FINAL
DRAFT

INTERNATIONAL
STANDARD

**ISO/IEC
FDIS
14165-331**

ISO/IEC JTC **1**

Secretariat: **ANSI**

Voting begins on:
2004-06-17

Voting terminates on:
2004-08-17

Information technology — Fibre Channel —

Part 331:
Virtual Interface (FC-VI)

Technologies de l'information — Canal de fibres —

Partie 331: Interface virtuelle (FC-VI)

Please see the administrative notes on page iii

RECIPIENTS OF THIS DRAFT ARE INVITED TO SUBMIT, WITH THEIR COMMENTS, NOTIFICATION OF ANY RELEVANT PATENT RIGHTS OF WHICH THEY ARE AWARE AND TO PROVIDE SUPPORTING DOCUMENTATION.

IN ADDITION TO THEIR EVALUATION AS BEING ACCEPTABLE FOR INDUSTRIAL, TECHNOLOGICAL, COMMERCIAL AND USER PURPOSES, DRAFT INTERNATIONAL STANDARDS MAY ON OCCASION HAVE TO BE CONSIDERED IN THE LIGHT OF THEIR POTENTIAL TO BECOME STANDARDS TO WHICH REFERENCE MAY BE MADE IN NATIONAL REGULATIONS.

Reference number
ISO/IEC FDIS 14165-331:2004(E)

© ISO/IEC 2004

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS

Not for Resale

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

Copyright notice

This ISO document is a Draft International Standard and is copyright-protected by ISO. Except as permitted under the applicable laws of the user's country, neither this ISO draft nor any extract from it may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, photocopying, recording or otherwise, without prior written permission being secured.

Requests for permission to reproduce should be addressed to either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Reproduction may be subject to royalty payments or a licensing agreement.

Violators may be prosecuted.

--`,,`,,`-`-`,,`,,`,`,,`---

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS

© ISO/IEC 2004 – All rights reserved

Not for Resale

In accordance with the provisions of Council Resolution 21/1986, this document is **circulated in the English language only**.

© ISO/IEC 2004 All rights reserved

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

PATENT STATEMENT

CAUTION: The developers of this standard have requested that holders of patents that may be required for the implementation of the standard, disclose such patents to the publisher. However, neither the developers nor the publisher have undertaken a patent search in order to identify which, if any, patents may apply to this standard. As of the date of publication of this standard, following calls for the identification of patents that may be required for the implementation of the standard, no such claims have been made. No further patent search is conducted by the developer or the publisher in respect to any standard it processes. No representation is made or implied that licenses are not required to avoid infringement in the use of this standard.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Contents

Page

1 Scope and purpose 2

2 Normative references 2
2.1 Normative references overview 2
2.2 Approved references 2
2.3 References under development 2
2.4 Other references 2
2.4.1 VI Architecture 1.0 reference specifications 2
2.4.2 Miscellaneous 2

3 Definitions and abbreviations 3
3.1 FC-VI definitions 3
3.2 VI Definitions 4
3.3 **Abbreviations 5**
3.4 Editorial conventions 5

4 Structure and concepts 7
4.1 Fibre channel structure and concepts 7
4.2 FC-VI structure and concepts 7

5 FC-VI protocol overview 10
5.1 FC-VI information units 10
5.2 FC-VI message transfer operation 10
5.2.1 FC-VI message transfer 10
5.2.2 FC-VI send message transfer operation 11
5.2.3 FC-VI RDMA write message transfer operation 12
5.2.4 FC-VI RDMA read message transfer operation 14
5.2.5 IU reception at an FC-VI edpoint 15
5.3 FC-VI connection setup operation 16
5.3.1 FC-VI client-server and peer-peer connection setup 16
5.3.2 FC-VI client-server connection setup 17
5.3.3 FC-VI Peer-to-Peer Connection Establishment 18
5.3.4 FC_VI concurrent peer-to-peer connection setup 20
5.3.5 FC-VI Disconnect Operation 22
5.4 Exchange ID reuse 23
5.5 Sequence ID reuse 24
5.6 Frame synonym detection 24
5.7 VI message length 25
5.8 FC-FS header usage for FC-VI 25
5.8.1 FC-FS header usage 25
5.8.2 CS_CTL field 26
5.8.3 TYPE field 26
5.8.4 F_CTL field 26
5.8.5 DF_CTL field 26
5.8.6 SEQ_CNT field 26
5.8.7 Parameter field 26
5.9 FC-VI device_header 27
5.9.1 FC-VI device_header description 27
5.9.2 FCVI_HANDLE field 27
5.9.3 FCVI_OPCODE field 28
5.9.4 FCVI_FLAGS field 28
5.9.4.1 FCVI_FLAGS field description 28

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

5.9.4.2 FCVI_FLAGS for message request IUs 28
5.9.4.3 FCVI_FLAGS for message response IUs 28
5.9.4.4 FCVI_FLAGS for connect request IUs 29
5.9.4.5 FCVI_FLAGS for connect response IUs 30
5.9.4.6 FCVI_FLAGS for disconnect IUs 31
5.9.5 Reserved fields 32
5.9.6 FCVI_MSG_ID field 32
5.9.7 FCVI_PARAMETER field 32
5.9.7.1 FCVI_PARAMETER field format 32
5.9.7.2 Connect response reason codes 34
5.9.7.2.1 Connect response non-error reason codes 34
5.9.7.2.2 Connect response error reason codes 34
5.9.7.3 Message response / disconnect reason codes 34
5.9.7.3.1 Descriptor error reason codes 34
5.9.7.3.2 Remote FC-VI port non-descriptor errors 35
5.9.7.3.3 Reserved for future expansion 35
5.9.7.3.4 Vendor unique reason codes 35
5.9.8 FCVI_RMT_VA field 35
5.9.9 FCVI_RMT_VA_HANDLE field 36
5.9.10 FCVI_TOT_LEN field / FCVI_CONNECTION_ID field 36

6 FC-VI Information Unit (IU) formats 37
6.1 FC-VI IU overview 37
6.2 FCVI_SEND_RQST IU 37
6.2.1 FCVI_SEND_RQST IU description 37
6.2.2 FCVI_SEND_RQST Device_Header information 37
6.3 FCVI_SEND_RESP IU 37
6.3.1 FCVI_SEND_RESP IU description 37
6.3.2 FCVI_SEND_RESP Device_Header information 37
6.4 FCVI_WRITE_RQST IU 38
6.4.1 FCVI_WRITE_RQST IU overview 38
6.4.2 FCVI_WRITE_RQST IU Device_Header information 38
6.5 FCVI_WRITE_RESP IU 38
6.5.1 FCVI_WRITE_RESP IU description 38
6.5.2 FCVI_WRITE_RESP IU Device_Header information 39
6.6 FCVI_READ_RQST IU 39
6.6.1 FCVI_READ_RQST IU description 39
6.6.2 FCVI_READ_RQST IU Device_Header information 39
6.7 FCVI_READ_RESP IU 40
6.7.1 FCVI_READ_RESP IU description 40
6.7.2 FCVI_READ_RESP IU Device_Header information 40
6.8 FCVI_CONNECT_RQST IU 40
6.8.1 FCVI_CONNECT_RQST IU description 40
6.8.2 FCVI_CONNECT_RQST Device_Header information 41
6.8.3 FCVI_CONNECT_RQST Payload Information 41
6.9 FCVI_CONNECT_RESP1 IU 42
6.9.1 FCVI_CONNECT_RESP1 IU description 42
6.9.2 FCVI_CONNECT_RESP1 Device_Header information 42
6.9.3 FCVI_CONNECT_RESP1 Payload Information 43
6.10 FCVI_CONNECT_RESP2 IU 44
6.10.1 FCVI_CONNECT_RESP2 IU description 44
6.10.2 FCVI_CONNECT_RESP2 Device_Header information 44
6.11 FCVI_CONNECT_RESP3 IU 44

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

6.11.1 FCVI_CONNECT_RESP3 IU description 44
6.11.2 FCVI_CONNECT_RESP3 Device_Header information 44
6.12 FCVI_DISCONNECT_RQST IU 45
6.12.1 FCVI_DISCONNECT_RQST IU description 45
6.12.2 FCVI_DISCONNECT_RQST Device_Header information 45
6.13 FCVI_DISCONNECT_RESP IU 46
6.13.1 FCVI_DISCONNECT_RESP IU description 46
6.13.2 FCVI_DISCONNECT_RESP Device_Header information 46

7 FC-VI Addressing and naming 47
7.1 FC-VI Addressing and naming overview 47
7.2 FCVI_NET_ADDRESS format 47
7.3 FCVI_ATTRIBUTES format 48
7.4 FC-VI address resolution 50
7.5 FARP ELS 51
7.6 Name server queries 52
7.7 Validation of host address to N_Port identifier mappings 52
7.7.1 Address mapping overview 52
7.7.2 Point-to-point topology 52
7.7.3 Private loop topology 52
7.7.4 Public loop topology 53
7.7.5 Fabric topology 53

8 FC-VI Error detection and recovery 54
8.1 FC-VI error detection and recovery overview 54
8.2 FC-VI endpoint states 54
8.3 FCVI_ULP_TIMEOUT definition 54
8.4 Message transfer error detection and recovery rules 55
8.4.1 Message error detection 55
8.4.2 Message transfer error recovery 55
8.5 Connection setup error detection and recovery rules 56
8.5.1 Connection setup error handling overview 56
8.5.2 Connection setup error detection 56
8.5.3 Connection setup error recovery 56
8.5.4 Connection setup originator retry rules 57
8.6 Disconnect operation error detection and recovery rules 57
8.6.1 Disconnect operation error handling overview 57
8.6.2 Disconnect operation error detection 57
8.6.3 Disconnect operation error recovery rules 58

Annex A: (informative) Concurrent matching peer requests examples 59
A.1 Annex A overview 59
A.2 Case 1 59
A.3 Case 2 60
A.4 Case 3 61
A.5 Case 4 61
A.6 Case 5 62
A.7 Case 8 63

Annex B: (informative) FC-VI Message Transfer Error Handling Examples 64
B.1 Annex B Overview 64
B.2 Message Transfer Error Handling Operation 64
B.2.1 Message Transfer Error Handling Operation Overview 64
B.2.2 Message Transfer Error Definitions 64

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

B.2.3 Error Detection and Recovery Rule Processing 65
B.2.4 Message Responder and Message Originator Error Recovery Actions 65
B.2.5 Message responder error detection actions 66
B.2.6 Message originator Class 2 error detection actions 68
B.2.6.1 Message originator Class 2 error detection overview 68
B.2.6.2 Message response timeout at message originator 69
B.3 Message transfer error detection and recovery examples 69
B.3.1 Error examples overview 69
B.3.2 Mrcv > Mexp error example 69
B.3.2.1 Mrcv > Mexp example description 69
B.3.2.2 Mrcv > Mexp: In-order fabric and unreliable 70
B.3.2.3 Mrcv > Mexp: In-order fabric and reliable delivery 70
B.3.2.4 Mrcv > Mexp: Out-of-order fabric 71
B.3.3 Mrcv = Mexp error example 71
B.3.3.1 Mrcv = Mexp example description 71
B.3.3.2 Mrcv = Mexp: In-order fabric and unreliable 72
B.3.3.3 Mrcv = Mexp: Out-of-order fabric 72

Annex C: (informative) Connection Setup Error Handling Examples 73
C.1 Annex C overview 73
C.2 Connection setup error handling definitions 73
C.3 Connect request originator and connect request responder rules 73
C.4 Connect request originator rules 73
C.5 Connect request responder rules 75
C.5.1 Connect request responder retry rules 76
C.6 Error detection and recovery examples for connection setup 77
C.6.1 Overview 77
C.6.2 FC-VI connection setup timers 78
C.6.3 VipConnectRequest completion 79
C.6.4 VipConnectAccept completion 79
C.6.5 Enabling message transmission and reception 79
C.6.6 Client timeout of VipConnectRequest 80
C.6.7 Lost FCVI_CONNECT_RQST IU 80
C.6.7.1 Lost FCVI_CONNECT_RQST IU example 80
C.6.7.2 Retried connection setup 81
C.6.8 Lost FCVI_CONNECT_RESP1 IU 82
C.6.9 Lost FCVI_CONNECT_RESP2 IU 83
C.6.9.1 Lost FCVI_CONNECT_RESP2 IU example 83
C.6.9.2 Server timing out connection setup 84
C.6.10 Lost FCVI_CONNECT_RESP3 IU 84
C.6.10.1 Lost FCVI_CONNECT_RESP3 IU example 84

Annex D: (informative) Disconnect operation error handling examples 86
D.1 Disconnect operation example description 86
D.2 FC-VI disconnect operation example 87

Annex E: (informative) Message streaming for reliable reception 89

Annex F: (informative) Enabling Message Transmission in the FC-VI NIC 90

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Tables

Page
1 FC-VI Information unit summary 10
2 Peer B actions based on connect responses from peer A 23
3 16-byte FC-VI device_header 27
4 32-byte FC-VI device_header 27
5 FCVI_FLAGS Bit definitions for message request IUs 28
6 FCVI_FLAGS Bit definitions for message response IUs 29
7 FCVI_FLAGS Bit definitions for connect request IUs 29
8 FC-VI connection mode definition 30
9 FCVI_FLAGS Bit definitions for connect response IUs 30
10 FCVI_FLAGS Bit definitions for disconnect IUs 31
11 FCVI_PARAMETER field for connect response and disconnect IUs 32
12 Reason code for CONN_STS 33
13 FCVI_CONNECT_RQST IU payload format 41
14 FCVI_CONNECT_RESP1 IU Payload Format 43
15 FCVI_NET_ADDRESS Format 48
16 FCVI_ATTRIBUTES format 48
17 Format of FCVI_ATTR_FLAGS in FCVI_ATTRIBUTES 49
18 FCVI_QOS format 49
A.1 Peer B actions based on connect responses from peer A 59

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Figures

Page
1 FC-VI addressing objects 9
2 FC-VI send for Unreliable Delivery or Reliable Delivery 11
3 FC-VI send for Reliable Reception 12
4 FC-VI RDMA write for Unreliable Delivery or Reliable Delivery 13
5 FC-VI RDMA write for Reliable Reception 14
6 FC-VI RDMA read for Reliable Reception and Reliable Delivery 15
7 Concurrent Receive Streams at a FC-VI Endpoint 16
8 FC-VI client-server connection setup 17
9 Peer-to-peer connection setup 19
10 Peer-to-peer connection setup, concurrent matching peer requests 21
11 FC-VI disconnect operation 23
12 FC-FS header for send operation 25
A.1 Case 1 60
A.2 Case 2 60
A.3 Case 3 61
A.4 Case 4 62
A.5 Case 5 62
A.6 Case 6 63
B.1 Mrcv > Mexp 70
B.2 Mrcv = Mexp 72
C.1 Client-server connection setup 77
C.2 Client timeout of VipConnectRequest 80
C.3 Lost FCVI_CONNECT_RQST IU 81
C.4 Lost FCVI_CONNECT_RESP1 IU 82
C.5 Lost FCVI_CONNECT_RESP2 83
C.6 Lost FCVI_CONNECT_RESP3 IU 84
D.1 FC-VI disconnect operation 87

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Foreword

ISO (the International Organisation for Standardisation) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organisations, governmental and nongovernmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardisation.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 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.

International Standard ISO/IEC 14165-331 was prepared by Joint Technical Committee ISO/IEC JTC 1, Subcommittee SC 25, Interconnection of information technology equipment.

Introduction

This standard defines an upper-layer protocol within the domain of Fibre Channel, that is designed to permit efficient peer-to-peer or client-server messaging between nodes, and to comply with the Virtual Interface (VI) Architecture. Vendors that wish to implement devices that connect to FC-VI may follow the requirements of this and other normatively referenced standards to manufacture an FC-VI compliant device.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

1 Scope and purpose

This standard defines the Fibre Channel mapping protocol for the Virtual Interface (VI) Architecture (FC-VI). FC-VI defines the Fibre Channel Information Units in accordance with the VI Architecture model. FC-VI additionally defines how Fibre Channel services are used to perform the services required by the VI Architecture model of its network transport.

2 Normative references

2.1 Normative references overview

The following standards contain provisions that, through reference in the text, constitute provisions of this international standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this international standard are encouraged to investigate the possibility of applying the most recent editions of the standards listed below. Members of IEC and ISO maintain registers of currently valid International Standards.

2.2 Approved references

The following approved international and regional standards (ISO, IEC, CEN/CENELEC and ITUT) may be obtained from the international and regional organizations that control them.

None.

2.3 References under development

At the time of publication, the following referenced standards were still under development.

ISO/IEC 14165-251, *Fibre Channel - Framing and Signaling Interface (FC-FS)*

ISO/IEC 14165-413, *Fibre Channel - Generic Services-3 (FC-GS-3)*

ISO/IEC 14165-122, *Fibre Channel - Arbitrated Loop-2 (FC-AL-2)*

2.4 Other references

2.4.1 VI Architecture 1.0 reference specifications

The following references for VI Architecture are the product of Intel, Microsoft, and Compaq. The VI Architecture 1.0 specification is completely defined in these three documents. For information on the current status and availability of the document, contact the VI Developer Forum (VIDF) web site at www.vidf.org.

Virtual Interface Architecture Specification, V1.0 (VI-ARCH)

Virtual Interface (VI) Architecture Developer’s Guide, V1.0 (VI-DG)

Virtual Interface (VI) Architecture Developer’s Guide Error Table Supplement, V1.0 (VI-ERRTAB)

2.4.2 Miscellaneous

IP Version 6 Addressing Architecture, RFC 2373, July 1998 (RFC2373)

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

3 Definitions and abbreviations

3.1 FC-VI definitions

3.1.1 completing a descriptor: a VI Provider completes a Descriptor by updating the status field and setting the Done bit.

3.1.2 FC-VI connection: a VI Connection that is established and maintained between two FC-VI Ports.

3.1.3 FC-VI connectionpoint: the context used to listen for FC-VI Connection requests and responses within an FC-VI Port. It is bound to an IP address and a Discriminator.

3.1.4 FC-VI connection setup: an FC-VI operation that consists of a Sequence of FC-VI Connection IUs that establish an FC-VI Connection.

3.1.5 FC-VI disconnect: an FC-VI operation that consists of a Sequence of FC-VI Connection IUs that removes an FC-VI Connection or aborts a FC-VI Connection Setup.

3.1.6 FC-VI endpoint: the context for a VI within an FC-VI Port. Each end of an FC-VI Connection is an FC-VI Endpoint.

3.1.7 FC-VI message transfer: an FC-VI operation that consists of one or more FC-VI Message IUs to transfer a VI Message between FC-VI Ports.

3.1.8 FC-VI port: a Fibre Channel Port that is capable of FC-VI operation and complies with this specification.

3.1.9 FC-VI provider: the hardware and software services that implement the transport dependent functions of a VI Provider over a Fibre Channel transport conforming to this specification.

3.1.10 fully qualified message ID (FQMID): the tuple of {FCVI_HANDLE, FCVI_MSG_ID, Exchange Context (F_CTL:23)} that uniquely identifies and routes each received FC-VI IU to the correct FC-VI Endpoint context within a FC-VI Port.

3.1.11 host name: a symbolic name associated with a VI capable Node. The Host Name is represented as an ASCII character string to the VI Application.

3.1.12 in-order fabric: a Fibre Channel configuration where the order of frame arrival at a receiving Port is identical to the transmission order at the originating Port. An Arbitrated Loop is one example of an In-Order Fabric.

3.1.13 local: the entity (Endpoint, Connectionpoint, Provider, etc.) at this end of a FC-VI Connection.

3.1.14 out-of-order fabric: a Fibre Channel configuration where the order of frame arrival at a receiving Port may be different than the transmission order at the originating Port.

3.1.15 remote: the entity (Endpoint, Connectionpoint, Provider, etc.) at the other end of a FC-VI Connection.

3.1.16 VI connection: a connection between two VI Endpoints.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

3.1.17 VI endpoint: a pair of work queues and associated context visible to a VI Application. Also known as a VI.

3.1.18 VI message: VI Application data that is transferred between FC-VI Ports over a previously established FC-VI Connection.

3.2 VI Definitions

The following VI Terms used in this specification are defined in the VI Architecture Specification and the VI Architecture Developer's Guide. The user of this document is expected to be familiar with these terms.

Client-Server

Control Segment

Descriptor

Discriminator

Peer

Peer-to-Peer

RDMA

RDMA Read

RDMA Write

Reliability Level

Reliable Delivery

Reliable Reception

Send

Unreliable Delivery

VI

VI Address

VI Application

VI Handle

VI NIC

VI Provider

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

3.3 Abbreviations

D_ID	Destination_Identifier [ISO/IEC 14165-251]
ELS	Extended Link Service [ISO/IEC 14165-251]
FC	Fibre Channel [ISO/IEC 14165-251]
FC-FS	The architecture specified by the Fibre Channel standard [ISO/IEC 14165-251]
FC-PH	The architecture specified by the Fibre Channel standard [ISO/IEC 14165-251]
FCP	Fibre Channel Protocol for SCSI [ISO/IEC 14776-222]
FC-4	Fibre Channel Layer 4 mapping layer [ISO/IEC 14165-251]
IU	Information Unit [ISO/IEC 14165-251]
S_ID	Source_Identifier [ISO/IEC 14165-251]
ULP	Upper Layer Protocol [ISO/IEC 14165-251]

3.4 Editorial conventions

Definitions, conventions, abbreviations, acronyms and symbols applicable to this standard are provided, unless they are identical to that described in any referenced standard, in which case they are included by reference. Some definitions from the glossary or body of other standards are included here for easy reference.

In this Standard, a number of conditions, mechanisms, sequences, parameters, events, states, or similar terms are printed with the following conventions:

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

Such terms and words have special meaning and are defined in other standards. All terms and words not conforming to the convention noted above have the normal technical English meanings.

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

In case of any conflict between text, figure, table and state diagram, the state diagram, then table, then figure, and finally, text takes precedence. Exceptions to this convention are indicated in the appropriate subclauses.

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

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

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

Hexadecimal notation

Hexadecimal notation is used to represent fields. For example, a four-byte FCVI_HANDLE field containing a binary value of 00000000 11111111 10011000 11111010 is shown in hexadecimal format as ‘00FF98FAh’.

--`,,`-`-`,,`,,`,`,,`---

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Binary notation

Binary notation is used to represent fields. For example, a one-byte Match Address Code Point in a FARP-REQ containing a binary value of 0000 0010 is shown in binary format as ‘00000010b’.

Not Equal

The symbol “!=” means “not equal”.

--`,,`,,`-`-`,,`,,`,`,,`---

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

4 Structure and concepts

4.1 Fibre channel structure and concepts

Fibre Channel (FC) is logically a point-to-point serial data channel. FC architecture has been designed so that it may be implemented with high performance hardware that requires little real-time software management. The Fibre Channel Physical (FC-FS) transport described in ISO/IEC 14165-251 performs those functions required to transfer data from one N_Port to another. The FC-PH transport may be treated as a very powerful delivery service with information grouping and multiple defined classes of service.

A switching fabric allows communication among more than two N_Ports.

The Fibre Channel Arbitrated Loop (FC-AL) is an alternate N_Port interconnect architecture that uses FC mechanisms to transfer data between an NL_Port selected by the arbitration process and any of the other NL_Ports on the loop. Once communication is opened between two NL_Ports, standard FC frames are used to provide an FC-PH compliant delivery service. An Arbitrated Loop may be attached to a Fabric by an FL_Port.

An FC-4 mapping layer defines the steps required to perform the functions defined by a ULP. The protocol is defined in terms of the stream of FC IUs generated by a pair of N_Ports that support the FC-4. In this document, N_Ports and NL_Ports capable of supporting FC-VI transactions are collectively referred to as FC-VI Ports.

The number of Exchanges that may simultaneously be open between two FC-VI Ports is defined by the FC-PH implementation. The architectural limit for this value is 65,535. The maximum number of active Sequences that may simultaneously be open between two FC-VI Ports is limited by FC-PH to no more than 255. To allow management Exchanges to be originated, a certain number of extra Exchanges and at least one extra SEQ_ID should always be available.

FC-FS allows management protocols above the FC-PH interface to perform link management functions. The standard FC-PH primitive sequences, link management protocols, and basic and extended link services are used as required by FC-VI Endpoints.

4.2 FC-VI structure and concepts

FC-VI defines the mapping of VI Messages and VI Connections to FC-PH. FC-VI is based on the method of first establishing an FC-VI Connection between two FC-VI Ports, and then sending VI Messages over the FC-VI Connection. FC-VI defines the Information Units (IUs) that are required to establish and remove FC-VI Connections, and the IUs that are required to send VI Messages.

Each VI Message is transferred in one Exchange, and each FC-VI Connection is established in one Exchange (Client-Server) or two Exchanges (Peer-to-Peer). As required by FC-FS, each FC-VI Information Unit (IU) is constructed from one Sequence. One or more IUs are grouped together to form one Exchange. FC-VI defines FC-VI Connection IUs to establish and remove FC-VI Connections, and defines FC-VI Message IUs to send VI Messages.

FC-VI defines a FCVI_HANDLE to specify the destination VI for the VI Message. FC-VI Ports shall send FC-VI Connection IUs between them to establish the FCVI_HANDLEs. An FC-VI Port may choose its own FCVI_HANDLE for each of its FC-VI Endpoints. The FC-VI Port is required to use the FCVI_HANDLE specified by the other FC-VI Port when transmitting a VI Message. The FC-VI Provider may use any appropriate method to map a FCVI_HANDLE to a VI Handle as used by a VI Application, since FC-VI does not define a specific mapping between FCVI_HANDLEs and VI Handles. An FCVI_HANDLE shall be unique for all open FC-VI Connections for an FC-VI Port.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

FC-VI defines a FC-VI Message ID (FCVI_MSG_ID). The FCVI_MSG_ID is set to one by an FC-VI Port when it sends the first IU of its first VI Message on a new FC-VI Connection. The FC-VI Port that receives the VI Message shall echo the same FCVI_MSG_ID in any Response IU. Each FC-VI Port shall increment the FCVI_MSG_ID value by one for every VI Message sent within a VI to facilitate the detection of lost VI Messages at the receiver. A separate FCVI_MSG_ID count is kept for each direction of a VI. See Clause 8 for a complete description of FC-VI error detection and recovery.

VI Reliability Levels are defined with respect to VI Message delivery. FC-VI supports all three levels of reliability (Unreliable Delivery, Reliable Delivery, Reliable Reception) and all connection models (Client-Server, Peer-to-Peer) defined by the VI Architecture. FC-VI supports all VI data transfer models (Send, RDMA Write and RDMA Read) defined by the VI Architecture.

FC-VI supports unacknowledged (class 2) and acknowledged (class 3) classes of service. FC-VI supports Arbitrated Loop, Fabric, Loop attached Fabric, and Point-to-Point topologies. FC-VI supports an In-Order Fabric and an Out-of-Order Fabric.

Note 1: FC-VI may support additional classes of service other than class 2 and class 3. However, no consideration was given to other classes of service during the creation of this specification.

The VI Architecture does not define the format of how VI Addresses are mapped to a transport specific addresses. The FC-VI mapping of VI to Fibre Channel associates an IP address with a FC-VI Port. The method of assigning IP addresses to FC-VI Ports is outside the scope of this standard. FC-VI allows more than one IP address to be assigned to the same FC-VI Port. FC-VI also allows the same IP address to be assigned to more than one FC-VI Port. The format of an IP address for FC-VI shall use the format specified by RFC 2373.

FC-VI provides methods to map an N_Port Identifier to an IP address assigned to a FC-VI Port. The Fibre Channel Name Server of the Directory Service as defined in FC-GS-2 is the preferred method to map an N_Port Identifier to an IP address for Fabric topologies. FC-VI additionally specifies the Fibre Channel FARP ELS to resolve IP addresses to N_Port Identifiers for configurations that do not support a Name Server.

Figure 1 illustrates the addressing objects associated with an FC-VI Port.

Before VI Messages are sent, an FC-VI Connection shall be established between a pair of FC-VI Endpoints. An FC-VI Endpoint is an FC-4/ULP construct for a VI Endpoint. A VI Endpoint is created when a VI Application executes the VipCreateVI call (see VI-ARCH).

An FC-VI Port may need to discover the N_Port Identifier of the FC-VI Port to which it wishes to send a connection request to. This is accomplished by querying the Name Server or issuing a FARP-REQ. The process of establishing an FC-VI Connection is defined as an FC-VI Connection Setup. FC-VI defines a four phase sequencing of IUs within a single Exchange to establish a connection. The minimum number of phases to support the VI Architecture is three. FC-VI adds an additional phase to terminate the Connection Setup Exchange at the Exchange originator to facilitate error recovery.

An FC-VI Provider listens for incoming connection requests on an FC-VI Connectionpoint on behalf of a VI Application. An FC-VI Provider creates an FC-VI Connectionpoint when a VI Application executes a VipConnectWait for a Client-Server connection or VipConnectPeerRequest for a Peer-to-Peer connection (see VI-DG).

The FC-VI Provider listens for incoming connection requests on a specific IP address and a Discriminator bound to the FC-VI Connectionpoint. A Discriminator is typically unique per FC-VI Connectionpoint, but it is not required to be so. When the connection successfully completes, the connection originator FC-VI Endpoint is bound to the connection responder FC-VI Endpoint.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

An incoming Connect Request, FCVI_CONNECT_RQST, contains both the Local and Remote IP addresses and Discriminators. A connection match exists if the IP address and Discriminator in the FCVI_CONNECT_RQST match the IP address and Discriminator in the FC-VI Connectionpoint of the Connect Request responder. See Section 6.7 in the VI Architecture Developer's Guide for the exact matching rules for Connection Setup.

Each FC-VI Endpoint communicates a FCVI_HANDLE to the other FC-VI Endpoint during FC-VI Connection Setup. Once the FC-VI Connection is established, the FCVI_HANDLE is used to route all VI Messages to the proper FC-VI Endpoint within an FC-VI Port. IP addresses and Discriminators are not used for VI Message transfer. During an FC-VI Connection Setup, the FC-VI Provider shall guarantee that the Local FC-VI Endpoint is bound to the connection and enabled for reception of Messages before any Connect Response is issued.

The VI Architecture defines two methods to establish a connection. A Client-Server Connection Setup is similar to the Sockets model - a Server waits for connection requests from a Client. A Peer-to-Peer Connection Setup involves two Peers attempting to establish a connection with each other at about the same time.

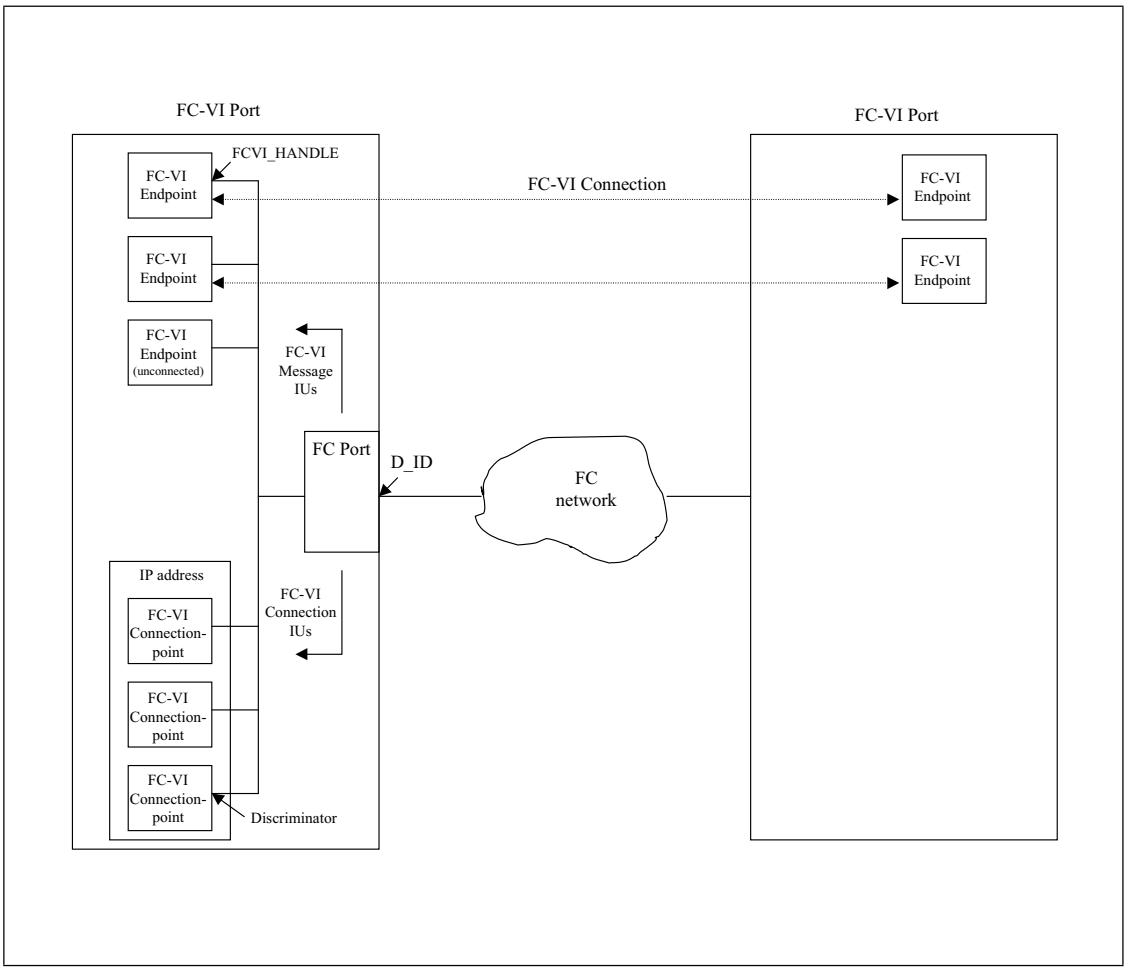


Figure 1 – FC-VI addressing objects

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

5 FC-VI protocol overview

5.1 FC-VI information units

The Information Units (IUs) used by FC-VI Ports and their characteristics are shown in Table 1. All FC-VI IUs shall be composed of Device_Data frames. All Connect Request/Response IUs are either Unsolicited Control or Solicited Control. This facilitates the FC-2 to route Message Request/Response IUs directly to specialized FC-VI data transfer hardware assists and route the Connect Request/Response IUs to a CPU software path.

The Table 1 column 'Mand/Opt' defines which IUs are Mandatory and which are Optional. The F/M/L column indicates whether the IU is the First, Middle or Last of the Exchange. FCVI_READ_RQST is optional since RDMA Read is optional. All Message Responses are optional, since they are only required for RDMA Read or Reliable Reception. All connection IUs are mandatory.

Table 1 – FC-VI Information unit summary

IU Category	FC-VI Information Unit Type	R_CTL	Device_Header Size	FC-VI Opcode	Payload Content	F/M/L	Mand / Opt
Message Request	FCVI_SEND_RQST	01h	32 Bytes	00h	Message Data	F/M/L	M
Message Request	FCVI_WRITE_RQST	01h	32 Bytes	01h	Message Data	F/M/L	M
Message Request	FCVI_READ_RQST	06h	32 Bytes	02h	none	F	O
Message Response	FCVI_SEND_RESP	07h	16 Bytes	08h	none	L	O
Message Response	FCVI_WRITE_RESP	07h	16 Bytes	09h	none	L	O
Message Response	FCVI_READ_RESP	01h	32 Bytes	0Ah	Message Data	M/L	O
Connect Request	FCVI_CONNECT_RQST	02h	32 Bytes	10h	Connect Info	F	M
Connect Request	FCVI_DISCONNECT_RQST	02h	32 Bytes	12h	none	F	M
Connect Response	FCVI_CONNECT_RESP1	03h	32 Bytes	18h	Connect Info	M	M
Connect Response	FCVI_CONNECT_RESP2	03h	32 Bytes	19h	none	M	M
Connect Response	FCVI_CONNECT_RESP3	03h	32 Bytes	1Ah	none	L	M
Connect Response	FCVI_DISCONNECT_RESP	03h	32 Bytes	1Bh	none	L	M

5.2 FC-VI message transfer operation

5.2.1 FC-VI message transfer

An FC-VI Message Transfer consists of a Message Request and Message Response, or a Message Request with no Message Response. A Message Request shall consist of one or more Message Request IUs of the same IU type. A Message Response shall consist of one or more Message Response IUs of the same IU type. If a response is required, the last request IU shall transfer Sequence Initiative. If no response is expected, the last request IU shall terminate the Exchange. If a response is required, the last response IU shall terminate the Exchange. Exchanges are dynamically created by the Fibre Channel transport mechanism to move the VI Message from one FC-VI Port to another. Utilizing the same Exchange for Message Request IUs and Message Response IUs (if needed) binds all Sequences and frames associated with one VI Message.

Note 2: The FC-VI mapping of FC-VI Connections to FC-FS Exchanges is dynamic to allow the number of FC-VI Connections to be independent from the number of Exchanges. The maximum number of FC-VI Connections is 2^{32}-1 per port, limited by the FCVI_HANDLE field in the FC-VI Device_Header.

A VI Message may be transferred with a FC-VI Send Message Transfer, a FC-VI Write Message Transfer or a FC-VI Read Message Transfer operation.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

The use of Continuously Increasing SEQ_CNT, Continuously Increasing Relative Offset, and total Message length facilitate the detection of lost frames at the Message responder within an Exchange. If the value of SEQ_CNT repeats within an Exchange, Continuously Increasing Relative Offset may be used to guarantee frame uniqueness within an Exchange and facilitate the detection of frame synonyms (identical SEQ_ID and SEQ_CNT values for frames within an Exchange). The value of Relative Offset is guaranteed to never repeat within an Exchange. See 5.5.

The following subclauses detail the different cases for VI Messages mapping into one FC Exchange. These subclauses only define error free operation for FC-VI operation. See Clause 8 for the definition of FC-VI error detection and recovery.

5.2.2 FC-VI send message transfer operation

A VI Message may be transferred using one or more FCVI_SEND_RQST IUs. The VI Message may be sent in Unreliable Delivery or Reliable Delivery mode, as shown in Figure 2, or Reliable Reception mode, as shown in Figure 3. The FC-VI Provider may send one or multiple FCVI_SEND_RQST IUs to transmit the VI Message. No Response shall be issued for Unreliable Delivery or Reliable Delivery modes. A response consisting of one FCVI_SEND_RESP IU is required for Reliable Reception mode. Each figure below represents one FC Exchange, and each arrow represents one FC Sequence or IU.

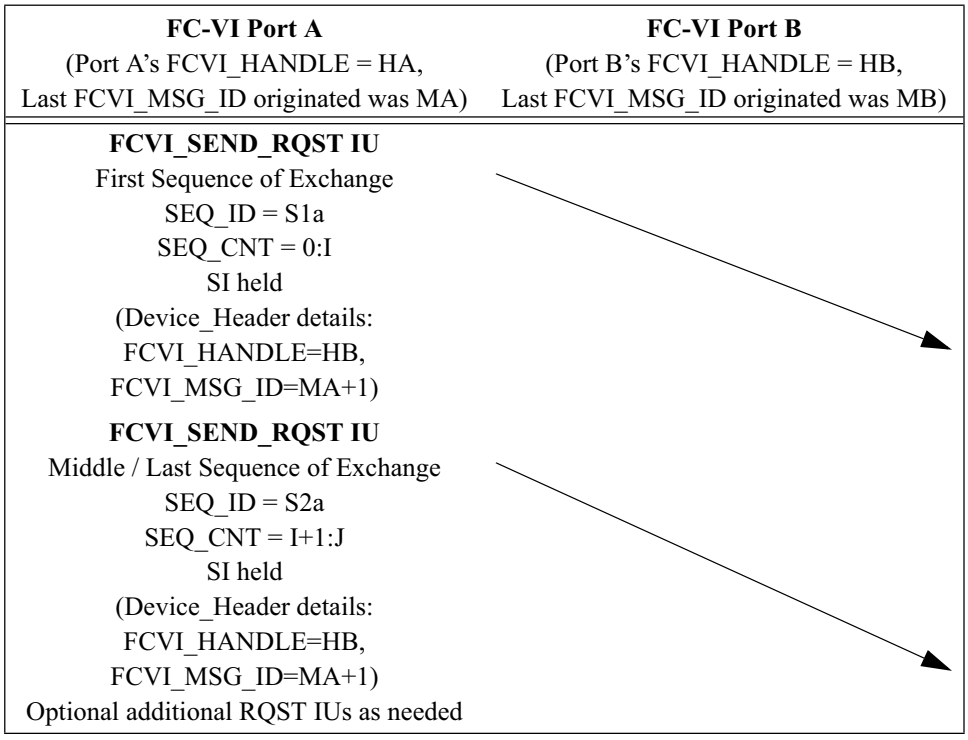


Figure 2 – FC-VI send for Unreliable Delivery or Reliable Delivery

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

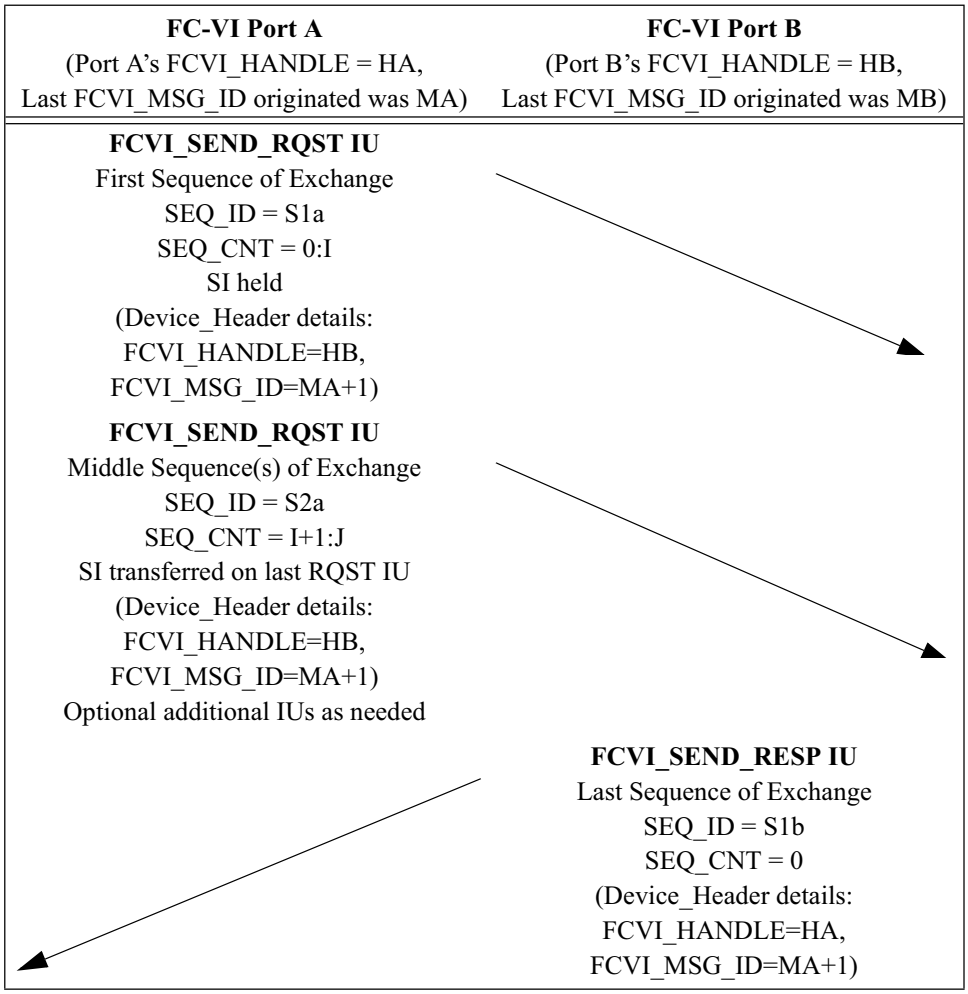


Figure 3 – FC-VI send for Reliable Reception

5.2.3 FC-VI RDMA write message transfer operation

A VI Message may be transferred using one or more FCVI_WRITE_RQST IUs. The VI Message may be sent in Unreliable Delivery or Reliable Delivery mode, as shown in Figure 4, or sent in Reliable Reception mode, as shown in Figure 5. The FC-VI Provider may send one or multiple FCVI_WRITE_RQST IUs to deliver the VI Message. No response shall be issued for Unreliable Delivery or Reliable Delivery modes. A response consisting of one FCVI_WRITE_RESP IU is required

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

for Reliable Reception mode. Each figure below represents one FC Exchange, and each arrow represents one FC Sequence.

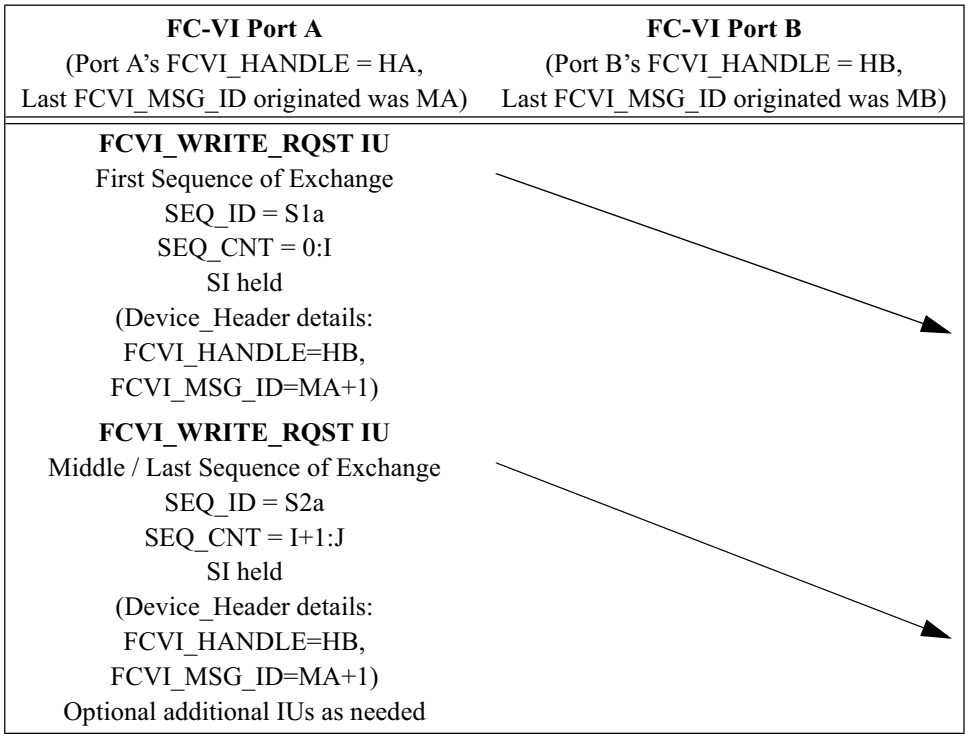


Figure 4 – FC-VI RDMA write for Unreliable Delivery or Reliable Delivery

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

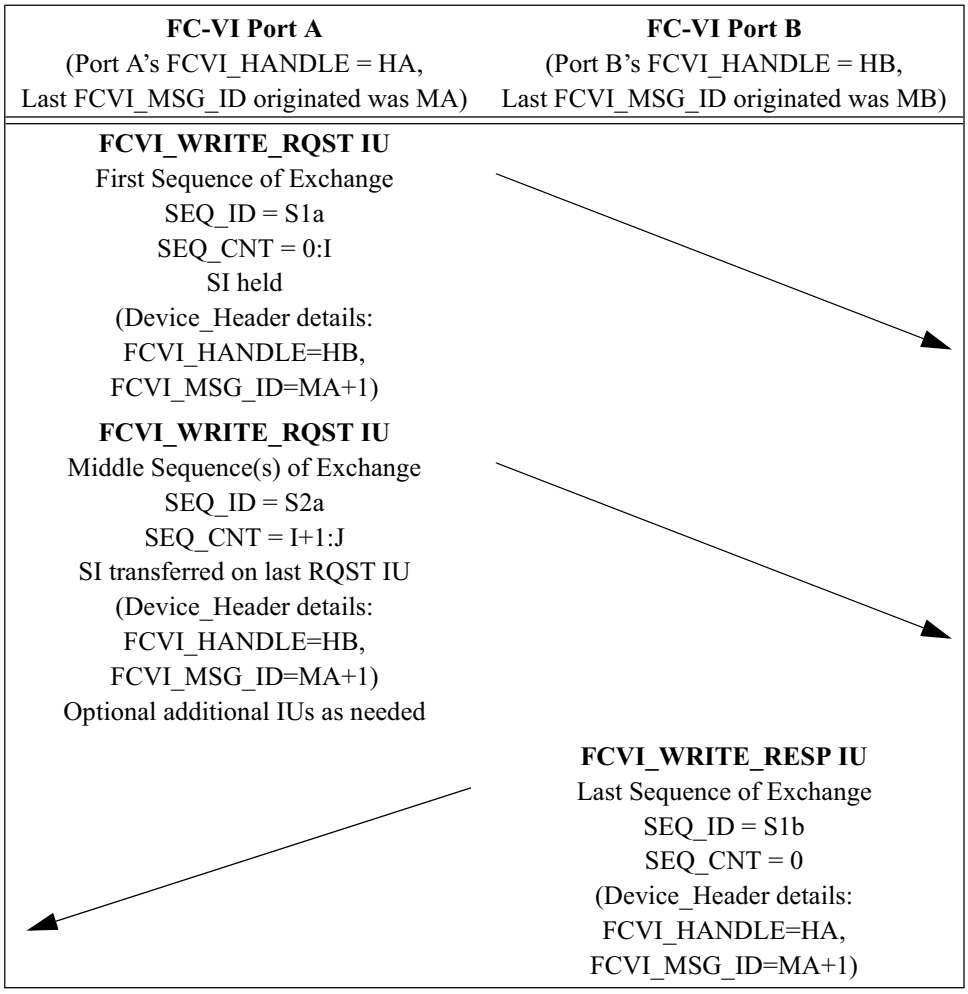


Figure 5 – FC-VI RDMA write for Reliable Reception

5.2.4 FC-VI RDMA read message transfer operation

A VI Message may be transferred with one FCVI_READ_RQST IU and one or more FCVI_READ_RESP IUs. The VI Message my be transferred in Reliable Delivery or Reliable Reception mode, as shown in Figure 6. The FC-VI Provider may send one or multiple FCVI_READ_RESP IUs. A response is always required. Each figure below represents one FC Exchange, and each arrow represents one FC Sequence.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

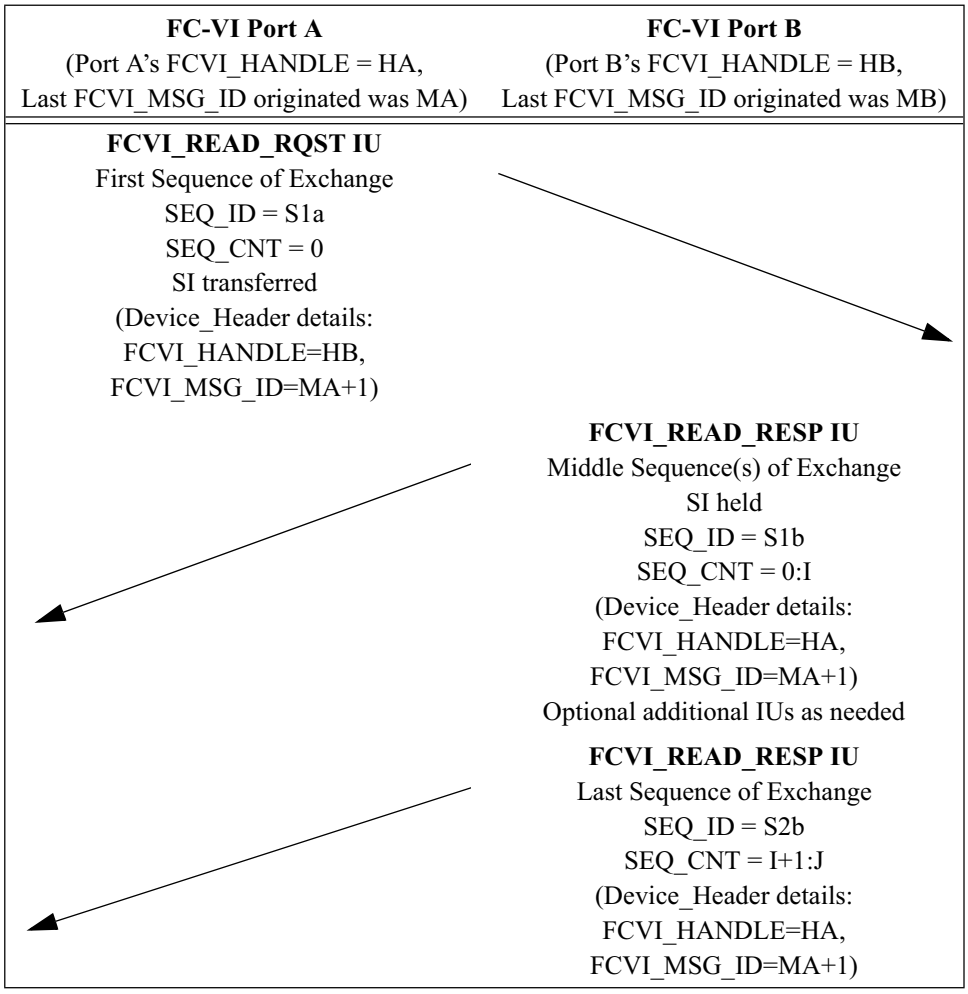


Figure 6 – FC-VI RDMA read for Reliable Reception and Reliable Delivery

5.2.5 IU reception at an FC-VI edpoint

A FC-VI Provider that supports RDMA Reads or Reliable Reception shall be capable of managing two potentially concurrent receive streams for each Local FC-VI Endpoint - a requestor stream and a responder stream. A requestor stream consists of solicited RDMA Read responses and solicited Reliable Reception responses for Messages originated by the Local FC-VI Endpoint. A responder stream consists of unsolicited Sends, RDMA Read requests, and RDMA Writes originated by the Remote FC-VI Endpoint.

An example of concurrent requestor and responder receive streams is shown in Figure 7. The Local Endpoint at FC-VI Port X has originated a RDMA Read Message (Message MA) and is receiving the RDMA Read response data from FC-VI Port Y. At the same time, the Remote Endpoint in FC-VI Port Y has originated a RDMA Write Message (Message MB) to the Local Endpoint in FC-VI Port X. Since the FC-VI Providers in FC-VI Ports X and Y assign FCVI_HANDLEs independently, it is possible that MA may be equal to MB.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

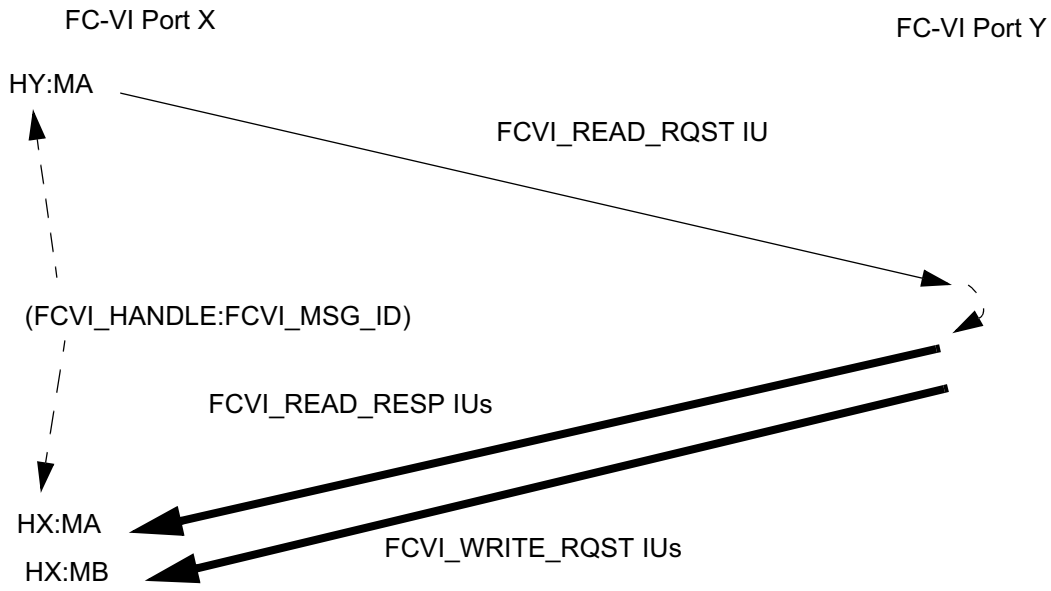


Figure 7 – Concurrent Receive Streams at a FC-VI Endpoint

Thus, the Local Endpoint in FC-VI Port X needs another mechanism other than the tuple {FCVI_HANDLE:FCVI_MSG_ID} to demultiplex the two receive streams to the correct FC-VI Endpoint context. The FC-VI Provider may use the Fully Qualified Message ID (FQMID), which is defined by the tuple {FCVI_HANDLE, FCVI_MSG_ID, Exchange Context (F_CTL:23)}, to route received IUs to the correct FC-VI Endpoint context. The Exchange Context allows the FC-VI Provider to differentiate between solicited responses and unsolicited requests targeted for the same Endpoint. An alternative is to use the Exchange ID and not the FCVI_HANDLE to demultiplex multiple requestor streams (RDMA Read responses, Read responses and Write responses) to the correct FC-VI Endpoint at the Exchange originator.

5.3 FC-VI connection setup operation

5.3.1 FC-VI client-server and peer-peer connection setup

For a Peer-to-Peer Connection Setup, each Peer is a connection originator and a connection responder. For Client-Server, only the Client is the connection originator, while the Server is the connection responder.

Each Connect Request and Connect Response is a separate FC-VI Connection IU. Sequence Initiative is always passed for each connection IU. The FCVI_CONNECT_RQST is the first Sequence of the Exchange, the FCVI_CONNECT_RESP1 and FCVI_CONNECT_RESP2 IUs are the middle Sequences of the Exchange, and FCVI_CONNECT_RESP3 IU is the last Sequence of the Exchange.

This subclause and the following subclauses detail all of the different cases for FC-VI Connections being established or removed. These subclauses only define error free operation for FC-VI operation. See Clause 8 for the definition of FC-VI error detection and recovery.

--`,`-`-`,,`,`,`,,`---
Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

5.3.2 FC-VI client-server connection setup

Figure 8 illustrates FC-VI Connection Setup for Client-Server. In Figure 8 the Client is Port A while the Server is Port B. The FCVI_FLAGS = 0 in the FCVI_CONNECT_RESP1 IU indicates that the Server has accepted the Client's connection request. The FCVI_FLAGS = 0 in the FCVI_CONNECT_RESP2 indicates that the Client acknowledges the Server's connect accept. The

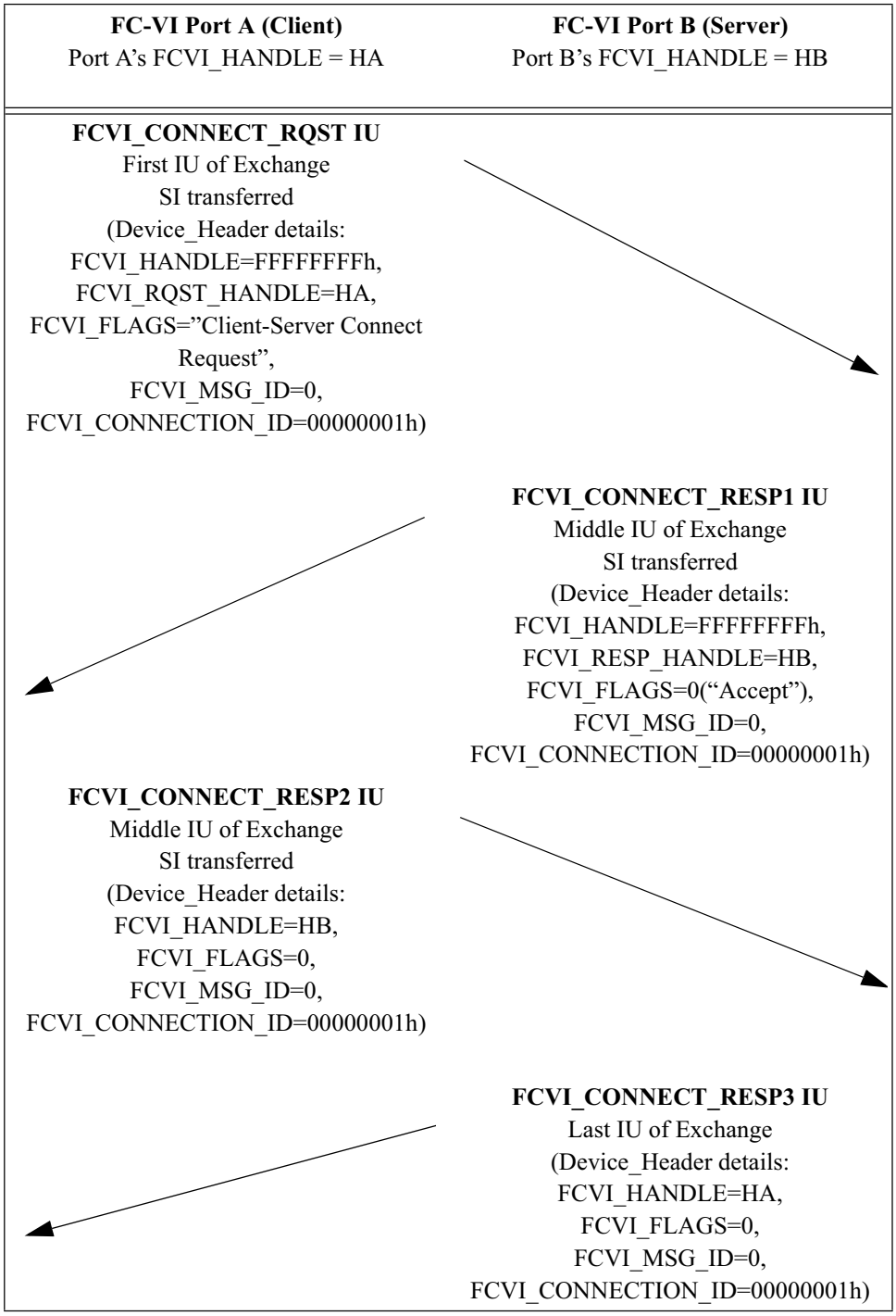


Figure 8 – FC-VI client-server connection setup

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

FCVI_FLAGS = 0 in the FCVI_CONNECT_RESP3 indicates that the Server acknowledges the Client's acknowledgment.

The Connect Request originator shall set the FCVI_HANDLE field in the Device_Header of the FCVI_CONNECT_RQST IU equal to the FFFFFFFFh, which indicates the responder handle is not yet assigned. The Connect Request originator shall set the FCVI_RQST_HANDLE field in the Device_Header of the FCVI_CONNECT_RQST IU equal to the FCVI_HANDLE of the Local FC-VI Endpoint. The Connect Request responder shall set the FCVI_HANDLE in the FCVI_CONNECT_RESP3 IU equal to this value received in the FCVI_RQST_HANDLE field in the Device_Header of the FCVI_CONNECT_RQST IU.

The Connect Request responder shall set the FCVI_HANDLE field in the Device_Header of the FCVI_CONNECT_RESP1 IU equal to the FFFFFFFFh, which indicates the requestor handle is not yet assigned. The Connect Request responder shall set the FCVI_RESP_HANDLE field in the Device_Header of the FCVI_CONNECT_RESP1 IU equal to the FCVI_HANDLE of the Local FC-VI Endpoint. The Connect Request originator shall set the FCVI_HANDLE in the FCVI_CONNECT_RESP2 IU equal to this value received in the FCVI_RESP_HANDLE field in the Device_Header of the FCVI_CONNECT_RESP1 IU.

The FCVI_CONNECTION_ID shall be used by the Connect Request originator to assign a unique identifier to all open Connection Setup operations. The FCVI_MSG_ID shall be set to zero for all Connection Setup IUs.

5.3.3 FC-VI Peer-to-Peer Connection Establishment

Figure 9 illustrates the FC-VI Connection Setup for Peer-to-Peer when one Peer issues a Connect Request before the other Peer has begun the Connection Setup. In Figure 9, Peer 1 is Port A, while Peer 2 is Port B. The VI Application in Peer 2 has not yet issued a Peer Connect Request when the FCVI_CONNECT_RQST IU is received from Peer 1. Peer 2 replies with a CONN_STS set to one and a Reason Code of "No Waiting Remote Connectionpoint". At some later point Peer 2 issues a matching FCVI_CONNECT_RQST IU and a FC-VI Connection is established between Port A and Port B.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

.

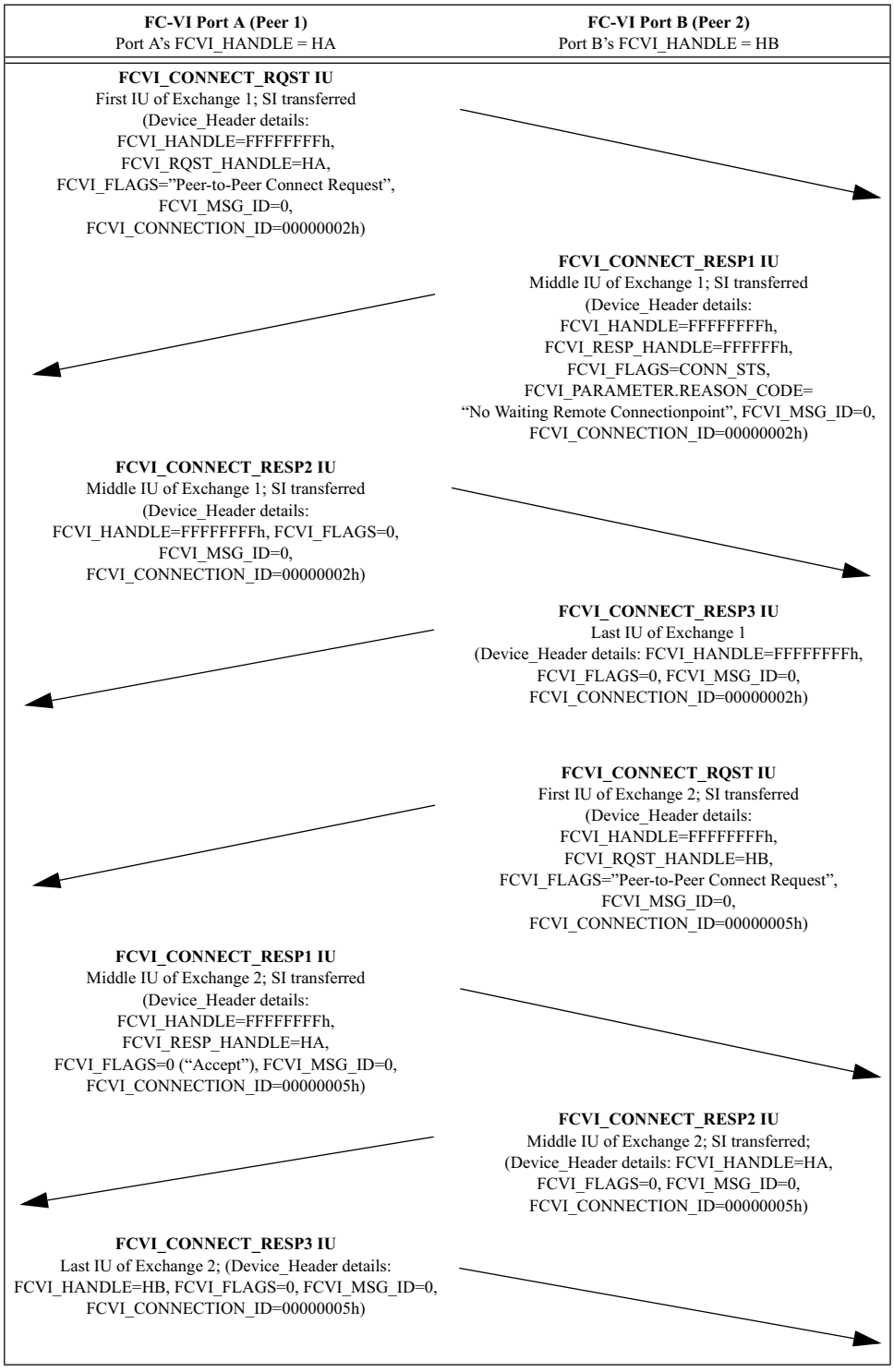


Figure 9 – Peer-to-peer connection setup

For Exchange 1, the Connect Request responder shall set the FCVI_RESP_HANDLE field in the Device_Header of the FCVI_CONNECT_RESP1 IU (FCVI_CONNECTION_ID = 00000002h) equal to FFFFFFFFh, since there is no waiting Peer at the Connect Request responder. The Connect Re-

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

quest originator shall set the FCVI_HANDLE in the Device_Header in the FCVI_CONNECT_RESP2 IU in Exchange 1 equal to FFFFFFFFh, the value of FCVI_RESP_HANDLE. The Connect Request responder shall set the FCVI_HANDLE in the Device_Header in the FCVI_CONNECT_RESP3 IU in Exchange 1 equal to FFFFFFFFh. The unassigned FCVI_HANDLES in FCVI_CONNECT_RESP2 and FCVI_CONNECT_RESP3 indicate the Connection Setup did not establish a connection.

For Exchange 2, the Connect Request responder shall set the FCVI_RESP_HANDLE field in the Device_Header of the FCVI_CONNECT_RESP1 IU (FCVI_CONNECTION_ID = 00000005h) equal to FFFFFFFFh. The Connect Request originator shall set the FCVI_HANDLE in the Device_Header in the FCVI_CONNECT_RESP2 IU in Exchange 2 equal to “HA”h, the value of FCVI_RESP_HANDLE, since there now is a waiting Peer at the Connect Request responder. The Connect Request responder shall set the FCVI_HANDLE in the Device_Header in the FCVI_CONNECT_RESP3 IU in Exchange 2 equal to “HB”h, the value of FCVI_RQST_HANDLE.

5.3.4 FC_VI concurrent peer-to-peer connection setup

Figure 10 illustrates FC-VI Connection Setup for Peer-to-Peer when each Peer issues a matching FCVI_CONNECT_RQST IU before a FCVI_CONNECT_RESP1 IU is received. This situation is defined as Concurrent Matching Peer-to-Peer Requests.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

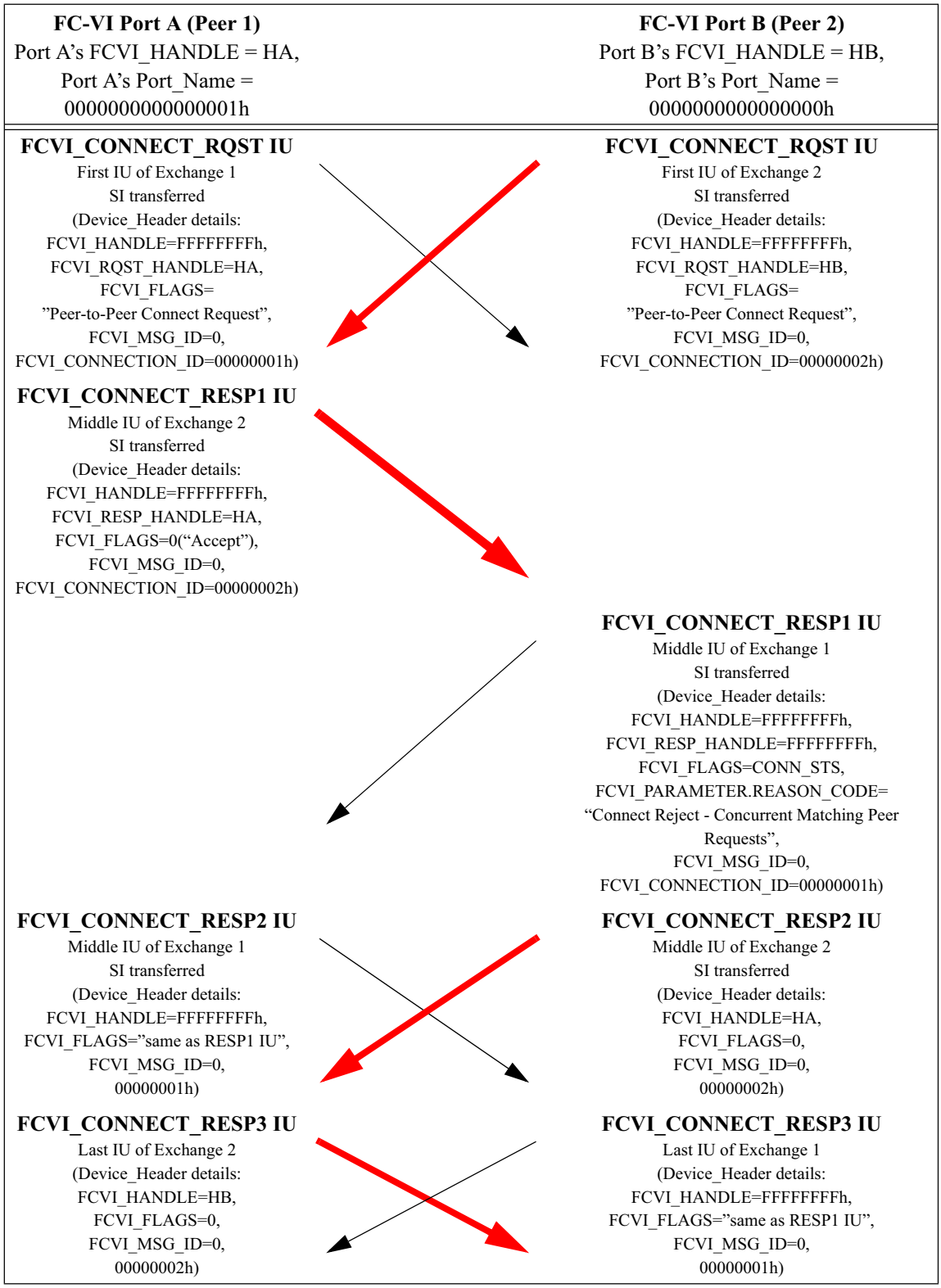


Figure 10 – Peer-to-peer connection setup, concurrent matching peer requests

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

For Concurrent Matching Peer-to-Peer Requests, an arbitration protocol shall be used to guarantee that only one Peer accepts a matching Connect Request and successfully completes the Connection Setup. A comparison between the Port_Names of the connection originators is used for the arbitration protocol. The Peer with the numerically lower Port_Name shall win the arbitration and rejects the Peer's Connect Request. The Peer with the numerically higher Port_Name shall lose the arbitration and accept the Peer's Connect Request. Thus, the Peer with the numerically lower Port_Name successfully completes the Connection Setup. The following rules define the arbitration protocol for Concurrent Matching Peer-to-Peer Requests:

a) If my Port_Name is greater than my Peer's Port_Name, send a FCVI_CONNECT_RESP1 of "Accept" (CONN_STS = 0 in FCVI_FLAGS).

b) Else, if my Port_Name is less than the Peer's Port_Name, wait for the Peer's FCVI_CONNECT_RESP1.

   1) If the Peer's FCVI_CONNECT_RESP1 is "No Waiting Remote Connectionpoint" or "No Discriminator Match", send a FCVI_CONNECT_RESP1 of "Connect Accept".

   2) Else, if the Peer's FCVI_CONNECT_RESP1 is "Connect Accept", send a FCVI_CONNECT_RESP1 with CONN_STS equal to one with a Reason Code of "Connect Reject - Concurrent Matching Peer Requests".

   3) Else, if the Peer's FCVI_CONNECT_RESP1 is "Connect Reject - Concurrent Matching Peer Requests", send a FCVI_CONNECT_RESP1 with CONN_STS equal to one with a Reason Code of "Connect Reject - Protocol Error".

   4) Else, send a FCVI_CONNECT_RESP1 with CONN_STS equal to one with a Reason Code of "Connect Reject".

c) Else, if my Port_Name equal to Peer's Port_Name, send a FCVI_CONNECT_RESP1 with CONN_STS equal to one with a Reason Code of "Connect Reject - Protocol Error".

To support loopback of Peer-to-Peer Connection Setup, a FC-VI Provider shall recognize that the Connect Request originator and the Connect Request responder are the same Port and simulate the Port_Name comparison with an implementation dependent mechanism (i.e., the ability to arbitrate between concurrent matching requests).

Table 2 defines the correct FCVI_CONNECT_RESP1 from the Peer (Peer B) with the lower Port_Name for Concurrent Matching Peer-to-Peer Requests for each possible FCVI_CONNECT_RESP1 from the Remote Peer with the higher Port_Name (Peer A). See Annex A for examples of Concurrent Matching Peer-to-Peer Requests.

Figure 2 is an example of Concurrent Matching Peer-to-Peer Requests. Peer 1 is Port A, while Peer 2 is Port B. The thicker arrows correspond to IUs for Exchange 2, while the thinner arrows correspond to Exchange 1. In the case illustrated, Peer 1's Port_Name is numerically greater than Peer 2's Port_Name. Peer 1 accepts the connection request while Peer 2 rejects it. Thus, Peer 2 shall successfully complete the connection.

5.3.5 FC-VI Disconnect Operation

Figure 11 shows the FC-VI protocol to perform a FC-VI Disconnect. A VI Application decides when it is appropriate to disconnect, according to the VI Architecture specifications. A FC-VI Disconnect operation may be initiated from either FC-VI Port by sending a FCVI_DISCONNECT_RQST IU. The

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Table 2 – Peer B actions based on connect responses from peer A

Response from Peer A (higher Port_Name)	Peer B (lower Port_Name) Replies With:
Connect Accept	Connect Reject - Concurrent Matching Peer Requests
No Discriminator Match	Connect Accept
No Waiting Remote Connectionpoint	Connect Accept
Connect Reject - Concurrent Matching Peer Requests	Connect Reject - Protocol Error
Connect Reject - Transport Error	Connect Reject
Connect Reject - Protocol Error	Connect Reject
Any other response	Connect Reject

other FC-VI Port shall respond with a FCVI_DISCONNECT_RESP IU. The FC-VI Provider that controls the FC-VI Port may also issue a Disconnect under certain cases.

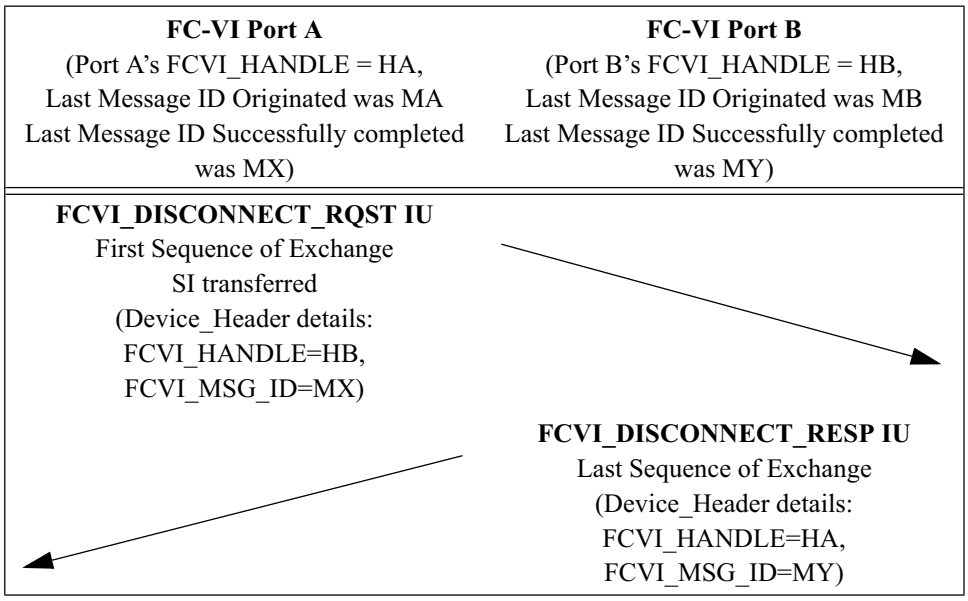


Figure 11 – FC-VI disconnect operation

5.4 Exchange ID reuse

FC-VI requires Exchange termination on the last IU, either request or response, of the Message. Each Message transfer starts a new Exchange. Since FCVI_MSG_ID is guaranteed to be unique for each Message, a Message Originator may immediately reuse an Exchange ID of a previously terminated Exchange.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

5.5 Sequence ID reuse

As required by FC-FS, all FC-VI Information Units shall be represented by one FC Sequence. The Sequence ID (SEQ_ID) may be reused by the Sequence Initiator as soon as the Sequence is closed.[2] In Class 3, the Sequence Initiator shall consider a Sequence closed when the last frame of the Sequence is sent. In Class 2, the Sequence Initiator shall consider a Sequence closed when all ACKs for all frames in the Sequence have been received. For either class 2 or class 3 a Sequence ID may be reused when the Sequence ID has been recovered after a Sequence error.

5.6 Frame synonym detection

FC-VI frames shall use a Continuously Increasing Sequence Count for each Exchange to assist detecting potentially missing frames. Sequence Count shall not repeat within one IU. An FC-VI Port shall set Word 1, bit 17 - SEQ_CNT (S) - to one in the N_Port Common Service Parameters when originating a PLOGI. An FC-VI Port shall set Word 1, bit17 in the ACC to a PLOGI.

If the SEQ_CNT (S) bit is not set in the PLOGI or ACC, an FC-VI Port shall reject any Connect Request from the Port that sent the PLOGI or ACC with the (S) bit not set by transmitting a FCVI_CONNECT_RESP1 with CONN_STS set to one and a Reason Code of “Invalid Service Parm”. In addition, an FC_VI Port shall not originate a FC-VI Connection Setup with a Port that sent a PLOGI or ACC with the (S) bit not set.

To facilitate frame synonym detection at the Sequence Recipient, FC-VI requires that the Sequence Initiator shall use Continuously Increasing Relative Offset. All FC-VI compliant N_Ports shall indicate support for Continuously Increasing Relative Offset by setting the Continuously Increasing (C) bit to one in the Common Service Parameters of both the PLOGI and ACC (Word 1, bit 30). The Relative Offset of the first byte of the VI Message payload in the first frame of the next Sequence to transmit shall be +1 greater than the maximum relative offset of any byte of the VI Message payload in the previous Sequence.

If the Continuously Increasing (C) bit is not set the PLOGI or ACC, an FC-VI Port shall reject any Connect Request from the Port that sent the PLOGI or ACC with the (C) bit not set by transmitting a FCVI_CONNECT_RESP1 with CONN_STS set to one and a Reason Code of “Invalid Service Parm”. In addition, an FC_VI Port shall not originate a FC-VI Connection Setup with a Port that sent a PLOGI or ACC with the (C) bit not set.

2. Note that PH Clause 18.6 requires the use of at least two different SEQ_IDs for Sequence steaming.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Figure 12 shows an example of FC-2 header operation during an FC-VI Exchange. Figure 12 assumes a uniform frame size.

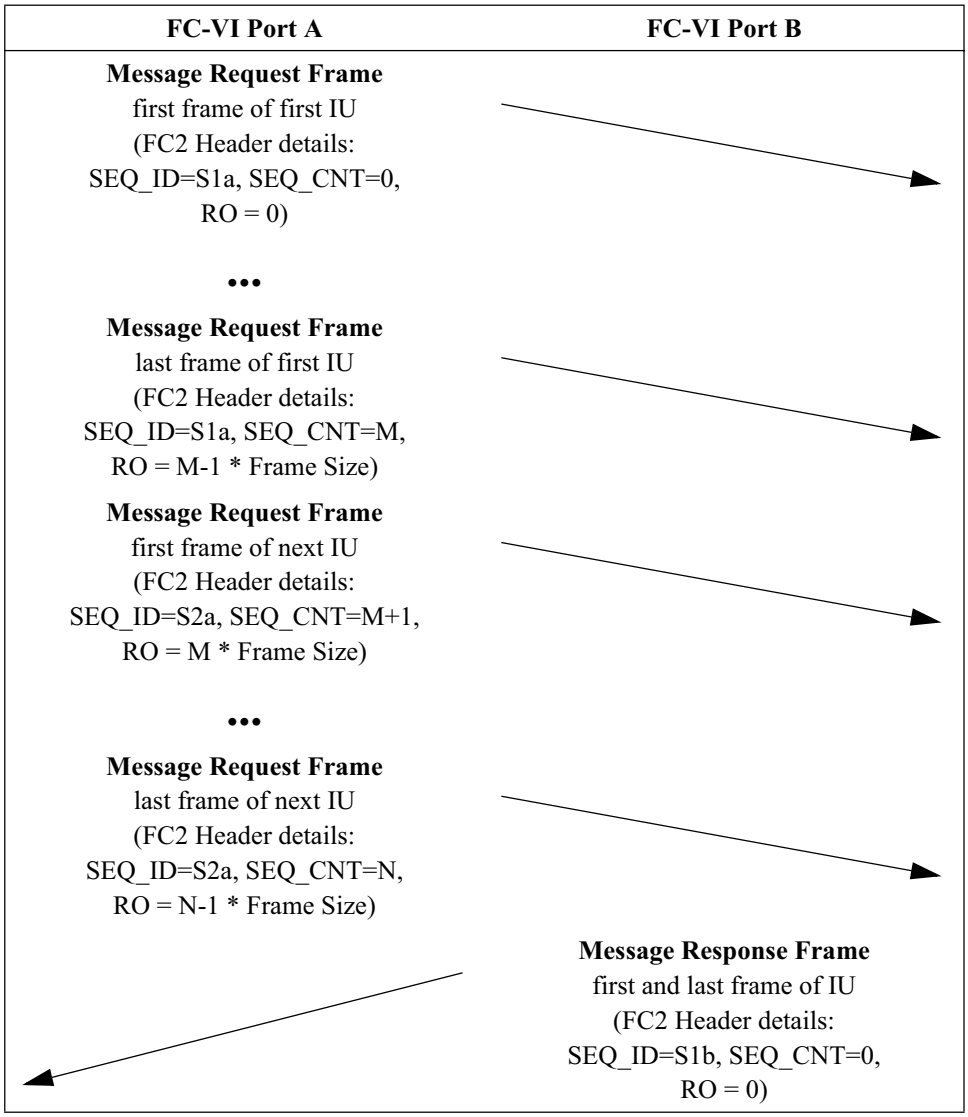


Figure 12 – FC-FS header for send operation

5.7 VI message length

VI Messages may be a maximum length of 2^{32} -1 bytes, which is the maximum VI Message size specified by the VI Architecture.

5.8 FC-FS header usage for FC-VI

5.8.1 FC-FS header usage

For FC-VI, fields in the frame header use the standard FC-FS definitions. The following explanations of the FC header fields provide information about the use of those fields to implement FC-VI functionality. If a field is not defined below, then its usage is the same as defined by FC-FS.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

5.8.2 CS_CTL field

FC-VI Providers that choose to make use of the PREF bit defined in the CS_CTL field of the frame-header (see FC-FS) shall conform to the following policy: Use of the Preference Function shall be on a per FC-VI Connection basis. If a particular FC-VI Connection is to make use of the Preference Function then every Fibre Channel frame for a VI Message sent on that FC-VI Connection shall have the PREF bit set to 1 in the CS_CTL field. If a particular FC-VI Connection is not making use of the Preference Function then each and every Fibre Channel frame for a VI Message sent on that FC-VI Connection shall have the PREF bit set to 0 in the CS_CTL field. Each FC-VI Endpoint of an FC-VI Connection may independently make use of the Preference Function. It is not required that both FC-VI Endpoints simultaneously set the PREF bit to one. The Preference Function shall only be used for FC-VI Message Transfer operations.

5.8.3 TYPE field

The TYPE field shall be 58h for all frames of any FC-VI IU.

5.8.4 F_CTL field

For Message Requests that do not require a Message Response, the Last Sequence bit shall be set to one in the last Message Request IU. For Message Requests that do require a Message Response, the Sequence Initiative bit shall be set to one in the last Message Request IU. The Last Sequence bit shall be set to one in the last Message Response IU.

For FC-VI Connection Setup IUs, the Sequence Initiative bit shall be set to one in a FCVI_CONNECT_RQST IU, a FCVI_CONNECT_RESP1 IU and a FCVI_CONNECT_RESP2 IU. The Last Sequence bit shall be set to one in a FCVI_CONNECT_RESP3 IU. The Sequence Initiative bit shall be set to one in a FCVI_DISCONNECT_RQST IU. The Last Sequence bit shall be set to one in a FCVI_DISCONNECT_RESP IU.

5.8.5 DF_CTL field

The DF_CTL field indicates any optional headers that may be present. FC-VI IUs require either a 16 or 32 byte Device_Header. The Device_Header is present in all frames of any FC-VI IU. The size of the Device_Header associated with each IU is defined in Table 1, in the DH Size column.

5.8.6 SEQ_CNT field

The SEQ_CNT field indicates the frame order within the Sequence as defined by FC-FS. FC-VI requires that the SEQ_CNT shall be continuously increasing within an Exchange, independent of whether the Sequence is streamed or non-streamed. The Sequence Initiator shall not reuse the same SEQ_CNT within a single Sequence.

See 5.5.

5.8.7 Parameter field

The Parameter field shall indicate the Relative Offset of the first byte of each frame's payload, referenced to a starting value of 0. FC-VI requires that the Relative Offset shall be continuously increasing within an Exchange. For RDMA operations, the Relative Offset value of each frame shall be added to the FCVI_RMT_VA field in the Device_Header to determine the starting address for the VI Application buffer read or write.

See 5.5.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

5.9 FC-VI device_header

5.9.1 FC-VI device_header description

Every FC-VI frame shall contain a Device_Header. For FCVI_SEND_RESP and FCVI_WRITE_RESP IUs, a 16 byte Device_Header is used, as shown in Table 3. For all other IUs, a 32 byte Device_Header is used, as shown in Table 4.

Table 3 – 16-byte FC-VI device_header

Bits Word	31– 24 Byte 0	23–16 Byte 1	15–08 Byte 2	07–00 Byte 3
0	FCVI_HANDLE			
1	FCVI_OPCODE	FCVI_FLAGS	Reserved	
2	FCVI_MSG_ID			
3	FCVI_PARAMETER			

Table 4 – 32-byte FC-VI device_header

Bits Word	31– 24 Byte 0	23–16 Byte 1	15–08 Byte 2	07–00 Byte 3
0	FCVI_HANDLE			
1	FCVI_OPCODE	FCVI_FLAGS	Reserved	
2	FCVI_MSG_ID			
3	FCVI_PARAMETER			
4	(MSB) FCVI_RMT_VA			
5				(LSB)
6	FCVI_RMT_VA_HANDLE			
7	FCVI_TOT_LEN / FCVI_CONNECTION_ID			

5.9.2 FCVI_HANDLE field

The FCVI_HANDLE field is used by a FC-VI Port to uniquely identify the FC-VI Endpoint for the incoming FC-VI IU. Each FC-VI Port shall put the correct FCVI_HANDLE value into every IU transmitted on an FC-VI Connection. The FCVI_HANDLE value for each FC-VI Endpoint of an FC-VI Connection is established during FC-VI Connection Setup. Each FC-VI Port chooses its own FCVI_HANDLE value for the other FC-VI Port to use. Values from 00000000h to FFFFFFFEh are valid FCVI_HANDLEs.

The value of FFFFFFFFh is used to indicate a not yet assigned FCVI_HANDLE. It shall be used for the first two IUs sent during FC-VI Connection Setup, FCVI_CONNECT_RQST and FCVI_CONNECT_RESP1. It shall also be used for a FCVI_DISCONNECT_RQST sent during FC-VI Connection Setup when the FCVI_HANDLE is not yet for the Remote Endpoint of the connection.

--`,,`,,`-`-`,,`,,`,`,,`---
Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

5.9.3 FCVI_OPCODE field

The FCVI_OPCODE field completely specifies the FC-VI IU type and format. Table 1 indicates the FCVI_OPCODE assigned to the various FC-VI IUs. There are four different IU categories defined. The FC-VI Connect Request/Response IUs are used for creating and destroying FC-VI Connections between FC-VI Ports. The VI Message Request/Response IUs are used for Message Transfer only after the FC-VI Connection is established.

5.9.4 FCVI_FLAGS field

5.9.4.1 FCVI_FLAGS field description

The FCVI_FLAGS field contains control flags for the FC-VI operation. Reserved bits shall always be zero. Table 5 define the FCVI_FLAGS settings for Message Request IUs; Table 6 defines the FCVI_FLAGS settings for Message Response IUs; Table 7 defines the FCVI_FLAGS settings for Connect Request IUs; Table 9 defines the FCVI_FLAGS settings for Connect Response IUs; Table 10 define the FCVI_FLAGS settings for Disconnect IUs.

5.9.4.2 FCVI_FLAGS for message request IUs

The FCVI_FLAGS bit positions and definitions for all Message Request IUs are shown in Table 5.

Table 5 – FCVI_FLAGS Bit definitions for message request IUs

Bit Position	Bit Name	Bit Definition
7	Reserved	0
6	Reserved	0
5	Reserved	0
4	Reserved	0
3	Reserved	0
2	Reserved	0
1	Reserved	0
0	IMM_DATA	FCVI_PARAMETER field is Immediate Data when one

If IMM_DATA (bit zero) in FCVI_FLAGS is equal to one, the FCVI_PARAMETER field in the FC-VI Device_Header (Word 3) shall contain Immediate Data. If IMM_DATA in FCVI_FLAGS is equal to zero, the FCVI_PARAMETER field in the FC-VI Device_Header (Word 3) is reserved. IMM_DATA may only be set for a FCVI_SEND_RQST or a FCVI_WRITE_RQST IU.

5.9.4.3 FCVI_FLAGS for message response IUs

The bit positions and definitions for all Message Response IUs are shown in Table 6.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Table 6 – FCVI_FLAGS Bit definitions for message response IUs

Bit Position	Bit Name	Bit Definition
7	Reserved	0
6	Reserved	0
5	Reserved	0
4	Reserved	0
3	TRANS_ERR	Transport Error when one.
2	PROT_ERR	RDMA Protection Error when one.
1	DESC_ERR	Remote Descriptor Error when one.
0	RESP_ERR	Message Transfer successful when zero, errored when one.

For Message Response IUs, the RESP_ERR bit is set when the VI Message responder indicates a failure of the Message Transfer. The reason for the failure is indicated by TRANS_ERR, PROT_ERR and DESC_ERR. The definitions of the TRANS_ERR, PROT_ERR and DESC_ERR bits are defined in 5.9.7.3.1, Descriptor Error Reason Codes.

5.9.4.4 FCVI_FLAGS for connect request IUs

The bit positions and definitions for all Connect Request IUs are shown in Table 7.

Table 7 – FCVI_FLAGS Bit definitions for connect request IUs

Bit Position	Bit Name	Bit Definition
7	Reserved	0
6	Reserved	0
5	Reserved	0
4	RETRY	Retried Connection Setup when one
3	FCVI_CONN_INFO	If one, the Connect Request contains 256 bytes of FC-VI Provider Connect Info If zero, the 256 bytes of FC-VI Provider Connect Info is not present
2	CONN_MODE2	Connection Mode bit 2
1	CONN_MODE1	Connection Mode bit 1
0	CONN_MODE0	Connection Mode bit 0

For Connect Request IUs, the Connection Mode bits 2:0 shall define the connection mode requested. The defined values are shown in Table 8.

If FCVI_CONN_INFO is set to one, the payload of the FCVI_CONNECT_RQST IU contains 256 bytes of connection information from the Local FC-VI Provider in the FCVI_CONNECT_INFO field in the FCVI_CONNECT_RQST payload. If FCVI_CONN_INFO is set to zero, the 256 bytes of connection information are not present in the FCVI_CONNECT_RQST IU.

Connection information may be passed between FC-VI Providers in the FCVI_CONN_INFO field. There is currently no defined VI Architecture interface to allow VI Applications to specify connection information during Connection Setup.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

If RETRY is set to one, the FCVI_CONNECT_RQST IU is part of a retried Connection Setup for a particular value of FCVI_CONNECTION_ID. If RETRY is set to zero, the FCVI_CONNECT_RQST IU part of an initial Connection Setup for a particular value of FCVI_CONNECTION_ID. The Connect Request originator shall set the RETRY bit in the FCVI_CONNECT_RQST IU when a Connection Setup is being retried.

Table 8 – FC-VI connection mode definition

CONN_MODE[2:0]	Connection Mode
111	0
110	0
101	0
100	0
011	0
010	Peer-to-Peer Connect Request
001	Client-Server Connect Request
000	0

In Table 8, for Connect Request IUs a Connection Mode of 001b indicates a Client-Server Connect Request, while a Connection Mode of 010b indicates a Peer-to-Peer Connect Request.

5.9.4.5 FCVI_FLAGS for connect response IUs

The bit positions and definitions for all Connect Response IUs are shown in Table 9.

Table 9 – FCVI_FLAGS Bit definitions for connect response IUs

Bit Position	Bit Name	Bit Definition
7	Reserved	0
6	Reserved	0
5	Reserved	0
4	Reserved	0
3	Reserved	0
2	RETRY	Retried Connection Setup when one
1	FCVI_CONN_INFO	If one, the Connect Response contains 256 bytes of FC-VI Provider Connect Info If set to zero, the 256 bytes of FC-VI Provider Connect Info is not present
0	CONN_STS	Connection Setup status is present when one.

The CONN_STS bit is set to one when a Connection Setup status is available. It may be set in the FCVI_CONNECT_RESP1 IU, FCVI_CONNECT_RESP2 IU or the FCVI_CONNECT_RESP3 IU. Thus, CONN_STS may be set for all Connection IUs except FCVI_CONNECT_RQST. If CONN_STS is set, the FCVI_PARAMETER field shall indicate the additional status information. See Table 11 and Table 12 for the definition of Connection Setup status.

For a FCVI_CONNECT_RESP1 IU, if FCVI_CONN_INFO is set to one, the payload of the FCVI_CONNECT_RESP1 IU contains 256 bytes of connection information from the Remote FC-VI

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Provider in the FCVI_CONNECT_INFO field. If FCVI_CONN_INFO is set to zero, the 256 bytes of connection information are not present in the FCVI_CONNECT_RESP IU.

If RETRY is set to one, the Connect Response IU is part of a retried Connection Setup for a particular value of FCVI_CONNECTION_ID. If RETRY is set to zero, the Connect Response IU part of an initial Connection Setup for a particular value of FCVI_CONNECTION_ID. RETRY shall be set to one for all IUs sent on a retried Connection Setup.

The Connect Setup responder shall set the RETRY bit in the FCVI_CONNECT_RESP1 and FCVI_CONNECT_RESP3 IUs equal to the RETRY bit it received in the FCVI_CONNECT_RQST IU within an Exchange. The Connect Setup requestor shall set the RETRY bit in the FCVI_CONNECT_RESP2 IU equal to the RETRY bit it sent in the FCVI_CONNECT_RQST IU within an Exchange.

5.9.4.6 FCVI_FLAGS for disconnect IUs

The bit positions and definitions for all Disconnect IUs is shown in Table 10.

Table 10 – FCVI_FLAGS Bit definitions for disconnect IUs

Bit Position	Bit Name	Bit Definition
7	Reserved	0
6	Reserved	0
5	Reserved	0
4	Reserved	0
3	Reserved	0
2	CONN_SETUP_ABORT	If one, a Connection Setup is being aborted If zero, an established connection is being aborted.
1	VI_APP_DISCON	If one, the VI Application has issued a Disconnect request If zero, the FC-VI Provider has issued a Disconnect request
0	CONN_STS	Connection Setup status is present when one.

The CONN_STS bit is set to one when additional connection status information is available. It may be set in the FCVI_DISCONNECT_RQST IU or the FCVI_DISCONNECT_RESP IU. If CONN_STS is set, the FCVI_PARAMETER field shall indicate the additional status information. See Table 11 and Table 12 for the definition of connection status. If the CONN_STS bit is set to zero, no additional connection status information is available.

For a FCVI_DISCONNECT_RQST IU, if VI_APP_DISCON (bit one in FCVI_FLAGS) is set to one, a VI Application has issued the Disconnect Request. For a FCVI_DISCONNECT_RQST IU, if VI_APP_DISCON is set to zero, the FC-VI Provider has issued the Disconnect Request. In a FCVI_DISCONNECT_RESP IU the VI_APP_DISCON shall be set equal the VI_APP_DISCON in the FCVI_DISCONNECT_RQST IU for this Exchange.

If the CONN_SETUP_ABORT bit is set to one, the Disconnect request origniator is aborting a Connection Setup. If the CONN_SETUP_ABORT bit is set to zero, the Disconnect request origniator is aborting an established connection. For a FCVI_DISCONNECT_RESP IU, CONN_SETUP_ABORT shall be the same as that received in the FCVI_DISCONNECT_RQST IU for this Exchange.

--`,,`-`-`,,`,,`,`,,`---

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

5.9.5 Reserved fields

All bits in any Reserved field shall be set to zero by the originator. The responder shall ignore all bit values in any Reserved field.

5.9.6 FCVI_MSG_ID field

The FCVI_MSG_ID field is used to uniquely identify a VI Message within a FC-VI Endpoint on a FC-VI Connection. The FCVI_MSG_ID field facilitates the Message responder detecting lost Messages. Each Message originator maintains a FCVI_MSG_ID count for Messages sent on each of its open FC-VI Connections. The FCVI_MSG_ID shall be assigned by the originator for all Request IUs. The originator shall use a value of 00000001h on the first Message of the VI. Subsequent Messages shall increment the FCVI_MSG_ID by one. When the FCVI_MSG_ID reaches FFFFFFFFh, the originator shall wrap the FCVI_MSG_ID to 0000000h on the next Message. The responder shall return the originator's FCVI_MSG_ID value on all corresponding Response IUs for that Message.

To prevent Message frame synonyms, the FC-VI Provider shall not wrap FCVI_MSG_ID within R_A_TOV.

5.9.7 FCVI_PARAMETER field

5.9.7.1 FCVI_PARAMETER field format

The format of the FCVI_PARAMETER field is specified by the FCVI_FLAGS field.

For VI Message Request IUs (except FCVI_READ_RQST), if the FCVI_FLAGS field has the IMM_DATA bit asserted, then the FCVI_PARAMETER field contains Immediate Data from the Descriptor. Otherwise, the FCVI_PARAMETER field shall be zero for Message Request IUs.

For Message Response IUs, the Parameter field shall be set to zero.

For Connect Request IUs, the Parameter field shall be set to zero.

For Connect Response and Disconnect IUs, if the CONN_STS bit is set in the FCVI_FLAGS field, the FCVI_PARAMETER field contains additional connection status information as shown in Table 11.

Table 11 – FCVI_PARAMETER field for connect response and disconnect IUs

Bits Word	31– 24 Byte 0	23–16 Byte 1	15–08 Byte 2	07–00 Byte 3
0	Reserved	Reason Code	Reason Explanation	Vendor Unique

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

The Reason Codes are defined in Table 12. A Reason Explanation of 00h indicates no additional explanation for that Reason Code; all other values are reserved.

Table 12 – Reason code for CONN_STS

Code Point	Reason Code
00h	Reserved
01h	No Discriminator Match
02h	Connect Timeout
03h	No Waiting Remote Connectionpoint
04h	Connect Reject
05h	Connect Reject - Concurrent Matching Peer Requests
06h to 1Fh	Reserved
20h	Connect Reject - Invalid Service Parm
21h	Connect Reject - Protocol Error
22h	Connect Reject - Transport Error
23h to 3Fh	Reserved
40h	Transport Error
41h	RDMA Protection Error
42h	Remote Descriptor Error
43h	Remote RDMA Write Protection Error
44h	Remote RDMA Write Data Error
45h	Remote RDMA Write Packet Abort
46h	Remote RDMA Transport Error
47h	Remote RDMA Read Protection Error
48h	Protocol Error
49h	Connection Setup Timeout
4Ah	Connection Does Not Exist
4Bh	Disconnect Accept
4Ch to 5Fh	Reserved
60h to EFh	Reserved
F0h to FFh	Vendor Unique Error

Reason Codes 01h to 1Fh are non-error response codes for Connect Response IUs (FCVI_CONNECT_RESP1, FCVI_CONNECT_RESP2 and FCVI_CONNECT_RESP3). Reason Codes 20h to 3Fh are error response codes for Connect Response IUs. Reason Codes 40h to 5Fh are valid for Disconnect IUs (FCVI_DISCONNECT_RQST IUs or FCVI_DISCONNECT_RESP IUs). Reason Codes 60h to 6Fh are reserved. Reason Codes 60h to 6Fh are Vendor Unique. The Vendor Unique field may be used by FC-VI Providers to specify additional Reason Codes.

--`,,`-`-`,,`,,`,`,,`---
Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

5.9.7.2 Connect response reason codes

5.9.7.2.1 Connect response non-error reason codes

The following Reason Codes are valid for all non-error status for Connect Response IUs.

No Discriminator Match: There is no Remote Connectionpoint that matches the Discriminator in the Connect Request.

Connect Timeout: The VI Application at the Connect Request originator has timed out waiting for a Connect Response.

No Waiting Remote Connectionpoint: There is no Connectionpoint waiting at the Remote FC-VI Provider.

Connect Reject: The Remote Endpoint has rejected the Connect Request.

Connect Reject - Concurrent Matching Peer Requests: The Remote FC-VI Provider has detected a Concurrent Matching Peer Request and has rejected the Connect Request.

Reserved: Reserved for future Connect Response non-error codes.

5.9.7.2.2 Connect response error reason codes

The following Reason Codes are valid for all error status for Connect Response IUs.

Connect Reject - Invalid Service Param: The Remote FC-VI Provider has rejected the Connect Request due to invalid PLOGI service parameters (see 5.5).

Connect Reject - Protocol Error: The Remote FC-VI Provider has detected a protocol error and has rejected the Connect Request.

Connect Reject - Transport Error: The Remote FC-VI Provider has detected a transport error and has rejected the Connect Request.

Reserved: Reserved for future Connect Response error codes.

5.9.7.3 Message response / disconnect reason codes

5.9.7.3.1 Descriptor error reason codes

The VI Architecture (see Clause 2.3.1 in VI-ARCH), defines the following Descriptor error codes. The FC-VI Provider may directly map the following FC-VI error codes to the corresponding bits in the Status Field of the Descriptor Control Segment (See Clause 5.2 in VI-DG).

The following Reason Codes are valid for all Message Response IUs and all Disconnect IUs.

Transport Error: The Local or Remote FC-VI Provider has detected a Fibre Channel transport error.

RDMA Protection Error: A protection error for an RDMA operation was detected by the Remote FC-VI Provider.

Remote Descriptor Error: The Remote FC-VI Provider has detected a Descriptor error. Possible errors include no receive Descriptor posted, or length, format or protection error on Descriptor access.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

5.9.7.3.2 Remote FC-VI port non-descriptor errors

The VI Architecture defines non-Descriptor error codes (except for protocol error) as seen by the Local VI Application when reported in an asynchronous error notification (See Section 3.8.1 in VI-DG). FC-VI allows the Disconnect originator to report the originator's non-Descriptor error codes in the FCVI_DISCONNECT_RQST IU, or the Disconnect responder to report the responder's non-Descriptor error codes in the FCVI_DISCONNECT_RESP IU.

Note 3: The VI Architecture does not provide an interface to report these Remote errors to the Local VI Application in either the Descriptor or an asynchronous error notification. The Local FC-VI Provider may record these errors for reporting to a VI management application, or it may make the information available to a VI Application through a proprietary interface.

The following Reason Codes are valid for all Disconnect IUs.

Remote RDMA Write Protection Error: The Remote FC-VI Provider detected a protection error on an incoming RDMA Write operation.

Remote RDMA Write Data Error: The Remote FC-VI Provider detected a data corruption error on an incoming RDMA Write operation.

Remote RDMA Write Packet Abort Error: The Remote FC-VI Provider detected a partial data packet on an incoming RDMA Write operation.

Remote RDMA Transport Error: The Remote FC-VI Provider detected a Fibre Channel transport error on an incoming RDMA operation.

Remote RDMA Read Protection Error: The Remote FC-VI Provider detected a protection error on an incoming RDMA Read operation.

Protocol Error: The Remote FC-VI Provider detected a protocol error.

Connection Setup Timeout: The Connect Request originator has detected a timeout waiting for a FCVI_CONNECT_RESP3 IU.

Connection Does Not Exist: The connection or Connection Setup does not exist at the Disconnect responder.

Disconnect Accept: The Disconnect responder is accepting the Disconnect request and is aborting the connection or the Connection Setup.

Reserved: Reserved for future Message Response or Disconnect error codes.

5.9.7.3.3 Reserved for future expansion

Reserved: Reserved for future connection status or error codes.

5.9.7.3.4 Vendor unique reason codes

Vendor Unique: Vendor unique error codes.

5.9.8 FCVI_RMT_VA field

The FCVI_RMT_VA field specifies the base virtual address for the Remote memory region for the FCVI_RDMA_READ_RQST IU, the FCVI_RDMA_WRITE_RQST IU or the

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

FCVI_RDMA_READ_RESP IU. It is reserved for all other IUs. The FCVI_RMT_VA is a 64 bit quantity. The MSB (Most Significant Byte) appears in Word 4, Byte 0, while the LSB (Least Significant Byte) appears in Word 5, Byte 3. The value of FCVI_RMT_VA is identical for all IUs for a Message.

5.9.9 FCVI_RMT_VA_HANDLE field

The FCVI_RMT_VA_HANDLE field specifies the handle for the Remote memory region for the FCVI_RDMA_READ_RQST IU, the FCVI_RDMA_WRITE_RQST IU or the FCVI_RDMA_READ_RESP IU. It is reserved for all other IUs.

5.9.10 FCVI_TOT_LEN field / FCVI_CONNECTION_ID field

The FCVI_TOT_LEN field specifies the VI Message total data length for all FC-VI Message Request IUs. It indicates the total Message length for FCVI_SEND_RQST IUs, FCVI_WRITE_RQST IUs, and FCVI_READ_RQST IUs.

The FCVI_CONNECTION_ID field is used for all Connection IUs. It shall be used by the Connect Request originator to assign a unique identifier to all open Connection Setup operations.

For all Connect Response IUs, the FCVI_CONNECTION_ID field shall be set equal to the FCVI_CONNECTION_ID field from the FCVI_CONNECT_RQST for that Exchange.

If a Connection Setup is aborted, the FCVI_CONNECTION_ID field in the FCVI_DISCONNECT_RQST shall be set equal to the FCVI_CONNECTION_ID field from the FCVI_CONNECT_RQST. The FCVI_CONNECTION_ID in the Disconnect response shall be set equal to the FCVI_CONNECTION_ID field from the FCVI_DISCONNECT_RQST.

A value of FFFFFFFFh is reserved for an unassigned or unknown FCVI_CONNECTION_ID.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

6 FC-VI Information Unit (IU) formats

6.1 FC-VI IU overview

The following subclauses describes the Device_Header and payload contents of the IUs transferred during an FC-VI operation. The Information Units used by FC-VI Ports and their characteristics are described in 5.1 of this document.

6.2 FCVI_SEND_RQST IU

6.2.1 FCVI_SEND_RQST IU description

The FCVI_SEND_RQST IU shall transfer all or part of a VI Message to the Remote FC-VI Endpoint. One or more FCVI_SEND_RQST IUs may be sent per Exchange. Refer to Figure 2 or Figure 3 for a definition of the complete Exchange. The Payload of each frame in this IU carries VI Application data. For Unreliable Delivery and Reliable Delivery, the last FCVI_SEND_RQST IU shall terminate the Exchange. For Reliable Reception, the last FCVI_SEND_RQST IU shall transfer Sequence Initiative.

6.2.2 FCVI_SEND_RQST Device_Header information

The FCVI_SEND_RQST IU shall contain a 32 byte FC-VI Device_Header as shown in Table 4.

The FCVI_HANDLE shall be set to the FCVI_HANDLE communicated by the Remote FC-VI Endpoint during FC-VI Connection Setup.

The FCVI_OPCODE field is defined in Table 1.

The FCVI_FLAGS field is defined in Table 5. Sending VI Immediate Data is optional with this IU. If Immediate Data is sent, the IMM_DATA bit in the FCVI_FLAGS field shall be one, and the FCVI_PARAMETER field shall contain the Immediate Data value. Otherwise, the FCVI_PARAMETER and FCVI_FLAGS fields shall be zero.

The FCVI_MSG_ID field shall be incremented by one from the previously sent Message on this connection.

The FCVI_RMT_VA and FCVI_RMT_VA_HANDLE shall be set to zero.

The FCVI_TOT_LEN field is set to the total length of the VI Message. The FC-VI Provider shall set FCVI_TOT_LEN equal to the Length field in the Control Segment of the Descriptor (see Section 4.3 in VI-DG).

6.3 FCVI_SEND_RESP IU

6.3.1 FCVI_SEND_RESP IU description

The FCVI_SEND_RESP IU shall be sent in response to a FCVI_SEND_RQST for Reliable Reception. At most one FCVI_SEND_RESP IU shall be sent per VI Message. This IU is not allowed for Unreliable Delivery or Reliable Delivery. Refer to Figure 3 for a definition of the complete Exchange. The Payload of this IU shall be empty. This IU shall contain only one frame. The FCVI_SEND_RESP IU shall terminate the Exchange.

6.3.2 FCVI_SEND_RESP Device_Header information

The FCVI_SEND_RESP IU shall contain a 16 byte FC-VI Device_Header as shown in Table 3.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

The FCVI_HANDLE shall be set to the FCVI_HANDLE communicated by the Remote FC-VI Endpoint during FC-VI Connection Setup.

The FCVI_OPCODE field is defined in Table 1.

The FCVI_FLAGS field is defined by Table 6.

The FCVI_PARAMETER field shall be zero.

6.4 FCVI_WRITE_RQST IU

6.4.1 FCVI_WRITE_RQST IU overview

The FCVI_WRITE_RQST IU shall write all or part of a VI Message to the Remote FC-VI Endpoint. The VI Message data is written to the Remote virtual address generated by adding the FCVI_RMT_VA field in the Device_Header and the Relative Offset field in the frame header. One or more FCVI_WRITE_RQST IUs may be sent per Exchange. Refer to Figure 4 or Figure 5 for a definition of the complete Exchange. The Payload of each frame in this IU carries VI Application data. For Unreliable Delivery and Reliable Delivery, the last FCVI_WRITE_RQST IU shall terminate the Exchange. For Reliable Reception, the last FCVI_WRITE_RQST IU shall transfer Sequence Initiative.

6.4.2 FCVI_WRITE_RQST IU Device_Header information

The FCVI_WRITE_RQST IU shall contain a 32 byte FC-VI Device_Header as shown in Table 4.

The FCVI_HANDLE shall be set to the FCVI_HANDLE communicated by the Remote FC-VI Endpoint during FC-VI Connection Setup.

The FCVI_OPCODE field is defined in Table 1.

The FCVI_FLAGS field is defined in Table 5. Sending VI Immediate Data is optional with this IU. If Immediate Data is sent, then the IMM_DATA bit in the FCVI_FLAGS field shall be one, and the FCVI_PARAMETER field shall contain the Immediate Data value. Otherwise, the FCVI_PARAMETER and FCVI_FLAGS fields shall be zero.

The FCVI_MSG_ID field shall be incremented by one from the previously sent Message on the connection.

The FCVI_RMT_VA field shall be set to the base address of the Remote virtual memory region where the VI Message shall be written to. This value shall be the same for all Request IUs in the Message.

The FCVI_RMT_VA_HANDLE is the handle for the Remote virtual memory region where the VI Message shall be written to. This value shall be the same for all request IUs in the Message.

The FCVI_TOT_LEN field is set to the total length of the VI Message. The FC-VI Provider shall set FCVI_TOT_LEN equal to the Length field in the Control Segment of the Descriptor (see Section 4.3 in VI-DG). This value shall be the same for all request IUs in the Message.

6.5 FCVI_WRITE_RESP IU

6.5.1 FCVI_WRITE_RESP IU description

The FCVI_WRITE_RESP IU shall be sent in response to a FCVI_WRITE_RQST for Reliable Reception. It is not allowed for Unreliable Delivery or Reliable Delivery. At most one FCVI_WRITE_RESP

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

IU shall be sent per VI Message. Refer to Figure 5 for a definition of the complete Exchange. The Payload of this IU shall be empty. This IU shall contain only one frame. The FCVI_WRITE_RESP IU shall terminate the Exchange.

6.5.2 FCVI_WRITE_RESP IU Device_Header information

The FCVI_WRITE_RESP IU shall contain a 16 byte FC-VI Device_Header as shown in Table 3.

The FCVI_HANDLE shall be set to the FCVI_HANDLE communicated by the Remote FC-VI Endpoint during FC-VI Connection Setup.

The FCVI_OPCODE field is defined in Table 1.

The FCVI_FLAGS field is defined in Table 6.

The FCVI_MSG_ID field shall be the same as that received in the FCVI_WRITE_RQST IU.

The FCVI_PARAMETER field shall be zero.

6.6 FCVI_READ_RQST IU

6.6.1 FCVI_READ_RQST IU description

The FCVI_READ_RQST IU shall request the read of a VI Message from the Remote FC-VI Endpoint This IU is allowed only for Reliable Delivery or Reliable Reception. The VI Message data is read from the Remote memory region beginning at the virtual address specified in the FCVI_RMT_VA field of the Device_Header. The Payload of this IU shall be empty. This IU shall contain only one frame. Sequence Initiative shall be passed at the completion of this IU. Refer to Figure 6 for a definition of the complete Exchange.

6.6.2 FCVI_READ_RQST IU Device_Header information

The FCVI_READ_RQST IU shall contain a 32 byte FC-VI Device_Header as shown in Table 4.

The FCVI_HANDLE shall be set to the FCVI_HANDLE communicated by the Remote FC-VI Endpoint during FC-VI Connection Setup.

The FCVI_OPCODE field is defined in Table 1.

The FCVI_FLAGS field is defined in Table 5. Sending VI Immediate Data is not allowed with this IU. The IMM_DATA bit in the FCVI_FLAGS field shall be zero, and the FCVI_PARAMETER field shall be zero.

The FCVI_MSG_ID field shall be incremented by one from the previously sent Message on the connection.

The FCVI_RMT_VA field shall be set to the base address of the Remote virtual memory region where the VI Message shall be read from.

The FCVI_RMT_VA_HANDLE is the handle for the Remote virtual memory region where the VI Message shall be read from.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

The FCVI_TOT_LEN field is set to the total length of the VI Message. The FC-VI Provider shall set FCVI_TOT_LEN equal to the Length field in the Control Segment of the Descriptor (see Section 4.3 in VI-DG).

6.7 FCVI_READ_RESP IU

6.7.1 FCVI_READ_RESP IU description

The FCVI_READ_RESP IU shall be sent in response to a FCVI_READ_RQST for Reliable Delivery and Reliable Reception. It shall transfer all or part of a VI Message to the Remote FC-VI Endpoint. The VI Message data is read from the Remote virtual address generated by adding the FCVI_RMT_VA field in the Device_Header and the Relative Offset field in the frame header. One or more FCVI_READ_RESP IUs may be sent per Exchange. The last FCVI_READ_RESP IU shall terminate the Exchange. Refer to Figure 6 for a definition of the complete Exchange. The Payload of each frame in this IU carries VI Application data.

6.7.2 FCVI_READ_RESP IU Device_Header information

The FCVI_READ_RESP IU shall contain a 32 byte FC-VI Device_Header as shown in Table 4.

The FCVI_HANDLE shall be set to the FCVI_HANDLE communicated by the Remote FC-VI Endpoint during FC-VI Connection Setup.

The FCVI_OPCODE field is defined in Table 1.

The FCVI_FLAGS field is defined by Table 6.

The MSG_ID field shall be the same as that received in the FCVI_READ_RQST IU.

The FCVI_PARAMETER field shall be zero.

The FCVI_RMT_VA field shall be set to the base address of the Remote virtual memory region where the VI Message is read from. This value shall be the same as the FCVI_RMT_VA field in the FCVI_READ_RQST IU. This value shall be the same for all Response IUs in the Message.

The FCVI_RMT_VA_HANDLE is the handle for the Remote virtual memory region where the VI Message is read from. This value shall be the same as the FCVI_RMT_VA_HANDLE field in the FCVI_READ_RQST IU. This value shall be the same for all Response IUs in the Message.

The FCVI_TOT_LEN field is set to the FCVI_TOT_LEN received in the FCVI_READ_RQST IU. This value shall be the same for all Response IUs in the Message.

6.8 FCVI_CONNECT_RQST IU

6.8.1 FCVI_CONNECT_RQST IU description

The FCVI_CONNECT_RQST IU shall request that an FC-VI Connection be established. Refer to Figure 8, Figure 9, or Figure 10 for a definition of the complete Exchange. The Payload for the FCVI_CONNECT_RQST IU is shown in Table 13. Sequence Initiative shall be passed at the completion of this IU. The payload length for the FCVI_CONNECT_RQST IU is 596 bytes if the FCVI_CONNECT_INFO is present, or 340 bytes if the FCVI_CONNECT_INFO is not present.

--`,,`-`-`,,`,,`,`,,`---
Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

6.8.2 FCVI_CONNECT_RQST Device_Header information

The FCVI_CONNECT_RQST IU shall contain a 32 byte FC-VI Device_Header as shown in Table 4.

The FCVI_HANDLE is unassigned.

The FCVI_OPCODE field is defined in Table 1.

The FCVI_FLAGS field are defined in Table 7.

The FCVI_MSG_ID field shall be set to zero.

The FCVI_PARAMETER field shall be zero.

The FCVI_RMT_VA shall be set to zero

The FCVI_RMT_VA_HANDLE shall be set to zero.

The FCVI_CONNECTION_ID shall be set to a unique value by the Connect Request originator.

6.8.3 FCVI_CONNECT_RQST Payload Information

Table 13 shows the payload for the FCVI_CONNECT_RQST IU.

Table 13 – FCVI_CONNECT_RQST IU payload format

Bits Word	31– 24 Byte 0	23–16 Byte 1	15–08 Byte 2	07–00 Byte 3
0	Reserved			
1	Reserved		FCVI_REVISION	
2	FCVI_RQST_HANDLE			
3	MSB			
4-38	FCVI_LOC_ADDR			
39				LSB
40	MSB			
41-75	FCVI_REM_ADDR			
76				LSB
77	MSB			
78-83	FCVI_LOC_ATTRS			
84				LSB
85	MSB			
86-147	FCVI_CONNECT_INFO (optional)			
148				LSB

--`,,`-`-`,,`,,`,`,,`---

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

The FCVI_REVISION field shall be set to 0001h to indicate the first revision of this specification.

The FCVI_RQST_HANDLE field shall be chosen by the Connect Request originator of this IU. The Connect Request responder shall use this same FCVI_RQST_HANDLE in the FCVI_HANDLE field when transferring VI Messages on this FC-VI Connection.

The FCVI_LOC_ADDR field is the FCVI_NET_ADDRESS for the FC-VI Connectionpoint of the Connect Request originator. Table 15 defines the format of this field.

The FCVI_REM_ADDR field is the FCVI_NET_ADDRESS for the FC-VI Connectionpoint of the Connect Request responder. Table 15 defines the format of this field.

The FCVI_LOC_ATTRS field is the VI attributes for the FC-VI Endpoint of the Connect Request originator. Table 16 defines the format of this field.

If the FCVI_CONN_INFO is set to one in the FCVI_FLAGS field of the Connect Request, the FCVI_CONNECT_INFO field contains 256 bytes of optional Local FC-VI Provider supplied connection information passed from the Connect Request originator to the Connect Request responder. If the FCVI_CONN_INFO is set to zero, the FCVI_CONNECT_INFO field is not present.

6.9 FCVI_CONNECT_RESP1 IU

6.9.1 FCVI_CONNECT_RESP1 IU description

The FCVI_CONNECT_RESP1 IU shall be sent in response to a FCVI_CONNECT_RQST IU. Refer to Figure 8, Figure 9, or Figure 10 for a definition of the complete Exchange. The Payload for the FCVI_CONNECT_RESP1 IU is shown in Table 14. Sequence Initiative shall be passed at the completion of this IU. The payload length for the FCVI_CONNECT_RESP1 IU is 596 bytes if the FCVI_CONNECT_INFO is present, or 340 bytes if the FCVI_CONNECT_INFO is not present. The format of the FCVI_CONNECT_RESP1 IU is identical to the format of the FCVI_CONNECT_RQST IU.

6.9.2 FCVI_CONNECT_RESP1 Device_Header information

The FCVI_CONNECT_RESP1 IU shall contain a 32 byte FC-VI Device_Header as shown in Table 4.

The FCVI_HANDLE field is unassigned.

The FCVI_OPCODE field is defined in Table 1.

The FCVI_FLAGS field shall be defined in Table 9.

The FCVI_MSG_ID field shall be set to zero.

If the FCVI_FLAGS field is zero, then the FCVI_PARAMETER field shall be zero. If the FCVI_FLAGS field is not zero (i.e., the CONN_STS bit is set), then the FCVI_PARAMETER field shall contain connection status information as shown in Table 11 and Table 12.

The FCVI_RMT_VA shall be set to zero

The FCVI_RMT_VA_HANDLE shall be set to zero.

The FCVI_CONNECTION_ID shall be the same as that received in the FCVI_CONNECT_RQST IU for this Exchange.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

6.9.3 FCVI_CONNECT_RESP1 Payload Information

Table 14 – FCVI_CONNECT_RESP1 IU Payload Format

Bits Word	31– 24 Byte 0	23–16 Byte 1	15–08 Byte 2	07–00 Byte 3
0	Reserved			
1	Reserved		FCVI_REVISION	
2	FCVI_RESP_HANDLE			
3	MSB			
4-38	FCVI_LOC_ADDR			
39				LSB
40	MSB			
41-75	FCVI_REM_ADDR			
76				LSB
77	MSB			
78-83	FCVI_LOC_ATTRS			
84				LSB
85	MSB			
86-147	FCVI_CONNECT_INFO (optional)			
148				LSB

The FCVI_REVISION field shall be set to 0001h to indicate the first revision of this specification.

The FCVI_RESP_HANDLE field shall be chosen by the Connect Request responder, the originator of this IU. The Connect Request originator shall use the value specified in the FCVI_RESP_HANDLE field for the FCVI_HANDLE field in all Message request IUs sent to the Connect Request responder on this FC-VI Connection.

The FCVI_LOC_ADDR field is the FCVI_NET_ADDRESS for the FC-VI Connectionpoint of the Connect Request responder. Table 15 defines the format of this field.

The FCVI_REM_ADDR field is the FCVI_NET_ADDRESS for the FC-VI Connectionpoint of the Connect Request originator. Table 15 defines the format of this field.

The FCVI_LOC_ATTRS field is the VI attributes for the FC-VI Endpoint of the Connect Request responder. Table 16 defines the format of this field.

If the FCVI_CONN_INFO is set to one in the FCVI_FLAGS field of the FCVI_CONNECT_RESP1, the FCVI_CONNECT_INFO field contains 256 bytes of optional FC-VI Provider supplied connection information passed from the Connect Request responder to the Connect Request originator. If the FCVI_CONN_INFO is set to zero, the FCVI_CONNECT_INFO field is not present.

--`,,`-`-`,,`,,`,`,,`---

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

6.10 FCVI_CONNECT_RESP2 IU

6.10.1 FCVI_CONNECT_RESP2 IU description

The FCVI_CONNECT_RESP2 IU shall be sent in response to a FCVI_CONNECT_RESP1 IU. Refer to Figure 8, Figure 9, or Figure 10 for a definition of the complete Exchange. The payload for the FCVI_CONNECT_RESP2 IU shall be empty. Sequence Initiative shall be passed at the completion of this IU. This IU shall contain only one frame.

6.10.2 FCVI_CONNECT_RESP2 Device_Header information

The FCVI_CONNECT_RESP2 IU shall contain a 32 byte FC-VI Device_Header as shown in Table 4.

The FCVI_HANDLE field shall be set to the FCVI_RESP_HANDLE received in the payload of the FCVI_CONNECT_RESP1 IU.

The FCVI_OPCODE field is defined in Table 1.

The FCVI_FLAGS field is defined in Table 9.

The FCVI_MSG_ID field shall be set to zero.

If the FCVI_FLAGS field is zero, then the FCVI_PARAMETER field shall be zero. If the FCVI_FLAGS field is not zero (i.e., the CONN_STS bit is set), then the FCVI_PARAMETER field shall contain connection status information as shown in Table 11 and Table 12.

The FCVI_RMT_VA shall be set to zero

The FCVI_RMT_VA_HANDLE shall be set to zero.

The FCVI_CONNECTION_ID shall be the same as that sent in the FCVI_CONNECT_RQST IU for this Exchange.

6.11 FCVI_CONNECT_RESP3 IU

6.11.1 FCVI_CONNECT_RESP3 IU description

The FCVI_CONNECT_RESP3 IU shall be sent in response to a FCVI_CONNECT_RESP2 IU. Refer to Figure 8, Figure 9, or Figure 10 for a definition of the complete Exchange. The payload for the FCVI_CONNECT_RESP3 IU shall be empty. This IU shall terminate the Exchange. This IU shall contain only one frame.

6.11.2 FCVI_CONNECT_RESP3 Device_Header information

The FCVI_CONNECT_RESP3 IU shall contain a 32 byte FC-VI Device_Header as shown in Table 4.

The FCVI_HANDLE field shall be set to the FCVI_RQST_HANDLE received in the payload of the FCVI_CONNECT_RQST IU.

The FCVI_OPCODE field is defined in Table 1.

The FCVI_FLAGS field is defined in Table 9.

The FCVI_MSG_ID field shall be set to zero.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

If the FCVI_FLAGS field is zero, then the FCVI_PARAMETER field shall be zero. If the FCVI_FLAGS field is not zero (i.e., the CONN_STS bit is set), then the FCVI_PARAMETER field shall contain connection status information as shown in Table 11 and Table 12.

The FCVI_RMT_VA shall be set to zero

The FCVI_RMT_VA_HANDLE shall be set to zero.

The FCVI_CONNECTION_ID shall be the same as that received in the FCVI_CONNECT_RQST IU for this Exchange.

6.12 FCVI_DISCONNECT_RQST IU

6.12.1 FCVI_DISCONNECT_RQST IU description

The FCVI_DISCONNECT_RQST IU shall request an FC-VI Connection be removed or a Connection Setup be aborted. Refer to Figure 11 for a definition of the complete Exchange. The Payload for the FCVI_DISCONNECT_RQST IU shall be empty. Sequence Initiative shall be passed at the completion of this IU. This IU shall contain only one frame.

6.12.2 FCVI_DISCONNECT_RQST Device_Header information

The FCVI_DISCONNECT_RQST IU shall contain a 32 byte FC-VI Device_Header as shown in Table 4.

The FCVI_HANDLE shall be set to the FCVI_HANDLE communicated by the Remote FC-VI Endpoint during FC-VI Connection Setup. If a Connection Setup is being aborted and the FCVI_HANDLE of the Remote FC-VI Endpoint is unknown, the FCVI_HANDLE shall be set to FFFFFFFFh.

The FCVI_OPCODE field is defined in Table 1.

The FCVI_FLAGS field is defined in Table 9. If the CONN_STS bit is set to one, the FCVI_PARAMETER field contains additional connection status information from the Disconnect responder as defined in Table 11 and Table 12. If the CONN_STS bit is set to zero, the FCVI_PARAMETER is set to zero. If the Disconnect originator is aborting a Connection Setup, the CONN_SETUP_ABORT shall be set to one. If the Disconnect originator is aborting an established FC-VI Connection, the CONN_SETUP_ABORT shall be set to zero. If the VI Application issues the Disconnect request, the VI_APP_DISCON shall be set to one. If the VI Provider issues the Disconnect request, the VI_APP_DISCON shall be set to zero.

If the Disconnect originator is aborting a Connection Setup, the FCVI_MSG_ID field shall be set to zero. If the Disconnect originator is aborting an established connection, the FCVI_MSG_ID field shall be set to the FCVI_MSG_ID of the last Message successfully completed by the Disconnect originator.

The FCVI_RMT_VA shall be set to zero

The FCVI_RMT_VA_HANDLE shall be set to zero.

If the Disconnect originator is aborting a Connection Setup, the FCVI_CONNECTION_ID shall be the same as that sent or received in the FCVI_CONNECT_RQST IU for this Exchange. If the Disconnect originator is aborting an established connection, the FCVI_CONNECTION_ID shall be zero.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

6.13 FCVI_DISCONNECT_RESP IU

6.13.1 FCVI_DISCONNECT_RESP IU description

The FCVI_DISCONNECT_RESP IU shall be sent in response to a FCVI_DISCONNECT_RQST IU. Refer to Figure 11 for a definition of the complete Exchange. The Payload for the FCVI_DISCONNECT_RESP IU shall be empty. This IU shall contain only one frame, and the Exchange is terminated at the completion of this IU.

6.13.2 FCVI_DISCONNECT_RESP Device_Header information

The FCVI_DISCONNECT_RESP IU shall contain a 32 byte FC-VI Device_Header as shown in Table 4.

The FCVI_HANDLE shall be set to the FCVI_HANDLE communicated by the Remote FC-VI Endpoint during FC-VI Connection Setup. If a Connection Setup is being aborted and the FCVI_HANDLE of the Remote FC-VI Endpoint is unknown, the FCVI_HANDLE shall be set to FFFFFFFFh.

The FCVI_OPCODE field is defined in Table 1.

The FCVI_FLAGS field is defined in Table 9. The CONN_SETUP_ABORT and VI_APP_DISCON flags shall be the same as that received in the FCVI_DISCONNECT_RQST IU for this Exchange. If the CONN_STS bit is set to one, the FCVI_PARAMETER field contains additional connection status information from the Disconnect responder as defined in Table 11 and Table 12. If the CONN_STS bit is set to zero, the FCVI_PARAMETER is set to zero.

If a Connection Setup is being aborted (CONN_SETUP_ABORT is set to one), the FCVI_MSG_ID field shall be set to zero. If an established connection is being aborted (CONN_SETUP_ABORT is set to zero), the FCVI_MSG_ID field shall be set to the FCVI_MSG_ID of the last Message successfully completed by the Disconnect responder.

The FCVI_RMT_VA shall be set to zero

The FCVI_RMT_VA_HANDLE shall be set to zero.

The FCVI_CONNECTION_ID shall be the same as that received in the FCVI_DISCONNECT_RQST IU for this Exchange. This field may only be non-zero when a Connection Setup is being aborted.

--`,,`,,`-`-`,,`,,`,`,,`---

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

7 FC-VI Addressing and naming

7.1 FC-VI Addressing and naming overview

This Clause defines address assignment for FC-VI Ports and the translation between VI Architecture addresses and FC-VI address constructs.

7.2 FCVI_NET_ADDRESS format

The VI Architecture does not define the format of a VI Address. Instead, the VI Architecture defines an abstract VI Address structure for a VI NIC, which contains a Host Address and a Discriminator. This VI Address structure is called a VIP_NET_ADDRESS and is defined by VI Architecture as follows:

```
typedef struct {
      VIP_UINT16 HostAddressLen;
      VIP_UINT16 DiscriminatorLen;
   VIP_UINT8   HostAddress[1];
} VIP_NET_ADDRESS;
```

The VIP_NET_ADDRESS is visible to a VI Application and it is used to establish a VI connection with a Remote VI Endpoint. The VI Architecture assumes that the VI Provider maps a VIP_NET_ADDRESS to a transport specific addressing structure. FC-VI maps a VIP_NET_ADDRESS to the FCVI_NET_ADDRESS format.

The Host Address for FC-VI shall be an IPv6 address (see RFC2373). The VI Architecture defines an interface to a generic name service to map from Host Names to Host Addresses. Since the Host Address is an IP address, FC-VI facilitates the use of DNS to map between Host Names and IP addresses. Thus, FC-VI implicitly assumes that each FC-VI Host or Node has the capability to access DNS, typically by an Ethernet connection running TCP/IP.

The specific format of the Discriminator portion of the VI Address is outside the scope of this specification, as it is intended to be application specific. However, The VI Architecture specifies that the Discriminator portion may be of variable width from 0 to MaxDiscriminatorLen bytes as indicated in the VI NIC Attributes. FC-VI defines a 128 byte field in FCVI_NET_ADDRESS field for the Discriminator. An FC-VI Provider shall indicate a MaxDiscriminatorLen in the VI NIC Attributes of no more than 128 bytes.

The actual byte length allocated within the Connect Request and response payload for the FCVI_NET_ADDRESS shall be 2 bytes (Reserved) + 1 byte (HostAddressLen) + 1 byte (DiscriminatorLen) + 16 bytes (IP Addr) + 128 bytes (Discriminator) = 148 bytes. For FC-VI, the value of HostAddressLen shall be 16 bytes. The value of DiscriminatorLen may be any value between 16 and 128 bytes. The Discriminator is always allocated as a 128 byte field in FCVI_NET_ADDRESS.

Table 15 illustrates the format used within FC-VI Connection IUs to convey a FCVI_NET_ADDRESS. The HOST_ADD field shall be the Host Address in IPv6 format. The MSB (Most Significant Byte) of the HOST_ADD is contained in Byte 0, Word 1. The LSB (Least Significant Byte) of the HOST_ADD is contained in Byte 3, Word 4. The HOST_ADD_LEN shall be equal to 10h. The DISCRIM field shall be the Discriminator, whose actual length is specified by the DISCRIM_LEN field. The maximum allowed value of DISCRIM_LEN is 80h, while the minimum value is 10h. The MSB (Most Significant Byte) of the DISCRIM is contained in Byte 0, Word 5. The LSB (Least Significant Byte) of the DISCRIM is contained in Byte 3, Word 36.

--`,,`,,`-`-`,,`,,`,`,,`---
Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Table 15 – FCVI_NET_ADDRESS Format

Bits Word	31– 24 Byte 0	23–16 Byte 1	15–08 Byte 2	07–00 Byte 3
0	Reserved		HOST_ADD_LEN	DISCRIM_LEN
1	MSB			
2-3	HOST_ADD			
4				LSB
5	MSB			
6-35	DISCRIM			
36				LSB

7.3 FCVI_ATTRIBUTES format

Table 16 illustrates the format used within FC-VI Connection IUs to convey FCVI_ATTRIBUTES. For a Connect Request, the FCVI_ATTRIBUTES convey the VI Attributes of the connection originator. For a FCVI_CONNECT_RESP1, the FCVI_ATTRIBUTES convey the VI Attributes of the connection responder.

Table 16 – FCVI_ATTRIBUTES format

Bits Word	31– 24 Byte 0	23–16 Byte 1	15–08 Byte 2	07–00 Byte 3
0	Reserved		FCVI_RELIABILITY_LVL	FCVI_ATTR_FLAGS
1	FCVI_MAX_TRANS_SIZE			
2	MSB			
3-5	FCVI_QOS			
6				LSB

FCVI_RELIABILITY_LVL: The Reliability Level of the VI. If set to 01h, the Reliability Level is Unreliable Delivery. If set to 02h, the Reliability Level is Reliable Delivery. If set to 03h, the Reliability Level is Reliable Reception. All other values are reserved.

FCVI_ATTR_FLAGS: The VI Attributes flags are defined in Table 17.

FCVI_MAX_TRANS_SIZE: The maximum Message size in bytes for this VI.

FCVI_QOS: The VI QOS parameters are defined in Table 18.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Table 17 – Format of FCVI_ATTR_FLAGS in FCVI_ATTRIBUTES

Bit Position	Bit Name	Bit Definition
7	Reserved	0
6	Reserved	0
5	Reserved	0
4	Reserved	0
3	Reserved	0
2	Reserved	0
1	EnRdmaWr	VI enabled for RDMA Write when one.
0	EnRdmaRd	VI enabled for RDMA Read when one.

Table 17 defines the format of the FCVI_ATTR_FLAGS field in the FCVI_ARRTIBUTES field (see Table 16). If EnRDMAWr is set to one, the VI is enabled to receive RDMA Write requests. If EnRDMAWr is set to zero, the VI is not enabled to receive RDMA Write requests. If EnRDMARd is set to one, the VI is enabled to receive RDMA Read requests. If EnRDMARd is set to zero, the VI is not enabled to receive RDMA Read requests.

Table 18 – FCVI_QOS format

Bits Word	31– 24 Byte 0	23–16 Byte 1	15–08 Byte 2	07–00 Byte 3
0	FCVI_PREF	Reserved		
1	FCVI_MAX_BANDWIDTH			
2	FCVI_MIN_BANDWIDTH			
3	FCVI_MAX_DELAY			
4	FCVI_PIPELINE_DEPTH			

Table 18 defines the format for FCVI_QOS field in FCVI_ATTRIBUTES field (see Table 16).

FCVI_PREF: If FCVI_PREF is set to 80h, the FC-VI Provider is requesting that all frames sent on this VI use the FC-FS Preference Function. When FCVI_PREF is set to 80h, the FC-VI Provider shall always send a Message frame with the CS_CTL field specifying the Preference Function for this VI. If FCVI_PREF is set to 00h, the FC-VI Provider is requesting that all frames sent on this VI shall not use the FC-FS Preference Function. When FCVI_PREF is set to 00h, the FC-VI Provider shall never send a Message frame with the CS_CTL field specifying the Preference Function for this VI. All other values are reserved.

FCVI_MAX_BANDWIDTH: Specifies the maximum bandwidth for the VI in bytes per second. If non-zero, the FC-VI Provider is guaranteeing the maximum bandwidth between the VI Endpoints of a connection shall be no more than FCVI_MAX_BANDWIDTH bytes per second. If zero, the FC-VI Provider is not guaranteeing the maximum bandwidth between the VI Endpoints of a connection. For this version of the FC-VI specification, this field shall always be zero.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

FCVI_MIN_BANDWIDTH: Specifies the minimum bandwidth for the VI in bytes per second. If non-zero, the FC-VI Provider is requiring the minimum bandwidth between the VI Endpoints of a connection shall be at least FCVI_MIN_BANDWIDTH bytes per second. If zero, the FC-VI Provider is not guaranteeing the minimum bandwidth between the VI Endpoints of a connection. For this version of the FC-VI specification, this field shall always be zero.

FCVI_MAX_DELAY: Specifies the maximum delay for the VI in microseconds. If non-zero, the FC-VI Provider is requiring the maximum delay between the VI Endpoints of a connection be no more than FCVI_MAX_DELAY microseconds. If zero, the FC-VI Provider is not guaranteeing the maximum delay between the VI Endpoints of a connection. For this version of the FC-VI specification, this field shall always be zero.

FCVI_PIPELINE_DEPTH: Specifies the maximum number of Messages that an originator may send without a confirmation of delivery indication from the Message responder. The pipeline depth may represent credits, command queue slots, unacknowledged messages or some other FC-VI Provider specific construct. If FCVI_PIPELINE_DEPTH is set to zero, then only one Message may be open at any one time in this VI. If FCVI_PIPELINE_DEPTH is set to some positive integer N, then at most N + 1 messages may be open at any one time on this VI.

A Message is open when the FC-VI Provider begins processing the Message Descriptor. The Message is closed when the Message Descriptor is completed by the FC-VI Provider.

7.4 FC-VI address resolution

A VI Application uses a Host Address to specify which VI Host it wishes to connect to during Connection Setup. FC-VI defines the Host Address as an IP address. The FCVI_CONNECT_RQST IU requires the N_Port Identifier for the destination FC-VI Port before a Connect Request may be originated.

FC-VI requires all FC-VI Ports to support FARP to map between IP addresses and N_Port Identifiers. All compliant implementations of FC-VI Ports shall be capable of initiating FARP-REQ to request an IP address to N_Port ID mapping for an FC-VI Port. All compliant implementations of FC-VI Ports shall recognize and respond to a FARP-REQ with a FARP-REPLY if the IP address in the FARP-REQ matches the IP address of the responding FC-VI Port. Alternative methods for FC-VI address resolution are allowed, as long as FARP is supported.

If a FC-VI Port is attached to a Fabric that supports a Name Server, the Name Server should be used in lieu of FARP to resolve IP address to N_Port Identifier mappings.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

7.5 FARP ELS

When issuing a FARP-REQ, an FC-VI Port shall

1) Set the Match Address Code Points equal to 00000100b (Match on IP address of responder).

2) Set the Responder Action equal to 02h, requesting the responder originate a FARP-REPLY to the FARP-REQ originator if the Match is successful.

3) Set the Requesting N_Port Identifier equal to the N_Port Identifier of the FARP-REQ originator.

4) Set the Responding N_Port Identifier field equal to zero.

5) Set the Requesting N_Port Port_Name equal to the Port_Name of the FARP-REQ originator.

6) Set the Responding N_Port Port_Name equal to zero.

7) Set the Requesting N_Port Node_Name equal to the Node_Name of the FARP-REQ originator.

8) Set the Responding N_Port Node_Name equal to zero.

9) Set the Requesting N_Port IP address equal to the IP address of the FARP-REQ originator.

10) Set the Responding N_Port IP address equal to the IP address of the FC-VI Port the FARP-REQ originator wishes to establish a FC-VI Connection with.

When receiving a FARP-REQ, an FC-VI Port shall perform the following three tests:

1) Compare for equality its FC-VI Port IP address to the Responding N_Port IP address in the FARP request payload

2) Test if the Match on IP address of responder code point, 00000100b, is set in the FARP-REQ Match Address Code Points

3) Test if the Responder Action in the FARP-REQ to is set to 02h, requesting the responder to originate a FARP-REPLY with the Requesting N_Port Identifier if the Match is successful.

If any test fails, the FC-VI Provider at the FARP-REQ recipient shall take no action.[3] If all three tests pass, the FC-VI Provider at the FARP-REQ recipient shall originate a FARP-REPLY to the FARP-REQ originator by formatting the FARP-REPLY payload as follows:

1) Set the Match Address Code Points equal to the Match Address Code Points in the FARP-REQ (00000100b).

2) Set the Requesting N_Port Identifier equal to the N_Port Identifier of the FARP-REQ originator.

3) Set the Responder Action equal to the Responder Action in the FARP-REQ (02h).

4) Set the Responding N_Port Identifier equal to its N_Port Identifier.

---

3. Other ULPs behind the Port, such as FC-IP, may take action on the FARP-REQ if the FC-VI Provider does not act on the FARP-REQ.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

5) Set the Requesting N_Port Port_Name equal to the Port_Name of the FARP-REQ originator.

6) Set the Responding N_Port Node_Name equal to the Port_Name of the FARP-REPLY originator.

7) Set the Requesting N_Port Node_Name equal to the Node_Name of the FARP-REQ originator.

8) Set the Responding N_Port Node_Name equal to the Node_Name of the FARP-REPLY originator.

9) Set the Requesting N_Port IP address equal to the IP address of the FARP-REQ originator.

10) Set the Responding N_Port IP address equal to the IP address of the FARP-REPLY originator.

7.6 Name server queries

An FC-VI Port that is attached to a Fabric which supports the Name Server may map N_Port Identifiers to IP addresses by issuing a GID_IPP Request to the Name Server. The Name Server is defined in FC-GS-2 and FC-GS-3. The originator of the GID_IPP Name Server Query supplies the IP address for which the N Port Identifier is sought.

The Name Server may return a list of N Port Identifiers associated with one IP address. The FC-VI Provider which receives more than one N Port Identifier to a GID_IPP request may use an implementation dependent method to select one of the N Port Identifiers to use as the destination for a Connection Setup request.

7.7 Validation of host address to N_Port identifier mappings

7.7.1 Address mapping overview

At all times, the <Host Address, N_Port Identifier> mapping shall be valid before use by an FC-VI Provider. There are many events that can invalidate this mapping. After a FC link interruption occurs, the N_Port Identifiers of all other FC-VI Ports that have previously completed PLOGI (N_Port Login) with an FC-VI Port may have changed, and its own N_Port Identifier may have changed. Because of this, validation of the address mapping is required after a LIP in a loop topology or after NOS/OLS in a point-to-point or Fabric topology. In addition, mappings may have changed due to remote link events in a Fabric topology. N_Port Identifiers shall not change as a result of Link Reset (LR), thus validation is not required.

If a logout is sent to or received from a Remote FC-VI Port, all FC-VI Connections shall be terminated with that Remote FC-VI Port.

7.7.2 Point-to-point topology

In a point-to-point topology, validation is not required after LR. NOS/OLS causes implicit Logout of the other FC-VI Port. After a NOS/OLS, each FC-VI Port shall perform a PLOGI before any communication with the other FC-VI Port may occur.

7.7.3 Private loop topology

After a LIP in a private loop topology, an FC-VI Port shall not transmit any FC-VI IU's to an FC-VI Port until the <Host Address, N_Port Identifier> mappings of the logged in FC-VI Port has been validated. The validation shall consist of completing an ADISC (sending or receiving ADISC ACC) to all logged in FC-VI Ports after a LIP. If the three addresses returned in the ACC to the ADISC - N_Port Identifier, Port_Name, Node_Name - exactly match the values of the Remote FC-VI Port prior to the LIP, then any active Exchanges may continue with that FC-VI Port. If any of the three addresses have

--`,,`,-`-`,,`,,`,`,,`---

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

changed, then the Remote FC-VI Port must be explicitly logged out. If a FC-VI Port's N_Port Identifier changes after a LIP, then all <Host Address, N_Port Identifier> mappings shall be invalidated and an explicit LOGO shall be sent to each logged in FC-VI Port.

7.7.4 Public loop topology

In a public loop topology, a FC-VI Port shall wait for the receipt of a FAN ELS after a LIP. If the Port_Name and Node_Name of the Fabric FL_Port contained in the FAN ELS response exactly match the values before the LIP, and if the AL_PA obtained by the FC-VI Port is the same as the one before the LIP, the FC-VI Port may resume all Exchanges. If not, then FLOGI (Fabric Login) shall be performed with the Fabric, all <Host Address, N_Port Identifier> mappings shall be invalidated, and all logged in FC-VI Ports shall be explicitly logged out. A public loop device shall perform private loop authentication to any nodes on the local loop which have an Area + Domain Address = 00-00-XXh by issuing an ADISC to each private FC-VI Port and performing address validation as specified for a private loop topology in 7.7.3.

7.7.5 Fabric topology

In a point-to-point topology, no validation is required after LR (link reset). After NOS/OLS, a FC-VI Port shall perform FLOGI. If, after FLOGI, the N_Port Identifier of the FC-VI Port, the Port_Name of the Fabric, and the Node_Name of the Fabric are the same as before the NOS/OLS, then the FC-VI Port may resume all Exchanges. If not, then FLOGI (Fabric Login) shall be performed with the Fabric, all <Host Address, N_Port Identifier> mappings shall be invalidated, and all logged in FC-VI Ports shall be explicitly logged out.

A Fabric attached FC-VI Port may elect to register for Registered State Change Notification as defined in FC-FS to receive notification for the loss of connectivity (link, switch or node failures) to other Fabric attached FC-VI Ports.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

8 FC-VI Error detection and recovery

8.1 FC-VI error detection and recovery overview

This Clause defines the rules for FC-VI error detection and recovery. Two sets of rules are defined. One set is defined for Message Transfer operations, while a second set is defined for Connection Setup operations. The rules support acknowledged (class 2) and unacknowledged (class 3) classes of service, as well as In-Order and Out-of-Order Fabrics.

Certain errors may cause the VI Endpoint to transition to the Error state. Once in the Error state, the FC-VI Endpoint stops all inbound and outbound traffic. However, a VI Application decides when it is appropriate to disconnect the VI Connection, according to the VI Architecture specifications.

A VipDisconnect is used by the VI Application to transition the Local VI Endpoint back to the Idle state and to signal the Disconnect to the Remote VI Endpoint. A VipDisconnect causes the FC-VI Provider to transmit a FC-VI Disconnect IU. A VipDisconnect may be issued by either the Local or Remote VI Application at any time for a given connection.

8.2 FC-VI endpoint states

The VI Architecture defines the valid states for a VI Endpoint. FC-VI Endpoints adopt these same states and adds a fifth state - Pending Connect Retry. These states are defined below.

a) Idle - FC-VI Endpoint state after a VI is created. Idle state transitions to Pending Connect state when a Connect Request is transmitted.

b) Pending Connect - FC-VI Endpoint state after a Connect Request is transmitted. If the Connection Setup succeeds, transition to the Connected state. If the VI Application times out the Connect Request or the Connect Response indicates a non-error condition (see 5.9.7.2.1, Connect Response Non-error Reason Codes), transition to the Idle state. If the Connection Setup fails due to a transport error, transition to the Pending Connect Retry state. Otherwise, transition to the Error state.

c) Pending Connect Retry - FC-VI Endpoint state for retrying failed Connect Requests. If Connection Setup succeeds, transition to Connected state. If Connection Setup does not succeed[4], transition to the Idle state. Otherwise, transition to the Error state.

d) Connected - Connect Request completed successfully. Transition to Idle state after successful Disconnect. Upon error, transition to the Error state

e) Error - A Disconnect is required to transition back to the Idle state.

8.3 FCVI_ULP_TIMEOUT definition

FC-VI Providers shall wait for a Message Response for up to FCVI_ULP_TIMEOUT.

FC-VI Providers that are originating a Connect Request shall wait for a FCVI_CONNECT_RESP3 IU for the same Exchange for up to FCVI_ULP_TIMEOUT.

---

4. The Connect Response status was No Discriminator Match, No Waiting Remote Connectionpoint, or Connect Reject.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

FC-VI Providers that have received a FCVI_CONNECT_RQST IU shall wait for a FCVI_CONNECT_RESP2 IU for the same Exchange for up to two times FCVI_ULP_TIMEOUT.

FC-VI Providers that are originating a FCVI_DISCONNECT_RESP IU shall wait for a FCVI_DISCONNECT_RESP IU for up to FCVI_ULP_TIMEOUT.

The value of FCVI_ULP_TIMEOUT shall be set to R_A_TOV.

8.4 Message transfer error detection and recovery rules

8.4.1 Message error detection

The following rules support error detection for Unreliable Delivery, Reliable Delivery, and Reliable Reception for Sends, RDMA Writes and RDMA Reads. The detected error is a Message Error. For Reliable Delivery and Reliable Reception, an FC-VI Provider shall deliver all Messages in order to the VI Application.

a) If a Message Response is required, the Message originator shall detect a Message Error if FCVI_ULP_TIMEOUT expires for the Message Response.

b) For In-Order Fabrics and Unreliable Delivery, the Message responder may detect a Message Error if any frames are received out-of-order.

c) For In-Order Fabrics and either Reliable Delivery or Reliable Reception, the Message responder shall detect a Message Error if any frames or Messages are received out-of-order.

d) For Out-Of-Order Fabrics, the Message responder shall detect a Message Error when a missing frame error is detected. See FC-FS.

Note 4: If the prior sequence within a Message was received without error, a Message responder may wait an implementation dependent amount of time to receive the next Sequence, since Fibre Channel does not require the Sequence Initiator to transmit a subsequent Sequence within any defined time period. If the next Sequence is not received before this implementation dependent amount of time has expired a Message Error shall be detected.

e) For Out-Of-Order Fabrics, a Message responder shall detect a Message Error for missing Messages.

Note 5: A Message responder should detect a missing Message by implementation dependent timeout methods, similar to Note 1 above.

8.4.2 Message transfer error recovery

The following rules are used to perform error recovery on a Local FC-VI Endpoint.

a) If the Message originator detected a FCVI_ULP_TIMEOUT on a Message response, it shall abort the Exchange associated with the Message using ABTS.

b) If the Message originator receives an ACK with Abort Sequence Condition bits set to 01b (Abort Sequence, Perform ABTS) in F_CTL, it shall abort the Exchange associated with the Message using ABTS.

c) For Unreliable Delivery, if a Message Error is detected, the FC-VI Provider may transition the Local FC-VI Endpoint to the Error state.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

d) For Unreliable Delivery, if a Message Error is detected by a Message responder and the Local FC-VI Endpoint is not transitioned to the Error state, the FC-VI Provider shall discard any frames for the FCVI_MSG_ID corresponding to the Message Error for R_A_TOV.

e) For Reliable Delivery and Reliable Reception, if a Message Error is detected, the FC-VI Provider shall transition the Local FC-VI Endpoint to the Error state.

f) If the Local FC-VI Endpoint has transitioned to the Error state, the FC-VI Provider shall discard all Message frames received with the FCVI_HANDLE of the FC-VI Connection for R_A_TOV. An FCVI_HANDLE shall not be reused for a new FC-VI Connection unless a Disconnect has been successfully completed (i.e., a Disconnect response has been received) and R_A_TOV has elapsed.

g) If an implicit or explicit Port logout is sent or received from an FC-VI Port, the FC-VI Provider may reuse any FCVI_HANDLEs and clear any Message Errors associated with the logged out FC-VI Port.

h) For class 2 service, if a Message Error is detected, the Message responder shall transmit an ACK with the Abort Sequence Condition bits set to 01b (Abort Sequence, Perform ABTS) in F_CTL.

8.5 Connection setup error detection and recovery rules

8.5.1 Connection setup error handling overview

The following rules support error detection and recovery for FC-VI Connection Setup. See Clause 5.3 for the FC-VI Connection Setup protocol description. Detected errors include a Connection Setup Exchange Error and a Connection Setup Timeout Error.

8.5.2 Connection setup error detection

a) If a Connect Request originator or responder detects any Sequence errors on an Exchange corresponding to a Connection Setup, a Connection Setup Exchange Error shall be detected.

b) If a Connect Request originator does not receive a FCVI_CONNECT_RESP3 within FCVI_ULP_TIMEOUT from transmitting a FCVI_CONNECT_RQST within the same Exchange, a Connection Setup Timeout Error shall be detected.

c) If a Connect Request responder does not receive a FCVI_CONNECT_RESP2 within two times FCVI_ULP_TIMEOUT from transmitting a FCVI_CONNECT_RESP1 in the same Exchange, a Connection Setup Timeout Error shall be detected.

8.5.3 Connection setup error recovery

a) If a Connection Setup Exchange Error or a Connection Setup Timeout Error is detected by the Connect Request originator, it shall abort the Exchange using ABTS and discard all frames received with the corresponding FCVI_CONNECTION_ID for R_A_TOV.

   1) If the Local FC-VI Endpoint is in the Pending Connect Retry state, the Endpoint shall be transitioned to the Error state.

   2) If the Local FC-VI Endpoint is in the Pending Connect state, the Endpoint shall be transitioned to the Pending Connect Retry state and the Connection Setup shall be retried.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

3) If the Local FC-VI Endpoint is in the Connect state, the Endpoint shall be transitioned to the Error state.

Note 6: If a FCVI_CONNECT_RESP2 or a a FCVI_CONNECT_RESP3 is lost, the Connect Request originator may be in the Connect state. The connection cannot be retried transparent to the VI Application at the Connect Request originator (rule a-3 in 8.5.3).

b) An FCVI_CONNECTION_ID for a failed FC-VI Connection Setup shall not be reused by a Connect Request originator unless a Disconnect has been successfully completed (i.e., a Disconnect response has been sent or received) and R_A_TOV has elapsed since the last use of the FCVI_CONNECTION_ID in a Connect Request Message.

c) If an implicit or explicit Port logout is sent or received from an FC-VI Port, an FC-VI Provider may reuse any FCVI_CONNECTION_ID associated with the logged out FC-VI Port.

d) If a Connection Setup Exchange Error or a Connection Setup Timeout Error is detected by the Connect Request responder, the Local FC-VI Endpoint shall be transitioned to the Error state.

8.5.4 Connection setup originator retry rules

The Connect Request originator shall perform a single retry for a failed Connection Setup request under the conditions defined in this subclause. If the retry fails, the Local FC-VI Endpoint is transitioned to the Error state. The following rules support Connect Request retries.

a) A Connect Request originator shall retry a previous Connection Setup if the Local FC-VI Endpoint is in the Pending Connect state by issuing a new FCVI_CONNECT_RQST on a new Exchange with the original FCVI_HANDLE, a new FCVI_CONNECTION_ID, and the RETRY bit set to one in the FCVI_FLAGS field. The RETRY bit must be set in all the IUs sent for this FCVI_CONNECTION_ID. The Endpoint state shall be transitioned to the Pending Connect Retry state.

b) A Connect Request responder that receives a FCVI_CONNECT_RQST or a FCVI_CONNECT_RESP2 with the RETRY bit set shall respond in accordance with the protocol in Clause 5.3, except that the RETRY bit shall be set in all correspondingly transmitted IUs.

8.6 Disconnect operation error detection and recovery rules

8.6.1 Disconnect operation error handling overview

The following rules support error detection and recovery for the FC-VI Disconnect operation. See Clause 5.3.5 for the FC-VI Disconnect operation protocol description. Detected errors include a Disconnect Operation Exchange Error and a Disconnect Operation Timeout Error.

8.6.2 Disconnect operation error detection

a) If a Disconnect Request originator or responder detects any Sequence errors on an Exchange corresponding to a Disconnection operation, a Disconnect Operation Exchange Error shall be detected.

b) If a Disconnect Request originator does not receive a FCVI_DISCONNECT_RESP for a FCVI_DISCONNECT_RQST in the same Exchange within FCVI_ULP_TIMEOUT a Disconnect Operation Timeout Error shall be detected.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

8.6.3 Disconnect operation error recovery rules

a) If a Disconnect Operation Exchange or Timeout Error is detected by a Disconnect Request originator, it shall abort the Exchange as specified in FC-FS, the FC-VI Endpoint shall be transitioned to the Error state and complete the VipDisconnect with a return code of VIP_NOT_REACHABLE.

b) If a Disconnect Operation Exchange Error is detected by a Disconnect Request responder, the Disconnect Request responder shall ignore the Disconnect Request.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Annex A:
(Normative)

Concurrent matching peer requests examples

A.1 Annex A overview

This annex shows Peer-to-Peer Connection Setup examples for Concurrent Matching Peer-to-Peer Requests. A Concurrent Matching Peer-to-Peer Request occurs when a Peer has issued a FCVI_CONNECT_RQST IU and receives a FCVI_CONNECT_RQST IU from another Peer and the Discriminators and IP addresses in the two Connect Requests match (the Peers are attempting to connect with each other). To prevent both Peers from completing the connection, the Peer with the lower Port_Name completes the connection while the Peer with the higher Port_Name has its Connect Request rejected.

Table A.A.1 is a version of Table 2 annotated with a “Case #” column. An example is used to illustrate each row or Case in the table. The examples assume an Out-of-Order Fabric to illustrate that the FCVI Connection Setup protocol operates correctly for Out-of-Order Fabrics.

Table A.1 – Peer B actions based on connect responses from peer A

Case #	Response from Peer A (higher Port_Name)	Peer B (lower Port_Name) Replies With:
1	Connect Accept	Connect Reject - Concurrent Matching Peer Requests
2	No Discriminator Match	Connect Accept
3	No Waiting Remote Connectionpoint	Connect Accept
4	Connect Reject - Concurrent Matching Peer Requests	Connect Reject - Protocol Error
5	Connect Reject - Transport Error	Connect Reject
6	Connect Reject - Protocol Error	Connect Reject
7	Any other response	Connect Reject - Protocol Error

For all examples shown in this annex, Peer B on the left of the figure has a lower Port_Name than Peer A on the right side of the figure. All examples eliminate FCVI_CONNECT_RESP2 and FCVI_CONNECT_RESP3 for clarity.

A.2 Case 1

Figure A.A.1 shows an example of Case 1 in Table A.A.1.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

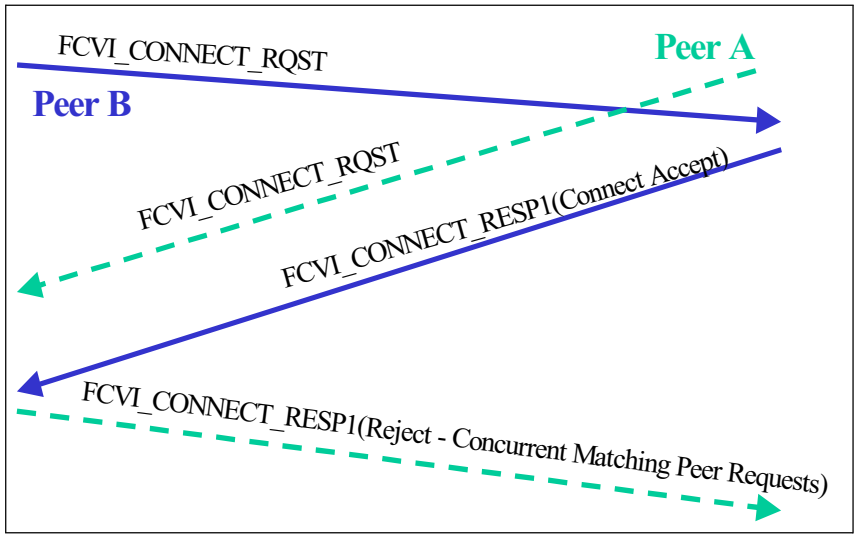


Figure A.1 – Case 1

Peer B receives a matching FCVI_CONNECT_RQST before it receives a FCVI_CONNECT_RESP1 to its own FCVI_CONNECT_RQST. Peer B waits for the FCVI_CONNECT_RESP1 from Peer A. The response of “Connect Accept” indicates Peer A has accepted Peer B’s Connect Request, since Peer A’s Port_Name is greater than Peer B’s Port_Name.

Peer B then rejects Peer A’s Connect Request by transmitting a FCVI_CONNECT_RESP1 with a CONN_STS Reason Code of “Connect Reject - Concurrent Matching Peer Requests”. Peer B successfully completes the connection to Peer A.

A.3 Case 2

Figure A.A.2 shows an example of Case 2 in Table A.A.1.

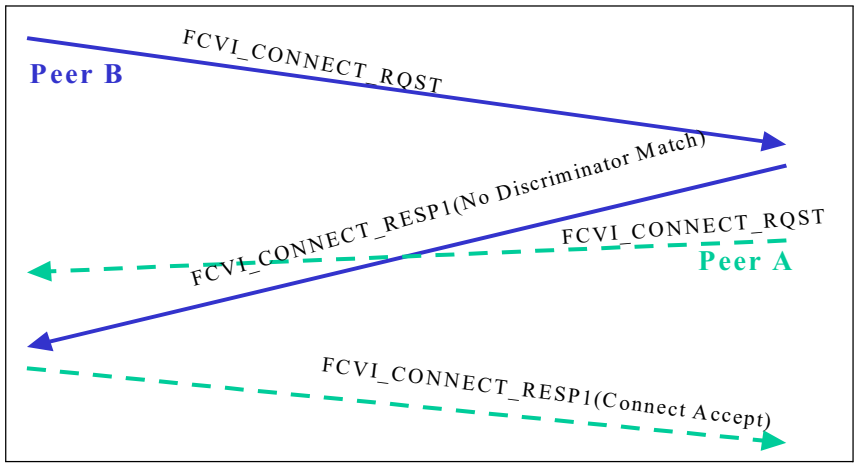


Figure A.2 – Case 2

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Peer B receives a matching FCVI_CONNECT_RQST before it receives a FCVI_CONNECT_RESP1 to its own FCVI_CONNECT_RQST. Peer B waits for the FCVI_CONNECT_RESP1 from Peer A. The response of “No Discriminator Match” signifies that Peer A’s FCVI_CONNECT_RESP1 was transmitted before Peer A’s FCVI_CONNECT_RQST was transmitted, but they were reordered in the Fabric. There is a Peer waiting, but on a different Connectionpoint set up from another FCVI_CONNECT_RQST. Peer A’s reordered FCVI_CONNECT_RQST is for the matching Connectionpoint.

Peer B accepts Peer A’s Connect Request by transmitting a FCVI_CONNECT_RESP1 with a CONN_STS set to zero, which signifies a “Connect Accept”. Peer A successfully completes the connection to Peer B. Otherwise, if Peer B rejects Peer A, neither side completes the Connection Setup.

A.4 Case 3

Figure A.A.3 shows an example of Case 3 in Table A.A.1.

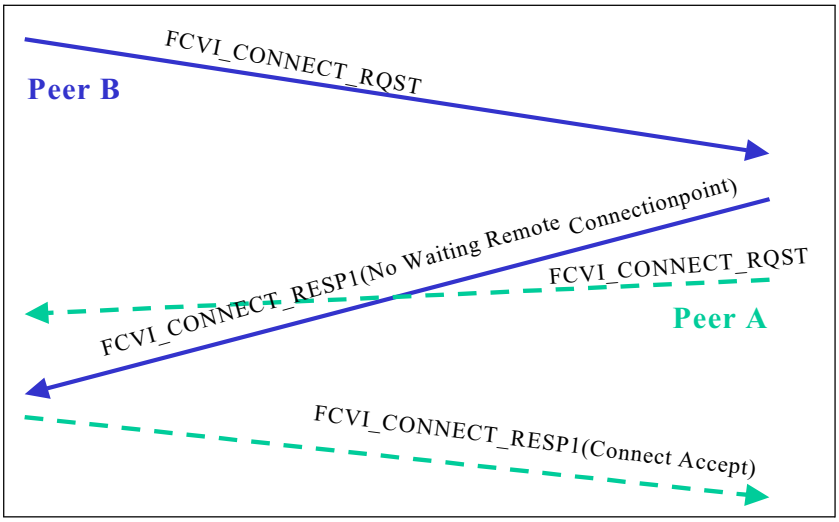


Figure A.3 – Case 3

Peer B receives a matching FCVI_CONNECT_RQST before it receives a FCVI_CONNECT_RESP1 to its own FCVI_CONNECT_RQST. Peer B waits for the FCVI_CONNECT_RESP1 from Peer A. The response of “No Waiting Remote Connectionpoint” signifies that Peer A’s FCVI_CONNECT_RESP1 was transmitted before Peer A’s FCVI_CONNECT_RQST was transmitted, but they were reordered in the Fabric. Peer A attempts a matching connection with Peer B just after it transmitted the “No Waiting Remote Connectionpoint”.
Peer B accepts Peer A’s Connect Request by transmitting a FCVI_CONNECT_RESP1 with a CONN_STS set to zero, which signifies a “Connect Accept”. Peer A successfully completes the connection to Peer B. Otherwise, if Peer B rejects Peer A, neither side completes the Connection Setup.

Case 3 is very similar to Case 2.

A.5 Case 4

Figure A.A.4 shows an example of Case 4 in Table A.A.1.

--`,,`-`-`,,`,,`,`,,`---
Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

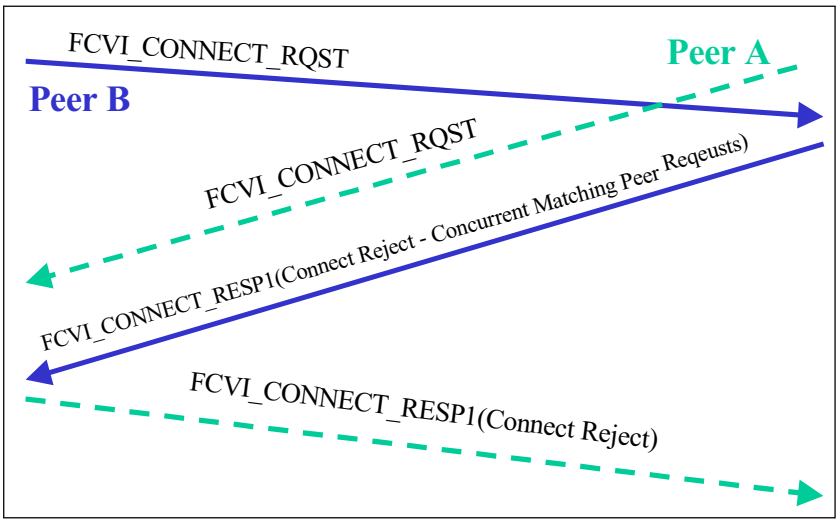


Figure A.4 – Case 4

Peer B receives a matching FCVI_CONNECT_RQST before it receives a FCVI_CONNECT_RESP1 to its own FCVI_CONNECT_RQST. Peer B waits for the FCVI_CONNECT_RESP1 from Peer A. The response of “Connect Reject - Concurrent Matching Peer Requests” signifies that Peer A believes it has the lower Port_Name, which is clearly a protocol error.

Peer B rejects Peer A’s Connect Request by transmitting a FCVI_CONNECT_RESP1 with a CONN_STS Reason Code “Connect Reject”. Neither side completes the Connection Setup.

A.6 Case 5

Figure A.A.5 shows an example of Case 7 in Table A.A.1.

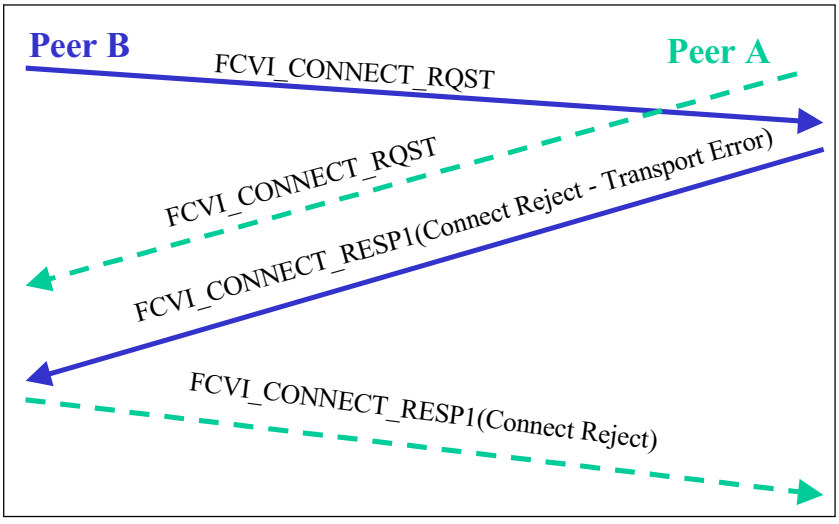


Figure A.5 – Case 5

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Peer B receives a matching FCVI_CONNECT_RQST before it receives a FCVI_CONNECT_RESP1 to its own FCVI_CONNECT_RQST. Peer B waits for the FCVI_CONNECT_RESP1 from Peer A. The response of “Connect Reject - Transport Error” signifies that Peer A has detected a transport error with the Connection Setup.

Peer B rejects Peer A’s Connect Request by transmitting a FCVI_CONNECT_RESP1 with a CONN_STS Reason Code “Connect Reject”. Neither side completes the Connection Setup.

A.7 Case 8

Figure A.A.6 shows an example of Case 8 in Table A.A.1.

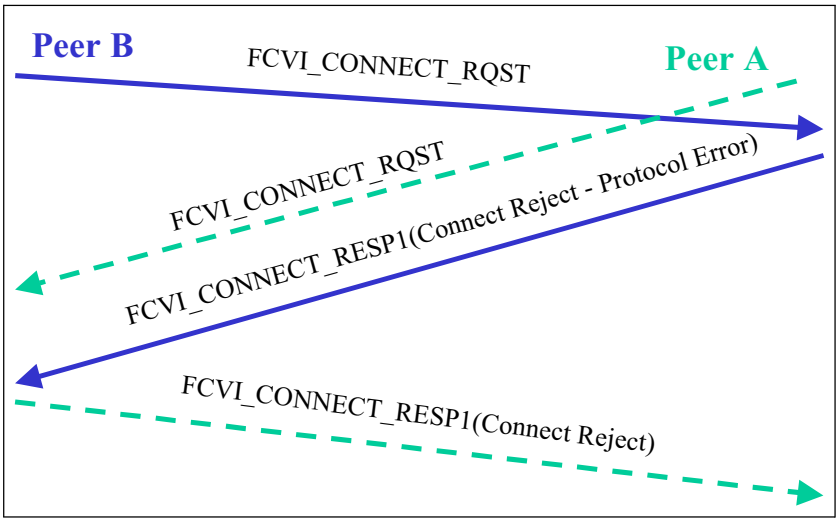


Figure A.6 – Case 6

Peer B receives a matching FCVI_CONNECT_RQST before it receives a FCVI_CONNECT_RESP1 to its own FCVI_CONNECT_RQST. Peer B waits for the FCVI_CONNECT_RESP1 from Peer A. The response of “Connect Reject - Protocol Error” signifies that Peer A is rejecting Peer B’s Connection Setup request due to a protocol error.

Peer B rejects Peer A’s Connect Request by transmitting a FCVI_CONNECT_RESP1 with a CONN_STS Reason Code “Connect Reject”. Neither side completes the Connection Setup.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Annex B:
(Informative)

FC-VI Message Transfer Error Handling Examples

B.1 Annex B Overview

This Annex shows an example of an implementation for FC-VI error detection and recovery. R_A_TOV is used to time the arrival of missing frames or Messages. Other implementations are possible.

B.2 Message Transfer Error Handling Operation

B.2.1 Message Transfer Error Handling Operation Overview

The following rules support error detection and recovery for Sends, RDMA Writes and RDMA Reads. This set of rules supports acknowledged (class 2) and unacknowledged (class 3) service, as well as In-Order Fabrics and Out-of-Order Fabrics.

B.2.2 Message Transfer Error Definitions

Define “M” as a Message identified by the FCVI_MSG_ID component of a Fully Qualified Message ID (FQMID). For a Message responder, Message M is open if at least one Message payload frame has been received and if any associated Descriptor has not been completed. In addition, Message M is considered open at the responder if no frames have arrived but M is less than Mrcv (see definition for Mrcv below). For a Message originator, Message M is open if at least one Message frame has been sent and if any associated Descriptor has not been completed.

Define "Mexp" as the next expected FCVI_MSG_ID (modulo FCVI_MSG_ID) for a FQMID. There is only be one value of Mexp per FCVI_HANDLE.

Define "Mrcv" as the FCVI_MSG_ID of the last received frame for a FQMID.

Define "SCexp" as the next expected SEQ_CNT value for each open Message M for a FQMID. For a new Message for which no frames have arrived, SCexp = 0. SCexp is incremented (modulo SEQ_CNT) after each frame is processed. For an In-Order Fabric, there is only one value for SCexp per FCVI_HANDLE, since only one Message may be open per VI at the Message responder. For an Out-of-Order Fabric, there may be many open Messages at the Message responder, where each open Message has its own value for SCexp.

Define "SCrcv" as the SEQ_CNT of the last received frame for a FQMID.

Define “Merr” as a Message assembly error for Message M identified by a FCVI_MSG_ID for a FQMID. Merr is set only for Unreliable Delivery.

Define “Herr” as a connection error for FCVI_HANDLE H for a FQMID. An Herr is set for FCVI_HANDLE H when an error is reported for Reliable Delivery or Reliable Reception. Certain catastrophic errors, such as a Protocol Error, can cause Herr to be set for Unreliable Delivery. An R_A_TOV timer is started when Herr is set.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

B.2.3 Error Detection and Recovery Rule Processing

Rules 1 through 6 are processed whenever a Device_Data frame is received. Rules 7 and 8e are processed when a timeout occurs, which is asynchronous to frame reception. If no rule evaluates to true (i.e., is processed), the received frame is processed normally.

B.2.4 Message Responder and Message Originator Error Recovery Actions

Rule 1 is processed by both a Message Originator and a Message Responder for Device_Data frames when an unrecoverable error has been detected and a Device_Data frame is received. Rule 2 may also be processed if Rule 1 is not in effect (i.e., Herr is not set for this FCVI-HANDLE). Rules 1 and 2 are used to discard either all IUs for a connection (rule 1) or to discard all IUs for a particular Message (rule 2). Rule 3 is only processed when Merr is first set for a FCVI_MESSAGE_ID. Any mechanism used to recover from errors are beyond the scope of the error detection and recovery method illustrated in this clause.

A Message Originator uses rules 1 and 2 to process received Message Response frames (FCVI_SEND_RESP, FCVI_WRITE_RESP, FCVI_READ_RESP). A Message Responder uses rules 1 and 2 to process received Message Request frames (FCVI_SEND_RQST, FCVI_WRITE_RQST, FCVI_READ_RQST).

If rule 1 is in effect, Herr can only be cleared after R_A_TOV if the Connection has been closed with an acknowledged Disconnect (either a Disconnect response has been sent or received). Otherwise, Herr will remain indefinitely set for the FCVI_HANDLE (until logout or node reset).

Rule 1: If Herr is set, then

a) Discard any received frames for FCVI_HANDLE H. If class 2 and the Message Responder, then transmit an ACK with the Abort Sequence Condition bits set to 01b (Abort Sequence, Perform ABTS) in F_CTL.

b) If a Disconnect has been successfully completed for this connection (i.e., a Disconnect response has been received) and R_A_TOV has elapsed, clear Herr

c) Else if a LOGO, either implicit or explicit, has been issued to or received from the Remote FC-VI Port, clear Herr

Rule 2: If Merr is set, then

a) Discard any received frames for Message M. If class 2 and the Message Responder, then transmit an ACK with the Abort Sequence Condition bits set to 01b (Abort Sequence, Perform ABTS) in F_CTL.

b) If a LOGO, either implicit or explicit, has been issued to or received from the Remote FC-VI Port, clear Merr

c) Else clear Merr after R_A_TOV

Rule 3 is processed only by the Message Responder when Merr is first set.

Rule 3a: If any frames for M have been received prior to setting Merr for a Send or a RDMA Write with Immediate Data, then complete the Descriptor for M in error.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Rule 3b: If no Descriptor is completed in error for Merr, an FC-VI Provider may optionally log the error in an implementation dependent manner.

B.2.5 Message responder error detection actions

Rules 4 through 7 are used to detect Message errors at the Message Responder. No ordering is implied by the rule numbering. Rules 4, 5 and 6 are designed to maintain Message ordering within a VI connection by tracking the expected Message ID (Mexp) for the next Device_Data frame. Rule 4 (Mrcv > Mexp), rule 5 (Mrcv < Mexp), and rule 6 (Mrcv = Mexp) are mutually exclusive - only one rule will be processed.

Some implementations that compare frame header fields (S_ID, SEQ_ID, SEQ_CNT, Relative Offset) to determine if the received frame is the next expected frame within the current Sequence being assembled may use this frame header comparison method as an alternative to rule 6 (Mrcv = Mexp).

Rule 4a: If Mrcv > Mexp, if In-Order Fabric, and if the Reliability Level is Unreliable Delivery, then

a) For each Message M, Mexp <= M < Mrcv, set Merr

b) If SCrcv != 0, then

   1) Set Merr for Message Mrcv

   2) Set Mexp = Mrcv + 1

   3) Set SCexp = 0

c) If SCrcv = 0, then

   1) Set Mexp = Mrcv

   2) Set SCexp = SCrcv + 1

   3) Process the received frame

Rule 4b: If Mrcv > Mexp, if In-Order Fabric, and if the Reliability Level is Reliable Delivery or Reliable Reception, then

a) Transition the VI to the Error state

b) Set Herr

Rule 4c: If Mrcv > Mexp and if Out-of-Order Fabric, then

a) For each Message M, Mexp < M < Mrcv, start an R_A_TOV timer for all potentially missing frames for M

b) For Message Mexp, start an R_A_TOV timer for all frames greater than or equal to SCexp

c) For Message Mrcv, start an R_A_TOV timer for all frames less than SCrcv

d) Set SCexp = SC + 1

e) Set Mexp = Mrcv

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Rule 5a[5]**:** If Mrcv < Mexp, if In-Order Fabric, and if Merr is not set, then

a) Transition the VI to the Error state

b) Set Herr (this is a Protocol Error)

Rule 5b: If Mrcv < Mexp and if Out-of-Order Fabric, process the frame.

Rule 6a: If Mrcv = Mexp, a Sequence error is detected and the Reliability Level is Unreliable Delivery, then

a) Set Merr for Message Mexp and process Rule 7

b) Set Mexp = Mrcv + 1

c) Set SCexp = 0

Rule 6b: If Mrcv = Mexp, a Sequence error is detected and the Reliability Level is Reliable Delivery or Reliable Reception, then

a) Transition the VI to the Error state

b) Set Herr (this is a Protocol Error)

Rule 6c: If Mrcv = Mexp, if In-Order Fabric, if SCrcv != SCexp and SCexp != 0, and if the Reliability Level is Unreliable Delivery, then

a) Set Merr for Message Mexp

b) Set Mexp = Mrcv + 1

c) Set SCexp = 0

Rule 6d: If Mrcv = Mexp, if In-Order Fabric, if SCrcv != SCexp, and if the Reliability Level is Reliable Delivery or Reliable Reception, then

a) Transition the VI to the Error state

b) Set Herr

Rule 6e: If Mrcv = Mexp, if Out-of-Order Fabric, and if SCrcv > SCexp, then

a) Start an R_A_TOV timer for all potentially missing frames greater than or equal to SCexp and less than SCrcv in Message Mexp

b) Process the received frame

c) Set SCexp = SCrcv + 1

---

5. Since the Merr rule (Rule 2) has precedence, a frame where Mrcv is less than Mexp should not occur for In-Order Fabric, i.e., a missing frame or Message was not detected.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Rule 7a: If an R_A_TOV timeout occurs for Message M and the Reliability Level is Unreliable Delivery, then

a) Set Merr for Message Mexp

Rule 7b: If an R_A_TOV timeout occurs for Message M and the Reliability Level is Reliable Delivery or Reliable Reception, then

a) Transition the VI to the Error state

b) Set Herr

B.2.6 Message originator Class 2 error detection actions

B.2.6.1 Message originator Class 2 error detection overview

Rules 8a to 8d are processed whenever a Link_Control frame is received at the Message Originator in class 2. No ordering is implied by the rule numbering. If no rule evaluates to true, the Link_Control frame is processed normally. Rule 8e is processed whenever an E_D_TOV timeout occurs, which may be asynchronous to Link_Control frame reception.

Rule 8a: If an F_BSY or P_BSY is received in response to a data frame and if an In-Order Fabric, then

a) Abort the Exchange by transmitting an ABTS

b) If Unreliable Delivery, set Merr

c) Else (for Reliable Delivery or Reliable Reception)

   1) Transition the VI to the Error state

   2) Set Herr

Rule 8b: If an F_BSY or P_BSY is received in response to a data frame and if Out-of-Order Fabric, retransmit the data frame as described in FC-FS.

Rule 8c: If an F_RJT or P_RJT is received in response to a data frame, then

a) Abort the Exchange by transmitting an ABTS

b) If Unreliable, set Merr

c) Else (for Reliable Delivery or Reliable Reception)

   1) Transition the VI to the Error state

   2) Set Herr

Rule 8d: If an ACK is received with the Abort Sequence Condition bits in F_CTL set to 01'b (Abort Sequence, Perform ABTS), then

a) Abort the Exchange by transmitting an ABTS

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

b) If Unreliable, set Merr

c) Else (for Reliable Delivery or Reliable Reception)

1) Transition the VI to the Error state

2) Set Herr.

Rule 8e: If an E_D_TOV timeout occurs waiting for an ACK, then

a) Abort the Exchange by transmitting an ABTS

b) If Unreliable, set Merr

c) Else (for Reliable Delivery or Reliable Reception)

1) Transition the VI to the Error state

2) Set Herr.

B.2.6.2 Message response timeout at message originator

The Message Originator waits for up to FCVI_ULP_TIMEOUT for a Message Response and then detects a Message error.

Rule 9: If a Message Response IU (FCVI_SEND_RESP, FCVI_WRITE_RESP, FCVI_READ_RESP) IU is not received within FCVI_ULP_TIMEOUT after Sequence Initiative has been passed on the last Message Request IU for this Message, then

a) Abort the Exchange by transmitting an ABTS

b) Transition the VI to the Error state

c) Set Herr.

B.3 Message transfer error detection and recovery examples

B.3.1 Error examples overview

This subclause illustrates examples of Message transfer error detection and recovery for the rules illustrated in Clause B.2.

B.3.2 Mrcv > Mexp error example

B.3.2.1 Mrcv > Mexp example description

Figure B.B.1 shows an example of a stream of Messages being received on one FCVI_HANDLE labeled H. The receiver last received frame SEQ_CNT 26 (SCexp-1) for Message 5 (Mexp). The next expected Message frame is SEQ_CNT 27 for Message 5.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

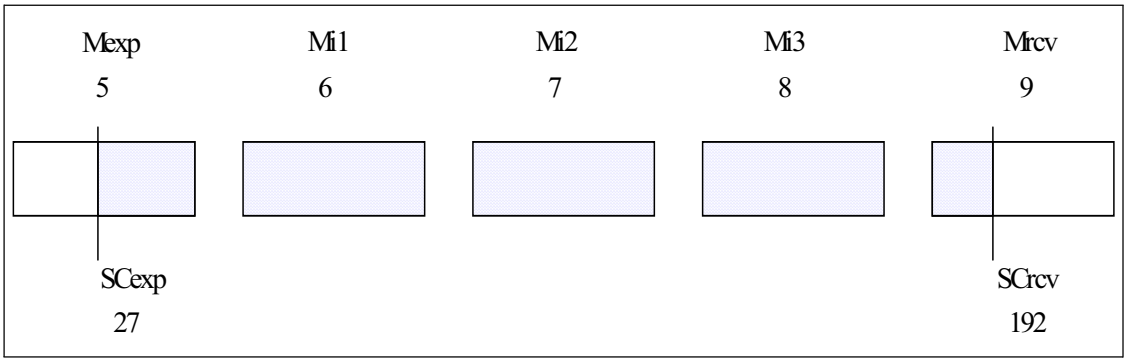


Figure B.1 – Mrcv > Mexp

When the next frame arrives, Mrcv is equal to 9, which is greater than Mexp (5), and SCrcv is equal to 192.

B.3.2.2 Mrcv > Mexp: In-order fabric and unreliable

Assume Figure B.B.1 an In-Order Fabric, class 3, a Reliability Level of Unreliable, and Message data is transmitted with a Send operation. From Annex B.2.5, Rule 4a applies.

Merr is set for Messages 5, 6, 7, and 8. Since SCrcv = 192, which is not equal to zero, Merr is also set for Message 9. The received frame is discarded. Mexp is advanced to Mrcv + 1 or 10. SCexp is set to zero, since the next expected non-errored frame is frame 0 for Message 10.

Since Merr has been set, Rule 3 from Annex B.2.5 is processed. Since frames have been received for Message 5 with a Send operation, the Descriptor for Message 5 is completed in error with a status of VIP_STATUS_TRANSPORT_ERROR. The Connection is allowed to persist, since the Reliability Level is Unreliable.

If no Descriptors were completed in error (for example, assume in the above example that Message 5 had completed and Mexp = 6, SCexp = 0), the FC-VI Provider may optionally log the error in an implementation dependent manner. There is no defined interface in the VI Architecture to signal an asynchronous error condition (VipErrorCallback) for a transport error (i.e., there is no error code of VIP_ERROR_TRANSPORT) to the VI Application.

Since Merr is set, rule 2 from B.2.5 is now in effect. Any subsequent frames that arrive for Messages 5 to 9 are discarded for a period of R_A_TOV. The receiver starts looking for frame 0 (i.e., SEQ_CNT = 0) of Message 10.

B.3.2.3 Mrcv > Mexp: In-order fabric and reliable delivery

Assume Figure B.B.1, an In-Order Fabric, class 3 and a Reliability Level of Reliable Delivery. The results would be the same if the Reliability Level was Reliable Reception. From B.2.5, Rule 4b applies.

The VI mapped to FCVI_HANDLE H is transitioned to the Error state. Herr is set for FCVI_HANDLE H. The received frame is discarded.

Since the VI is in the Error state, the VI Application should eventually transmit a VipDisconnect and tear down the connection. The Local FC-VI Provider may indicate the cause of the Disconnect to the Remote FC-VI Provider by setting CONN_STS to one with a Reason Code of “Transport Error” in the

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

FCVI_DISCONNECT_RQST IU if it is the Disconnect originator or FCVI_DISCONNECT_RESP IU if it is the Disconnect responder.

Since Herr is set, rule 1 from B.2.5 now applies. All subsequent received frames for H are discarded for a period of R_A_TOV. If a logout is issued to or received from the Remote FC-VI Port, Herr is cleared and H may be reused. Otherwise, Herr is cleared after R_A_TOV and a Disconnect has been successfully completed. If a Disconnect is not successfully completed (a Disconnect response was neither sent nor received), Herr will remain set until a logout is sent or received from the Remote FC-VI Port or the Local node is reset or power cycled.

B.3.2.4 Mrcv > Mexp: Out-of-order fabric

Assume Figure B.B.1, an Out-of-Order Fabric, class 3, and any Reliability Level. From Annex B.2.5, Rule 4c applies.

An R_A_TOV timer is started for all frames for Messages 6, 7 and 8. An R_A_TOV timer is started for all frames greater than or equal to SEQ_ID 27 (SCexp) for Message 5. An R_A_TOV timer is started for all frames less than SEQ_CNT 192 (SCrcv) for Message 9. SCexp is set equal to 193 (SCrcv + 1) and Mexp is set equal to 9 (Mrcv).

The series of frames from SEQ_CNT 27, Message 5 (SCexp, Mexp) to SEQ_ID 191, Message 9 (SCrcv - 1, Mrcv) creates an “epoch”, for which all frames in the epoch must arrive within R_A_TOV.

If any frame in the epoch arrives before R_A_TOV, Rule 4c applies and the received frame is processed.

If R_A_TOV expires and not all frames in the epoch have arrived, then Rule 7a applies for Unreliable Delivery and Rule 7b applies for Reliable Delivery or Reliable Reception.

If any frame in the epoch arrives after an R_A_TOV for a potentially missing frame, then Rule 1 (Herr is set) applies for Reliable Delivery and Reliable Reception and Rule 2 (Merr is set) applies for Unreliable Delivery.

B.3.3 Mrcv = Mexp error example

B.3.3.1 Mrcv = Mexp example description

Figure B.B.2 shows an example of frames being received for a single Message. The receiver last received frame SEQ_CNT 26 (SCexp-1) for Message 5 (Mexp) for FCVI_HANDLE H. The next expected Message frame is SEQ_CNT 27 for Message 5.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

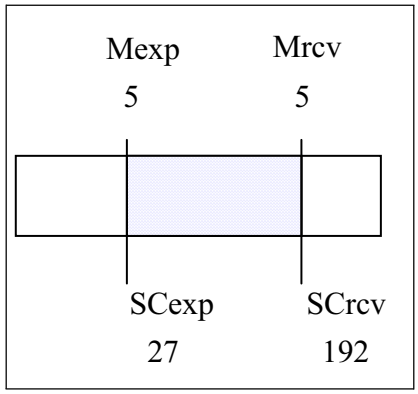


Figure B.2 – Mrcv = Mexp

The next frame to arrive is SCrcv = 192, Mrcv = 5.

B.3.3.2 Mrcv = Mexp: In-order fabric and unreliable

Assume Figure B.B.2, an In-Order Fabric, class 3, a Reliability Level of Unreliable, and Message data is transmitted with a Send operation. From B.2.5, Rule 6a or 6c is processed.

Merr is set for Message 5. The received frame is discarded. Mexp is set to 6 (Mexp + 1) and SCexp is set to zero, since next expected non-errored frame is frame 0 for Message 6. Rule 7a is processed and the Descriptor for Message M is completed in error.

Rule 2 in B.2.5 is now in effect for Message 5.

B.3.3.3 Mrcv = Mexp: Out-of-order fabric

Assume Figure B.B.2, an Out-of-Order Fabric, class 3, and any Reliability Level. From B.2.5, Rule 6e applies.

An R_A_TOV timer is started for all frames greater than or equal to SEQ_ID 27 (SCexp) and less than SEQ_CNT 192 (SCrcv) for Message 5. SCexp is set equal to 193 (SCrcv + 1) and Mexp is equal to 5 (Mrcv). The received frame is processed.

The series of frames from SEQ_CNT 27, Message 5 (SCexp, Mexp) to SEQ_ID 191, Message 5 (SCrev - 1, Mrcv) creates an “epoch”, for which all frames in the epoch must arrive within R_A_TOV.

If any frame in the epoch arrives before R_A_TOV, the frame is processed.

If R_A_TOV expires, Rule 7 applies. For Unreliable, Rule 7a applies and Merr is set for Message 5. Since Merr has been set, Rule 3a applies and the Descriptor for Message 5 is completed in error with a status of VIP_STATUS_TRANSPORT_ERROR. The Connection is allowed to persist, since the Reliability Level is Unreliable. Rule 2 is now in effect for Message M.

If R_A_TOV expires and the Reliability Level is Reliable Delivery or Reliable Reception, Rule 7b applies. The VI is transitioned to the Error state and Herr is set for FCVI_HANDLE H. Rule 1 is now in effect for FCVI_HANDLE H.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Annex C:
(Informative)

Connection Setup Error Handling Examples

C.1 Annex C overview

The following rules support error detection and recovery for FC-VI Connection Setup. This set of rules support acknowledged (class 2) and unacknowledged (class 3) service, as well as In-Order Fabrics and Out-of-Order Fabrics.

C.2 Connection setup error handling definitions

Define “CSerr” as a Connection Setup error for FCVI_CONNECTION_ID “ID”. A CSerr is set for FCVI_CONNECTION_ID ID when a Connection Setup does not successfully complete due to the following errors (see Table 12):

a) Connect Reject - Protocol Error

b) Connect Reject - Transport Error

c) Connection Setup Timeout

d) Protocol Error

e) Transport Error

C.3 Connect request originator and connect request responder rules

Rule 1: If CSerr is set, then the Local FC-VI Provider

a) Discards any received frames for FCVI_CONNECTION_ID ID

b) If a Disconnect has been successfully completed (i.e., a Disconnect response has been received) and R_A_TOV has elapsed, clears CSerr

c) Or if a LOGO, either implicit or explicit, has been issued to or received from the Remote FC-VI Port, clears CSerr

C.4 Connect request originator rules

Note 7: The Originator does not time the arrival of FCVI_CONNECT_RESP1. A lost FCVI_CONNECT_RESP1 will result in a timeout on the arrival of FCVI_CONNECT_RESP3.

Rule 2: If a FCVI_CONNECT_RESP3 is not received for a FCVI_CONNECT_RQST in the same Exchange within FCVI_ULP_TIMEOUT, and the Local FC-VI Endpoint is in the Idle state, the Local FC-VI Provider

a) Transmits an ABTS for the Exchange ID used in FCVI_CONNECT_RQST IU

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

b) Waits for the BA_ACC to ABTS for up to R_A_TOV. If the BA_ACC does not arrive within the timeout period, the FC-VI Provider may perform second level error recovery as specified in FC-FS

c) The FC-VI Provider may log a “Transport Error” in an implementation dependent manner.

Note 8: There is no supported interface defined in the VI Architecture to indicate a transport error associated with a VI Application timed out Connection Setup.

The Local FC-VI Endpoint may have transitioned from the Pending Connect to the Idle state for any of the following reasons:

a) The Local VI Application may have timed out waiting for a Connect Response

b) One of the following replies from the Remote Endpoint was received in the Connect Response

1) No waiting Remote Connectionpoint

2) No Discriminator match

3) Connect reject

The Local Endpoint may have transitioned to the Idle state from either the Pending Connect state or the Pending Connect retry state.

Rule 3: If a FCVI_CONNECT_RESP3 is not received for a FCVI_CONNECT_RQST in the same Exchange within FCVI_ULP_TIMEOUT, and the Local FC-VI Endpoint is in the Pending Connect state, the Local FC-VI Provider

a) Transmits an ABTS for the Exchange ID used in FCVI_CONNECT_RQST IU

b) Waits for the BA_ACC to ABTS for up to R_A_TOV. If the BA_ACC does not arrive within the timeout period, the FC-VI Provider may perform second level error recovery as specified in FC-FS

c) Retries the Connection Setup by transmitting the Connect Request in a new Exchange with the same FCVI_HANDLE, a new FCVI_CONNECTION_ID and the RETRY bit set to one in the FCVI_FLAGS field in the Device_Header of the Connect Request

d) Transitions the Local FC-VI Endpoint to the Pending Connect Retry state

Rule 4: If a FCVI_CONNECT_RESP3 is not received for a FCVI_CONNECT_RQST in the same Exchange within FCVI_ULP_TIMEOUT, and the Local FC-VI Endpoint is in the Pending Connect retry state, the Local FC-VI Provider

a) Transitions the VI to the Error state

b) Aborts the Exchange ID used in FCVI_CONNECT_RQST IU as defined in FC-FS

c) Sets CSerr for the FCVI_CONNECTION_ID sent in the Device_Header of the FCVI_CONNECT_RQST IU

d) Completes any Descriptors queued on the VI with an error of “Descriptor Flushed”

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

e) Returns the status (return code) of VIP_NOT_REACHABLE to the VI Application that issued the VipConnectRequest or the VipConnectPeerRequest

Rule 5: If a FCVI_CONNECT_RESP3 is not received for a FCVI_CONNECT_RQST in the same Exchange within FCVI_ULP_TIMEOUT, and the Local FC-VI Endpoint is in the Connected state, the Local FC-VI Provider

a) Transitions the VI to the Error state

b) Aborts the Exchange ID used in FCVI_CONNECT_RQST IU as defined in FC-FS

c) Sets CSerr for the FCVI_CONNECTION_ID sent in the Device_Header of the FCVI_CONNECT_RQST IU

d) Completes any Descriptors queued on the VI with an error of “Descriptor Flushed”

e) Returns the status (return code) of VIP_ERROR_CONN_LOST to the VI Application that issued the VipConnectRequest or the VipConnectPeerRequest

C.5 Connect request responder rules

Retrying a Peer-to-Peer Connection Setup is more complicated than retrying a Client-Server Connection Setup. As an example, the following situation may occur:

1) Peer 1 issues a FCVI_CONNECT_RQST IU and a FCVI_CONNECT_RESP1 IU never arrives (either IU is lost)

2) Peer 2 issues a Connect Request which to Peer 1 looks like a concurrent matching request

3) Peer 1 wins (higher Port Name) and accepts the Connect Request

4) Peer 1 times out, aborts the Exchange, and tears down the connection

The problem is tearing down the connection and starting over for a retry would be visible to the VI Application at either Endpoint. Besides, the connection is already set up. The solution is to require a Peer to ignore the timeout if a valid Peer-to-Peer connection exists. Rule 6a implements the solution.

Rule 6: If a FCVI_CONNECT_RESP2 is not received for a FCVI_CONNECT_RESP1 in the same Exchange within two times FCVI_ULP_TIMEOUT, the FC-VI Provider

a) If the Connection Setup is Peer-to-Peer and the Local VI Endpoint is in the Connected state, then

1) Complete the Connection Setup without terminating the connection

2) Optionally log a FCVI_ULP_TIMEOUT error

b) Else

1) Transition the VI to the Error state

2) Set CSerr for the FCVI_CONNECTION_ID received in the Device_Header of the FCVI_CONNECT_RQST IU.

--`,,`,,`-`-`,,`,,`,`,,`---

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

3) Complete any Descriptors queued on the VI with an error of “Descriptor Flushed”

C.5.1 Connect request responder retry rules

Rule 8: If the FC-VI Provider receives a FCVI_CONNECT_RQST IU, the RETRY bit is set in the FCVI_FLAGS field, and a Connection Setup is open for the FCVI_HANDLE indicated in the FCVI_RQST_HANDLE in the payload of the FCVI_CONNECT_RQST, the FC-VI Provider

a) Transmits the FCVI_CONNECT_RESP1, which is identical to any previously transmitted Connect Response IU for this FCVI_HANDLE except

   1) The RETRY bit in the FCVI_FLAGS field is set

   2) The Exchange ID is set to the Exchange ID of the retried FCVI_CONNECT_RQST

   3) The FCVI_CONNECTION_ID is set to the FCVI_CONNECTION_ID of the retried FCVI_CONNECT_RQST

   4) Set a Retry Flag for the FCVI_HANDLE indicated in FCVI_RQST_HANDLE

   5) Set CSerr for any previous FCVI_CONNECTION_ID for this FCVI_RQST_HANDLE

Rule 9: If the FC-VI Provider receives a FCVI_CONNECT_RQST IU, the RETRY bit is set in the FCVI_FLAGS field, and the Local Endpoint is in the Connected state, the FC-VI Provider then

a) Transmits the FCVI_CONNECT_RESP1, which is identical to any previously transmitted Connect Response IU for this FCVI_HANDLE[6] except

   1) The RETRY bit in the FCVI_FLAGS field is set

   2) The Exchange ID is set to the Exchange ID of the retried FCVI_CONNECT_RQST

   3) The FCVI_CONNECTION_ID is set to the FCVI_CONNECTION_ID of the retried FCVI_CONNECT_RQST

   4) Set a Retry Flag for the FCVI_HANDLE indicated in FCVI_RQST_HANDLE

Rule 10: If the FC-VI Provider receives a FCVI_CONNECT_RESP2 IU, the RETRY bit is set in the FCVI_FLAGS field, and the Local Endpoint is in the Connected state, the FC-VI Provider then

a) Transmits the FCVI_CONNECT_RESP3, which is identical to any previously transmitted Connect Response IU for this FCVI_HANDLE[6] except

   1) The RETRY bit in the FCVI_FLAGS field is set

   2) The Exchange ID is set to the Exchange ID of the retried FCVI_CONNECT_RESP2

   3) The FCVI_CONNECTION_ID is set to the FCVI_CONNECTION_ID of the retried FCVI_CONNECT_RESP2

---

6. Since the Connection Setup completed successfully, the CONN_STS in the FCVI_FLAGS in the FCVI_CONNECT_RESP1 is set to zero (“Connect Accept”).

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

4) Clear the Retry Flag for the FCVI_HANDLE indicated in FCVI_CONNECT_RESP2

Rule 11: If the FC-VI Provider receives a FCVI_CONNECT_RQST IU, the RETRY bit is set in the FCVI_FLAGS field, a Connection Setup is not open for the FCVI_HANDLE indicated in the FCVI_RQST_HANDLE in the payload of the FCVI_CONNECT_RQST, and a Connection does not exist for the same FCVI_HANDLE, the FC-VI Provider responds as if it is a new Connect Request, with the exception that the RETRY bit in the FCVI_FLAGS field is set for all Connect Responses for this Exchange.

Rule 12: If the FC-VI Provider is within a retried Connection Setup (Retry Flag set) for a particular FCVI_HANDLE and it receives a FCVI_CONNECT_RESP2 IU for the same FCVI_HANDLE with the correct FCVI_CONNECTION ID and Exchange ID, the FC-VI Provider responds with the appropriate FCVI_CONNECT_RESP3 IU and complete the Connection Setup at the Local Connectionpoint.

C.6 Error detection and recovery examples for connection setup

C.6.1 Overview

Figure A.C.1 shows a typical Client-Server FC-VI Connection Setup operation. For all figures in Clause C.6, solid vertical lines indicate when VI Application calls to the VI Provider complete relative to FC-VI IU transmission or reception. Dashed vertical lines indicate the duration of timer values specified in either the VI Application call or by the FC-VI Provider.

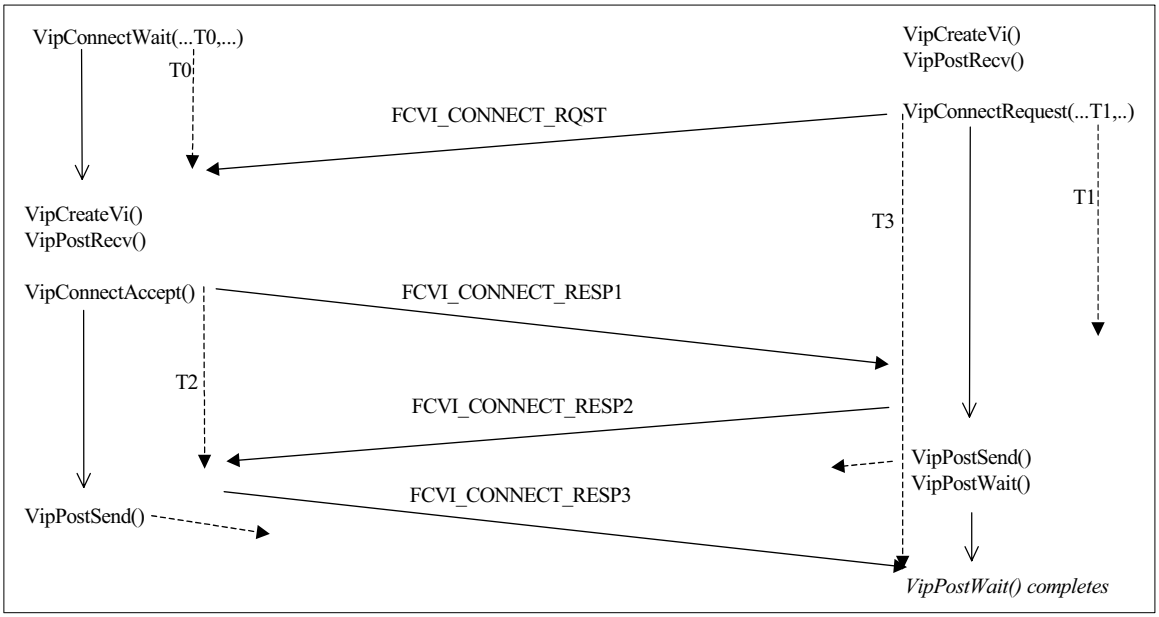


Figure C.1 – Client-server connection setup

In the example shown in Figure C.C.1, the Client on the right side of the figure creates a VI by issuing a VipCreateVI call to the VI Provider. The Client then posts a receive buffer with VipPostRecv and attempts to establish a connection by issuing a VipConnectRequest call. The Client specifies how long it is willing to wait for the Connection Setup to complete by setting the Timeout parameter of VipConnectRequest to “T1”. VipConnectRequest completes at the Client when

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

a) FCVI_CONNECT_RESP2 IU is sent

b) Timer T1 expires

c) Connection Setup has been retried without success.

The Client may issue VipPostSend call(s) after VipConnectRequest successfully completes. In this example, the Client waits for the completion of the VipPostSend by executing a VipPostWait call.

The Client FC-VI Provider should not complete any VipSendWait (or VipSendDone) calls until a FCVI_CONNECT_RESP3 is received or T3 (see definition in next subclause) expires and the Connection Setup fails. Since the Client goes to the Connected state (when FCVI_CONNECT_RESP2 is sent) before the Server (when FCVI_CONNECT_RESP3 is sent), the motivation to is throttle the Client from sending a significant number of Messages before the Server can send a Message.

In Figure C.C.1, the Server on the left side of the figure issues a VipConnectWait call to the VI Provider. The Server VI Application specifies how long it is willing to wait for a matching Connect Request by setting the Timeout parameter of VipConnectWait to “T0”. VipConnectWait completes at the Server when a FCVI_CONNECT_RQST IU is received with a matching Host Address and Discriminator (see Clause 6.7 in VI-ARCH) or T0 expires. In this particular example, the Server VI Application accepts the Connect Request and creates a VI with matching attributes with the VipCreateVI call, then queues a receive buffer with a VipPostRecv call. The Server then issues a VipConnectAccept call to accept the Connect Request. VipConnectAccept returns when a FCVI_CONNECT_RESP3 IU is sent or T2 (see definition in next subclause) expires. The Server may then issue VipPostSend calls if VipConnectAccept successfully completes.

A Peer-to-Peer Connection Setup is similar to the Client side of a Client-Server Connection Setup. One difference is VipConnectPeerRequest is non-blocking and returns immediately (i.e., before the FCVI_CONNECT_RESP1 is received). The VI Application uses VipConnectDone or VipConnectWait to determine when the initial connection phase completes. In the above example, VipConnectRequest would be replaced with the pair of calls VipConnectPeerRequest, VipConnectPeerWait. Another difference is that there is no ability for the Peer application to accept or reject an incoming Connect Request.

C.6.2 FC-VI connection setup timers

On the Server, “T0” is the Timeout parameter specified by a VI Application in the VipConnectWait call. If T0 expires, FCVI_CONNECT_RESP1 IU originated by the Server should not indicate a CONN_STS Reason Code of “Connect Timeout”, since the Server does not need to maintain state of all VipConnectWait calls that have ever timed out. A CONN_STS Reason Code of “No Waiting Remote Connectionpoint” is the correct response if the Server times out the VipConnectWait call.

On the Server, timer “T2” is set to two times FCVI_ULP_TIMEOUT and is specified by the Server FC-VI Provider. It is set when a FCVI_CONNECT_RESP1 is sent and cleared when a FCVI_CONNECT_RESP2 IU is received. If T2 expires, the Server FC-VI Providers completes the VipConnectAccept call with a return code of VIP_TIMEOUT, transition the VI to the Error state, and generates an asynchronous error notification with an Error Code of VIP_ERROR_CONN_LOST. Note that in the VI Architecture there is no mechanism to return a “transport error” for the VipConnectAccept call. The only solution is to return VIP_TIMEOUT (which may also mean the Client stopped waiting for a reply to its Connect Request) and then generates an asynchronous error notification of VIP_ERROR_CONN_LOST.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

On the Client, “T1” is the Timeout parameter specified by a VI Application in the VipConnectRequest call. If T1 expires before FCVI_CONNECT_RESP1 IU arrives, FCVI_CONNECT_RESP2 IU indicates a CONN_STS Reason Code of “Connect Timeout”. VipConnectAccept then completes with return code of VIP_TIMEOUT on the Server.

On the Client, timer “T3” is set to FCVI_ULP_TIMEOUT and is specified by the Client FC-VI Provider. It is set when FCVI_CONNECT_RQST IU is sent and cleared when FCVI_CONNECT_RESP3 IU is received. Since T1 is settable by the VI Application, its value may be greater than, less than or equal to T3. If T3 expires before T1 (the Client is still waiting for a Connect Response) and the Client is still waiting on a Connect Response (a reject or no waiting Connectionpoint was not received), the Connect Request is retried once by the FC-VI Provider. If T1 expires before T3 (the Client is no longer waiting for a Connect Response), the FC-VI Provider attempts to complete the Connection Setup handshake by waiting up to T3 seconds for the FCVI_CONNECT_RESP3 IU.

C.6.3 VipConnectRequest completion

The Client FC-VI Provider executes the following steps as a single atomic operation when completing VipConnectRequest:

1. Clear T1
2. transmit a FCVI_CONNECT_RESP2
3. complete VipConnectRequest

C.6.4 VipConnectAccept completion

The Server FC-VI Provider executes the following steps as a single atomic operation when completing VipConnectAccept:

1. Clear T2
2. transmit a FCVI_CONNECT_RESP3
3. complete VipConnectAccept

C.6.5 Enabling message transmission and reception

The Client is enabled for Message reception after a FCVI_CONNECT_RESP1 IU is received and before a FCVI_CONNECT_RESP2 IU is sent.

The Server is enabled for Message reception after a FCVI_CONNECT_RQST IU is received and before a FCVI_CONNECT_RESP1 IU is sent.

The Client is enabled for Message transmission after FCVI_CONNECT_RESP2 IU is sent and VipConnectRequest completes. Note that Message data may be sent by the Client before the FCVI_CONNECT_RESP3 IU is received. The Client FC-VI Provider does not complete any Descriptors until the Local Endpoint has transitioned to the Connected, Idle or Error states.

If T3 expires when the Local Endpoint is in the Pending Connect retry state, all posted Descriptors are completed in error. If T1 expires, the Local Endpoint transitions to the Idle state and all posted Descriptors are completed with an error of “Descriptor Flushed”.

The Server is enabled for Message transmission after a FCVI_CONNECT_RESP3 IU is sent and VipConnectAccept completes.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

C.6.6 Client timeout of VipConnectRequest

Figure C.C.2 illustrates the Client timing out the Connect Request before it completes (T1 < T3). In this case, the FCVI_CONNECT_RESP2 IU indicates a CONN_STS with a Reason Code of “Connect Timeout”. The VipConnectAccept call on the Server completes with a return code of VIP_TIMEOUT. Neither Client or Server see a completed connection and the Local Endpoint transitions to the Idle state.

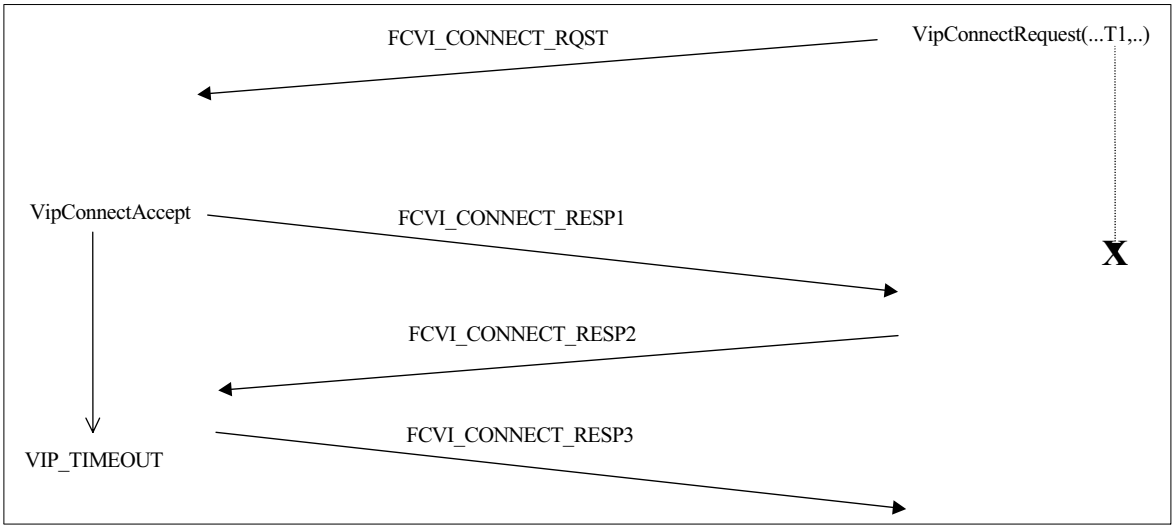


Figure C.2 – Client timeout of VipConnectRequest

C.6.7 Lost FCVI_CONNECT_RQST IU

C.6.7.1 Lost FCVI_CONNECT_RQST IU example

Figure C.C.3 shows an example of a lost FCVI_CONNECT_RQST IU.

The Client FC-VI Provider transmits a FCVI_CONNECT_RQST IU as a result of the Client VI Application issuing a VipConnectRequest call to the VI Provider.

Assume that T1 (VIP_TIMEOUT parameter in VipConnectRequest) < T3 (FCVI_ULP_TIMEOUT). The FCVI_CONNECT_RQST IU is lost and the Client VI Application times out waiting for the Connect Request to complete. VipConnectRequest completes with a return code of VIP_TIMEOUT. The Client FC-VI Provider transitions the Local Endpoint from the Connect Pending to the Idle state.

The Client FC-VI Provider waits for T3 seconds for the FCVI_CONNECT_RESP3 IU to arrive and times out. The Client FC-VI Provider then transmits an ABTS to recover the Exchange ID used for the Connection Setup. The Server FC-VI Provider has no Exchange or Connection Setup context, since it never received the FCVI_CONNECT_RQST IU. The Server FC-VI Provider replies by transmitting a BA_ACC (see “Special case - new Exchange” in 23.6.1.1 of FC-FS). The Client FC-VI Provider may optionally log a transport error.

The Client FC-VI Provider will not attempt a retry, since the Local VI Application is responsible for issuing a new Connect Request.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

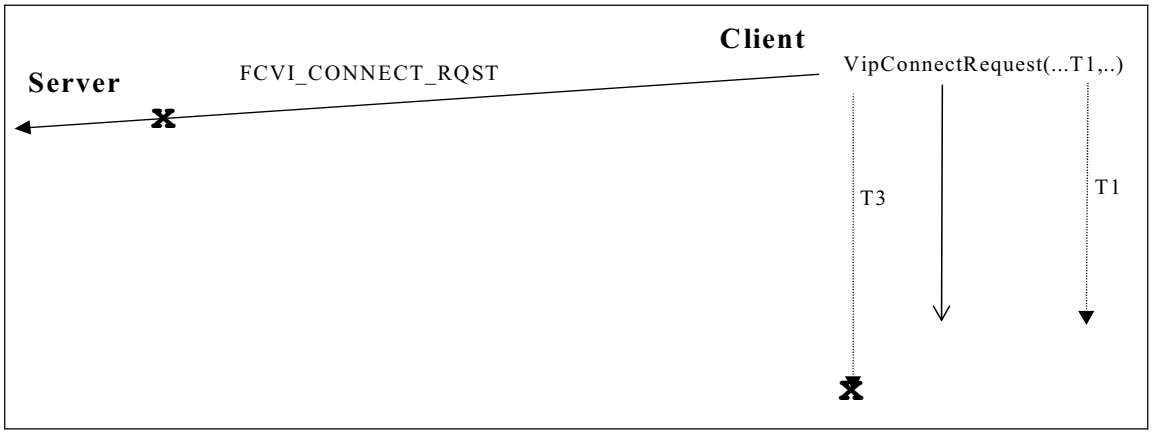


Figure C.3 – Lost FCVI_CONNECT_RQST IU

C.6.7.2 Retried connection setup

If T1 has not expired and T3 expires, the Client FC-VI Provider retries the Connection Setup and retransmits the FCVI_CONNECT_RQST IU. The RETRY bit is set to one in all Connection Setup IUs.

If the retried Connection Setup fails (T3 expires waiting for the retried FCVI_CONNECT_RESP3) and T1 has not expired, the Client FC-VI Provider transitions the Local Endpoint from the Connect Pending retry state to the Error state, clears the T1 timer and complete the VipConnectRequest with a return code of VIP_NOT_REACHABLE. The Client FC-VI Provider transmits an ABTS to recover the Exchange ID used for the retried Connection Setup.

If the Client VI Application issues a VipDisconnect[7] call to the VI Provider, the Client FC-VI Provider transmits a FCVI_DISCONNECT_RQST IU with the CONN_SETUP_ABORT flag set, the VIP_APP_DISCON flag set and the CONN_STS set in the FCVI_FLAGS field. The Reason Code for the error is set to “Connection Setup Timeout”. The FCVI_HANDLE in the Device_Header of the FCVI_DISCONNECT_RQST IU is set equal to the FCVI_RQST_HANDLE sent in the FCVI_CONNECT_RQST IU. The FCVI_CONNECTION_ID in the Device_Header is set equal to the FCVI_CONNECTION_ID in the Device_Header of the FCVI_CONNECT_RQST IU of the retried Connection Setup.

The Server FC-VI Provider responds by transmitting a FCVI_DISCONNECT_RESP IU with the CONN_SETUP_ABORT and VIP_APP_DISCON flags set equal to those in the FCVI_CONNECT_RQST IU. The CONN_STS set in the FCVI_FLAGS field. The Reason Code for the error is set to “Connection Does Not Exist”. The FCVI_HANDLE is set to FFFFFFFFh. The FCVI_CONNECTION_ID in the Device_Header of the FCVI_DISCONNECT_RESP IU is set equal to the FCVI_CONNECTION_ID in the Device_Header of the FCVI_DISCONNECT_RQST IU received on the same Exchange.

If the FCVI_DISCONNECT_RESP IU is received within FCVI_ULP_TIMEOUT seconds, the VipDisconnect call completes with a return code of VIP_SUCCESS. If the FCVI_DISCONNECT_RESP IU is not received within FCVI_ULP_TIMEOUT seconds, the Client FC-VI Provider may retry the Dis-

7. If the VI Application never issues a Disconnect, the VI remains in the Error state until the VI Application terminates or a VipCloseNIC call is issued.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

connect. If the retried FCVI_DISCONNECT_RESP IU is not received, the VipDisconnect call completes with a return code of VIP_NOT_REACHABLE and the Local Endpoint is not transitioned from the Error state.

Whenever a FCVI_ULP_TIMEOUT occurs waiting for a FCVI_DISCONNECT_RESP IU, the Client FC-VI Provider recovers the Exchange ID by transmitting an ABTS.

C.6.8 Lost FCVI_CONNECT_RESP1 IU

Figure C.C.4 shows an example of a lost FCVI_CONNECT_RESP1 IU.

The Client FC-VI Provider transmits a FCVI_CONNECT_RQST IU as a result of the Client VI Application issuing a VipConnectRequest call to the VI Provider. Assume that T1 (VIP_TIMEOUT parameter in VipConnectRequest) < T3 (FCVI_ULP_TIMEOUT). The Server FC-VI Provider responds by issuing a FCVI_CONNECT_RESP1 IU.

The FCVI_CONNECT_RESP1 IU is lost and the Client VI Application times out waiting T1 seconds for the Connect Request to complete. VipConnectRequest completes with a return code of VIP_TIMEOUT. The Client FC-VI Provider transitions the Local Endpoint from the Connect Pending to the Idle state.

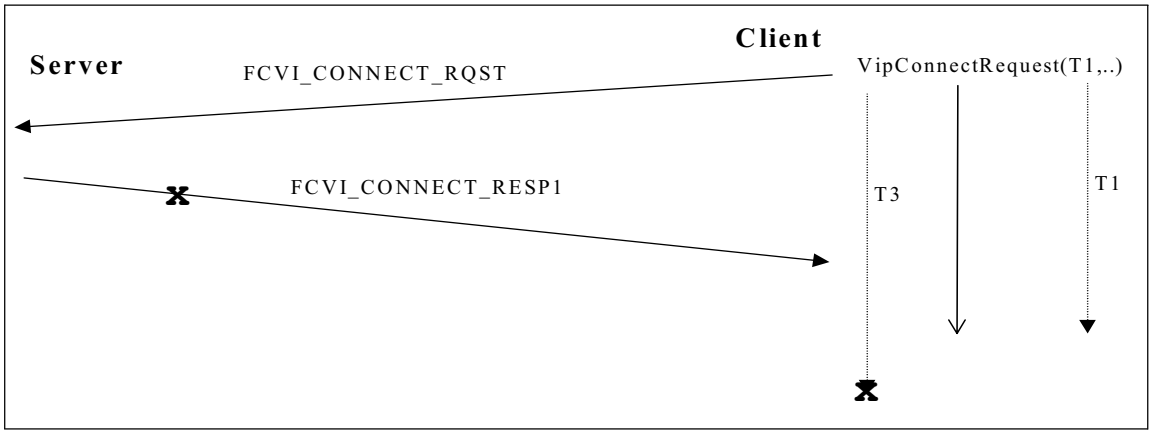


Figure C.4 – Lost FCVI_CONNECT_RESP1 IU

Since the Client cannot tell the difference between a lost FCVI_CONNECT_RQST and a lost FCVI_CONNECT_RESP1, the behavior is identical to the lost FCVI_CONNECT_RQST example described in the previous section, with the exception that the Server FC-VI Provider has Connection Setup context and will reply to a Disconnect request with a Reason Code other than “Connection Does Not Exist”.

C.6.9 Lost FCVI_CONNECT_RESP2 IU

C.6.9.1 Lost FCVI_CONNECT_RESP2 IU example

Figure C.C.5 shows an example of a lost FCVI_CONNECT_RESP2 IU.

--`,,`-`-`,,`,,`,`,,`---

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

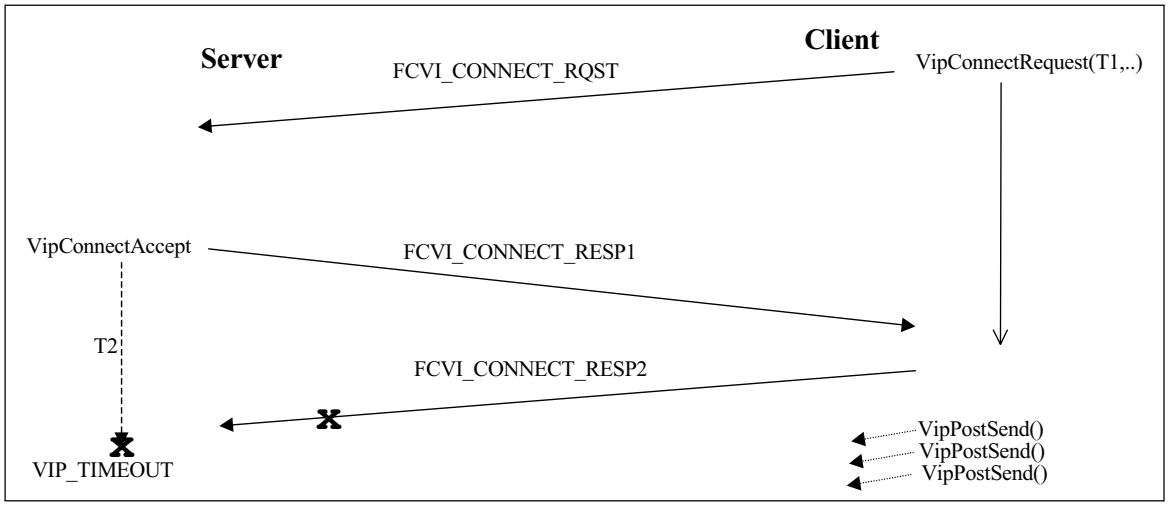


Figure C.5 – Lost FCVI_CONNECT_RESP2

The Client FC-VI Provider transmits the FCVI_CONNECT_RESP2 and completes VipConnectRequest.

Assume the FCVI_CONNECT_RESP1 indicates a “Connect Accept”. VipConnectRequest returns with a Return Code of VIP_SUCCESS. The Client Endpoint transitions to the Connected state. The Client VI Application may then transmit Messages by issuing VipPostSend calls.

The Client FC-VI Provider waits for T3 seconds for the FCVI_CONNECT_RESP3 IU to arrive and times out. The Client FC-VI Provider then transmits an ABTS to recover the Exchange ID used for the Connection Setup. The Client Endpoint is transitioned to the Error state. An asynchronous error notification of VIP_ERROR_CONN_LOST is generated by the Client FC-VI Provider. A Herr is set for the Local FCVI_HANDLE and frames are discarded for up to R_A_TOV.

If the Client VI Application issues a VipDisconnect call to the VI Provider, the Client FC-VI Provider transmits a FCVI_DISCONNECT_RQST IU with the CONN_SETUP_ABORT flag clear, the VIP_APP_DISCON flag set and the CONN_STS set in the FCVI_FLAGS field. The Reason Code for the error is set to “Transport Error”. The FCVI_HANDLE in the Device_Header of the FCVI_DISCONNECT_RQST IU is set equal to the FCVI_HANDLE for the connection. The FCVI_CONNECTION_ID in the Device_Header is set equal to FFFFFFFFh.

The Server FC-VI Provider responds by transmitting a FCVI_DISCONNECT_RESP IU with the CONN_SETUP_ABORT and VIP_APP_DISCON flags set equal to those received in the FCVI_DISCONNECT_RQST IU. The FCVI_CONNECTION_ID in the Device_Header of the FCVI_DISCONNECT_RESP IU is set equal to the FCVI_CONNECTION_ID in the Device_Header of the FCVI_DISCONNECT_RQST IU received on the same Exchange. A Herr is set for the Local FCVI_HANDLE sent as the FCVI_RESP_HANDLE and frames are discarded for up to R_A_TOV.

If the FCVI_DISCONNECT_RESP IU is received by the Client within FCVI_ULP_TIMEOUT seconds, the VipDisconnect call completes with a return code of VIP_SUCCESS. If the FCVI_DISCONNECT_RESP IU is not received within FCVI_ULP_TIMEOUT seconds, the Client FC-VI Provider may retry the Disconnect. If the retried FCVI_DISCONNECT_RESP IU is not received, the VipDisconnect call completes with a return code of VIP_NOT_REACHABLE and the state of the Client Endpoint is not changed from the Error state. Whenever a FCVI_ULP_TIMEOUT occurs wait-

--`,`-`-`,,`,,`,`,,`---

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

ing for a FCVI_DISCONNECT_RESP IU, the Client FC-VI Provider recovers the Exchange ID by transmitting an ABTS.

C.6.9.2 Server timing out connection setup

If FCVI_DISCONNECT_RQST IU arrives before T2 expires at the Server, T2 is cleared and the Server Endpoint is transitioned to the Error state. Since the Client was in the Connected state, the Disconnect request will indicate that a connection is being aborted (CONN_SETUP_ABORT flag clear). A Herr is set for the Local FCVI_HANDLE sent as the FCVI_RESP_HANDLE and frames are discarded for up to R_A_TOV.

If T2 expires before FCVI_DISCONNECT_RQST IU arrives, the Server Endpoint will be transitioned to the Error state and CSerr will be set for the FCVI_CONNECTION_ID received in the FCVI_CONNECT_RQST IU. Frames will be discarded for FCVI_CONNECTION_ID for up to R_A_TOV.

The Server VI Application may issue a VipDisconnect, possibly before the Client VI Application. Each Endpoint must issue a VipDisconnect to transition the Local Endpoint from the Error state to the Idle state.

C.6.10 Lost FCVI_CONNECT_RESP3 IU

C.6.10.1 Lost FCVI_CONNECT_RESP3 IU example

Figure C.C.6 shows an example of a lost FCVI_CONNECT_RESP3 IU.

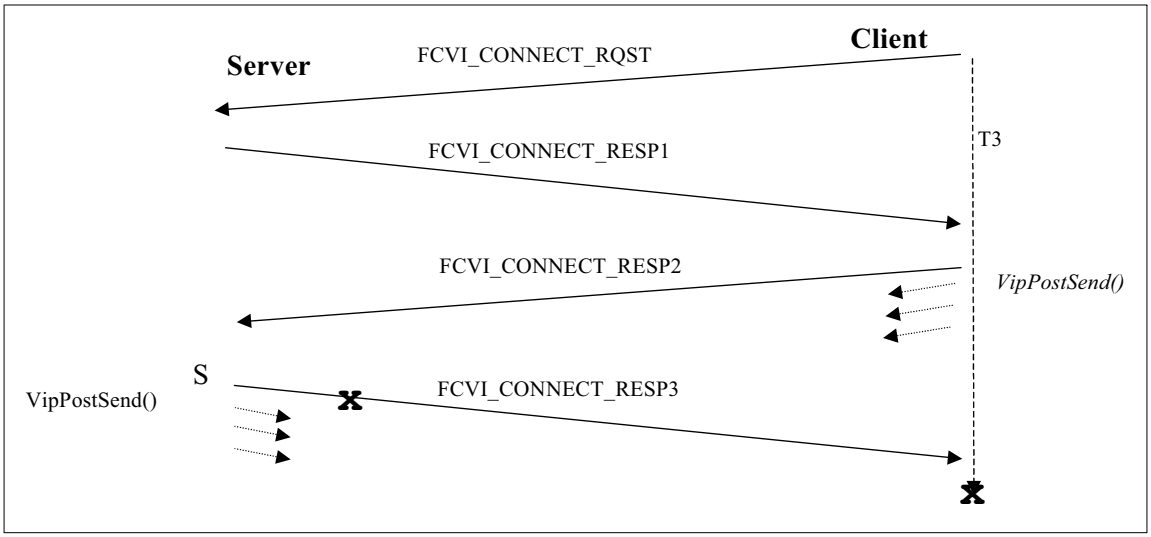


Figure C.6 – Lost FCVI_CONNECT_RESP3 IU

The Server FC-VI Provider transmits a FCVI_CONNECT_RESP3 IU and the VipConnectAccept completes with a return code of VIP_SUCCESS. The Server VI Application may then issue one or more VipPostSend calls to transfer Messages to the Client. Since the FCVI_CONNECT_RESP3 IU has been sent, the Server FC-VI Provider may successfully complete Send Descriptors for Unreliable Delivery or Reliable Delivery. The Client FC-VI Provider does not issue any Message Response IUs until Connection Setup has successfully completed (FCVI_CONNECT_RESP3 arrives).

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

The Client may transmit Message data as soon as the FCVI_CONNECT_RESP2 IU is sent. However, the Descriptors may not be completed until the FCVI_CONNECT_RESP3 IU arrives. The Client may receive Message Response IUs from the Server.

When T3 expires at the Client, the Client Endpoint will be transitioned to the Error state and an Herr will be set for the Client FCVI_HANDLE. When the Client VI Application issues a VipDisconnect call, the receipt of a FCVI_DISCONNECT_RQST IU will transition the Server Endpoint from the Connected to the Error state. The Server VI Application must then issue its own VipDisconnect to transition the Server Endpoint from the Error state to the Idle state.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Annex D:
(Informative)

Disconnect operation error handling examples

D.1 Disconnect operation example description

The following rules support error detection and recovery for FC-VI Disconnect operation. This set of rules support acknowledged (class 2) and unacknowledged (class 3) service, as well as In-Order Fabrics and Out-of-Order Fabrics.

Rule 1: If a VI Application issues a VipDisconnect call and an established connection is being aborted, the FC-VI Provider

a) Transmits a FCVI_DISCONNECT_RQST IU with FCVI_APP_DISCON set to one and the CONN_SETUP_ABORT flag set to zero. If the FC-VI Provider can provide information as to the cause of the connection abort, it may set CONN_STS to one and set the Reason Code appropriately (see Table 13). The FCVI_HANDLE in the Device_Header is set equal to the FCVI_HANDLE of the Remote Endpoint. The FCVI_MSG_ID in the Device_Header is set equal to the highest Message ID successfully completed by the Disconnect request originator. The FCVI_CONNECTION_ID in the Device_Header is set to a value of FFFFFFFFh.

b) Waits for the FCVI_DISCONNECT_RESP IU up to FCVI_ULP_TIMEOUT. If a FCVI_DISCONNECT_RESP IU is not received within the timeout period, the FC-VI Provider transmits an ABTS for the Exchange used for the FCVI_DISCONNECT_RQST IU. The FC-VI Provider may retry the FCVI_DISCONNECT_RQST IU.

c) If a FCVI_DISCONNECT_RESP IU is never received, leave the Local Endpoint in the Error state and complete the VipDisconnect with a return code of VIP_NOT_REACHABLE. Else, transition the Local Endpoint to the Idle state and complete the VipDisconnect with a return code of VIP_SUCCESS.

Rule 2: If a VI Application issues a VipDisconnect call as the result of a timeout during a Connection Setup (FCVI_CONNECT_RESP1 was never received and the Client did not time out the Connect Request), the FC-VI Provider

a) Transmits a FCVI_DISCONNECT_RQST IU with FCVI_APP_DISCON set to one, the CONN_SETUP_ABORT flag set to one, and CONN_STS set to one in the FCVI_FLAGS field with a Reason Code of “Connection Setup Timeout”. Set the FCVI_HANDLE equal to FFFFFFFFh. Set the FCVI_CONNECTION_ID in the Device_Header set to the FCVI_CONNECTION_ID sent in the FCVI_CONNECT_RQST IU.

b) Wait for the FCVI_DISCONNECT_RESP IU up to FCVI_ULP_TIMEOUT. If a FCVI_DISCONNECT_RESP IU is not received within the timeout period, the FC-VI Provider transmits an ABTS for the Exchange used for the FCVI_DISCONNECT_RESP IU. The FC-VI Provider may retry the FCVI_DISCONNECT_RQST IU.

c) Transition the VI to the Idle state if a FCVI_DISCONNECT_RESP IU is received.

Rule 3: If an FC-VI Provider receives a FCVI_DISCONNECT_RQST and the FCVI_HANDLE is not defined and the FCVI_CONNECTION_ID is unknown, the FC-VI Provider

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

a) Transmits a FCVI_DISCONNECT_RESP with the FCVI_HANDLE set to FFFFFFFFh in the Device_Header, the FCVI_CONNECTION_ID set to FFFFFFFFh in the Device_Header, the FCVI_APP_DISCON and CONN_SETUP_ABORT flags set equal to the value received in the FCVI_DISCONNECT_RQST, and the CONN_STS set to one in the FCVI_FLAGS field with a Reason Code of “Connection Does Not Exist”.

Rule 6: If an FC-VI Provider receives a FCVI_DISCONNECT_RQST, the FCVI_HANDLE is defined, the FCVI_CONNECTION_ID is unknown, and the CONN_SETUP_ABORT flag is set, the FC-VI Provider

a) Transmits a FCVI_DISCONNECT_RESP with the FCVI_HANDLE set to the FCVI_HANDLE of the Disconnect request originator, the FCVI_APP_DISCON and CONN_SETUP_ABORT flags set equal to the value received in the FCVI_DISCONNECT_RQST, and the CONN_STS set to one in the FCVI_FLAGS field with a Reason Code of “Protocol Error”.

D.2 FC-VI disconnect operation example

Figure D.D.1 illustrates a typical Disconnect of an established Connection between the two Endpoints of a FC-VI Connection. Assume the left Endpoint is the Local Endpoint and the right Endpoint is the Remote Endpoint. Assume that an error has occurred and the Local Endpoint has transitioned to the Error state.

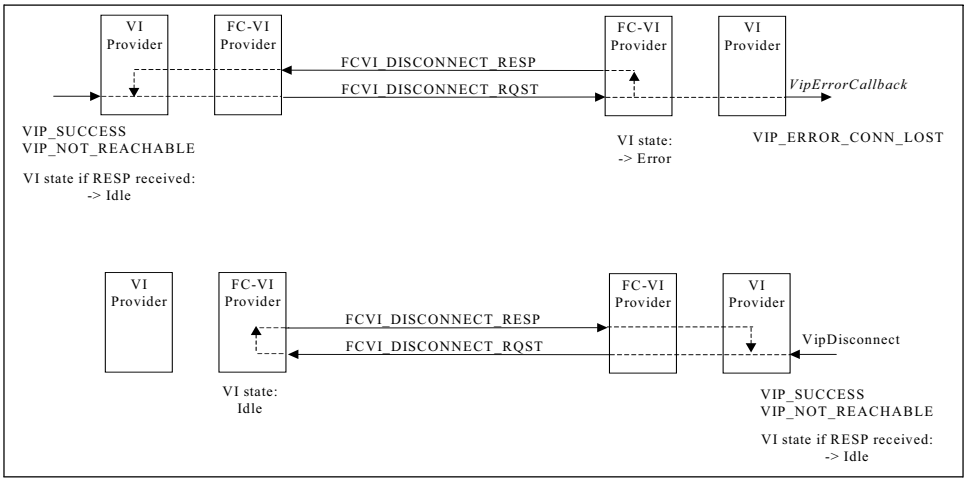


Figure D.1 – FC-VI disconnect operation

In the top of Figure D.D.1, the Local VI Application issues a VipDisconnect to the VI Provider, which causes the FC-VI Provider to transmit a FCVI_DISCONNECT_RQST IU.

The receipt of the FCVI_DISCONNECT_RQST IU by the Remote FC-VI Provider causes a VipErrorCallback with an error code of VIP_ERROR_CONN_LOST to be issued to the Remote VI Application. The Remote Endpoint is transitioned to the Error state.

The Remote FC-VI Provider transmits a FCVI_DISCONNECT_RESP IU. The VipDisconnect completes at the Local Endpoint with a return code of VIP_SUCCESS after receipt of the FCVI_DISCONNECT_RESP IU. If a FCVI_DISCONNECT_RESP never arrives, the VipDisconnect

--`,,`-`-`,,`,,`,`,,`---
Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

completes with a return code of VIP_NOT_REACHABLE. The Local Endpoint remains in the Error State if the FCVI_DISCONNECT_RESP IU never arrives.

At some point the Remote VI Application should issue a VipDisconnect to transition its Endpoint to the Idle state, as shown in the bottom of Figure D.1. This causes the Remote FC-VI Provider to transmit a FCVI_DISCONNECT_RQST IU. The Local FC-VI Provider replies with a FCVI_DISCONNECT_RESP IU. Since the Local Endpoint is already in the Idle state, no VipErrorCallback is issued to the Local VI Application. The receipt of a FCVI_DISCONNECT_RESP IU transitions the Local Endpoint to the Idle state.

Under different circumstances, both the Local and Remote Endpoints may issue a VipDisconnect at about the same time. Each Endpoint would already be in the Error State when the FCVI_DISCONNECT_RESP IU arrives, and no VipErrorCallback is issued on either Endpoint.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Annex E:
(Informative)

Message streaming for reliable reception

Similar to Unreliable Delivery and Reliable Delivery, Messages may be streamed for Reliable Reception. A Message is considered streamed when subsequent Messages have begun processing[8] before all prior Messages have been completed on a VI.

The failure semantics of Reliable Reception, as defined by the VI Architecture, dictate that if a failure occurs on a particular Message, no subsequent Descriptors will be successfully completed for that VI. This allows subsequent Descriptors to be processed by the FC-VI Provider, but they may not be completed until all previous Descriptors have been completed for that VI. If a failure occurs on a given Descriptor, all subsequent Descriptors will be completed in error and the VI connection will be terminated for that VI.

In a hypothetical error-free transport, if an error occurs in processing the Message at the responder, both the originator Descriptor and responder Descriptor for Message M will complete in error. All Descriptors for Messages less than M will have completed successfully on both the originator and responder, while all Message Descriptors greater than or equal to M will complete in error on both the originator and responder. Thus, the Message originator and responder remain in lock step with respect to Descriptor completions. This facilitates error recovery in a distributed environment such as SANs, since the originator knows exactly where the responder has failed and that no subsequent steps (i.e., Descriptors) is completed.

However, transport errors such as lost or corrupt frames may result in the Message responder successfully completing more Descriptors than the Message originator. For example, if a FCVI_RESP IU is lost, the responder may have successfully completed a Descriptor while the originator will complete the Descriptor in error for the same Message. If Messages are streamed to some depth N (i.e., N open Messages), then the responder may complete N more Descriptors than the originator for a VI if multiple transport errors occur.

A VI Application may limit the streaming depth by setting the FCVI_PIPELINE_DEPTH to a value N that limits the number of open Messages for a VI to no more than N Messages. The FCVI_PIPELINE_DEPTH is an attribute in the FCVI_QOS structure in the FCVI_ATTRIBUTES (see Table 18). This guarantees that, even in the presence of multiple transport errors, the responder can complete at most N more Descriptors than the originator.

Both the Disconnect request and the Disconnect response contain the FCVI_MSG_ID of the last Message successfully completed by the Remote Endpoint. A VI Application may use this value to determine at what point the Remote VI Application stopped processing a chain of Descriptors on a VI that terminated due to error. How a VI Application access this information is FC-VI Provider specific. There is currently no defined interface in the VI Architecture to accesses this information.

8. For example, some Message data has been transferred.

--`,,`,,`-`-`,,`,,`,`,,`---

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Annex F:
(Informative)

Enabling Message Transmission in the FC-VI NIC

This annex describes a mechanism to enable Message transmission in the FC-VI NIC during Connection Setup without the direct involvement of the FC-VI Connection Manager in the critical path.

The Connection Manager is typically implemented as a software module within the Kernel Agent. In the implementation model used in this Annex, the Connection Manager receives and processes all FC-VI Connection Setup IUs. On the Client, Message completion is enabled by the Connection Manager after it has received a FCVI_CONNECT_RESP3 IU, while Message reception is enabled before the FCVI_CONNECT_RESP2 IU is sent (see Figure A.C.1). Since the Server may send Messages immediately after a FCVI_CONNECT_RESP3 IU is sent, Client receives must be enabled before the FCVI_CONNECT_RESP3 IU is received at the Client. There is a race between the Client enabling Message reception and the Server sending Messages. The only way to avoid this race is to enable receives at the Client before the “Clear-To-Send” (CTS) indication is given to the Server, so receives must be enabled before the Client transmits the FCVI_CONNECT_RESP2 IU (the “CTS”).

However, this creates a potential problem if the FCVI_CONNECT_RESP3 IU never arrives. If this occurs, the Server may transmit Messages for up to FCVI_ULP_TIMEOUT before the Client times out on the Connection Setup and aborts the partial connection. The root of the problem is the Server sees the connection as established while the Client is waiting for the Connection Setup to complete.

Any connection protocol is subject to one Endpoint believing the connection is established while the other is waiting for the connection to complete if the last Connection Setup IU is lost. The Endpoint that transmitted the last IU believes the connection is established and is clear to send, while the Endpoint that is waiting for receipt of the last IU does not observe the connection as established and is prohibited from completing sends.

The mechanism proposed in this subclause significantly reduces the time during which a partially established connection may operate. For the Client, the FCVI_CONNECT_RESP3 IU is used as the Clear-To-Send (CTS) signal by the FC-VI NIC. The FCVI_CONNECT_RESP3 IU is received by the FC-VI NIC as the first “Message” on the VI. However, it is routed to the receive queue for the Connection Manager instead of the receive queue for the VI. The first real received Message on the VI is routed to the correct receive context for the VI. The FCVI_CONNECT_RESP3 IU has a FCVI_MSG_ID of zero and the first Message has a FCVI_MSG_ID of one. For an In-Order Fabric, the FC-VI NIC immediately detects a lost Message if the FCVI_CONNECT_RESP3 IU is lost. For an Out-of-Order Fabric, the FC-VI NIC detects a lost Message within R_A_TOV.

The successful receipt of CTS informs the Client FC-VI NIC that it may complete any Descriptors that have been processed. On the Server side, the FCVI_CONNECT_RESP2 IU cannot effectively be used as a CTS signal by the Server FC-VI NIC. The Server Connection Manager must create and transmit a FCVI_CONNECT_RESP3 IU and complete the VipConnectAccept before the Server transmits any Messages, so it must be involved in the receive path for FCVI_CONNECT_RESP3 IU.

Copyright International Organization for Standardization
Reproduced by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale