Inside a Modern Fibre Channel Architecture – Part 2

Live Webcast
October 27, 2021
10:00 AM PT/1:00 PM ET

FIBRE CHANNEL INDUSTRY ASSOCIATION

Today's Speakers

David Rodgers

FCIA BoD

Teledyne LeCroy

Patty Driever

FCIA BoD

Distinguished Engineer

Storage and Networking Solutions & Strategy

IBM

David Peterson

Principal Engineer

Broadcom

Craig Carlson

FCIA BoD

Senior Technologist

Marvell

About the FCIA

“The Fibre Channel Industry Association (FCIA) is a mutual benefit, non-profit, international organization of manufacturers, system integrators, developers, vendors, and industry professionals, and end users.”

About the Fibre Channel Industry Association (FCIA)

25+ Years
Promoting Fibre Channel Technology

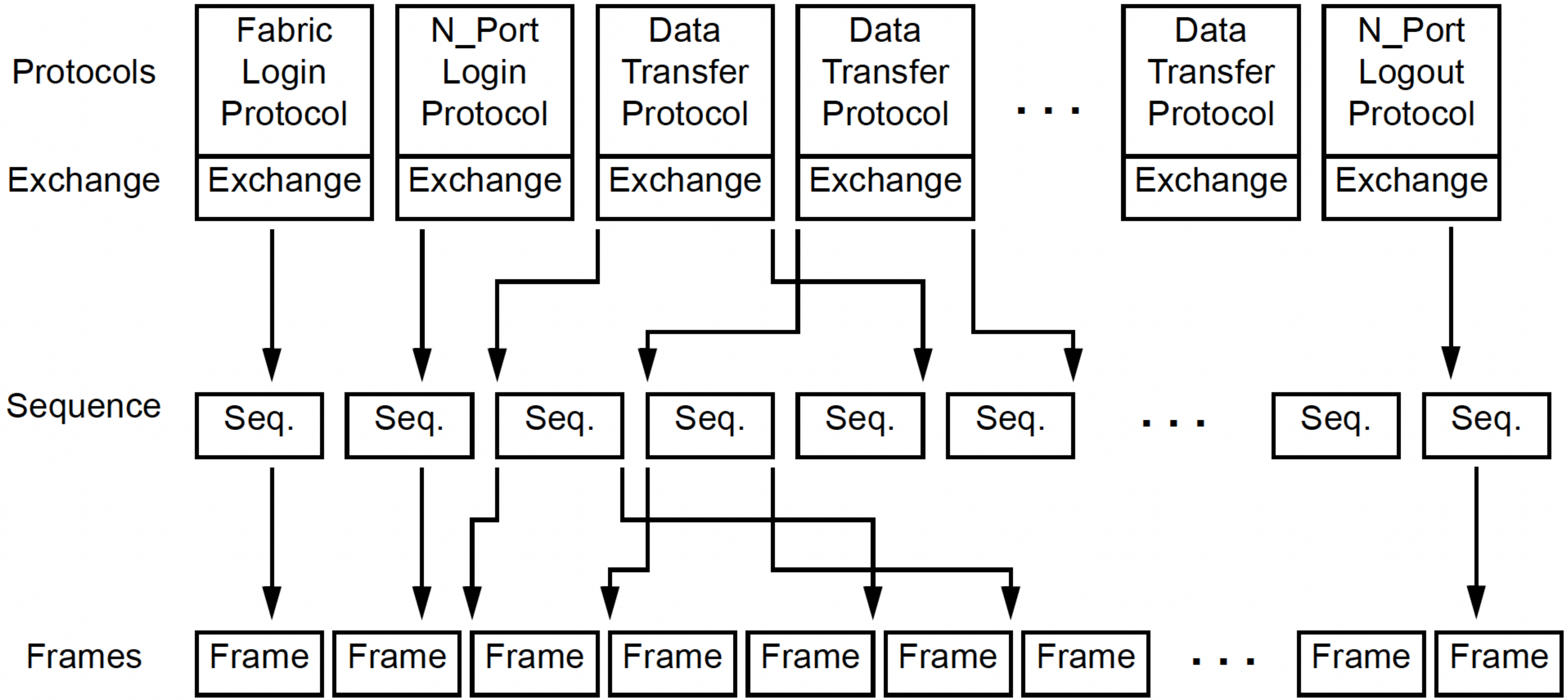
Industry Leading
Member Companies

142M+ FC Ports
Shipped Since 2001

Agenda

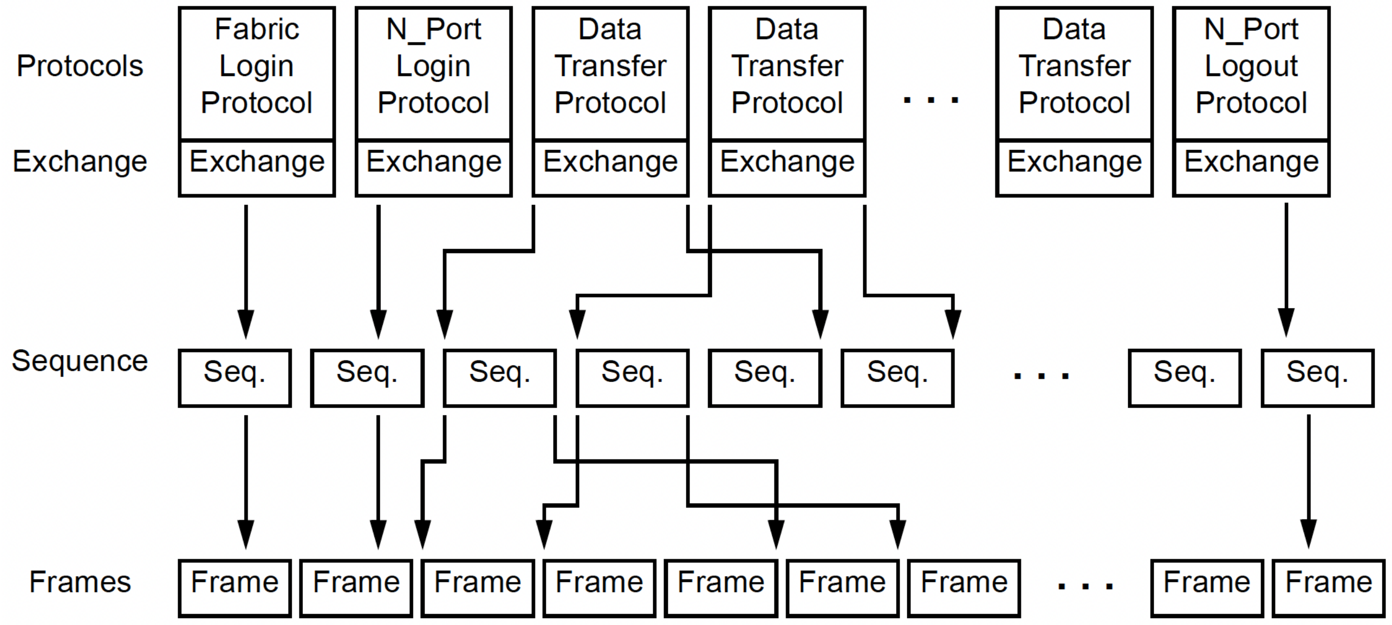
- Building blocks and their hierarchy
- Frames, Sequences, Exchanges, Protocols
- Information Units
- Segmentation and reassembly
- Error detection and recovery
- Current enhancements

Building Blocks



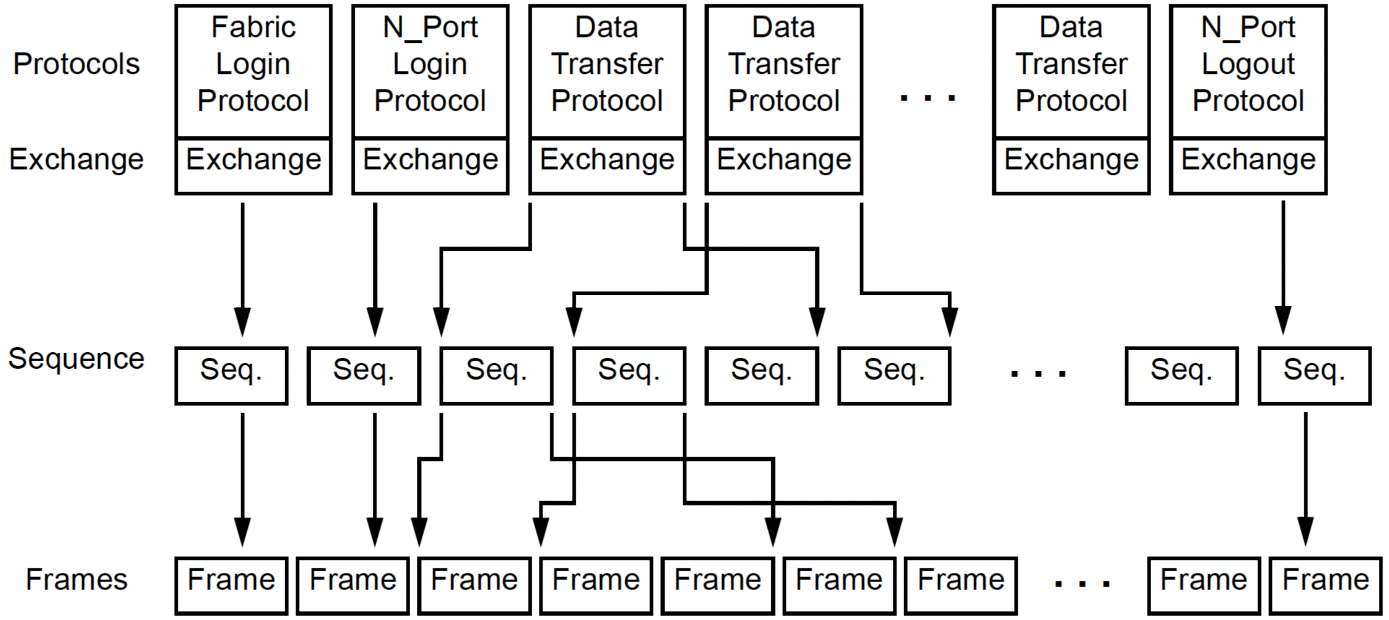
Frames, Sequences, Exchanges, Protocols

- Frame contains the information (payload) to be transferred
- Sequence is made up of one or more Data frames and if applicable, corresponding responses
- Exchange is made up of one or more Sequences flowing in a single direction from the Originator of the Exchange to the Responder or in both directions between the Originator and the Responder
- Prior to use by a ULP for data transfer, Fibre Channel has to be setup for the operating environment



Frames, Sequences, Exchanges, Protocols

- Fibre Channel operating environment is setup by performing Fabric Login and N_Port Login
- Once these two Logins are performed, an FC-4 may start using Fibre Channel until one or both of these Logins are invalidated
- Each Login uses an Exchange as the mechanism to accomplish the login function
- Data transfer is performed using an Exchange as the mechanism with the related FC-4 translating the ULP protocol to FC-2 protocol

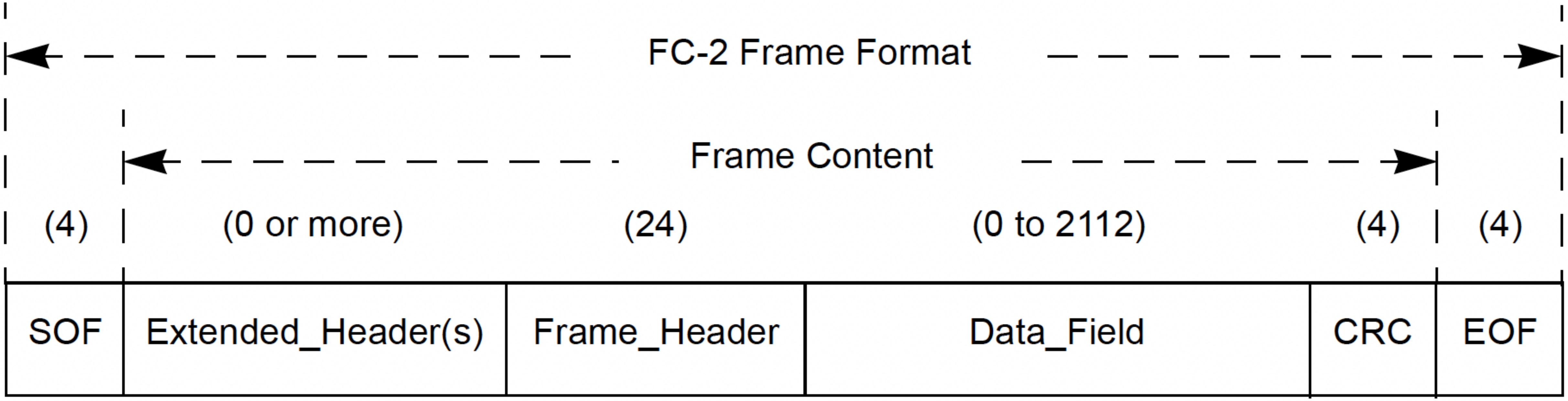


Fibre Channel Frame

- Categorized as Data frames and Link_Control frames
  - Data frames are classified as
    - Link_Data frames
    - Device_Data frames
    - Video_Data frames
  - Link_Control frames are classified as
    - Acknowledge (ACK) frames
    - Link_Response (Busy and Reject) frames
    - Link_Control command frames

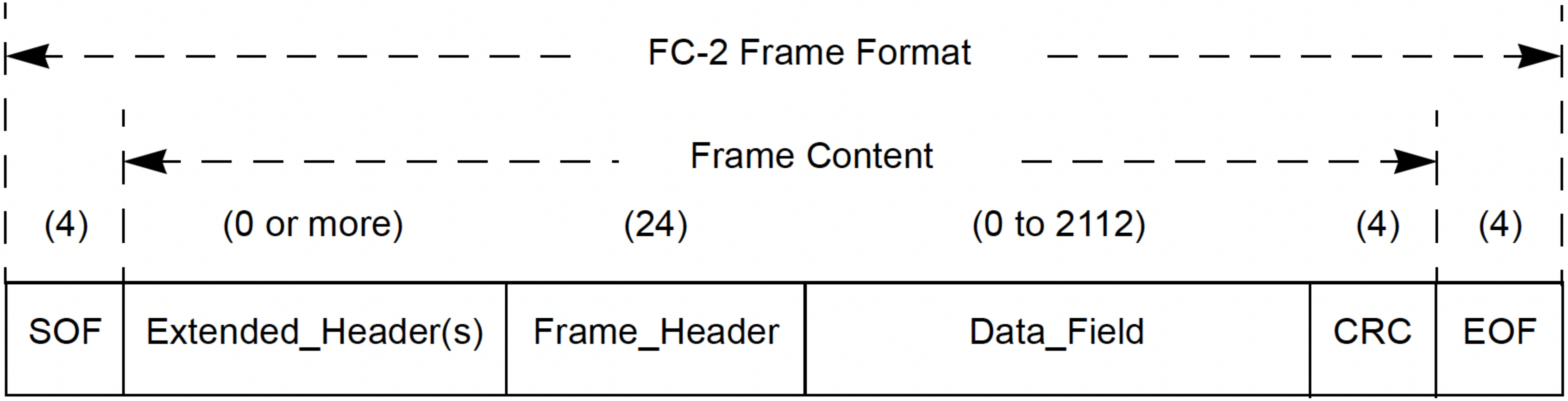
Fibre Channel Frame

- Based on a common frame format



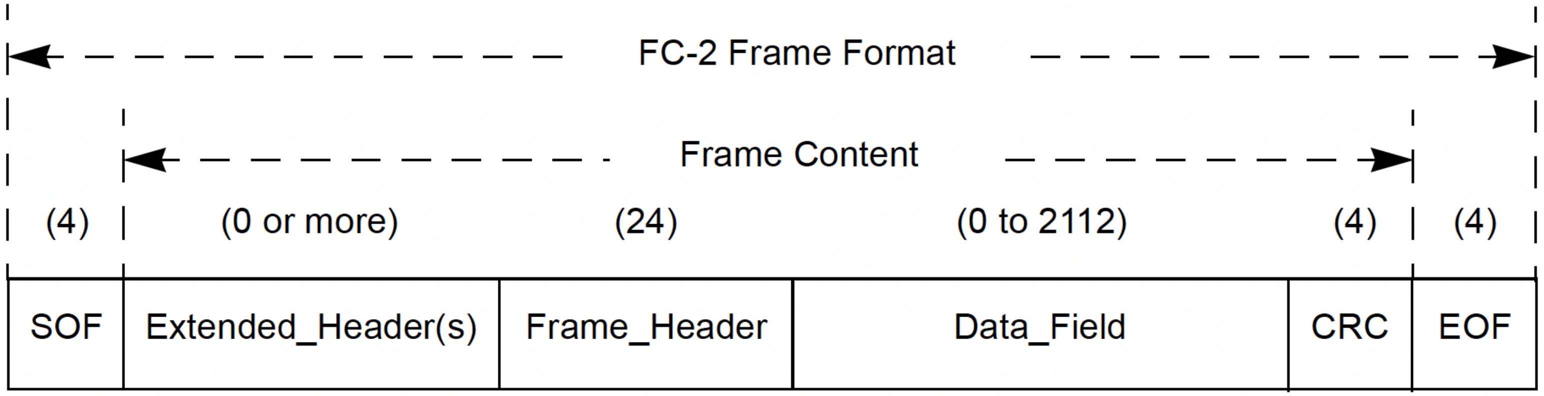
Fibre Channel Frame

- SOF – Start of Frame delimiter precedes the Frame Content
- EOF – End of Frame delimiter follows the Frame Content



Fibre Channel Frame

- Extended_Headers
  - VFT_Header (Virtual Fabric Tagging Header)
    - Allows VN_Ports to share the same physical link while connected to different Virtual Fabrics
  - IFR_Header (Inter-Fabric Routing Header)
    - Used at every Inter-Fabric Router to route the frame toward the destination fabric
  - Enc_Header (Encapsulation Header)
    - Used to transmit frames between Inter-Fabric Routers when connected through intermediate Fabrics that do not support the IFR_Header



Fibre Channel Frame

- Frame_Header

Bits Word	31 .. 24	23 .. 16	15 .. 08	07 .. 00
0	R_CTL	D_ID		
1	CS_CTL/Priority	S_ID		
2	TYPE	F_CTL		
3	SEQ_ID	DF_CTL	SEQ_CNT	
4	OX_ID		RX_ID	
5	Parameter			

Fibre Channel Frame

- Routing Control (R_CTL)
  - Contains ROUTING field and INFORMATION field
    - Device_Data frames
    - Extended Link Services (see FC-LS-5)
    - FC-4 Link_Data (see relevant FC-4 standard, e.g., FCP, FC-NVMe, SB-6)
    - Video_Data (see FC-AV and ARINC 818)
    - Extended_Headers
    - Basic Link Services
    - Link_Control Frame
    - Extended Routing (no standard usage is specified)

Fibre Channel Frame

- Address Identifiers (D_ID, S_ID)
  - D_ID contains address identifier of the destination N_Port
  - S_ID contains address identifier of the source N_Port
  - Each N_Port has a native N_Port_ID that is unique within the address domain of a Fabric

Fibre Channel Frame

- Class Specific Control (CS_CTL)/Priority
  - Meaning controlled by CS_CTL/Priority Enable bit (F_CTL, bit 17)
    - If CS_CTL/Priority Enable bit is set to zero, then bits are CS_CTL information

Bits	Abbr.	Meaning
31	PREF	0 = Frame is delivered with no Preference 1 = Frame may be delivered with Preference
30		Reserved for additional Preference function
29-24	DSCP	Differentiated Services Code Point

    - If CS_CTL/Priority Enable bit is set to one, bits are Priority information

Word 1, bit(s)	Meaning
31-25	Priority
24	Tagging Extension

Fibre Channel Frame

- TYPE
  - Identifies the protocol of the frame content for Data frames
    - Link Service
      - Basic Link Services
      - Extended Link Services
    - Video_Data
    - FC-4
      - Fibre Channel Protocol (FCP)
      - Single Byte Command Code Set (SBCCS)
      - NVMe over Fibre Channel (FC-NVMe)
      - Many others such as IPv4/v6

Fibre Channel Frame

- Frame Control (F_CTL)
  - Contains control information relating to the frame content
    - Exchange Context
    - Sequence Context
    - First_Sequence
    - Last_Sequence
    - End_Sequence
    - CS_CTL/Priority
    - Sequence Initiative
    - ACK_Form
    - Abort Sequence Condition
    - Relative offset present
    - Exchange reassembly
    - Fill Bytes

Fibre Channel Frame

- Sequence_ID (SEQ_ID)
  - Assigned by the Sequence Initiator
  - If SEQ_ID unique per Exchange bit is set to zero in the PLOGI request or PLOGI LS_ACC, then the SEQ_ID is unique among all concurrently open Sequences between the Sequence Initiator and the Sequence Recipient, independent of the Exchange ID
  - If SEQ_ID unique per Exchange bit is set to one in the PLOGI request and PLOGI LS_ACC, then the SEQ_ID is unique among all concurrently open Sequences with the same Exchange ID
  - Both Sequence Initiator and the Sequence Recipient track the status of frames within the Sequence using fields within the Sequence_Qualifier

Fibre Channel Frame

- Data Field Control (DF_CTL)
  - Specifies the presence of optional headers at the beginning of the Data_Field

Word 3, Bit(s)	Optional Header	Applicability
23	Reserved	all frames
22	0 = Neither ESP_Header nor ESP_Trailer 1 = Both ESP_Header and ESP_Trailer	all frames
21	0 = No Network_Header 1 = Network_Header	Device_Data and Video_Data frames
20	Obsolete	
19-18	Reserved	all frames
17-16	00b = No Device_Header 01b = 16 Byte Device_Header 10b = 32 Byte Device_Header 11b = 64 Byte Device_Header	Device_Data and Video_Data frames

  - *Note optional header(s) reduce the size of user data

Fibre Channel Frame

- Data Field Control (DF_CTL)
  - ESP_Header
    - Adheres to RFC 4303 except for Integrity Check Value (ICV) coverage
    - How to use ESP in FC is specified in FC-SP-2

Word \ Bits	31 .. 24	23 .. 16	15 .. 08	07 .. 00
0	R_CTL	D_ID		
1	CS_CTL / Priority	S_ID		
2	TYPE	F_CTL		
3	SEQ_ID	DF_CTL	SEQ_CNT	
4	OX_ID		RX_ID	
5	Parameter			
6	Security Parameter Index (SPI)			
7	ESP Sequence Number			
8 .. M	Other Optional Headers (if present)			
M+1 .. N	Payload (variable length)			
		Fill Bytes (if present)		
N+1 .. P	ESP Padding (2-254 bytes)			
			Pad Length	Not meaningful
P+1 .. Q	Integrity Check Value			
Q+1	CRC			

Fibre Channel Frame

- Data Field Control (DF_CTL)
  - – Application_Header
    - Constructed as 16 byte Device_Header
    - Support determined via N_Port Login (PLOGI)

Bits / Word	31 .. 00
0	Destination Application Identifier
1	Source Application Identifier
2	Reserved
3	Reserved

Fibre Channel Frame

- Sequence count (SEQ_CNT)
  - – Indicates the sequential order of Data frame transmission within a single Sequence or multiple consecutive Sequences for the same Exchange
- Originator Exchange_ID (OX_ID)
  - – Identifies the Exchange_ID assigned by the Originator of the Exchange
  - – Originator Exchange Status Block associated with the OX_ID is used to track the progress of a series of Sequences that comprises an Exchange

Fibre Channel Frame

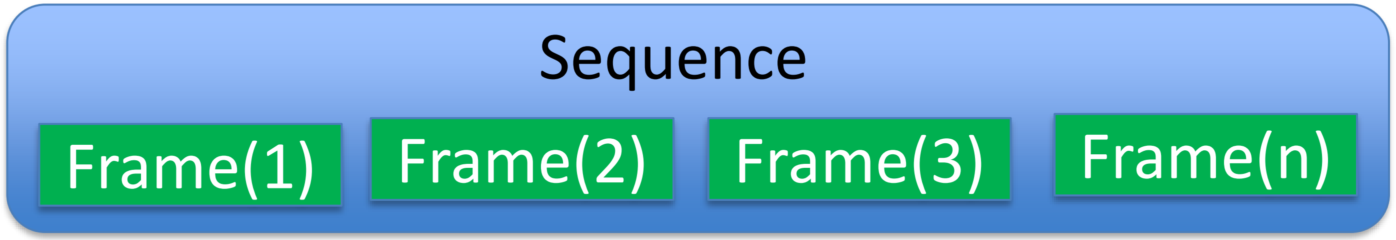
- Responder Exchange_ID (RX_ID)
  - Assigned by Responder to provide a unique, locally meaningful identifier at the Responder for an Exchange established by an Originator identified by an OX_ID
  - Value of FFFFh indicates the RX_ID is unassigned
  - Responder Exchange Status Block associated with the RX_ID is used to track the progress of a series of Sequences that compose an Exchange

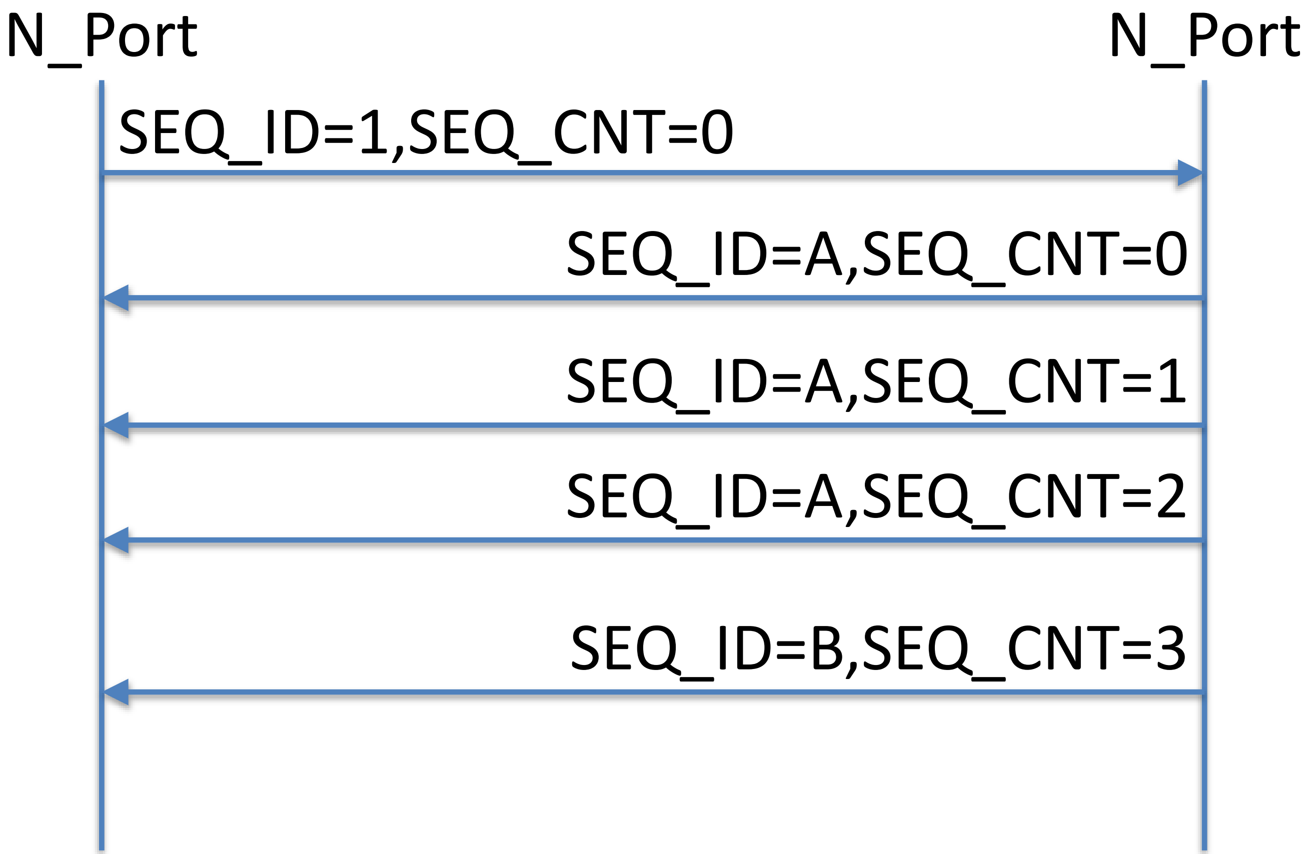
Fibre Channel Frame

- Parameter
  - Different meanings based on frame type
  - For Link_Control frames, it is used to carry information specific to the individual Link_Control frame
  - For Data frames with the Relative offset present bit set to 1, it contains the relative displacement of the first byte of the payload of the frame from the base address as specified by the ULP
  - For Data frames with the Relative offset present bit set to 0, it is interpreted in a protocol specific manner that may depend on the type of Information Unit carried by the frame
    - For example, Task retry identification as specified in FCP-5

Fibre Channel Sequence

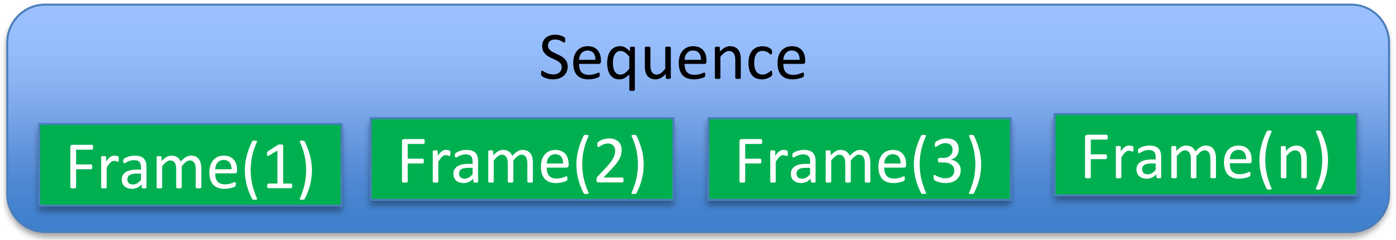
- Sequence – set of one or more Data frames

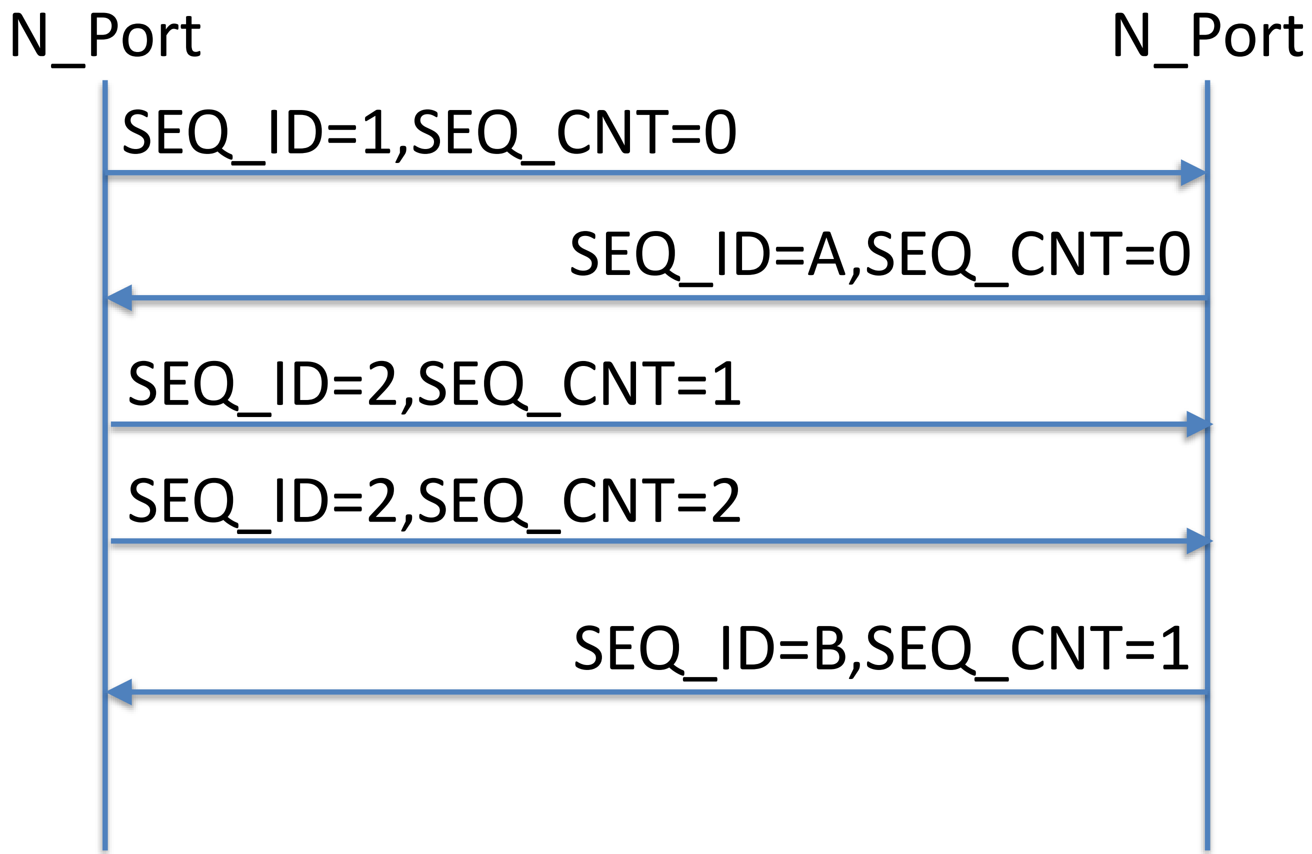




Fibre Channel Sequence

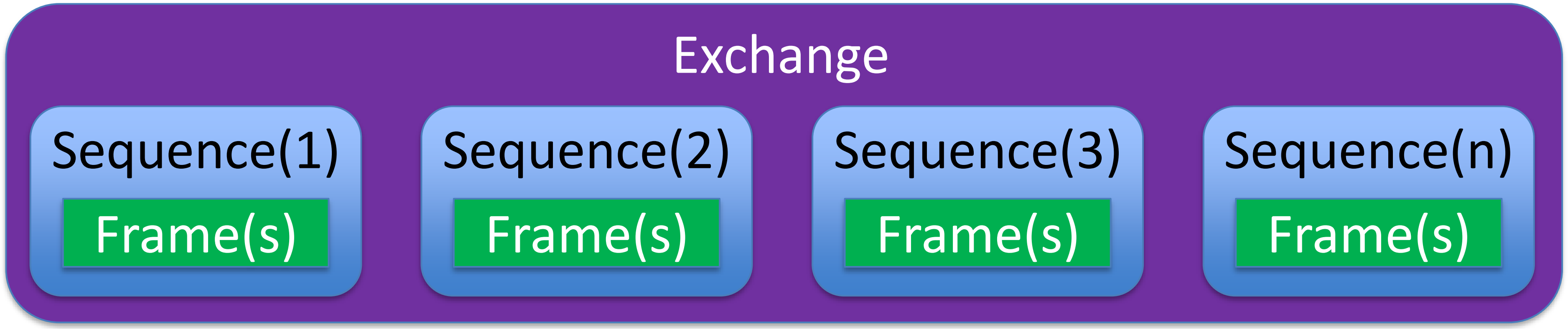
- Sequence – set of one or more Data frames





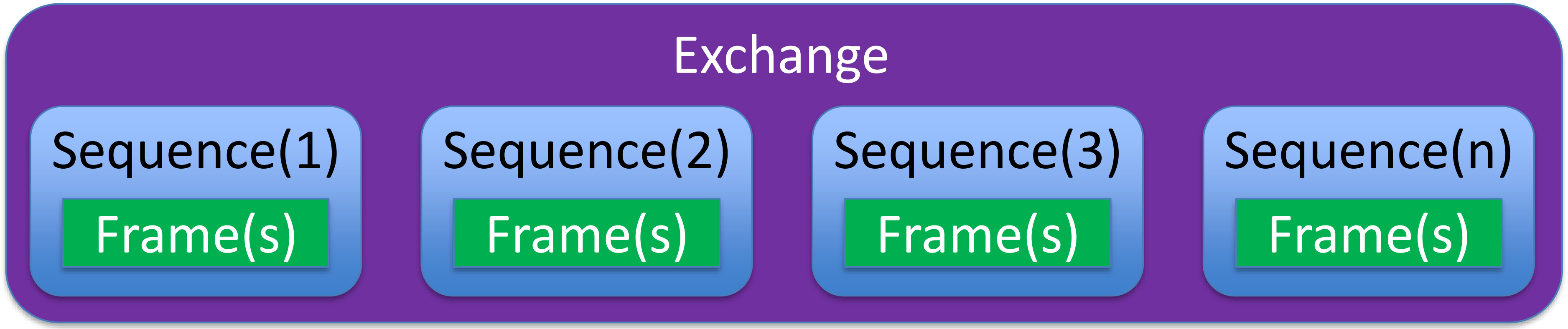
Fibre Channel Exchange

- Exchange – a set of one or more related Sequences for a single operation
  - Fundamental mechanism for coordinating the interchange of information and data between two N_Ports
  - All Data transmission is part of an Exchange



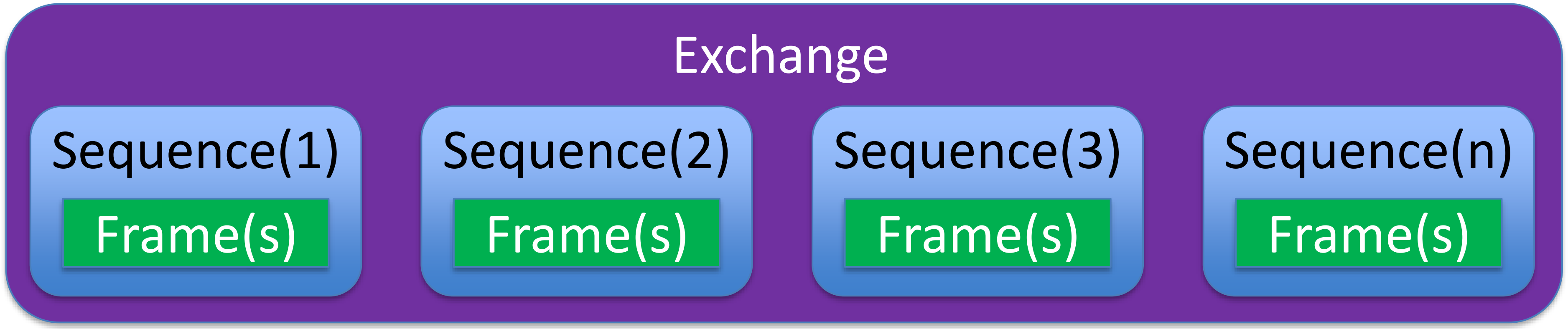
Fibre Channel Exchange

- Exchange – a set of one or more related Sequences for a single operation
  - All frames within an Exchange are delivered in order
  - In the case of a change in a Fabric topology (e.g., a link is removed from or is introduced into a Fabric topology), in order delivery may be temporarily suspended and out of order delivery may occur



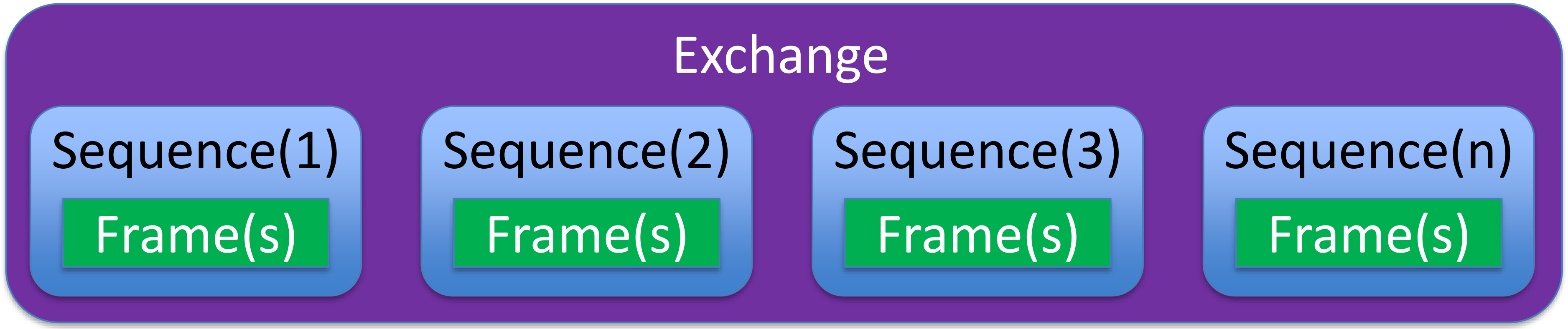
Fibre Channel Exchange

- Exchange – a set of one or more related Sequences for a single operation
  - Sequences for the same Exchange may flow in the same or opposite direction between a pair of N_Ports, but not simultaneously
    - Data flows in one direction at a time within an Exchange for a single N_Port pair



Fibre Channel Exchange

- Exchange – a set of one or more related Sequences for a single operation
  - May be unidirectional or bi-directional
  - Within a single Exchange only one Sequence is active at any given time for a single initiating N_Port
    - Sequence Initiator completes transmission of Data frames for a Sequence before initiating another Sequence for the same Exchange



Fibre Channel Protocols

- Primitive Sequence protocols
  - – Based on Primitive Sequences and specified for Link Failure, Link Initialization, Link Reset, and Online to Offline transition
- Fabric Login protocol
  - – N_Port exchanges Service Parameters with the Fabric
  - – Creates the first VN_Port associated with the PN_Port and the Fabric
  - – Is an explicit procedure (FLOGI ELS) that completes successfully
    - Sent in an Exchange that completes with an LS_ACC

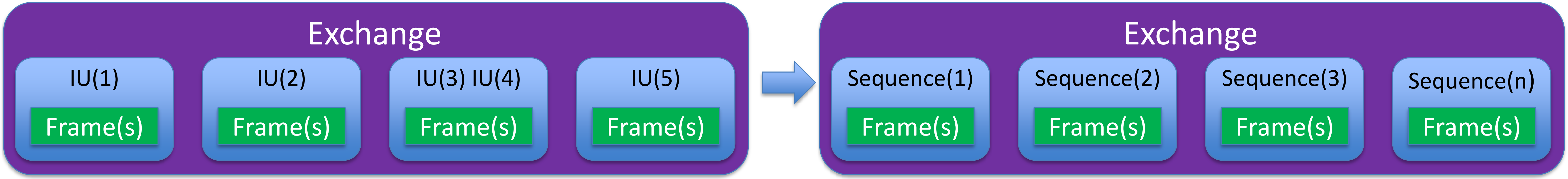
Fibre Channel Protocols

- Additional N_Port_ID protocol
  - Creates additional VN_Ports associated with the PN_Port and the Fabric
  - Is an explicit procedure (FLOGI ELS) that completes successfully
    - Sent in an Exchange that completes with an LS_ACC
- N_Port Login protocol
  - Before performing data transfer, an N_Port exchanges Service Parameters with another N_Port
  - Is an explicit procedure (PLOGI ELS) that completes successfully
    - Sent in an Exchange that completes with an LS_ACC

Fibre Channel Protocols

- Data transfer protocols
  - ULP data is transferred using data transfer protocols specified in FC-4 standards
    - See FCP-5, FC-SB-6, FC-NVMe-2, etc
- Logout protocol
  - Removes Service Parameters from another N_Port or the Fabric
  - Is an explicit procedure (LOGO ELS) that completes successfully
    - Sent in an Exchange that completes with an LS_ACC

FC-4 Information Units (IUs)



- ULP data blocks are mapped to FC-4 Information Units (IUs)
  - This mapping is specific to each FC-4 level standard (e.g. FCP-4, FC-SB-6, FC-NVMe-2)
- FC-4 IUs are mapped to Sequences
  - IUs do not span Sequences
- Sequences are transported in Fibre Channel frames
  - IUs may consist of multiple frames

Routing Control (R_CTL)

- One byte R_CTL field in the frame header is sub-divided into two 4-bit entities:
  - ROUTING: For all FC-4 IUs the ROUTING subfield = (0h)
  - INFORMATION: used at discretion of the FC-4 protocol

R_CTL		Frame Type
ROUTING	INFORMATION	
0h		Device_Data
2h		Extended Link Services
3h		FC-4 Link Data
4h		Video Data
5h		Extended Headers
8h		Basic Link Services
Ch		Link Control
Fh		Extended Routing
Others		Reserved

Device Data Information Categories	
0h	Uncategorized
1h	Solicited Data
2h	Unsolicited Control
3h	Solicited Control
4h	Unsolicited Data
5h	Data Descriptor
6h	Unsolicited Command
7h	Command Status
8h	Extended Command Status
9h	Extended Unsolicited Control
Ah	Extended Solicited Control
Others	

Examples of FC-4 IUs

FCP

R_CTL	IU Type
06h	CMD_IU
05h	XFER_RDY
01h	DATA_IU
07h	RSP_IU
03h	FC_CONF

Information Category

Unsolicited Command

Data Descriptor

Solicited Data

Command Status

Solicited Control

Extended Command Status

FC-NVMe

R_CTL	IU Type
06h	NVMe_CMD_IU
05h	XFER_RDY
01h	DATA_IU
07h	RSP_IU
03h	NVMe_CONF IU
08h	ERSP_IU

FC-SB

R_CTL	IU Type
06h	CMD_IU
05h	XFER_RDY
01h	DATA_IU
07h	RSP_IU
03h	CONFIRM

FCP TYPE ’08’

- FCP → host bus adapter chips with hardware assists and firmware accelerators
  - FCP = transport
  - FCP ≠ SCSI

- Two other ULPs have adopted use of FCP Transport
  - FC-SB Transport Mode
  - FC-NVMe

- Re-use the FC-2 Header TYPE value (08h) and a subset of SCSI-FCP IUs
  - CMD_IU
  - RSP_IU
  - XFER_RDY

FCP and FC-SB Transport Mode CMD_IUs

FCP CMD_IU

Rctl = FC4 Cmd	DID		
CS_CTL = 00	SID		
TYPE = 08			
SEQ_id	DF_CTL	SEQ Count	
OXID		RXID	
PARA			
LUN			
CMD REF ID	Priority / Task Attr	Task Mgmt Flgs	Add CDB Len / R W
Variable Len CDB	CTRL	Misc CDB Info	
Misc CDB Info			Addn'l CDB Len
Service Action Code		Service Action Specific	
Service Action Specific			
Additional FCP CDB (if Any)			
FCP Data Length			

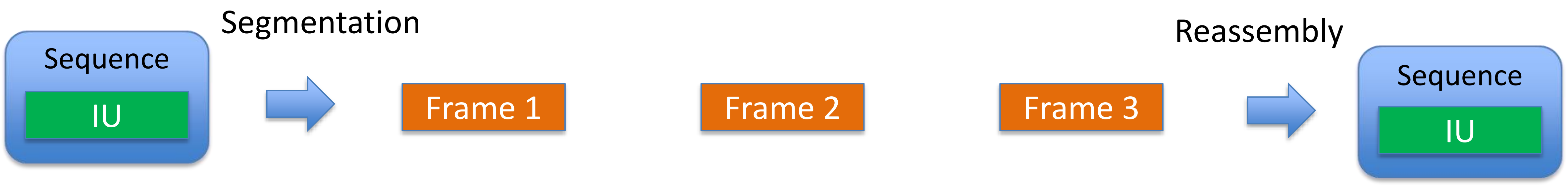
FC-SB Transport Mode CMD_IU

FC-2 Frame Header	Rctl = FC4 Cmd	DID		
	CS_CTL = 00	SID		
	Type = 08			
	SEQ_id	DF_CTL	SEQ Count	
	ODID		RXID	
	PARA			
Target device	FE	CH Image ID	Reserved	CU Image ID
	Device Address			
Command Descriptor Block	00 00 00			L1 R W
	7F	00	Reserved	
	Reserved			L2
	Service Action Code (1FFE, 1FFF)			I/O Priority
	Reserved			
Additional CDB area	DCW 0			
	DCW 0			
	DCW 1			
	DCW 1			
	LRC			
	Data Length			

FC-SB Legend

FC-2 Frame Header	Traanport Command Area Header (TCAH)
SB4 Header	Transport Command Area
Transport Command Header (TCH)	LRC
	Data Length

Segmentation & Reassembly



Two Methods:

1) SEQ_CNT
2) Relative Offset

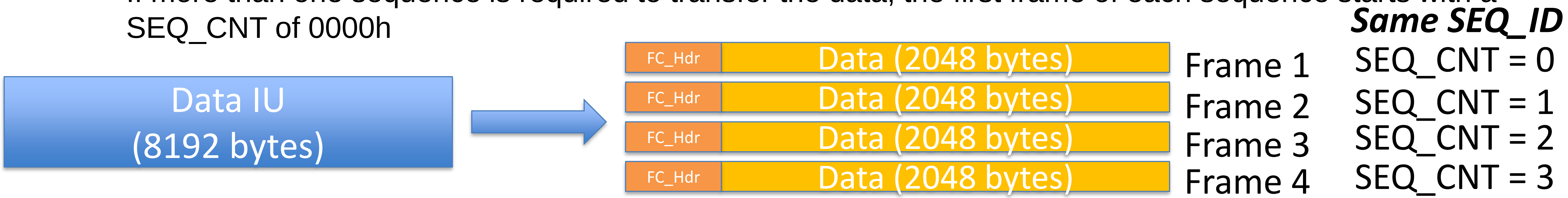
Specified on a per Sequence basis

'Relative Offset Present' bit in F_CTL of FC_Header

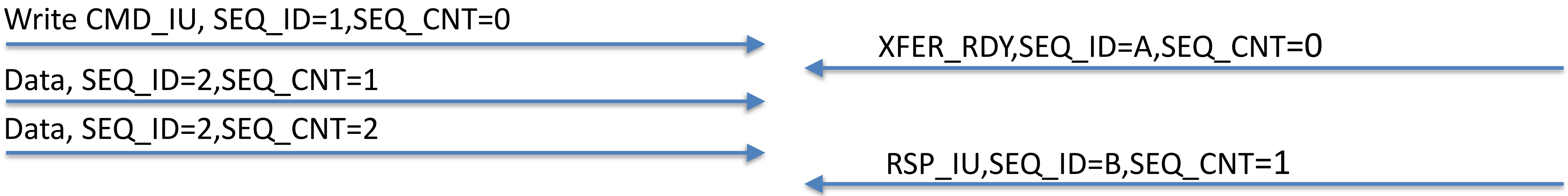
- 0 → Use SEQ_CNT
- 1→ Use Relative Offset

Using SEQ_CNT

- 2-byte field identifies the sequential order of frames within a sequence or multiple sequences of the same exchange
- The SEQ_CNT value for each frame of a sequence must be unique
  - Range is 0 to 65536, can wrap back to 0
- Using 'per sequence' method:
  - If more than one sequence is required to transfer the data, the first frame of each sequence starts with a SEQ_CNT of 0000h



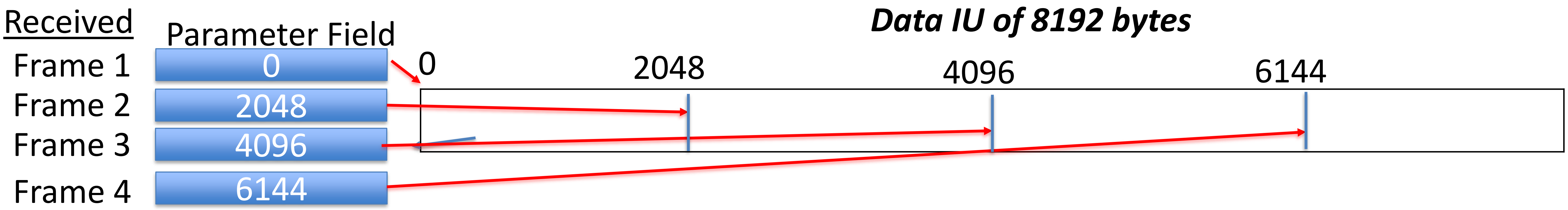
- Using 'per exchange' method:
  - Sequence count increments sequentially across sequence boundaries within an exchange



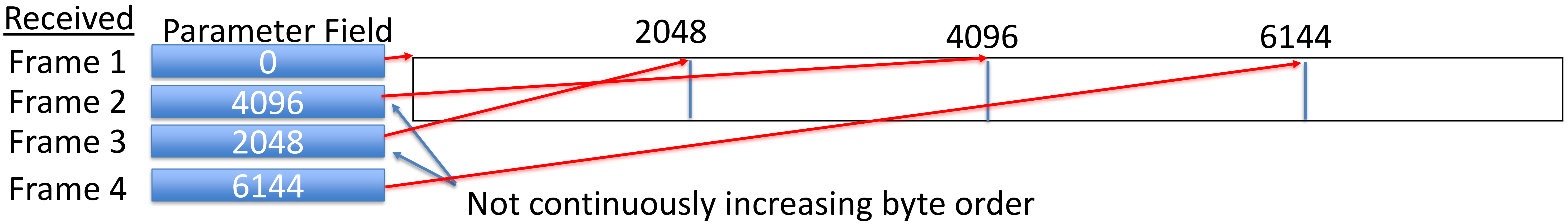
Using Relative Offset

- The Parameter field in the Frame_Header (word 5) specifies the ‘relative offset’ value
- Indicates where within the entire FC data sequence the data in this frame should be placed
- Example:

Remember: the ‘Relative Offset Present’ bit is also specified in each frame



- Optionally Nx_Ports may also indicate support for ‘Random Relative Offset’



Error Detection & Recovery

- Frame errors
  - Missing frames
    - Detected as the sequence times out
  - Corrupted frames
    - Discarded ----error detected at Sequence level
  - Sequence may be aborted at the Sequence level
  - May also cause Exchange errors
  - Error recovery may be performed on the failing Sequence or Exchange with the involvement of the sending upper level

- Link-level errors
  - Result from basic signal characteristics being in question
    - Loss-of-Signal
    - Loss-of-Synchronization and
    - Several link timeout errors that indicate no frame activity.
  - Recovery of Link-level errors involves transmission and reception of Primitive Sequences
    - May introduce Sequence errors that may be resolved after recovery at the link-level.

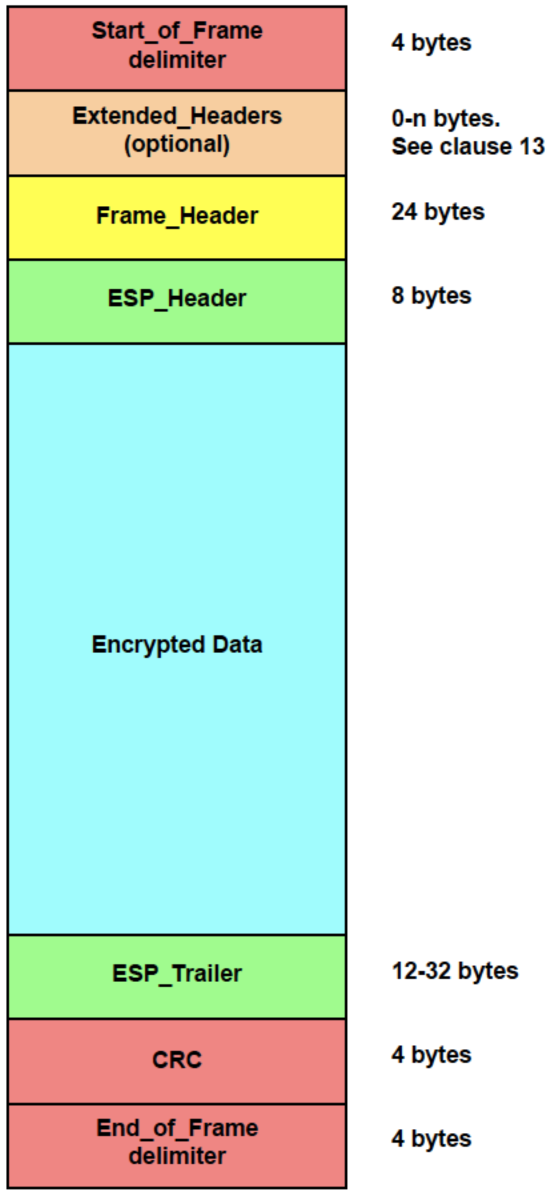
Current Enhancements

- Fibre Channel Security Protocol
- New Speeds
- Sequence Level Error Recovery (SLER)
- Fabric Notification(s)
- VMIDs

FC Security

- Authentication and Encryption have been part of the FC Standards since 2012
  - Defined in FC-SP-2 Standard
  - Uses the same ESP security header as TCP/IP
  - Provides for optional end-to-end authentication and encryption at the transport level (for data on the wire)

ESP Frame Example



- ESP Header (as defined in TCP/IP) is defined as an FC optional header
- Payload is encrypted
- ESP Trailer (as defined in TCP/IP)

Security Update

- TLS 1.0 and TLS 1.1 have been deprecated in favor of TLS 1.2 and the new TLS 1.3
- New Security Project started
  - – FC-SP-2 Amendment 2
  - – Main task is to update the TLS revisions
    - Maintain Fibre Channel security with current standards

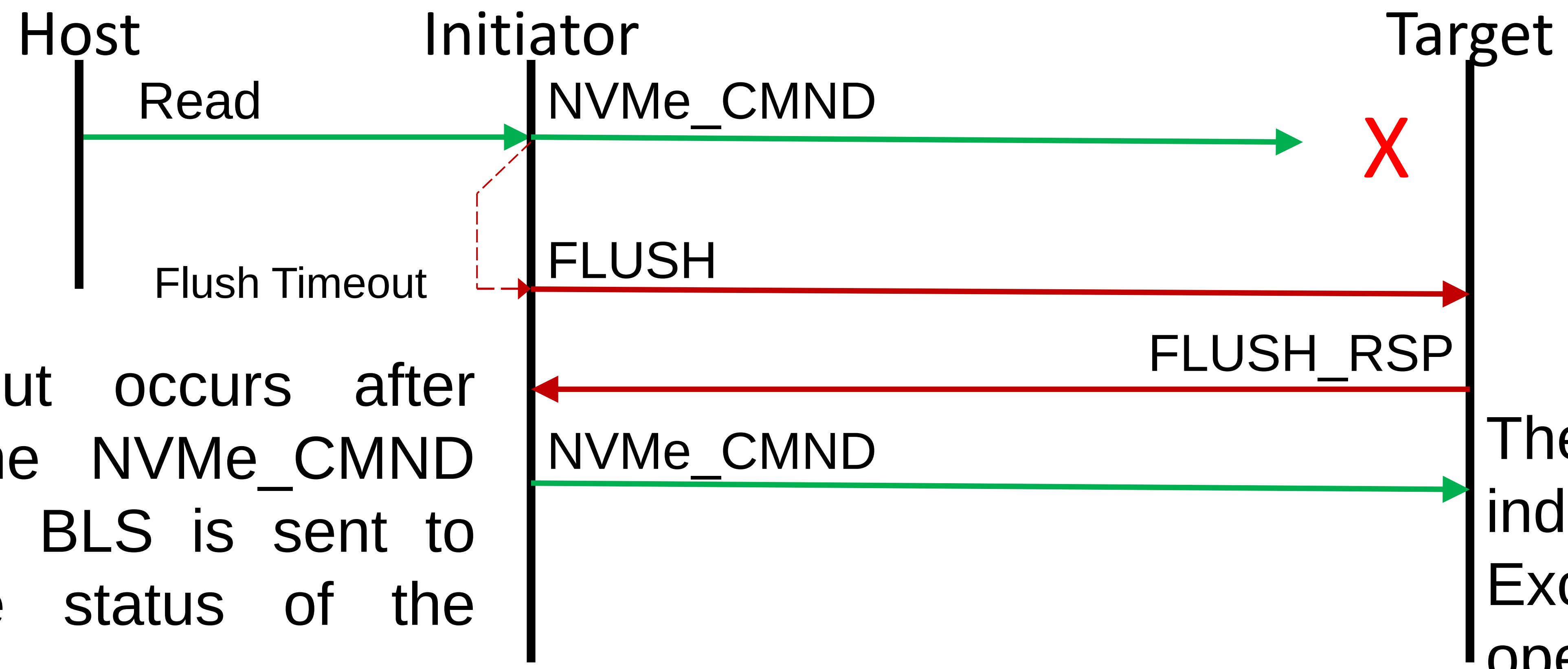
New Speeds

- FC-PI-8 standard being developed
  - Defines 128GFC
    - Based on 100GbE speed as defined in 802.3 standards
  - Continues characteristics of previous Fibre Channel speeds
    - Plug-and-play backwards compatible with at least the last 2 speeds
      - Automatic speed negotiation adjusts to speed
    - Similar distance and cabling requirements as 64GFC
    - Standard completion targeted for early 2022

SLER

- Sequence Level Error Recovery was introduced in FC-NVMe-2
  - Allows for fast, transport level, recovery from frame errors
  - In most cases The Upper Level Protocol doesn’t know any error occurred
- FCP-5, the new revision of FCP (the standard that maps SCSI onto Fibre Channel) is currently in approval process
  - SLER has been ported to SCSI as well
  - While SCSI already does its own error recovery, SLER in FCP-5 is an option mechanism that allows for faster recovery

Lost Command

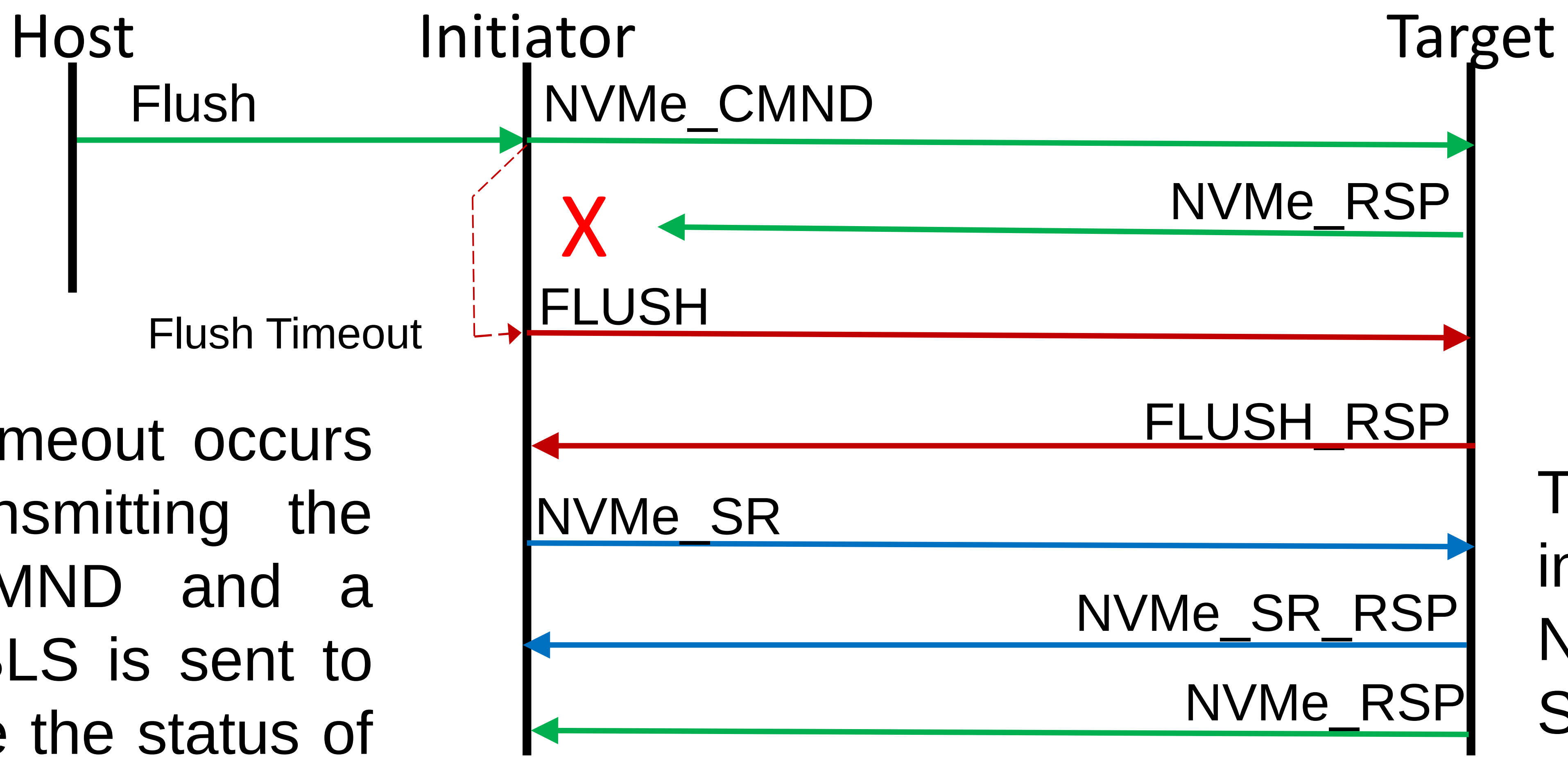


FLUSH timeout occurs after transmitting the NVMe_CMND and a FLUSH BLS is sent to determine the status of the Exchange.

The FLUSH_RSP indicates the Exchange is not open at the target, and the Exchange is closed.

The NVMe_CMND IU is retransmitted using the same OX_ID, SLER qualifier, and CSN.

Lost Response



FLUSH timeout occurs after transmitting the NVMe_CMND and a FLUSH BLS is sent to determine the status of the Exchange.

The FLUSH_RSP indicates the initiator NVMe_Port holds Sequence Initiative and the Exchange is open.

The initiator NVMe_Port transmits an NVMe_SR IU specifying the NVMe_RSP be resent.

Fabric Notifications

Fabric Notifications

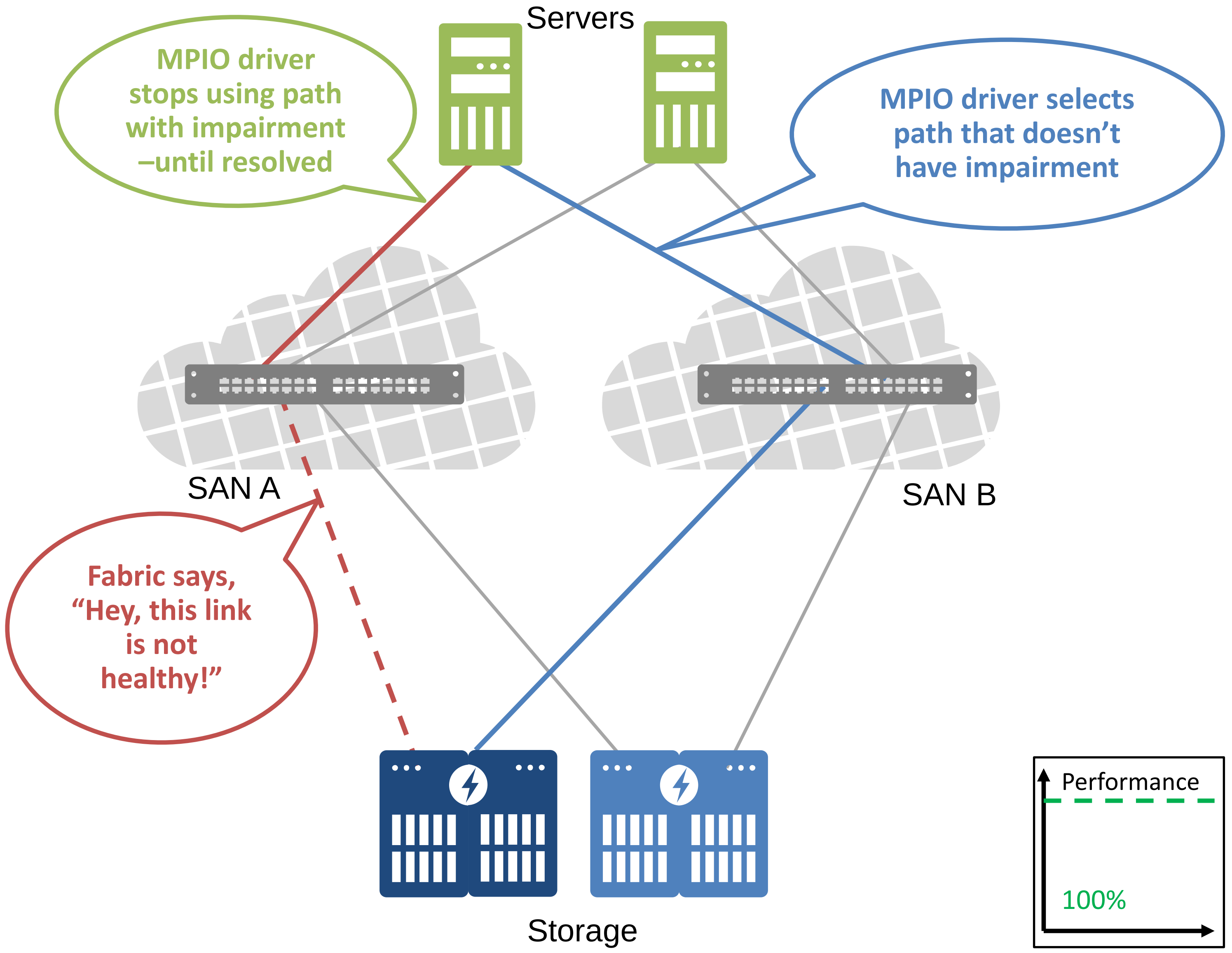
- Notifications and signals
  - Generated by the fabric
  - Inform devices of impairments

Notifications

- Reporting: Events sent to registered devices
- Diagnostics: Helps efficiently evaluate errors
- Operation: Extended Link Services (ELS)

Signals

- Signaling: Report resource depletion to registered devices
- Diagnostics: Transmitter indicates resource usage
- Operation: Link level Primitive Signal



Fabric Notification

History

November 2014

– Fibre Channel ecosystem investigations

2015-2017

– Research and experimentation

2018

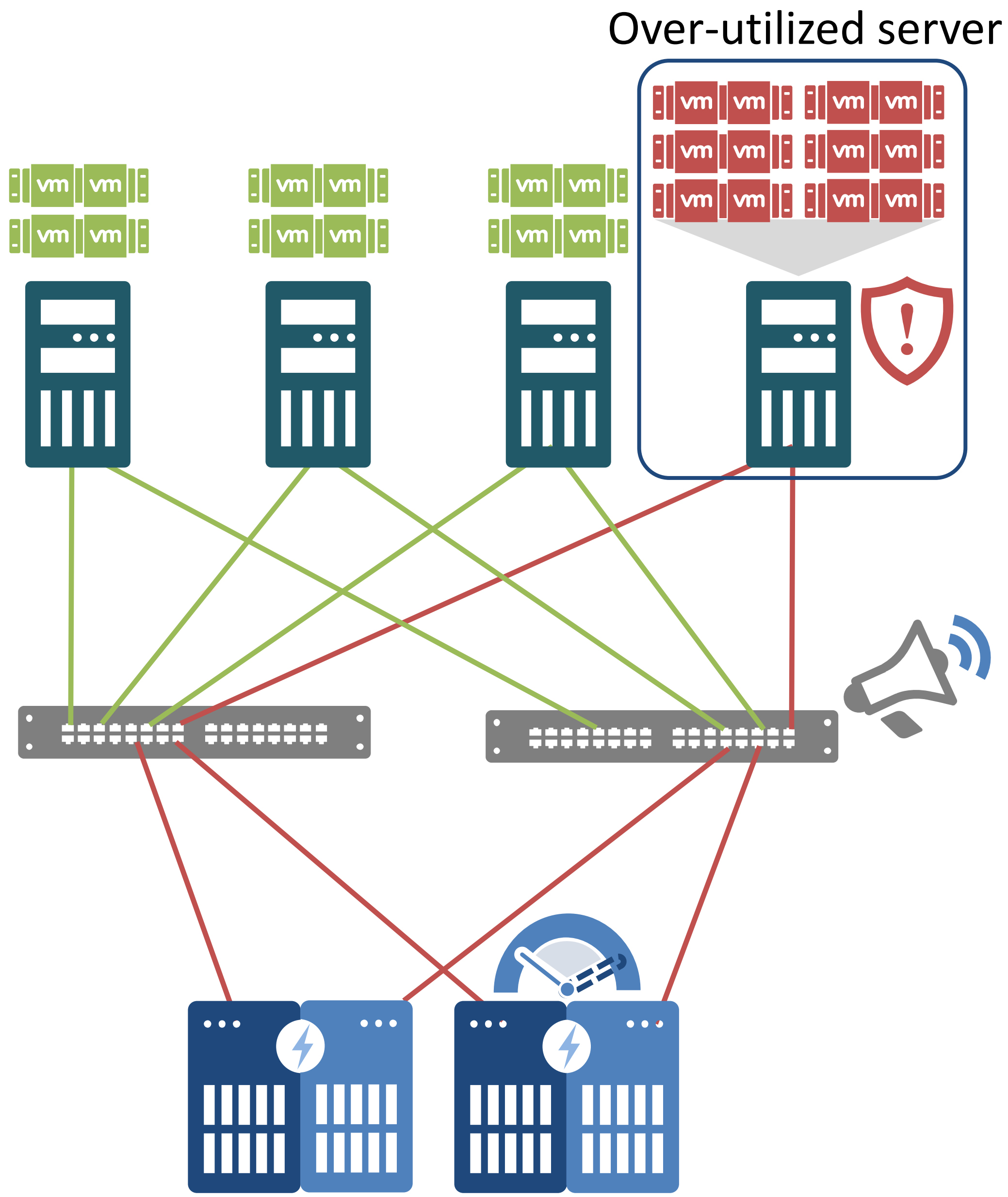
– Fibre Channel ecosystem collaboration

– Standardization starts

2019-2021

– Accepted into the T11 Standards

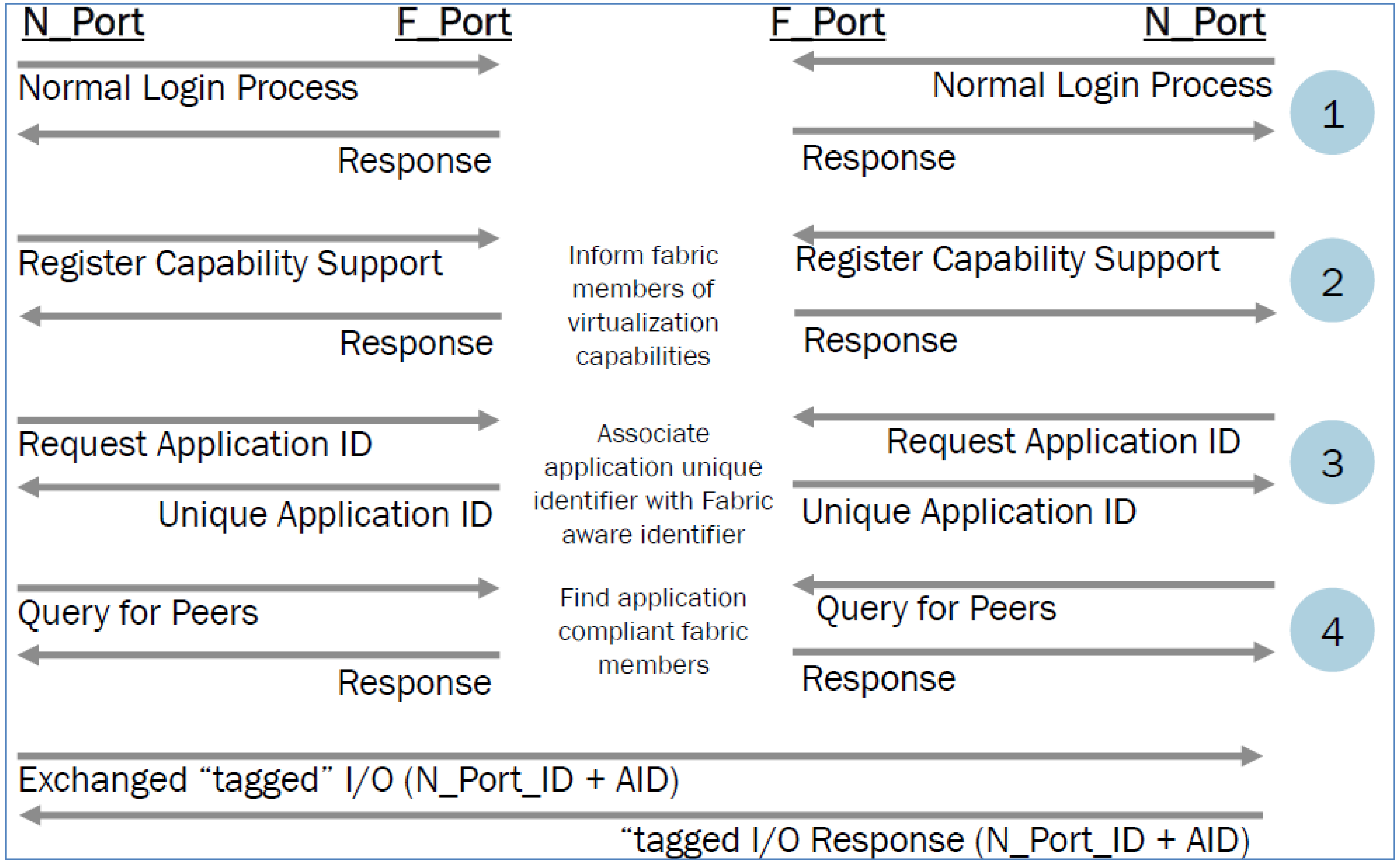
- FC-FS-6: Congestion Signals (r0.3)
- FC-LS-5: Notifications (r5.01)
- FC-SW-8: Fabric detection and generation (r1.01)



Application_Header

Virtual Machine example

- Normal Fabric Login
- N_Port Login
  - Enhanced for quick detection of supporting devices
- FC-4 Type Registration
  - Unique type for Application Services
- VM Registration
  - Allocation of tags for each VM
- FC-4 Type Query
  - Identify VM peers
- VM Tagged Flows
  - Fabric unique identifier



After this Webcast

- Please rate this event – we value your feedback
- We will post a Q&A blog at http://fibrechannel.org/ with answers to the questions we received today
- Follow us on Twitter @FCIAnews for updates on future FCIA webcasts
- Our library of FCIA on-demand webcasts http://fibrechannel.org/webcasts/ includes:
  - Fibre Channel Fundamentals
  - FC-NVMe
  - Long Distance Fibre Channel
  - Fibre Channel Speedmap
  - FCIP (Extension): Data Protection and Business Continuity
  - Fibre Channel Performance
  - FICON
  - Fibre Channel Cabling
  - 64GFC
  - FC Zoning Basics

Thank You

FIBRE CHANNEL INDUSTRY ASSOCIATION