

INTERNATIONAL STANDARD

ISO/IEC
14165-331

First edition
2007-07

Information technology – Fibre channel – Part 331: Virtual interface (FC-VI)



Reference number
ISO/IEC 14165-331:2007(E)



THIS PUBLICATION IS COPYRIGHT PROTECTED

Copyright © 2007 ISO/IEC, Geneva, Switzerland

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either IEC or IEC's member National Committee in the country of the requester.

If you have any questions about ISO/IEC copyright or have an enquiry about obtaining additional rights to this publication, please contact the address below or your local IEC member National Committee for further information.

IEC Central Office
3, rue de Varembe
CH-1211 Geneva 20
Switzerland
Email: inmail@iec.ch
Web: www.iec.ch

About the IEC

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

About IEC publications

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

- Catalogue of IEC publications: www.iec.ch/searchpub

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 It also gives information on projects, withdrawn and replaced publications.

- IEC Just Published: www.iec.ch/online_news/justpub

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

- Customer Service Centre: www.iec.ch/webstore/custserv

If you wish to give us your feedback on this publication or need further assistance, please visit the Customer Service Centre FAQ or contact us:

Email: csc@iec.ch
Tel.: +41 22 919 02 11
Fax: +41 22 919 03 00

INTERNATIONAL STANDARD

ISO/IEC 14165-331

First edition
2007-07

Information technology – Fibre channel – Part 331: Virtual interface (FC-VI)



PRICE CODE

X

For price, see current catalogue

CONTENTS

FOREWORD	10
Introduction.....	11
1 Scope	12
2 Normative references	12
3 Terms, definitions and abbreviations	12
3.1 FC-VI terms and definitions	12
3.2 VI Definitions	14
3.3 Abbreviations	15
3.4 Editorial conventions	15
4 Structure and concepts	17
4.1 Fibre channel structure and concepts	17
4.2 FC-VI structure and concepts	17
5 FC-VI protocol overview	21
5.1 FC-VI information units	21
5.2 FC-VI message transfer operation	21
5.2.1 FC-VI message transfer	21
5.2.2 FC-VI send message transfer operation	22
5.2.3 FC-VI RDMA write message transfer operation	24
5.2.4 FC-VI RDMA read message transfer operation	26
5.2.5 IU reception at an FC-VI edpoint	27
5.3 FC-VI connection setup operation	29
5.3.1 FC-VI client-server and peer-peer connection setup	29
5.3.2 FC-VI client-server connection setup	29
5.3.3 FC-VI Peer-to-Peer Connection Establishment	31
5.3.4 FC_VI concurrent peer-to-peer connection setup	33
5.3.5 FC-VI Disconnect Operation	35
5.4 Exchange ID reuse	36
5.5 Sequence ID reuse	37
5.6 Frame synonym detection	37
5.7 VI message length	38
5.8 FC-FS header usage for FC-VI	39
5.8.1 FC-FS header usage	39
5.8.2 CS_CTL field	39
5.8.3 TYPE field	39
5.8.4 F_CTL field	39
5.8.5 DF_CTL field	39
5.8.6 SEQ_CNT field	39
5.8.7 Parameter field	40
5.9 FC-VI device_header	40
5.9.1 FC-VI device_header description	40
5.9.2 FCVI_HANDLE field	40
5.9.3 FCVI_OPCODE field	41
5.9.4 FCVI_FLAGS field	41
5.9.4.1 FCVI_FLAGS field description	41
5.9.4.2 FCVI_FLAGS for message request IUs	41
5.9.4.3 FCVI_FLAGS for message response IUs	42
5.9.4.4 FCVI_FLAGS for connect request IUs	42
5.9.4.5 FCVI_FLAGS for connect response IUs	43

5.9.4.6 FCVI_FLAGS for disconnect IUs	44
5.9.5 Reserved fields	45
5.9.6 FCVI_MSG_ID field	45
5.9.7 FCVI_PARAMETER field	46
5.9.7.1 FCVI_PARAMETER field format.....	46
5.9.7.2 Connect response reason codes	48
5.9.7.2.1 Connect response non-error reason codes.....	48
5.9.7.2.2 Connect response error reason codes.....	48
5.9.7.3 Message response / disconnect reason codes	48
5.9.7.3.1 Descriptor error reason codes.....	48
5.9.7.3.2 Remote FC-VI port non-descriptor errors.....	49
5.9.7.3.3 Reserved for future expansion	50
5.9.7.3.4 Vendor unique reason codes	50
5.9.8 FCVI_RMT_VA field	50
5.9.9 FCVI_RMT_VA_HANDLE field	50
5.9.10 FCVI_TOT_LEN field / FCVI_CONNECTION_ID field	50
6 FC-VI Information Unit (IU) formats	51
6.1 FC-VI IU overview	51
6.2 FCVI_SEND_RQST IU	51
6.2.1 FCVI_SEND_RQST IU description	51
6.2.2 FCVI_SEND_RQST Device_Header information	51
6.3 FCVI_SEND_RESP IU	51
6.3.1 FCVI_SEND_RESP IU description	51
6.3.2 FCVI_SEND_RESP Device_Header information	52
6.4 FCVI_WRITE_RQST IU	52
6.4.1 FCVI_WRITE_RQST IU overview	52
6.4.2 FCVI_WRITE_RQST IU Device_Header information	52
6.5 FCVI_WRITE_RESP IU	53
6.5.1 FCVI_WRITE_RESP IU description	53
6.5.2 FCVI_WRITE_RESP IU Device_Header information	53
6.6 FCVI_READ_RQST IU	53
6.6.1 FCVI_READ_RQST IU description	53
6.6.2 FCVI_READ_RQST IU Device_Header information	53
6.7 FCVI_READ_RESP IU	54
6.7.1 FCVI_READ_RESP IU description	54
6.7.2 FCVI_READ_RESP IU Device_Header information	54
6.8 FCVI_CONNECT_RQST IU	55
6.8.1 FCVI_CONNECT_RQST IU description	55
6.8.2 FCVI_CONNECT_RQST Device_Header information	55
6.8.3 FCVI_CONNECT_RQST Payload Information	55
6.9 FCVI_CONNECT_RESP1 IU	57
6.9.1 FCVI_CONNECT_RESP1 IU description	57
6.9.2 FCVI_CONNECT_RESP1 Device_Header information	57
6.9.3 FCVI_CONNECT_RESP1 Payload Information	58
6.10 FCVI_CONNECT_RESP2 IU	59
6.10.1 FCVI_CONNECT_RESP2 IU description	59
6.10.2 FCVI_CONNECT_RESP2 Device_Header information	59
6.11 FCVI_CONNECT_RESP3 IU	59
6.11.1 FCVI_CONNECT_RESP3 IU description	59
6.11.2 FCVI_CONNECT_RESP3 Device_Header information	59
6.12 FCVI_DISCONNECT_RQST IU	60
6.12.1 FCVI_DISCONNECT_RQST IU description	60

6.12.2 FCVI_DISCONNECT_RQST Device_Header information	60
6.13 FCVI_DISCONNECT_RESP IU	61
6.13.1 FCVI_DISCONNECT_RESP IU description	61
6.13.2 FCVI_DISCONNECT_RESP Device_Header information	61
7 FC-VI Addressing and naming	62
7.1 FC-VI Addressing and naming overview	62
7.2 FCVI_NET_ADDRESS format	62
7.3 FCVI_ATTRIBUTES format	63
7.4 FC-VI address resolution	65
7.5 FARP ELS	66
7.6 Name server queries	67
7.7 Validation of host address to N_Port Identifier mappings	67
7.7.1 Address mapping overview	67
7.7.2 Point-to-point topology	67
7.7.3 Private loop topology	67
7.7.4 Public loop topology	68
7.7.5 Fabric topology	68
8 FC-VI Error detection and recovery	69
8.1 FC-VI error detection and recovery overview	69
8.2 FC-VI endpoint states	69
8.3 FCVI_ULP_TIMEOUT definition	69
8.4 Message transfer error detection and recovery rules	70
8.4.1 Message error detection	70
8.4.2 Message transfer error recovery	70
8.5 Connection setup error detection and recovery rules	71
8.5.1 Connection setup error handling overview	71
8.5.2 Connection setup error detection	71
8.5.3 Connection setup error recovery	71
8.5.4 Connection setup originator retry rules	72
8.6 Disconnect operation error detection and recovery rules	72
8.6.1 Disconnect operation error handling overview	72
8.6.2 Disconnect operation error detection	72
8.6.3 Disconnect operation error recovery rules	73
Annex A (normative) Concurrent matching peer requests example.....	74
A.1 Overview.....	74
A.2 Case 1	75
A.3 Case 2	75
A.4 Case 3	76
A.5 Case 4	76
A.6 Case 5	77
A.7 Case 6	78
Annex B (informative) FC-VI message transfer error handling examples	79
B.1 Overview.....	79
B.2 Message transfer error handling operation.....	79
B.2.1 Message transfer error handling operation overview	79
B.2.2 Message transfer error definitions	79
B.2.3 Error Detection and Recovery Rule Processing	80
B.2.4 Message responder and message originator error recovery actions	80
B.2.5 Message responder error detection actions	81
B.2.6 Message originator Class 2 error detection actions	83
B.2.6.1 Message originator Class 2 error detection overview.....	83

B.2.6.2 Message response timeout at message originator	84
B.3 Message transfer error detection and recovery examples.....	84
B.3.1 Error examples overview	84
B.3.2 Mrcv > Mexp error example	85
B.3.2.1 Mrcv > Mexp example description	85
B.3.2.2 Mrcv > Mexp: In-order fabric and unreliable	85
B.3.2.3 Mrcv > Mexp: In-order fabric and reliable delivery.....	85
B.3.2.4 Mrcv > Mexp: Out-of-order fabric.....	86
B.3.3 Mrcv = Mexp error example	86
B.3.3.1 Mrcv = Mexp example description	86
B.3.3.2 Mrcv = Mexp: In-order fabric and unreliable	87
B.3.3.3 Mrcv = Mexp: Out-of-order fabric.....	87
Annex C (informative) Connection setup error handling examples Overview	89
C.1 Connection setup error handling definitions	89
C.2 Connect request originator and connect request responder rules.....	89
C.3 Connect request originator rules	89
C.4 Connect request responder rules	91
C.4.1 Connect request responder retry rules	92
C.5 Error detection and recovery examples for connection setup	93
C.5.1 Overview	93
C.5.2 FC-VI connection setup timers	94
C.5.3 VipConnectRequest completion	95
C.5.4 VipConnectAccept completion	95
C.5.5 Enabling message transmission and reception	95
C.5.6 Client timeout of VipConnectRequest	96
C.5.7 Lost FCVI_CONNECT_RQST IU	96
C.5.7.1 Lost FCVI_CONNECT_RQST IU example.....	96
C.5.7.2 Retried connection setup.....	97
C.5.8 Lost FCVI_CONNECT_RESP1 IU	98
C.5.9 Lost FCVI_CONNECT_RESP2 IU	99
C.5.9.1 Lost FCVI_CONNECT_RESP2 IU example.....	99
C.5.9.2 Server timing out connection setup	100
C.5.10 Lost FCVI_CONNECT_RESP3 IU	101
C.5.10.1 Lost FCVI_CONNECT_RESP3 IU example.....	101
Annex D (informative) Disconnect operation error handling examples	102
D.1 Disconnect operation example description.....	102
D.2 FC-VI disconnect operation example	103
Annex E (informative) Message streaming for reliable reception	105
Annex F (informative) Enabling Message transmission in the FC-VI NIC	106
Documents for VI Architecture (see Clause 2 for further explanation)	
Virtual Interface Architecture Specification, V1.0 (VI-ARCH)	111
Virtual Interface (VI) Architecture Developer's Guide, V1.0 (VI-DG)	195
Virtual Interface (VI) Architecture Developer's Guide Error Table Supplement, V1.0	291
IP Version 6 Addressing Architecture, RFC 2373, July 1998 (RFC2373)	321

Table 1 – FC-VI Information unit summary.....	21
Table 2 – Peer B actions based on connect responses from peer A	36
Table 3 – 16-byte FC-VI device_header.....	40
Table 4 – 32-byte FC-VI device_header.....	40
Table 5 – FCVI_FLAGS Bit definitions for message request IUs	41
Table 6 – FCVI_FLAGS Bit definitions for message response IUs	42
Table 7 – FCVI_FLAGS Bit definitions for connect request IUs	42
Table 8 – FC-VI connection mode definition	43
Table 9 – FCVI_FLAGS Bit definitions for connect response IUs	43
Table 10 – FCVI_FLAGS Bit definitions for disconnect IUs	44
Table 11 – FCVI_PARAMETER field for connect response and disconnect IUs	46
Table 12 – Reason code for CONN_STS.....	47
Table 13 – FCVI_CONNECT_RQST IU payload format	56
Table 14 – FCVI_CONNECT_RESP1 IU Payload Format.....	58
Table 15 – FCVI_NET_ADDRESS Format	63
Table 16 – FCVI_ATTRIBUTES format.....	63
Table 17 – Format of FCVI_ATTR_FLAGS in FCVI_ATTRIBUTES	64
Table 18 – FCVI_QOS format	64
Table A.1 – Peer B actions based on connect responses from peer A	74

Figure 1 – FC-VI addressing objects	20
Figure 2 – FC-VI send for Unreliable Delivery or Reliable Delivery	23
Figure 3 – FC-VI send for Reliable Reception	24
Figure 4 – FC-VI RDMA write for Unreliable Delivery or Reliable Delivery	25
Figure 5 – FC-VI RDMA write for Reliable Reception	26
Figure 6 – FC-VI RDMA read for Reliable Reception and Reliable Delivery	27
Figure 7 – Concurrent Receive Streams at a FC-VI Endpoint	28
Figure 8 – FC-VI client-server connection setup	30
Figure 9 – Peer-to-peer connection setup	32
Figure 10 – Peer-to-peer connection setup, concurrent matching peer requests	34
Figure 11 – FC-VI disconnect operation	36
Figure 12 – FC-FS header for send operation	38
Figure A.1 – Case 1	75
Figure A.2 – Case 2	75
Figure A.3 – Case 3	76
Figure A.4 – Case 4	77
Figure A.5 – Case 5	77
Figure A.6 – Case 6	78
Figure B.1 – $Mrcv > Mexp$	85
Figure B.2 – $Mrcv = Mexp$	87
Figure C.1 – Client-server connection setup	93
Figure C.2 – Client timeout of VipConnectRequest	96
Figure C.3 – Lost FCVI_CONNECT_RQST IU	97
Figure C.4 – Lost FCVI_CONNECT_RESP1 IU	98
Figure C.5 – Lost FCVI_CONNECT_RESP2	99
Figure C.6 – Lost FCVI_CONNECT_RESP3 IU	101
Figure D.1 – FC-VI disconnect operation	103

INFORMATION TECHNOLOGY – FIBRE CHANNEL –

Part 331: Virtual interface (FC-VI)

FOREWORD

- 1) ISO (International Organization for Standardization) and IEC (International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards. Their preparation is entrusted to technical committees; any ISO and IEC member body interested in the subject dealt with may participate in this preparatory work. International governmental and non-governmental organizations liaising with ISO and IEC also participate in this preparation.
- 2) In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.
- 3) The formal decisions or agreements of IEC and ISO on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC and ISO member bodies.
- 4) IEC, ISO and ISO/IEC publications have the form of recommendations for international use and are accepted by IEC and ISO member bodies in that sense. While all reasonable efforts are made to ensure that the technical content of IEC, ISO and ISO/IEC publications is accurate, IEC or ISO cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.
- 5) In order to promote international uniformity, IEC and ISO member bodies undertake to apply IEC, ISO and ISO/IEC publications transparently to the maximum extent possible in their national and regional publications. Any divergence between any ISO/IEC publication and the corresponding national or regional publication should be clearly indicated in the latter.
- 6) ISO and IEC provide no marking procedure to indicate their approval and cannot be rendered responsible for any equipment declared to be in conformity with an ISO/IEC publication.
- 7) All users should ensure that they have the latest edition of this publication.
- 8) No liability shall attach to IEC or ISO or its directors, employees, servants or agents including individual experts and members of their technical committees and IEC or ISO member bodies for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication of, use of, or reliance upon, this ISO/IEC publication or any other IEC, ISO or ISO/IEC publications.
- 9) Attention is drawn to the normative references cited in this publication. Use of the referenced publications is indispensable for the correct application of this publication.
- 10) Attention is drawn to the possibility that some of the elements of this International Standard may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

International Standard ISO/IEC 14165-331 was prepared by subcommittee 25: Interconnection of information technology equipment, of ISO/IEC joint technical committee 1: Information technology.

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

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

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

INTRODUCTION

This International Standard 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.

INFORMATION TECHNOLOGY – FIBRE CHANNEL –

Part 331: Virtual interface (FC-VI)

1 Scope

This part of ISO/IEC 14165 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

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced documents (including any amendments) applies.

ISO/IEC 14165-122, *Information technology – Fibre channel – Arbitrated Loop-2 (FC-AL-2)*

ISO/IEC 14165-251, *Information technology – Fibre channel – Framing and Signalling Interface (FC-FS)* (To be published)

ISO/IEC 14165-414, *Information technology – Fibre channel – Generic services-4 (FC-GS-4)*

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 the convenience of the reader they are added as supplementary documents.

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

IP Version 6 Addressing Architecture, RFC 2373, July 1998 (RFC2373)
(can be downloaded from the Internet)

3 Terms, definitions and abbreviations

3.1 FC-VI terms and 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 connection point

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 standard.

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 standard

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

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

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

3.1.16 VI connection

connection between two VI Endpoints

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 standard are defined in the VI Architecture Specification and the VI Architecture Developer's Guide, provided as supplementary documents.

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

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 sub-clauses.

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’.

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”.

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.

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 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 standard.

The VI Architecture does not define the format of how VI Addresses are mapped to 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 Vip-

ConnectWait 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.

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.

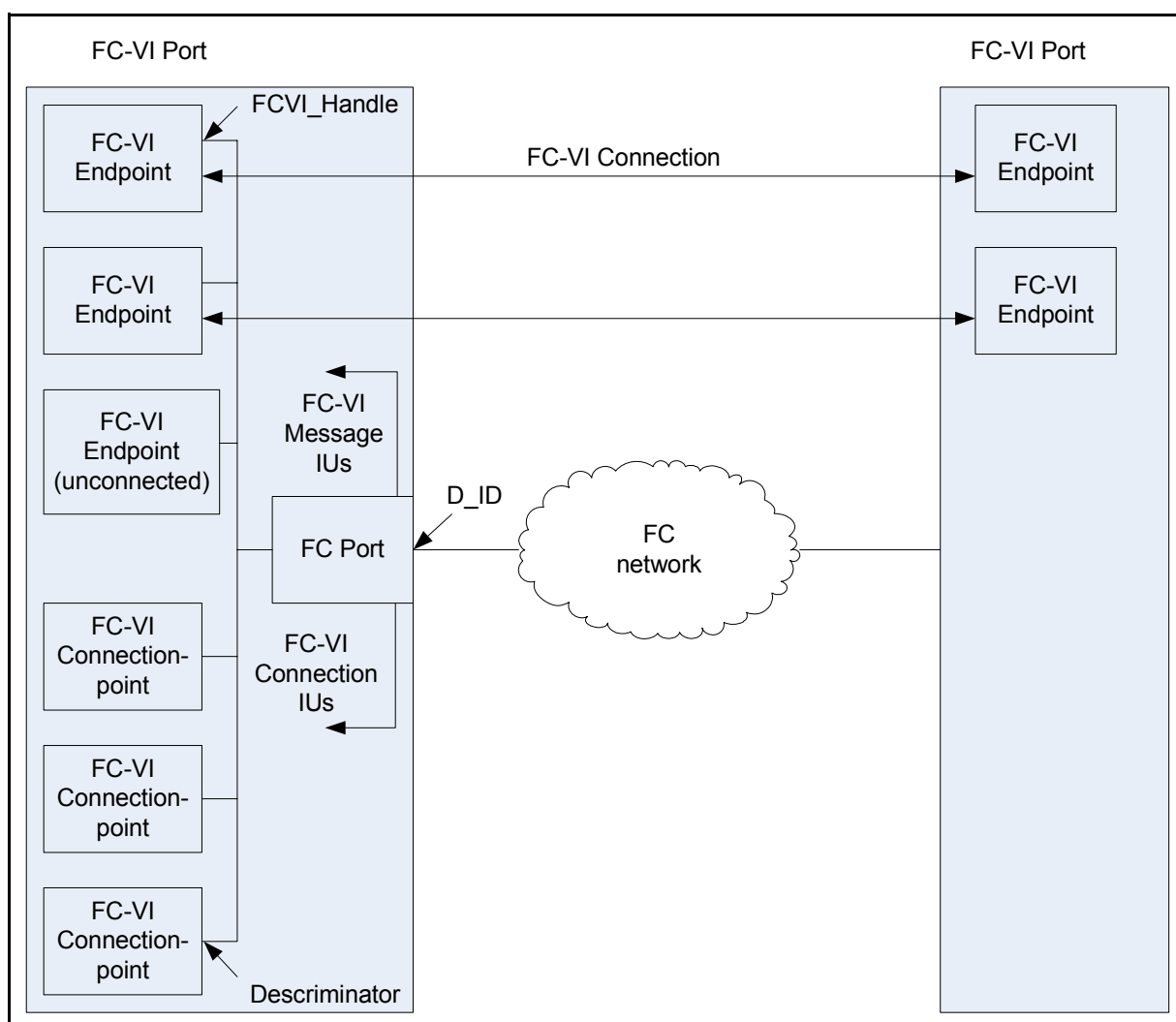


Figure 1 – FC-VI addressing objects

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 Ex-

change for Message Request IUs and Message Response IUs (if needed) binds all Sequences and frames associated with one VI Message.

NOTE 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.

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 re-

sponse 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.

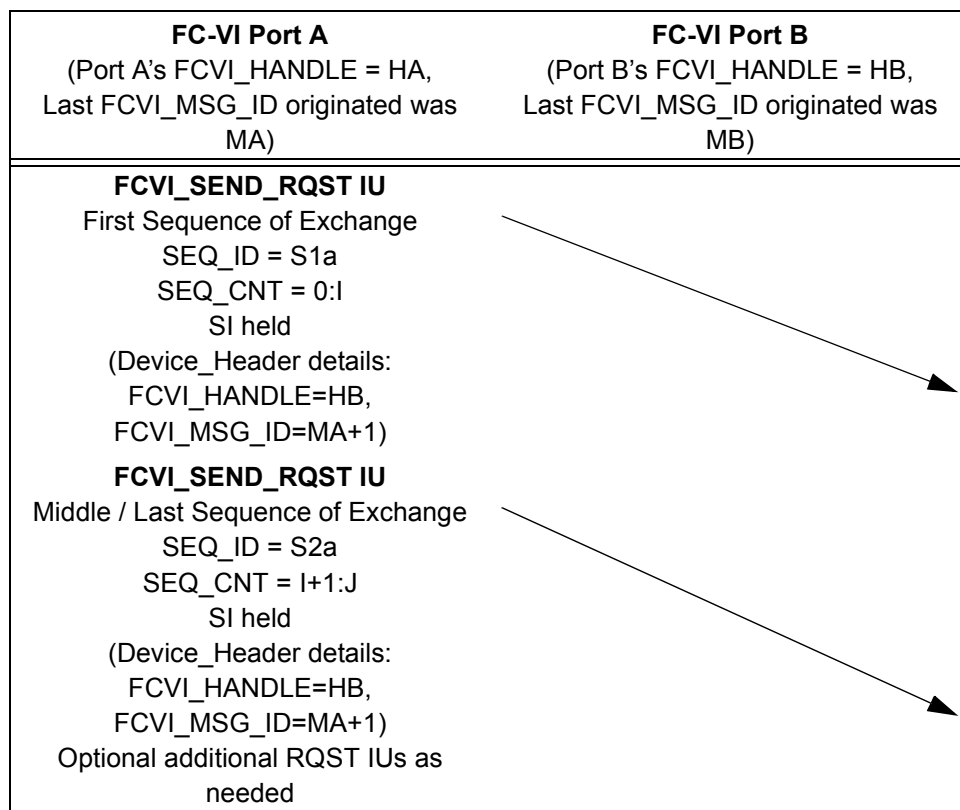


Figure 2 – FC-VI send for Unreliable Delivery or Reliable Delivery

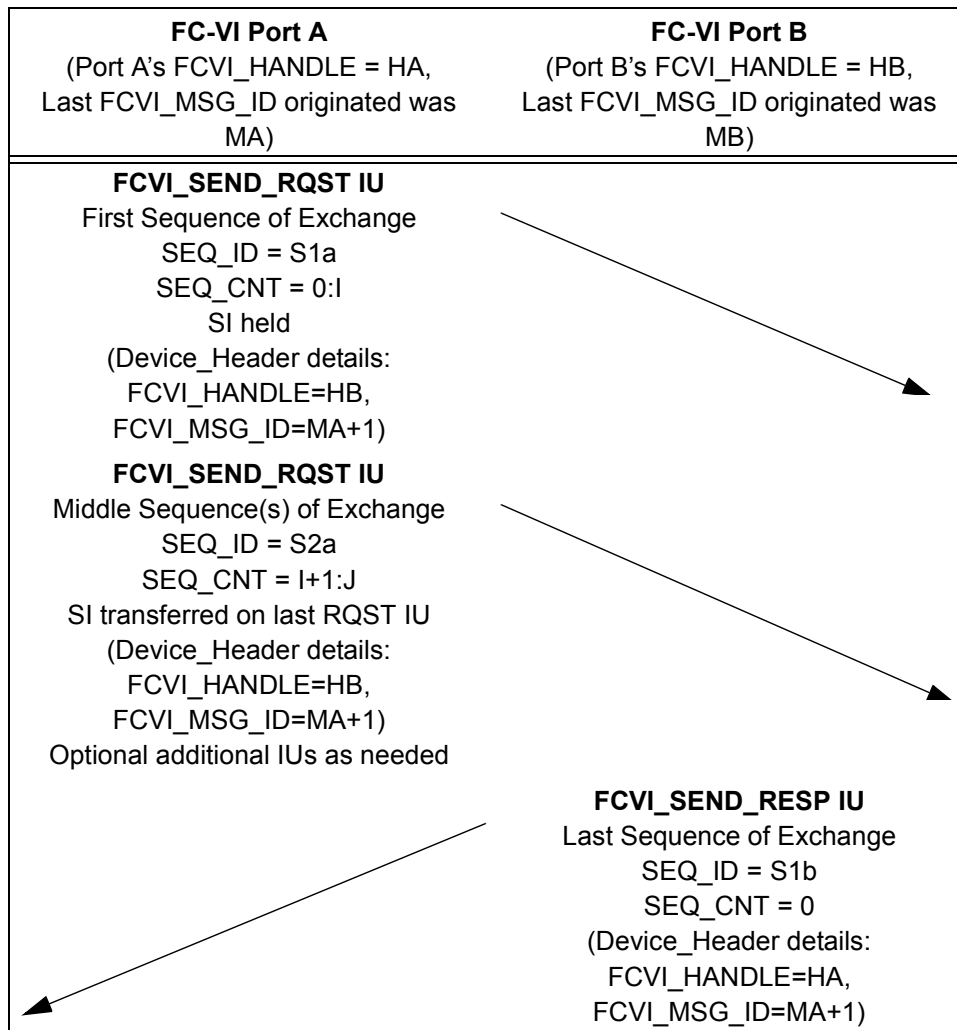


Figure 3 – FC-VI send for Reliable Reception

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

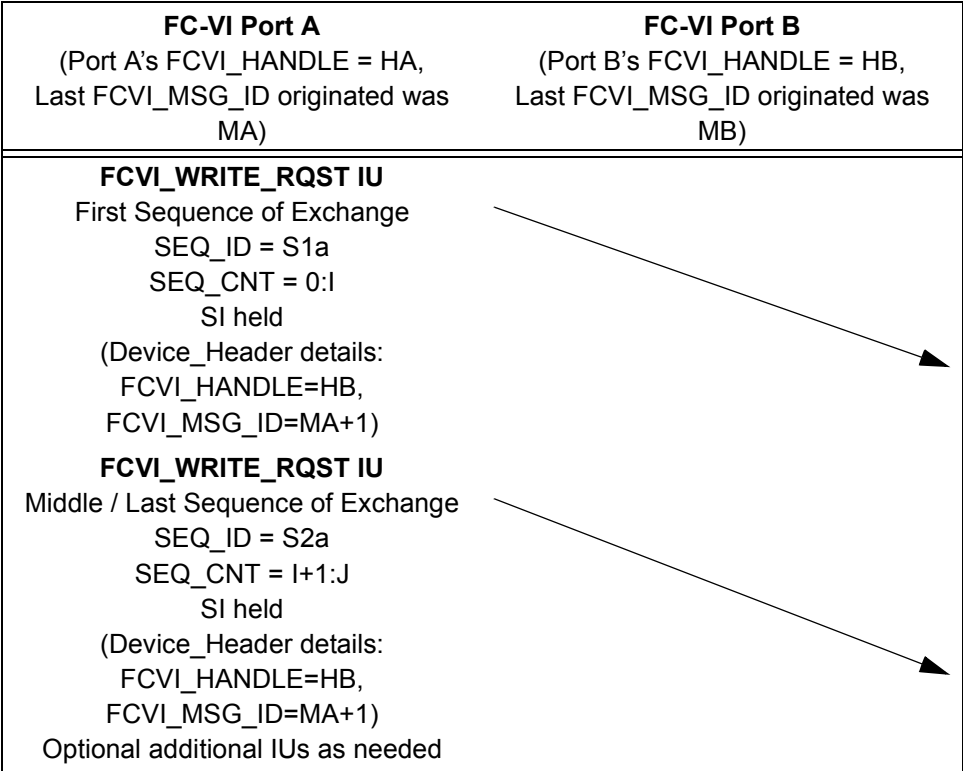


Figure 4 – FC-VI RDMA write for Unreliable Delivery or Reliable Delivery

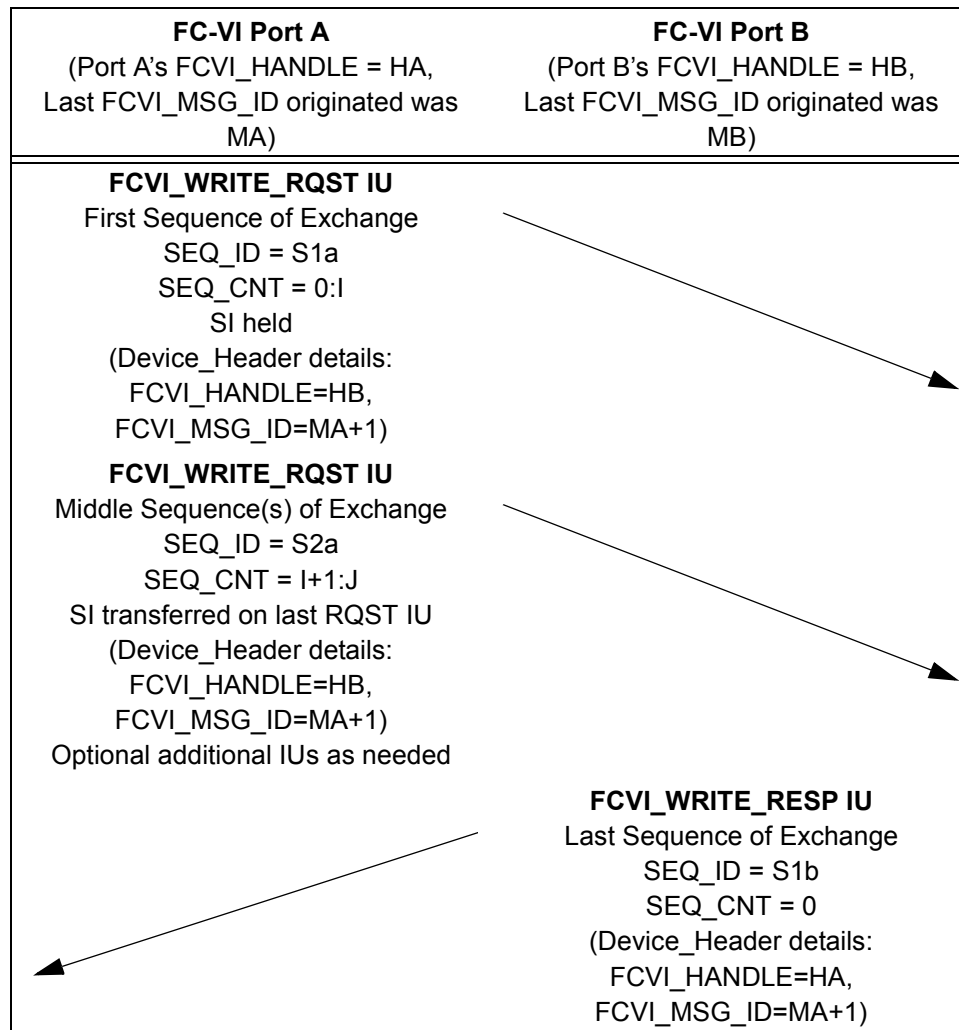


Figure 5 – FC-VI RDMA write for Reliable Reception

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 may 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.

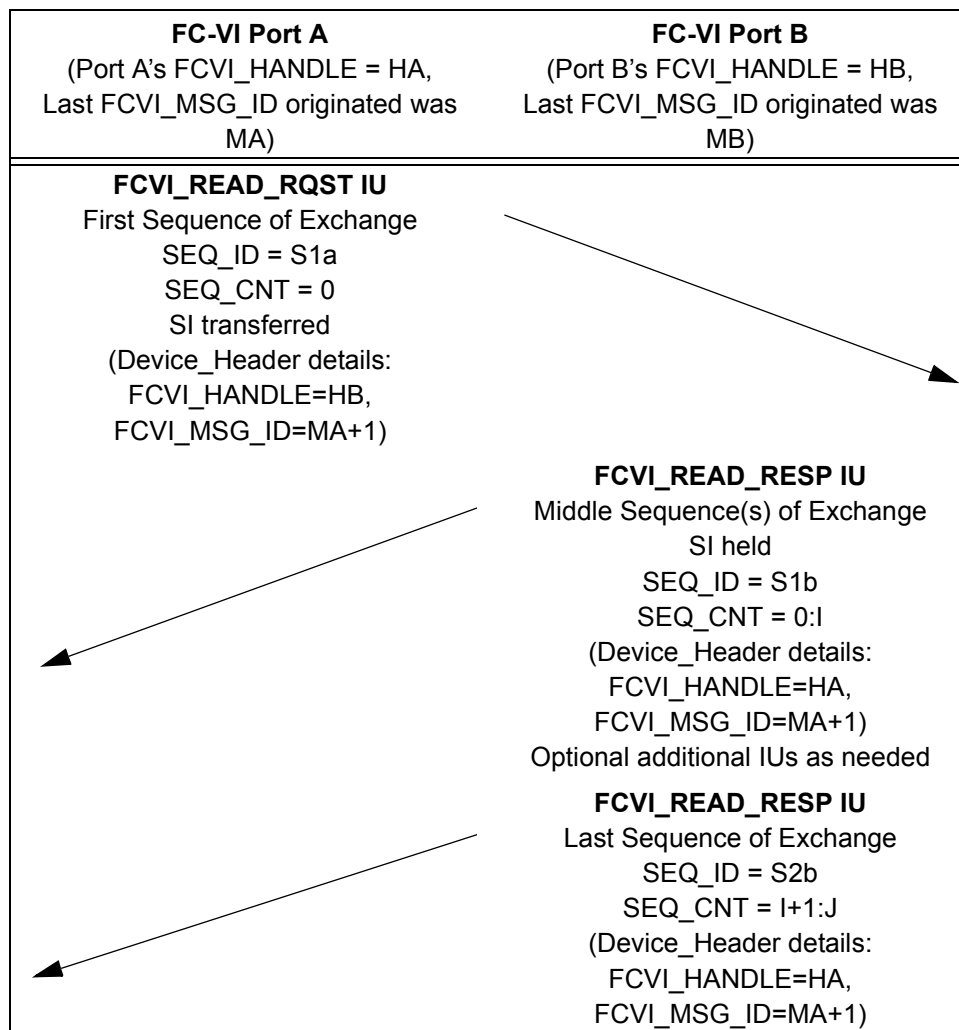


Figure 6 – FC-VI RDMA read for Reliable Reception and Reliable Delivery

5.2.5 IU reception at an FC-VI endpoint

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.

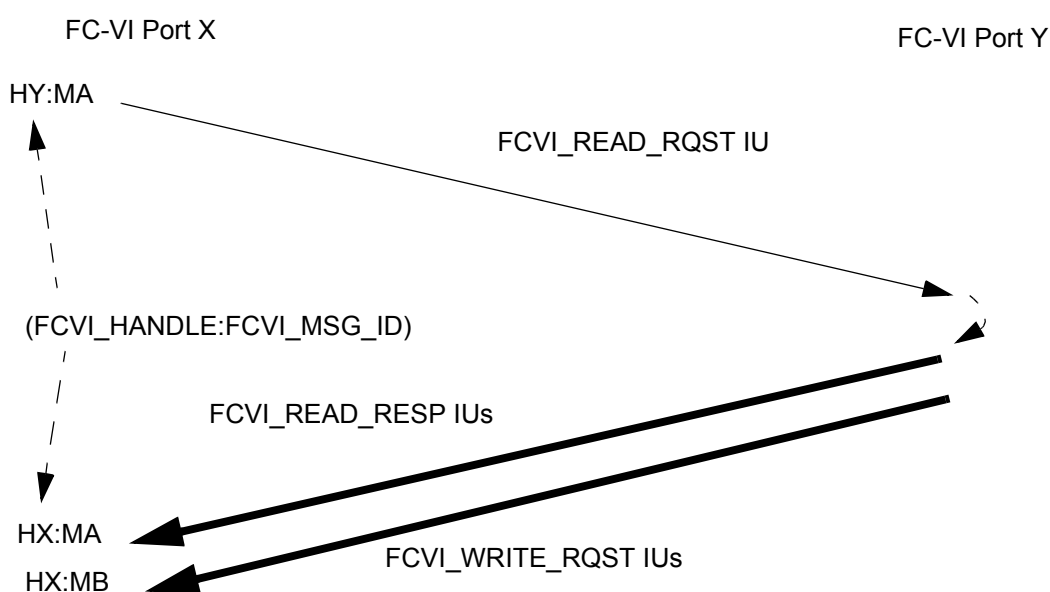


Figure 7 – Concurrent Receive Streams at a FC-VI Endpoint

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.

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

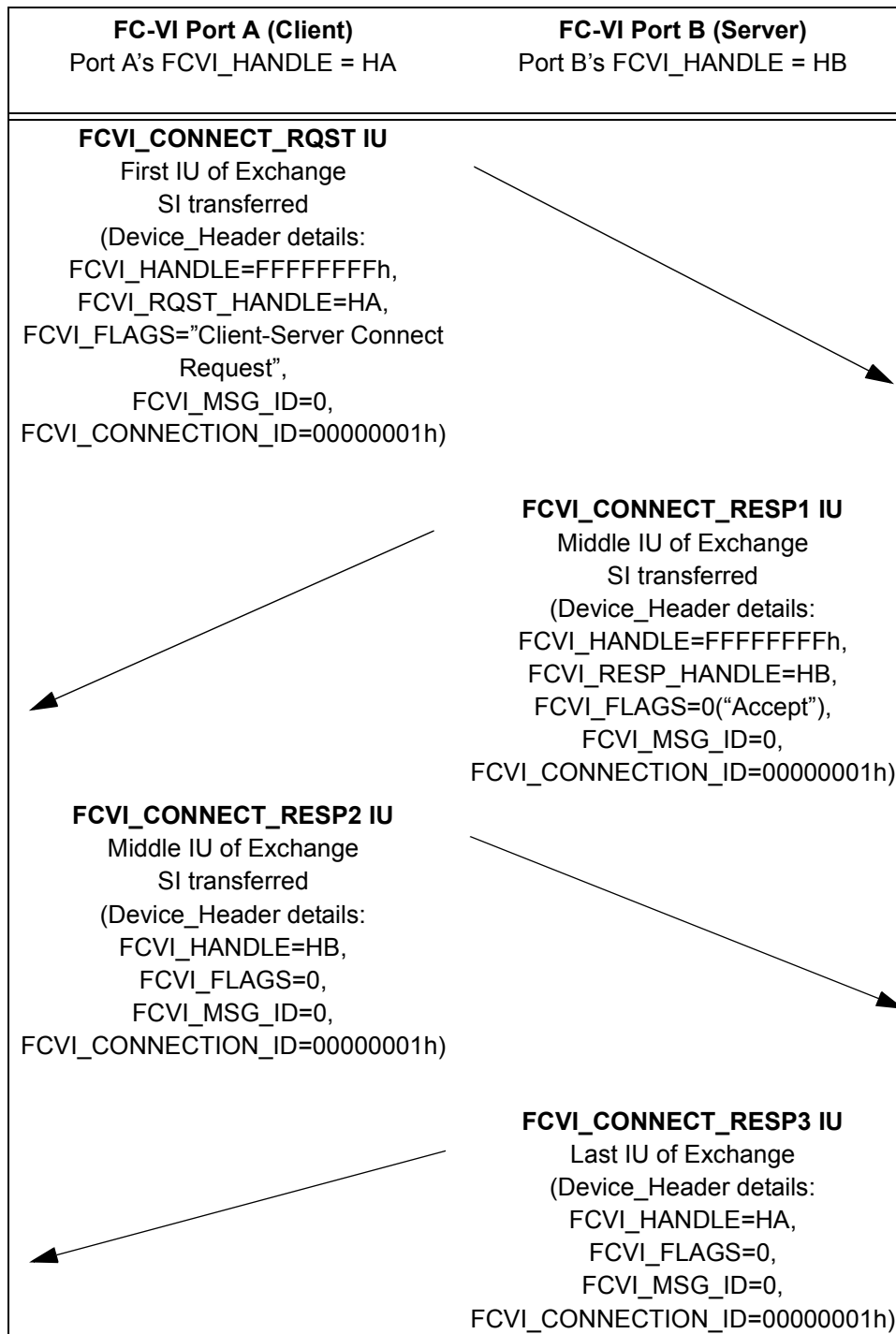


Figure 8 – FC-VI client-server connection setup

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.

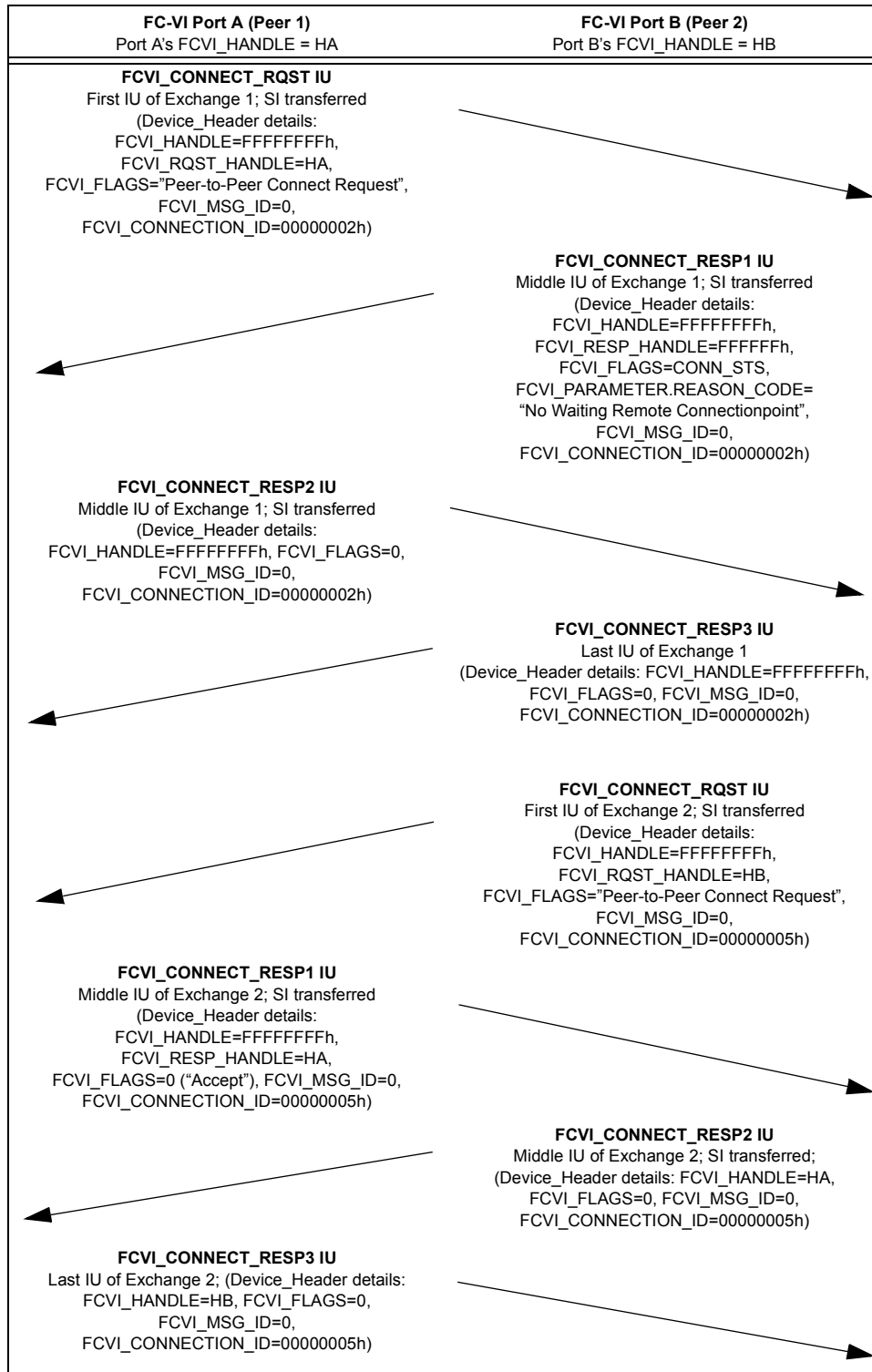


Figure 9 – Peer-to-peer connection setup

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 Request 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.

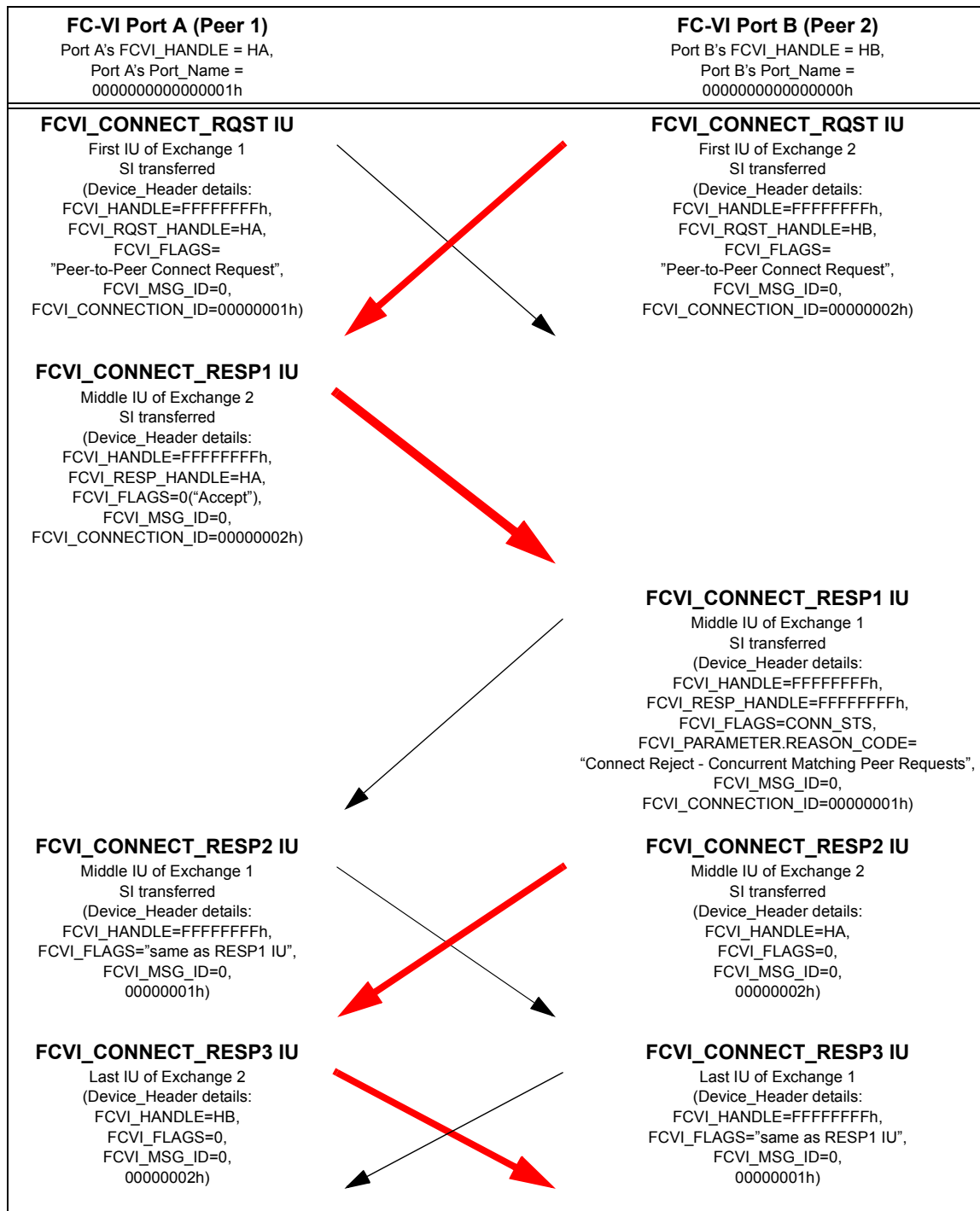


Figure 10 – Peer-to-peer connection setup, concurrent matching peer requests

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 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 in certain cases.

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

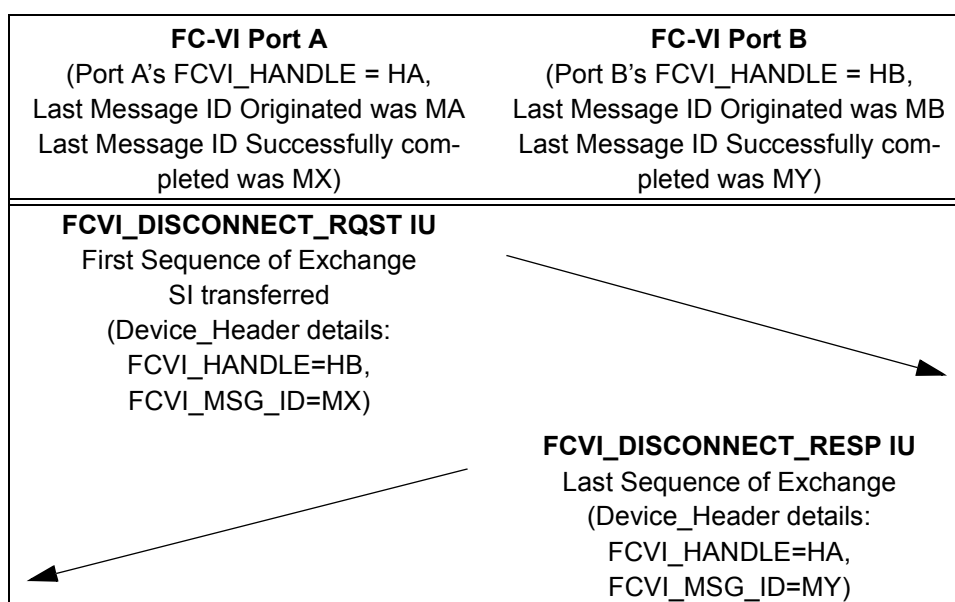


Figure 11 – FC-VI disconnect operation

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.

5.5 Sequence ID reuse

As required by ISO/IEC 14165-251 (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.

NOTE FC-FS requires the use of at least two different SEQ_IDs for Sequence streaming.

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, bit 17 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.

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

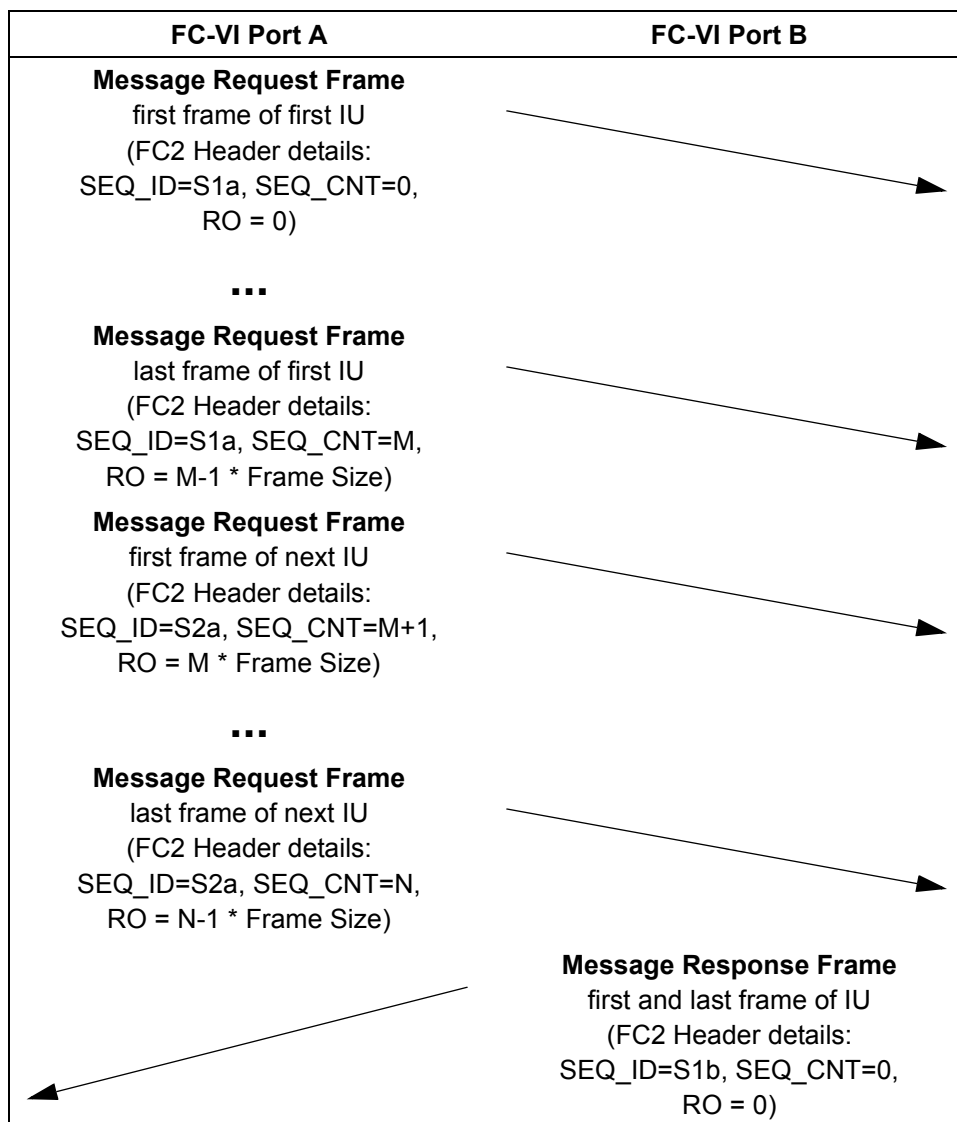


Figure 12 – FC-FS header for send operation

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.

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 frameheader (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.

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 to 24 Byte 0	23 to 16 Byte 1	15 to 08 Byte 2	07 to00 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 to 24 Byte 0	23 to 16 Byte 1	15 to 08 Byte 2	07 to 00 Byte 3	
0	FCVI_HANDLE				
1	FCVI_OPCODE	FCVI_FLAGS	Reserved		
2	FCVI_MSG_ID				
3	FCVI_PARAMETER				
4	(MSB)	FCVI_RMT_VA			(LSB)
5					
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 FFFFFFFFh 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.

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.

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_INF O	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.

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 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 originator is aborting a Connection Setup. If the CONN_SETUP_ABORT bit is set to zero, the Disconnect request originator 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.

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 to 24 Byte 0	23 to 16 Byte 1	15 to 08 Byte 2	07 to 00 Byte 3
0	Reserved	Reason Code	Reason Explanation	Vendor Unique

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.

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 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 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.