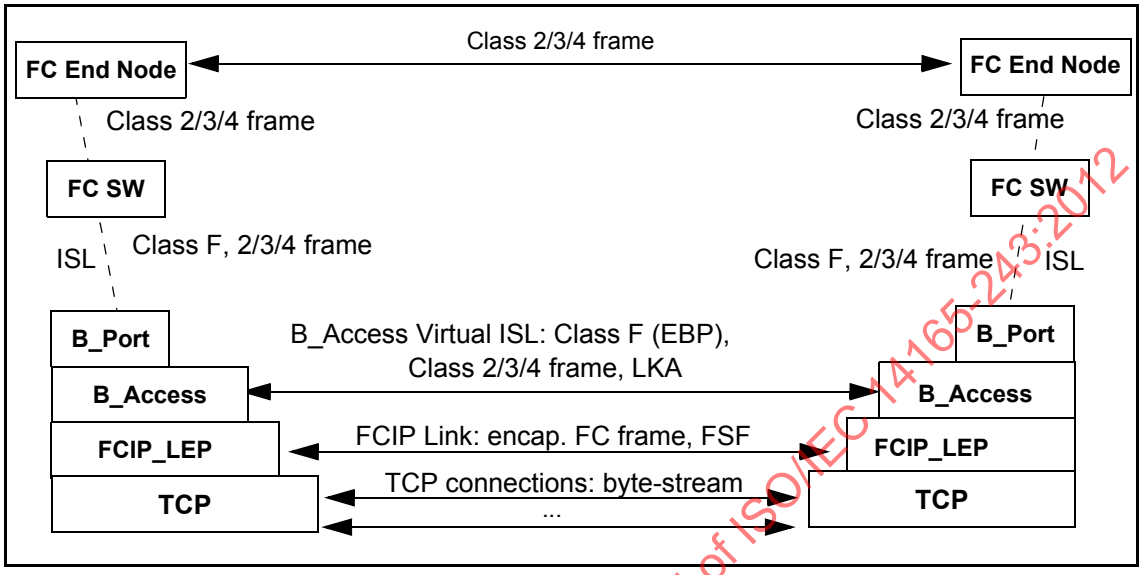


Figure 31 – Scope of B_Access Virtual ISL

The Exchange B_Access Parameters (EBP) Switch Fabric Internal Link Service (SW_ILS) is sent by a B_Access portal to a remote B_Access portal in order to establish operating link parameters and port capabilities for the B_Access Virtual ISL formed by the two B_Access portal peers. Successful acceptance of EBP SW_ILS shall be completed before the B_Ports begin switch port mode initialization.

Protocol:

a) Exchange B_Access Parameters (EBP) request Sequence; and

b) Reply Switch Fabric Internal Link Service Sequence.

Addressing: For use in switch port configuration, the S_ID field shall be set to FFFFFDh, indicating the Fabric Controller of the originating B_Access portal. The D_ID field shall be set to FFFFFDh, indicating the Fabric Controller of the destination B_Access portal.

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

Table 27 – EBP request payload

Item	Size Bytes	Remarks
28 01 00 00h	4	
R_A_TOV	4	Value in milliseconds
E_D_TOV	4	Value in milliseconds
K_A_TOV	4	Value in milliseconds
Requester B_Access_Name	8	
Class F Service Parameters	16	

Requester B_Access_Name: This field shall contain the B_Access_Name of the device that originated the EBP request.

R_A_TOV: This field shall be set to the value, in milliseconds, of R_A_TOV required by the FC-BB-3_IP device.

E_D_TOV: This field shall be set to the value, in milliseconds, of E_D_TOV required by the FC-BB-3_IP device.

K_A_TOV: This field shall be set to the value, in milliseconds, of K_A_TOV required by the FC-BB-3_IP device.

Class F Service Parameters: This field shall contain the B_Access Class F Service Parameters and its format is identical with its use in the ELP SW_ILS (see FC-SW-4).

Reply Switch Fabric Internal Link Service Sequence:

Service Reject (SW_RJT)

Accept (SW_ACC)

Accept payload

Payload: The format of the EBP Accept payload is shown in table 28.

Table 28 – EBP accept payload

Item	Size Bytes	Remarks
02 00 00 00h	4	
R_A_TOV	4	Value in milliseconds
E_D_TOV	4	Value in milliseconds
K_A_TOV	4	Value in milliseconds
Responder B_Access_Name	8	
Class F Service Parameters	16	

The fields in table 28 are the same as defined for table 27 except for the Responder B_Access_Name field.

Responder B_Access_Name: This field shall contain the B_Access_Name of the remote device that responds to the EBP request.

The SW_RJT Reply payload format is given in FC-SW-4. The EBP reject reason code explanation is shown in table 29.

Table 29 – EBP reject reason code explanation

Encoded Value (Bits 23-16)	Description
0000 0000	No additional explanation
0000 0001	Class F Service Parameter error
0000 0010	Invalid B_Access_Name
0000 0011	K_A_TOV mismatch
0000 0100	E_D_TOV mismatch
0000 0101	R_A_TOV mismatch
others	Reserved

9.4.3.3.2 B_Access initialization state machine

The B_Access initialization state machine is shown in figure 32.

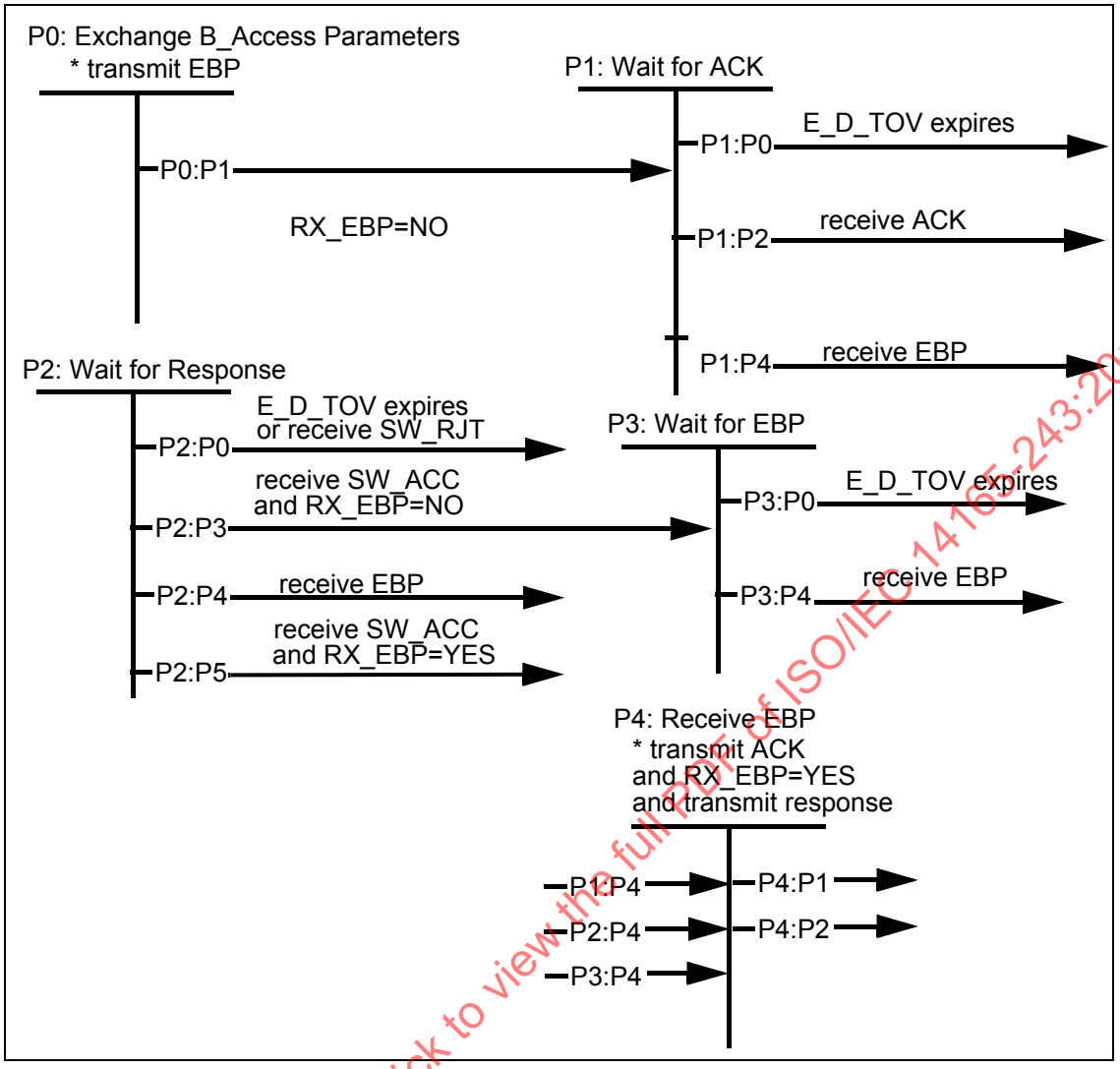


Figure 32 – B_Access initialization state machine

State P0: Exchange B_Access Parameters. This state marks the beginning of the B_Access initialization. Activity other than that described within the state machine is suspended until initialization is complete.

Transition P0:P1. The B_Access resets the RX_EBP flag.

State P1: Wait for ACK. In this state the B_Access waits until an ACK for the B_Access's transmitted EBP is received.

Transition P1:P0. This transition occurs when the B_Access has not received an ACK within E_D_TOV after the transmission of an EBP.

Transition P1:P2. This transition occurs when the B_Access receives an ACK before E_D_TOV expires.

Transition P1:P4. This transition occurs when the B_Access receives an EBP while waiting for an ACK.

State P2: Wait for Response. In this state the B_Access has received an ACK for its EBP and is waiting for a response.

Transition P2:P0. This transition occurs when the B_Access has not received a response within E_D_TOV after the transmission of an EBP or receives an SW_RJT.

Transition P2:P3. This transition occurs when the B_Access receives an SW_ACC and has not received an EBP.

Transition P2:P4. This transition occurs when the B_Access receives an EBP while waiting for a response.

Transition P2:P5. This transition occurs when the B_Access receives an SW_ACC and has received an EBP.

State P3: Wait for EBP. In this state the B_Access has received an ACK for its EBP and is waiting for an EBP.

Transition P3:P0. This transition occurs when the B_Access has not received an EBP within E_D_TOV of the transmission of an EBP.

Transition P3:P4. This transition occurs when a B_Access receives an EBP while waiting for a response.

State P4: Receive EBP. In this state the B_Access has received an EBP. The B_Access responds with an ACK and transmits an SW_ACC or SW_RJT depending upon whether or not the received configuration parameters contained within the EBP are acceptable. The B_Access sets RX_EBP to indicate an EBP has been received and is accepted.

Transition P4:P1. This transition occurs when a B_Access receives an EBP from its peer yet hasn't received an ACK for a previously transmitted EBP.

Transition P4:P2. This transition occurs when a B_Access receives an EBP from its peer yet hasn't received a response for a previously transmitted EBP.

9.4.3.4 B_Port Control and Service Module (CSM)

The B_Port CSM is identical to the E_Port CSM described in 9.3.4.4.

9.4.3.5 B_Port Platform Management Module (PMM)

The B_Port PMM is identical to the E_Port PMM described in 9.3.4.5.

9.4.4 IP Network Interface

The B_Port IP network interface is identical to the E_Port IP network interface described in 9.3.5 with a change in item (c), where a single IP address is per B_Access/FCIP_LEP pair.

9.5 FC-BB-3_IP Network Topologies

Figure 33 shows some example FC-BB-3_IP network topologies that exists between three FC-BB-3_IP sites:

a) FCIP Link 1 connects Sites 1 and 2 and consists of three TCP connections;

b) FCIP Link 2 connects Sites 1 and 2 and consists of two TCP connections. FCIP Link 2 however is distinct from Link 1 although it exists between the same two FC/FCIP Entity pairs (i.e., FC/FCIP_Entity_1 and FC/FCIP_Entity_2);

c) FCIP Link 3 connects Sites 1 and 3 and consists of two TCP connections. FCIP Link 3 exists between FC/FCIP_Entity_3 and FC/FCIP_Entity_5; and

d) FCIP Link 4 connects Sites 2 and 3 and consists of one TCP connection. FCIP Link 4 exists between FC/FCIP_Entity_4 and FC/FCIP_Entity_6.

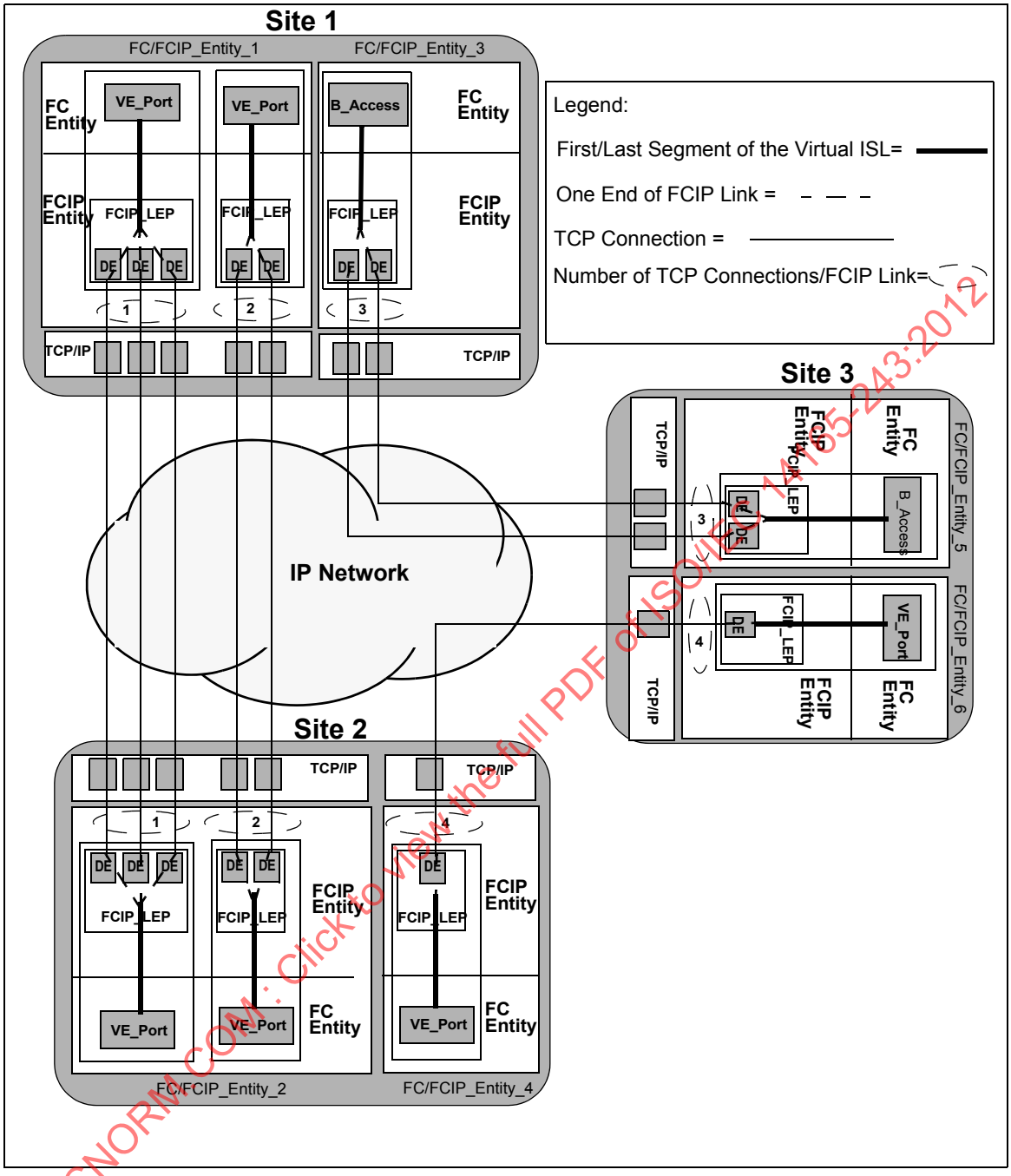


Figure 33 – FC-BB-3_IP network topologies

9.6 Mapping and message encapsulation using TCP/IP

9.6.1 Encapsulated frame structures

9.6.1.1 FC frame encapsulation structure

An encapsulated FC frame is carried in one or more TCP segments as shown in figure 34. Each segment's format is shown in table 30. The structure of each encapsulated FC frame is shown in table 31 and consists of

a FC Encapsulation Header and a byte-encoded SOF/EOF delimited Class 2, 3, 4, or F FC frame as described in RFC 3643.

Table 30 – TCP/IP Segment structure carrying encapsulated FC frame

Field	Sub-field	Size (Bytes)
IP Header		Min:20 Max:40
TCP Header		Min:20 Max:40
TCP payload	one or more portions of encapsulated FC frames	Min:64 Max:2 176

Table 31 – Encapsulated FC frame structure

Field	Size (Bytes)
FC Encapsulation Header	28
SOF (see Note below)	4
FC-Header	24
FC frame payload (includes optional header)	Min: 0 Max:2 112
CRC	4
EOF (see Note below)	4

FC frame encapsulation (see RFC 3643) describes the structures of the 4-byte SOF/EOF values fields and the FC Encapsulation Header. The FC Encapsulation Header consists of several fields: Protocol#, Version, pFlags, Flags, Frame Length, Timestamp, and CRC.

Protocol#: indicates the FCIP protocol.

Version: indicates the version number.

pFlags: defines flag bits FSF and Ch that distinguish encapsulated FC frames from FCIP originated or echoed control frames.

Flags: the CRCV bit value indicates if the contents of the CRC field are valid or invalid. For FC-BB-3_IP protocol the CRCV bit shall be zero (i.e., invalid).

Frame Length: contains the length of the entire FC encapsulated frame including the FC Encapsulation Header and the FC frame, including SOF and EOF words.

Timestamp: contains the time at which the FC encapsulated frame was sent as known to the sender. The format of integer and fraction timestamp word values is specified in Simple Network Time Protocol (SNTP) Version 4 (see RFC 2030). The contents of the timestamp integer and timestamp fraction words shall be set as described in 9.3.4.5.2.

CRC: for FC-BB-3_IP protocol the CRC field shall be set to zero.

9.6.1.2 Encapsulated FCIP Special Frame (FSF) structure

An encapsulated FCIP Special Frame (FSF) is carried as a TCP segment as shown in table 32. The structure of an encapsulated FSF is shown in table 33 and consists of an FC Encapsulation Header and an FCIP Special Frame.

Table 32 – TCP/IP Segment structure carrying encapsulated FSF

Field	Sub-field	Size (Bytes)
IP Header		Min:20 Max:40
TCP Header		Min:20 Max:40
TCP payload	encapsulated FSF	76

Table 33 – Encapsulated FSF structure

Field	Size (Bytes)
FC Encapsulation Header	28
FCIP Special Frame (FSF)	48

See 9.6.1.1 for a description of the FC Encapsulation Header structure and format.

The FSF structure is defined in FCIP (see RFC 3821) and consists of several fields: Source FC Fabric_Name, Source FC/FCIP Entity identifier, Connection Nonce, Connection Usage Flags, Connection Usage Code, Destination FC Fabric_Name, and K_A_TOV.

Source FC Fabric_Name: the identifier for the FC fabric associated with the FC/FCIP Entity pair that generates the FCIP Special Frame. If the FC fabric is an FC switch, then the field contains the Switch_Name.

Source FC/FCIP Entity identifier: a unique identifier for the FC/FCIP Entity pair that generates the FSF. The value is assigned by the FC fabric whose name appears in the Source FC Fabric_Name field.

Connection Nonce: contains a 64-bit random number generated to uniquely identify a single TCP connect request. in order to provide sufficient security for the nonce, the randomness recommendations described in RFC 3821 should be followed.

Connection Usage Flags: identifies the types of SOF values to be carried on the connection. All or none of the bits corresponding to Class F, 2, 3, or 4 may be set to one. If all of the bits are zero, then the types of FC frames intended to be carried on the connection has no specific relationship to SOF code.

Connection Usage Code: contains Fibre Channel defined information regarding the intended usage of the connection. The FCIP Entity uses the contents of the Connection Usage Flags and the Connection Usage Code fields to locate appropriate QoS settings in the shared database of TCP connection information and apply those settings to a newly formed connection. All values for this field are reserved.

Destination FC Fabric_Name: may contain the Fibre Channel identifier for the FC fabric associated with the FC/FCIP Entity pair that echoes, as opposed to generates, the FSF.

K_A_TOV: contains the FC Keep Alive Timeout value to be applied to the new TCP connection.

9.6.2 TCP/IP encapsulation

Figure 34 illustrates the TCP/IP encapsulation of an encapsulated FC frame. The TCP/IP encapsulation of an encapsulated FSF is similar.

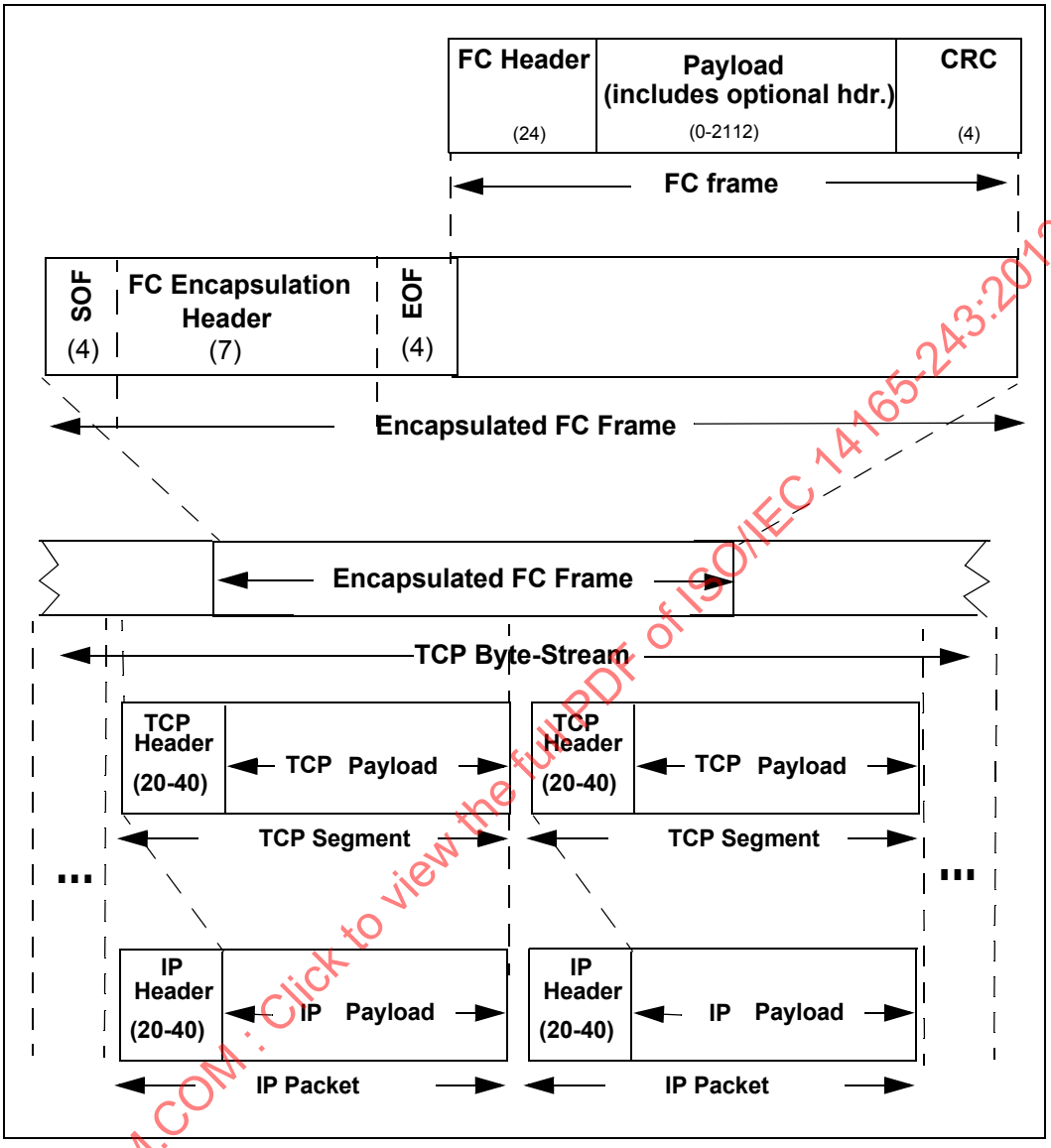


Figure 34 – TCP/IP encapsulation of an encapsulated FC frame

9.7 FC-BB-3_IP Protocol Procedures

9.7.1 Overview

This subclause describes the FC-BB-3_IP protocol procedures for platform management (see 9.7.2), connection management (see 9.7.3), and error detection and recovery (see 9.7.4). There are no specific procedures defined for housekeeping functions such as maintenance of error or event logs.

9.7.2 Procedures for platform management

9.7.2.1 Function

Platform management has three main functions, discovery, security, and time synchronization.

9.7.2.2 Procedures for discovery

Device discovery is one of the functions of the Platform Management Module (PMM). Each FC-BB-3_IP device is statically or dynamically configured with a list of IP addresses and other identifiers (e.g., N_Port_Names) cor-

responding to participating FC/FCIP Entities. If dynamic discovery of participating FC-BB-3_IP devices is supported, the function is performed using Service Location Protocol version 2 (see RFC 3822).

FC/FCIP Entities themselves do not actively participate in the discovery of FC source and destination identifiers. Discovery of FC addresses accessible via the FC/FCIP Entity is provided by techniques and protocols within the FC architecture as described in FC-FS-2 and FC-SW-4.

9.7.2.3 Procedures for extending FC-SP security

9.7.2.3.1 Authentication mechanisms

The Platform Management Module (PMM) is responsible for extending security at the FC level.

Entity authentication occurs at the FCIP and FC levels as illustrated in figure 29. Authentication mechanisms at the FCIP level are defined in FCIP (see RFC 3821). Authentication mechanisms at the FC level are defined in FC-SP.

During initialization of a Virtual ISL, each switch may authenticate the other switch with FC-SP authentication mechanisms. FC-BB-3_IP provides for extending the protection of FC-SP authentication to subsequently added TCP connections via either the ASF SW_ILS described in 9.7.2.3.2 or vendor-specific configuration information.

NOTE 28 - The unqualified use of the term Virtual ISL refers to both VE_Port Virtual ISL and B_Access Virtual ISL.

When an FCIP Entity receives a TCP connect request for an additional TCP connection to an existing FCIP Link to which FC-SP authentication has been applied, the FCIP Entity generates a request to the FC Entity to authenticate the additional TCP connection including at least the following information:

a) Connection Nonce;

b) Destination FC Fabric_Name;

c) Connection Usage Flags; and

d) Connection Usage Code.

If FC-SP authentication procedures are not being applied to the Virtual ISL, the FC Entity shall respond to the FCIP Entity indicating that the new TCP is authentic.

NOTE 29 - If the first TCP connection in a Virtual ISL is not authenticated using the applicable FC-SP procedures, no security is gained by authenticating other TCP connections.

NOTE 30 - The preferred security mechanism for the Public Internet IP network is the success or failure of an ASF SW_ILS.

9.7.2.3.2 Authenticate Special Frame (ASF)

The Authenticate Special Frame (ASF) Switch Fabric Internal Link Service (SW_ILS) is used by an FC Entity to authenticate additional TCP connections on existing FCIP Links. To authenticate a new TCP connection using the ASF SW_ILS, the FC Entity shall use the information provided by the FCIP Entity to transmit an ASF request on the Virtual ISL to which the new TCP connection is being added using a TCP connection in the Virtual ISL that has already been authenticated.

The FC Entity shall use the information from the new FSF request to populate the fields in the ASF request. The fields are the same as defined for FSF (see 9.6.1.2). The format of the ASF request payload is shown in table 34.

The FC Entity shall transmit the ASF over the previously authenticated TCP connection. This piggybacking technique authenticates additional TCP connections by riding on top of previously authenticated TCP connections.

An FC Entity that receives an ASF SW_ILS shall verify that the information in the request payload identifies a TCP connection initiated by that FC/FCIP Entity pair. If it verifies that this information is sound then the FC Entity shall respond with an SW_ACC (see table 35), otherwise it shall respond with an SW_RJT with a reason

code of “Unable to perform command request” and a reason code explanation of “Class F Service Parameter error”.

Protocol:

a) Authenticate Special Frame (ASF) request Sequence; and

b) Reply Switch Fabric Internal Link Service Sequence.

Addressing: The S_ID field shall be set to FFFFFDh, indicating the Fabric Controller of the originating FC Entity. The D_ID field shall be set to FFFFFDh, indicating the Fabric Controller of the receiving FC Entity.

Payload: The format of the ASF request payload is shown in table 34.

Table 34 – ASF request payload

Item	Size Bytes
28 03 00 00h	4
Destination FC Fabric_Name	8
Connection Nonce	8
Connection Usage Flags	1
Reserved	1
Connection Usage Code	2
Reserved	4

Destination FC Fabric_Name: This field is the Fabric_Name of the destination switch and is the Source FC Fabric_Name from the FSF frame.

Connection Nonce: This field is the Connection Nonce from the FSF request.

Connection Usage Flags: This field is the Connection Usage Flags from the FSF request and signifies the acceptance of these flags.

Connection Usage Code: This field is the Connection Usage Code from the FSF request and signifies the acceptance of these codes.

Reply Switch Fabric Internal Link Service Sequence:

a) Service Reject (SW_RJT); or

b) Accept (SW_ACC).

Accept payload.

Payload: The format of the ASF accept payload is shown in table 35.

Table 35 – ASF accept response payload

Item	Size Bytes
02 00 00 00h	4

9.7.3 Procedures for connection management

9.7.3.1 Function

The primary function of the Control and Services Module (CSM) is managing connections.

9.7.3.2 Procedures for link setup

In order to realize a Virtual ISL/FCIP Link between two FC-BB-3_IP endpoints, an FC-BB-3_IP device establishes TCP connection(s) with its peer FC-BB-3_IP device.

NOTE 31 - A Virtual ISL exists between two VE_Ports or two B_Access portals and an FCIP Link exists between two FCIP_LEPs. Conceptually, the procedures for establishing these two are identical.

It may also be useful to assign a pool of connections for transmission of high priority and control frames (e.g., Class F) on connections so they do not encounter head-of-line blocking behind Class 2, Class 3, or Class 4 traffic. The use of multiple connections and policies for distributing frames on these connections is described in 9.7.3.5.

A Virtual ISL/FCIP Link and the two FC-BB-3_IP device endpoints that are involved become operational only after the first TCP connection is established. The sequence of operations performed in order to establish a Virtual ISL/FCIP Link is as follows:

1) the FC-BB-3_IP device initializes its local resources to enable it to listen to TCP connection requests;

2) the FC-BB-3_IP device discovers the FC-BB-3_IP device endpoints to which it is able to establish a Virtual ISL/FCIP Link. The result of the discovery shall be, at the minimum, the IP address and the TCP port of the peer endpoint. The discovery process may rely on administrative configuration or on services such as SLPv2 as described in 9.7.2.2;

3) the processes defined by FCIP are used to establish TCP connections. FC level authentication of the first TCP connection is accomplished using the mechanisms and management controls described in FC-SP. To extend FC-SP authentication to additional TCP connections the mechanisms described in 9.7.2.3 shall be followed;

4) at this point, both endpoints have their respective VE_Port/FCIP_LEP pairs or B_Access/FCIP_LEP pairs established;

5) after connection establishment, the FC-BB-3_IP device constructs the encapsulated FC frames according to the methods described in 9.3.4.5.2.2;

6) at this point the Virtual ISL endpoints shall exchange FC virtual port initialization frames to enable and identify port operation. The E_Port port mode initialization state machine is described in FC-SW-4 and the B_Access portal initialization state machine is described in 9.4.3.3.2. Switch-to-switch authentication shall use FC-SP authentication mechanisms;

7) an FC-BB-3_IP device operates in E_Port or B_Port mode. When operating in E_Port mode, normal FSPF messages are exchanged and the switch port becomes operational. When operating in B_Port mode, it is expected that the external E_Ports may exchange FSPF messages over the Virtual ISL which result in the link becoming operational;

8) link costs are implementation-defined;

9) in certain deployments, a single FC-BB-3_IP device may establish Virtual ISLs/FCIP Links with multiple FC-BB-3_IP device endpoints. In this situation, the FC-BB-3_IP device shall manage TCP operational parameters independently for each Virtual ISL or FCIP Link. Also, the FC-BB-3_IP device VE_Port may perform the E_Port initialization independently, for each Virtual ISL/FCIP Link. The B_Access also may perform initialization independently, for each Virtual ISL/FCIP Link; and

10) the FC Entity may participate in determining allowed TCP connections, TCP connection parameters, quality of service usage, and security usage by modifying interactions with the FCIP Entity that are modeled as a shared database. See RFC 3821.

9.7.3.3 Procedures for data transfer

The procedures for data transfer are as follows:

a) the sending FC Entity shall deliver FC frames to the correct FCIP_LEP/FCIP_DE in the correct FCIP Entity;

b) each FC frame delivered to the FCIP_DE shall be accompanied by a time value synchronized with the clock maintained by the FC Entity at the other end of the FCIP Link (see 9.3.4.5.2.2); and

c) when FC frames exit FCIP_DE(s) via the FC Transmitter Portal(s), the FC Entity should forward them to the FC fabric. However, before forwarding the FC frame the FC Entity shall verify the end-to-end transit time as described in 9.3.4.5.2.3.

9.7.3.4 Procedures for FCIP Link disconnection

The FC Entity may require the FCIP Entity to perform TCP close requests (e.g., to perform a controlled shut-down of a link or to respond to high link error rates). If the FC Entity requests the closure of all TCP connections in an FCIP Link, the FCIP Link is disconnected.

When the FCIP Link is disconnected, notification of the disconnection shall be accomplished according to the procedures in 9.8.3.5.

9.7.3.5 Procedures for multiple connection management

A pair of FC-BB-3_IP device endpoints may establish a number of TCP connections between them. Since a Virtual ISL potentially maps a fairly large number of FC flows, where a flow is defined as a pair of Fibre Channel S_ID/D_ID addresses, it may not be practical to establish a separate TCP connection for each FC flow. However, once an FC flow is assigned to an FCIP_DE within the Virtual ISL, all FC frames of that flow shall be sent on that same FCIP_DE. This rule is in place to honor any in-order delivery guarantees that may have been made between the two end points of the FC flow.

When a TCP connect request is received and that request would add a new TCP connection to an existing FCIP_LEP, the procedures described in 9.7.2.3.1 shall be followed.

9.7.4 Procedures for error detection recovery

9.7.4.1 Procedures for handling invalid FC frames

Data corruption is detected at two different levels, TCP checksum and FC frame encapsulation errors. Data corruption detected at the TCP level shall be recovered via TCP data recovery mechanisms. The recovery for FC frame errors is described below. The TCP and FC frame recovery operations are performed independently.

Fibre Channel frame errors and the expected resolution of those errors are described in RFC 3821 and summarized below:

NOTE 32 - The behavior given below is that of the FCIP Entity.

a) all incoming frames on the FC receiver port are verified for correct header, proper format, valid length and valid CRC. A frame having an incorrect header or CRC shall be discarded or processed in accordance with the rules for the particular type of FC_Port;

b) all frames transmitted by the encapsulated frame tansmitter are valid FC encapsulations of valid FC frames with correct TCP check sums on the correct TCP/IP connection;

c) the FC frames contained in incoming encapsulated frames on the encapsulated frame receiver port are verified for a valid header, proper content, proper SOF and EOF values, and valid length. FC frames that are not valid according to those checks are managed according to the following rules:

   1) the frame may be discarded; or

   2) the frame may be transmitted in whole or in part by the FC transmitter port and ended with an EOF indicating that the content of the frame is invalid; and

d) if there is any discrepancy between statements in this subclause and RFC 3821, then RFC 3821 shall prevail.

9.7.4.2 Procedures for error recovery

The FC Entity shall recover from events that the FCIP Entity is unable to handle, such as:

a) loss of synchronization with FCIP frame headers from the encapsulated frame receiver portal requiring resetting the TCP connection; and

b) recovering from FCIP frames that are discarded as a result of synchronization problems.

The FC Entity may recover from connection failures.

Since FC Primitive Signals and Primitive Sequences are not exchanged between FCIP devices, there may be times when an FC frame is lost within the IP network. When this event occurs it is the responsibility of the communicating FC devices to detect and correct the errors based on the features defined in FC-FS-2.

In order to facilitate faster detection of loss of link connectivity, FC Entities shall make use of the Link Keep Alive (LKA) ELS (see FC-LS). The LKA ELS is exchanged across the Virtual ISL as shown in figure 28 (i.e., E_Port implementation) or figure 31 (i.e., B_Port implementation). The exact number of lost LKA heartbeats that forces the FC Entity to mark the link down is a configurable parameter with a default value of 2. Once the link has been marked down, the FC Entity shall attempt to re-establish the link via the FCIP Entity.

9.7.5 FC-BB-3_IP system parameters

9.7.5.1 FC timers

FC has two important timeouts, E_D_TOV and R_A_TOV.

E_D_TOV determines the life of an individual FC frame in any particular fabric element. The effects of E_D_TOV on the fabric as a whole are typically cumulative since each fabric element contains its own E_D_TOV timers for any frame received.

R_A_TOV determines the life of an individual FC frame in the fabric as a whole. For a fabric, R_A_TOV implies that no particular frame shall remain in, and thus be emitted from, the fabric after the timer expires.

K_A_TOV is a timer defined in this standard that is used by the Link Keep Alive (LKA) ELS (see FC-LS) as a trigger for issuing LKA. The LKA should be sent at least every K_A_TOV if no traffic has been sent and/or received on the connection. The default value for K_A_TOV is 1/2 E_D_TOV.

9.7.5.2 TCP timers

Given the multitude of current and probable TCP implementations, IETF Requests For Comments related to TCP, applications network requirements, etc., it is impossible to provide even rudimentary guidance in suggesting values for the tunable parameters associated with TCP.

9.7.5.3 Maximum number of attempts to complete an encapsulated FC frame transmission

This is an unspecified parameter and is implementation-specific.

9.7.5.4 Maximum number of outstanding encapsulated FC frames

This is an unspecified parameter and is implementation-specific.

9.8 FC-BB-3_IP service considerations

9.8.1 Latency delay

The time required for a frame to pass from one FC-BB-3_IP device to another across the IP network is variable and beyond the direct control of the FC/FCIP Entity pair. However, the IP network transit time affects the FC Entity's ability to meet FC timeout requirements (e.g., the R_A_TOV requirements of the fabric). Therefore, the FC Entity is required to use facilities provided by the FCIP Entity to compute the IP network transit time for frames. See 9.3.4.5.2.

Class F frames may be excepted from IP network transit time checking, however, all other classes of frames shall have their IP network transit time computed and checked. If a frame is found to have an IP network transit time that would cause the frame's lifetime in the fabric to exceed FC requirements, the FC Entity shall discard the frame.

9.8.2 Throughput

9.8.2.1 How timeouts affect throughput

Both FC and TCP timeouts affect throughput as follows:

a) small R_A_TOV values may cause encapsulated FC frames to be discarded frequently in the FCIP_DE necessitating FC end-node retransmissions;

b) large TCP timeouts may result in encapsulated FC frames becoming stale in the IP network, leading the FCIP_DE to discard them again necessitating FC end-node retransmissions; and

c) discarding encapsulated FC frames due to improper settings of timeout values and errors in the IP network lowers the effective throughput.

The FC/FCIP Entities have little or no control over TCP timeouts. The FC/FCIP Entities never initiate retransmissions, that is done either by TCP or by the FC end nodes.

9.8.2.2 How loss affects throughput

TCP retransmissions occur due to loss or corruption of TCP segments. If TCP retransmissions cause the allowed transit time to exceed a threshold, then encapsulated FC frames shall be discarded. Either case is likely to cause the effective throughput to be reduced.

9.8.2.3 Other factors that affect throughput

Throughput may be affected by a mismatch in the effective rates of data transfer across the FC and the IP network interfaces. This mismatch may occur due to differences in the physical line speeds at the FC network and the IP network interfaces or due to the fundamental difference in the two flow control mechanisms.

FC uses BB_Credit flow control and TCP uses a sliding window based flow control. FC-BB-3_IP does not specify the mechanism that aligns the two flow control schemes, although it is thought that performance may be affected if this aspect is not considered. The FC-BB-3_IP device needs to ensure that the TCP connections are able to handle the frame arrival rate from the FC fabric. The FC Entity shall work cooperatively with the FCIP Entity to manage flow control problems in either the IP network or FC fabric.

In order to achieve better TCP aggregate throughput properties in the face of packet losses, a pair of peer FC-BB-3_IP devices may use multiple DEs between them, and use appropriate policies for mapping FC frames to these connections.

9.8.3 Reliability

9.8.3.1 Loss of connectivity

The FC-BB-3_IP device has the capability of detecting loss of connectivity with its remote peer (see 9.7.4.2). Upon detecting a loss of connectivity, an FC-BB-3_IP device establishes a new connection, or uses an existing TCP connection to the same FC-BB-3_IP device endpoint. An FC-BB-3_IP device shall not retransmit an encapsulated FC frame on the new connection. This is to ensure exactly-once delivery semantics to the FC endpoint.

The FC Entity may test for failed TCP connections. Should such a test detect a failed TCP connection, the FC Entity shall disconnect that connection following the procedures in 9.7.3.4.

9.8.3.2 Loss of synchronization

The FC Entity shall recover from events that the FCIP Entity is unable to handle, such as:

a) loss of synchronization with FC-BB-3_IP encapsulated FC frame headers from the Encapsulated Frame Receiver Portal requiring resetting the TCP connection; and

b) recovering from FC-BB-3_IP encapsulated FC frames that are discarded as a result of synchronization problems (see RFC 3821).

9.8.3.3 Loss or corruption of TCP segments

TCP flow control and error control has mechanisms to detect lost or corrupted TCP segments. TCP retransmits the TCP segments that were lost or corrupted.

TCP flow control provides the ability to regulate the flow of data on the IP network interface based on the perceived IP network congestion conditions, potentially avoiding large losses of data.

9.8.3.4 Loss or corruption of FC frames

The FC interface of the FC-BB-3_IP device has no mechanisms to detect lost data but only to detect corrupted frames. Corrupted frames detected prior to transmission into the IP network, are discarded and not sent over the IP network.

FC BB_Credit flow control provides the ability to regulate the flow of data on the FC network interface with no loss.

9.8.3.5 FCIP error reporting

The FC Entity receives notifications from the FCIP Entity due to a number of errors detected by the FCIP Entity. As a result, the E_Port implementation of the FC Entity shall report those errors to the local FC switch element via the local VE_Port (see figure 26). Similarly the B_Port implementation shall report the error to the local B_Access (see figure 30). In addition, the FC Entity may pass these error reports to the local PMM for inclusion in a local event log.

The FC Entity shall convert the error message received from the FCIP Entity into a Registered Link Incident Report (RLIR) (see FC-LS). It is the RLIR that is forwarded from the FC Entity to either the VE_Port (see figure 26) or B_Access (see figure 30).

On receipt of the message from the FC Entity, VE_Port or B_Access shall immediately forward the RLIR to the Domain Controller of the Switch.

As a minimum the FC Entity shall accept the following information from the FCIP Entity:

a) loss of FC frame synchronization (see RFC 3821);

b) failure to setup TCP connection (see RFC 3821);

c) duplicate connect request (see RFC 3821) ;

d) TCP connect request timeout (see RFC 3821) ;

e) successful completion of FC Entity request to close TCP connection (see RFC 3821);

f) loss of TCP connectivity (see RFC 3821);

g) excessive number of dropped datagrams (see RFC 3821);

h) any confidentiality violations (see RFC 3821);

i) SA parameter mis-match (see RFC 3821); and

j) LKA timeout notification (see FC-LS).

The FC Entity shall generate and forward an RLIR to the management server for the following:

a) loss of FC frame synchronization (see RFC 3821);

b) failure to setup additional TCP connection (see RFC 3821); and

c) additional duplicate TCP connect request (see RFC 3821).

9.8.4 Quality of Service (QoS)

The FC-BB-3_IP protocol may use TCP/IP QoS features to support FC capabilities.

9.8.5 Delivery order

Each VE_Port/FCIP_LEP pair defines a separate FCIP Link. FCIP_DEs within an FCIP_LEP share the FCIP Link. Multiple FCIP_DEs between FCIP_LEPs introduce multiple traffic paths (e.g., Class F, Class 2/3/4). The order in which the FCIP_DEs are serviced on the FCIP Link is not specified. One possibility is providing different priority levels to each traffic path changing the overall delivery order.

The only delivery order guarantee provided by TCP is correctly ordered delivery of FC-BB-3_IP encapsulated FC frames between a pair of FCIP_DEs. The FC Entity is expected to specify and handle all other FC frame delivery ordering requirements.

NOTE 33 - The order of the FC frames sent by the encapsulated frame transmitter may not be the same as the order sent by the source FC end node. This is due to the fact that some types of FC login allow FC frames to be re-ordered in the FC fabric before reaching the FC receiver port.

9.8.6 IP multicast and broadcast

This standard does not make use of IP multicast and broadcast.

9.8.7 Security and authentication

The IETF security standards referenced by RFC 3821 provide numerous mechanisms for securing TCP connections between FC/FCIP Entity pairs (e.g., IPsec packet authentication and confidentiality). It is important to note that the public Internet IP network is subject to a large variety of security attacks, meaning that serious consideration should be given to enabling the full suite of security features described in FCIP whenever the public Internet IP network is to be used to transmit FCIP frames.

The TCP connection authentication mechanism described in 9.7.2.3.1 provides FC-BB-3 specific authentication for the second, third, etc., TCP connections in an FCIP Link and its associated Virtual ISL as long as the first TCP connection is authenticated using the mechanisms described in 9.7.2.3 and FC-SP.

10 FC-BB-3_GFPT Structure and Concepts

10.1 Applicability

Clause 4 discussed the FC-BB-3_GFPT reference model. This clause discusses the FC-BB-3_GFPT functional model.

10.2 FC-BB-3_GFPT overview

This clause discusses further aspects of FC-BB-3_GFPT operation, including initialization, flow control, and procedures for adaptation of FC information for transport using the Asynchronous Transparent Generic Framing Procedure (GFPT). Mapping FC-BB-3_GFPT into Asynchronous GFPT allows for applications where the WAN transport rate is less that the FC client data rate.

Figure 35 illustrates the protocol levels and layers involved in FC-BB-3_GFPT processes and devices.

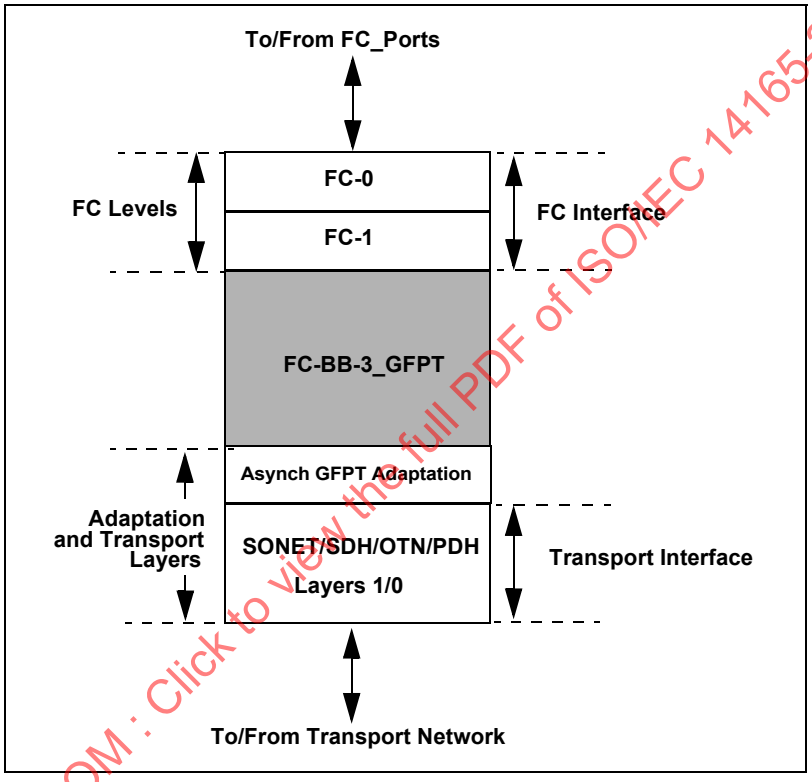


Figure 35 – FC-BB-3_GFPT protocol levels and layers

The FC-BB-3_GFPT level, the functions of which are defined by this standard, produces a stream of 8B/10B codewords in either direction. In the FC interface direction, it produces a full FC rate, synchronous codeword stream, constructed from the FC Ordered Sets (i.e., frames, Primitive Signals and Primitive Sequences) delivered by the GFPT adaptation level, and emulates a standard FC-2 level interfacing to a standard FC-1 level. In the WAN network direction, the FC-BB-3_GFPT level produces a filtered stream of codewords that includes all FC frames, selected Primitive Signals and selected Ordered Sets of Primitive Sequences that are forwarded by the FC-1 level, as well as GFPT_WAN Primitive Signals that are used for WAN flow control and management purposes. The asynchronous GFPT adaptation level and the transport (i.e., SONET/SDH/OTN/PDH) layers are described by ITU-T standards (see clause 2). Buffering of selected codewords and Ordered Sets, in both directions of propagation (i.e., from the FC-1 level, and from the GFPT adaptation level), is one function of the FC-BB-3_GFPT level components.

10.3 FC-BB-3_GFPT functional model

10.3.1 FC-BB-3_GFPT initialization

FC-BB-3_GFPT devices do not directly participate in FC link initialization or FC_Port initialization. FC link initialization and FC_Port initialization occurs between the attached FC_Port and the remote FC_Port. This is a key distinction between the FC-BB-3_GFPT device model and other FC-BB-3 models.

FC-BB-3_GFPT devices monitor and transport the FC_Port initialization Exchanges (i.e., ELP, FLOGI, and PLOGI) that take place between the attached FC_Port and the remote FC_Port, but only to capture and potentially modify the parameters that are relevant to link-level flow control.

If an FC-BB-3_GFPT device hosts more than one GFPT_WAN facility, and each GFPT_WAN facility includes the GFPT_WAN link and the corresponding client-facing FC_Ports on each device, a separate state machine operates on each GFPT_WAN facility. The state machines in each FC-BB-3_GFPT device on the same GFPT_WAN facility operate independently of one another.

10.3.2 FC-BB-3_GFPT initialization state machine

10.3.2.1 Initialization state machine keywords

The keywords used in the FC-BB-3_GFPT initialization state machine diagram (see figure 36) are specified in table 36.

Table 36 – FC-BB-3_GFPT initialization state machine keywords

Keyword	Description
WAN	Received from remote FC-BB-3_GFPT device.
FC	Received from attached FC_Port.
10B_ERR	GFPT-defined character used to represent an FC illegal codeword or running disparity error at ingress, or an irresolvable character error produced during GFPT_WAN transmission. Interpretation of these characters is described in 10.3.8.
FC-Error	8B/10B character error or running disparity error.

10.3.2.2 Initialization state machine

The FC-BB-3_GFPT initialization state machine is specified in figure 36.

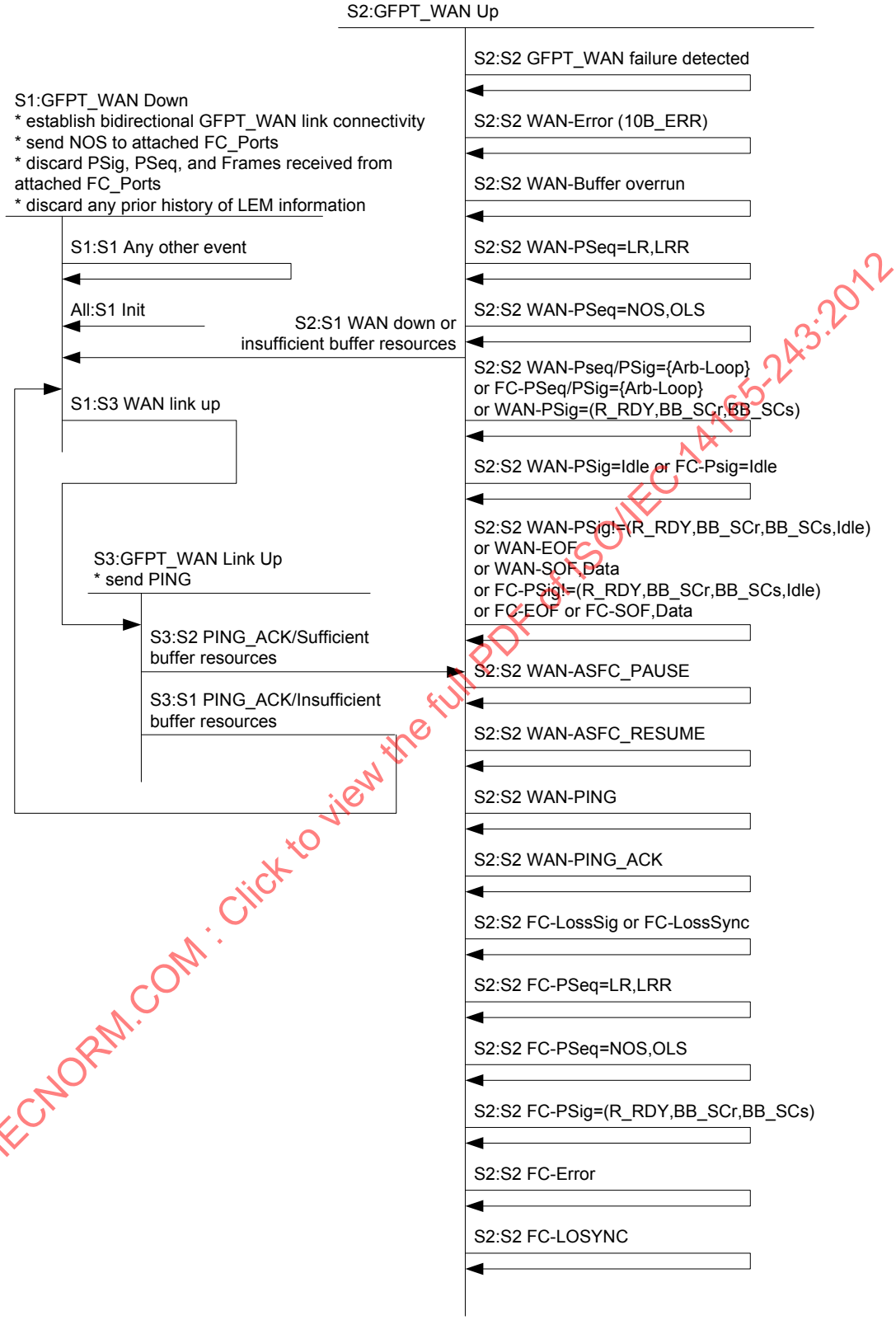


Figure 36 – FC-BB-3_GFPT initialization state machine

Transition All:S1 Init. This transition occurs when an initialization event occurs in a state where it is not already handled. An initialization event may be:

a) a power-on reset condition;

b) outside intervention requesting an initialization event; or

c) a vendor-specific initialization event.

State S1: GFPT_WAN Down. In this state, FC-BB-3_GFPT devices shall:

a) initiate establishment of GFPT_WAN links, including Transport Trail configuration and turn-up, and the establishment of a fully operational GFP server in both directions (see 10.3.8);

b) send the Not Operational (NOS) Primitive Sequence to attached FC_Ports, if any;

c) discard any Primitive Signals, Primitive Sequences, or frames received from attached FC_Ports; and

d) discard any prior history of Login Exchange Monitor (LEM) information (see 10.3.3).

Transition S1:S1 Any other event. This transition occurs when an event other than what is specified in state S1 occurs.

Transition S1:S3 WAN link up. This transition occurs when a WAN link up event is detected.

State S3: GFPT_WAN Link Up. In this state, the FC-BB-3_GFPT device shall transmit one or more PING Primitive Signals to the remote FC-BB-3_GFPT device.

Transition S3:S2 PING_ACK/Sufficient buffer resources. This transition occurs upon reception of at least one PING_ACK Primitive Signal from the remote FC-BB-3_GFPT device and sufficient buffer resources are available to prevent data loss (see 10.3.6). This transition is not coordinated between the FC-BB-3_GFPT devices at each end of the GFPT_WAN link.

Transition S3:S1 PING_ACK/Insufficient buffer resources. This transition occurs upon reception of at least one PING_ACK Primitive Signal from the remote FC-BB-3_GFPT device and WAN latency measurements indicate insufficient buffer resources are available within the device to prevent data loss (see 10.3.6).

State S2: GFPT_WAN Up. In this state, the GFPT_WAN link is normally (i.e., excluding Transport Trail protection events) operational and the attached FC_Ports at either end of the link are permitted to communicate with each other.

Validated (see 10.3.8) Primitive Sequences received from the attached FC_Ports shall be transmitted between attached FC_Ports by the FC-BB-3_GFPT devices according to the rate adaptation and other rules described in 10.3.8.

If a Primitive Sequence is received from an attached FC_Port while an FC-BB-3_GFPT device has ceased transmitting information across the GFPT_WAN link because it has received one or more ASFC_PAUSE Primitive Signals, the FC-BB-3_GFPT device shall:

a) flush its WAN-facing buffer;

b) clear the WAN pause condition; and

c) transmit the Primitive Sequence to the remote FC-BB-3_GFPT device.

If a Primitive Sequence is received from a remote FC-BB-3_GFPT device while an FC-BB-3_GFPT device is unable to transmit frames to an attached FC_Port because it has no available BB_Credit, the FC-BB-3_GFPT device shall:

a) flush its attached FC_Port-facing buffer; and

b) transmit the Primitive Sequence to the attached FC_Port.

When FC_BB_3_GFPT devices have no data to transmit to the WAN, the FC_BB-3_GFPT device shall follow the rules specified in ITU-T Rec. G.7041/Y.1303.

Transition S2:S1 WAN down or insufficient buffer resources. This transition occurs when:

a) a WAN down event is detected. A WAN down event occurs when GFPT_WAN recovery does not occur within WAN_HOLDOFF_TOV (see 10.3.9) of the time the GFPT_WAN failure is detected; or