



Education

FCoE Direct End-Node to End-Node (aka FCoE VN2VN)

John L Hufferd, Hufferd Enterprises

- ◆ The material contained in this tutorial is copyrighted by the SNIA unless otherwise noted.
- ◆ Member companies and individual members may use this material in presentations and literature under the following conditions:
 - ◆ Any slide or slides used must be reproduced in their entirety without modification
 - ◆ The SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.
- ◆ This presentation is a project of the SNIA Education Committee.
- ◆ Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.
- ◆ The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.

- A new concept has recently been accepted for standardization by the FC-BB-6 Working Group within the Fibre Channel (T11) standards committee; it is called FCoE VN2VN (aka Direct End-Node to End-Node)
- T11 previously standardized the FCoE specification (which defines the encapsulation of Fibre Channel frames within Ethernet Frames) and is currently extending that specification to permit FCoE connections **DIRECTLY** between FCoE End-Nodes
- This tutorial will show the fundamentals of the extended FCoE concept that permits it to operate without FC switches or FCoE Switches (aka FCFs) and will describe how it might be exploited in a Data Center environment

- **Introduction**
- **FCoE & FCoE VN2VN**
- **Architecture**
- **Discovery & Link Instantiation**
- **Topologies**
- **Scenarios**
- **Summary**

- This presentation provides an overview of a new proposed Standard called (herein) FCoE Direct End-Node to End-Node (aka FCoE VN2VN)
 - ◆ This is a Lossless Ethernet connection directly between Adapters' Virtual N_Ports
- One should think about FCoE VN2VN as placing the FCoE (FC) protocol on a Lossless Ethernet without the additional requirement of FC Switches or FCoE Switches known as FC Forwarders (FCFs)
 - ◆ Permits connections through only (Lossless) Ethernet Switches
 - ◆ Permits connections via a single wire Point-to-Point
- The protocol is being defined in the INCITS Fibre Channel (T11) technical committee (FC-BB-6 Ad Hoc Work Group)

FCoE & FCoE VN2VN

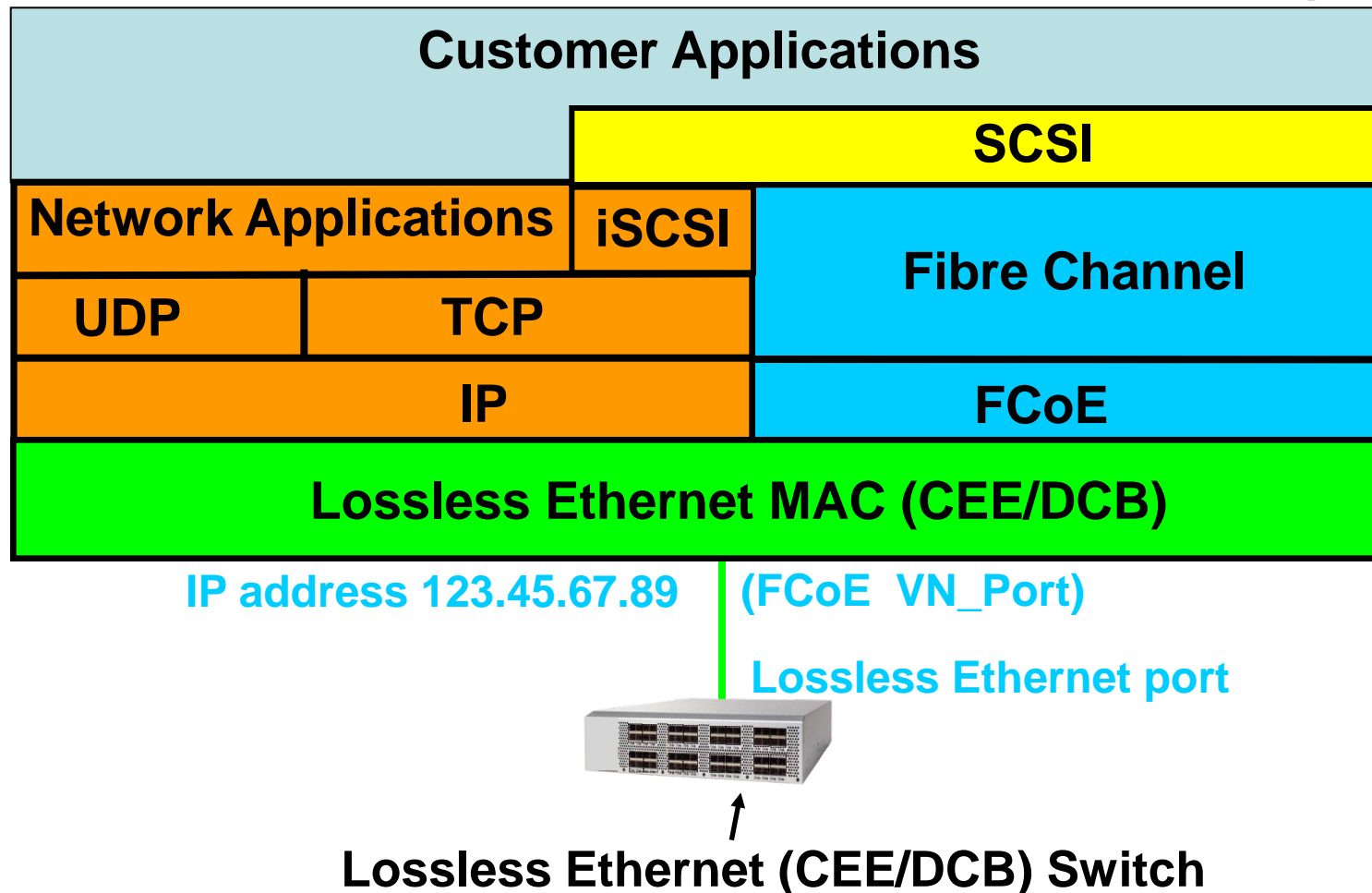
FCoE is an Alternative to FC

- **FCoE stands for FC over Ethernet**
- **FCoE was defined as an alternative network structure for carrying FC protocols**
- **FCoE requires specific Ethernet extensions to be implemented**
 - ◆ Lossless switches and fabrics (e.g., supporting IEEE 802.3 PAUSE) configurations are required
 - ◆ Jumbo frame support is strongly recommended (not a standard, but widely available)
 - ◆ **Deployments of FCoE should utilize the advances in Ethernet currently specified in IEEE 802.1, specifically:**
 - › Priority-based Flow Control (PFC) → 802.1Qbb
 - › Enhanced Transmission Selection (ETS) → 802.1Qaz
 - › DCB (capability) eXchange (DCBX) Protocol → 802.1Qaz
 - › Congestion Notification (802.1Qau),
 - ◆ **Possible** future → Multi-pathing (IETF– TRILL, IEEE 1aq, et.al.)
 - ◆ **These 802.1 advances are important for Converged Flows (Messaging, Clustering and Storage)**
This set of functions is called CEE – Converged Enhanced Ethernet (intended for a Data Center Environment) or (in the IEEE) DCB -- Data Center Bridging
- **FCoE Fabrics require an FCoE Lossless Ethernet Switch that understands & supports FC protocols – These Switches are called FCFs (FCoE Forwarders)**

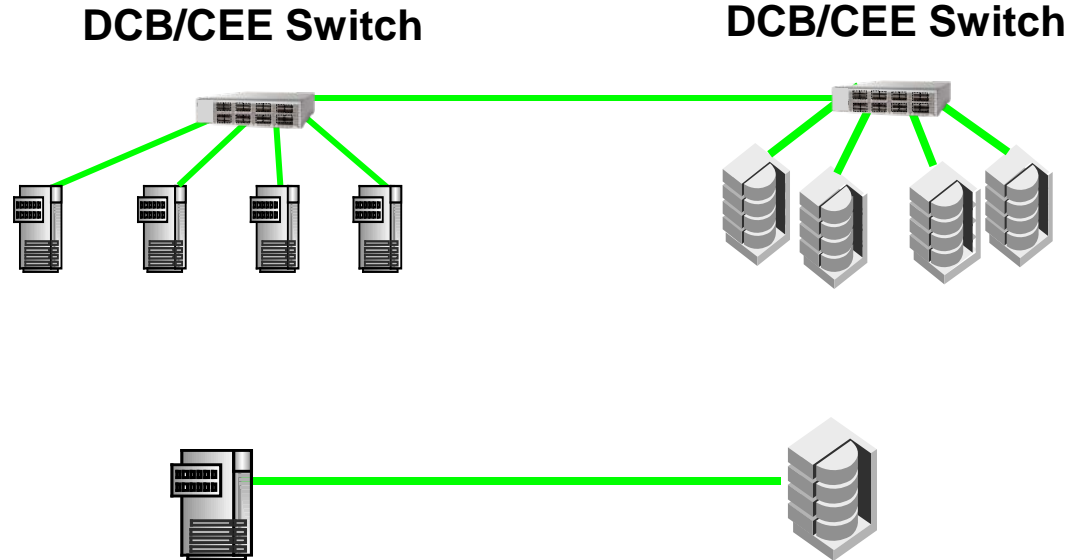
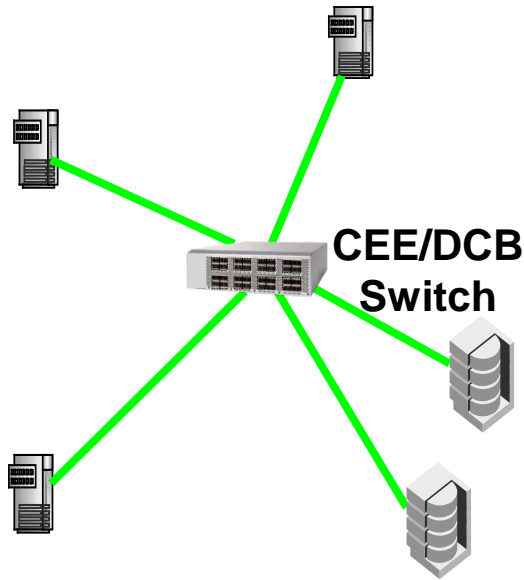
- FCoE VN2VN is a Lossless Ethernet connection between End-Node Adapters' VN_Ports
- Other than Ethernet Cables, only CEE/DCB Ethernet switches may exist between the End-Nodes (VN_Ports) – Therefore, the connection maybe either a:
 - ◆ Switched Lossless Ethernet connectionOr
 - ◆ Point to Point Lossless Ethernet Wire connection
- FCoE VN2VN permits FCoE networks to be built without any FC Switches or FCoE Switches (aka FCFs)
- FC Data Flow (& Packets) will flow End to End as if they were flowing over a direct (point to point) FC link
 - ◆ Must operate identically on a VN2VN connection as on a direct FC link
- **No Fibre Channel services or advanced features** (e.g. Name services, **Zoning**, virtual fabrics, IFR, security, etc.) are provided in the network
 - ◆ Use of VLANs, and ACLs, can help security issues
 - ◆ Security can also be enhanced via FC-SP

Connections to a CEE/DCB Switch

- Fibre Channel is carried over lossless Ethernet as a L3 protocol



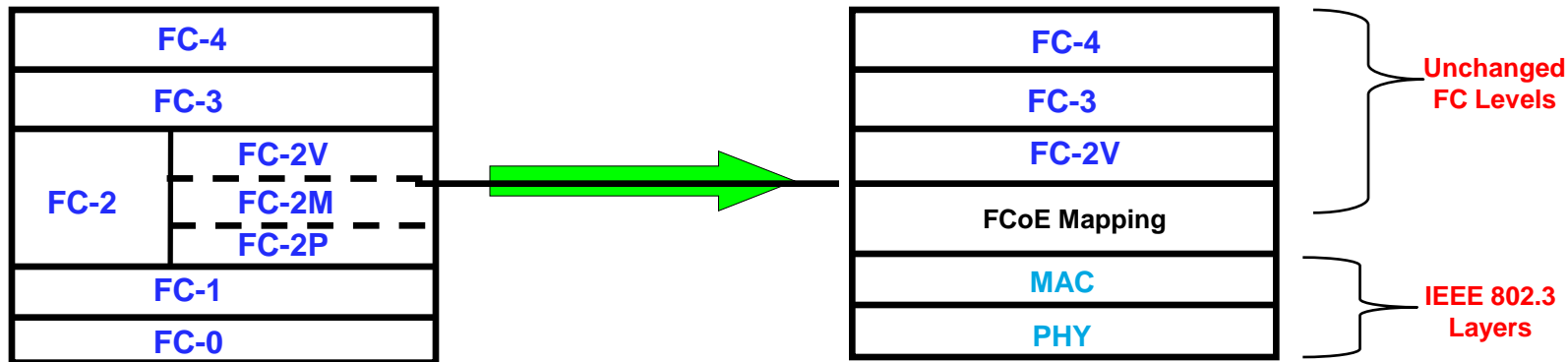
The Simple VN2VN Interconnect



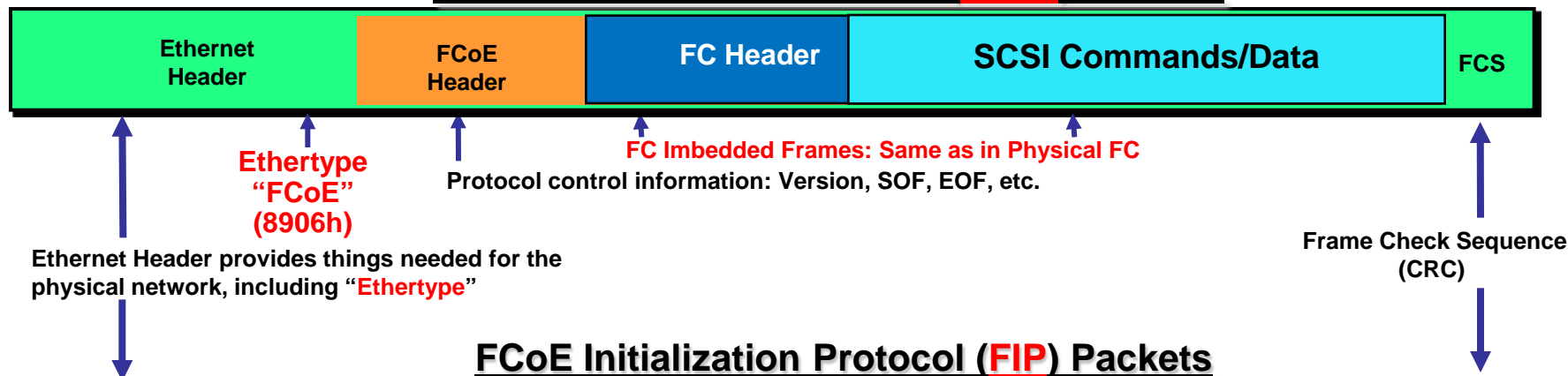
**An CEE/DCB – Switch(es) may connect a number of VN2VN capable VN_Ports together
Or
A single Wire with VN2VN capable End Nodes can be interconnected**

Architecture

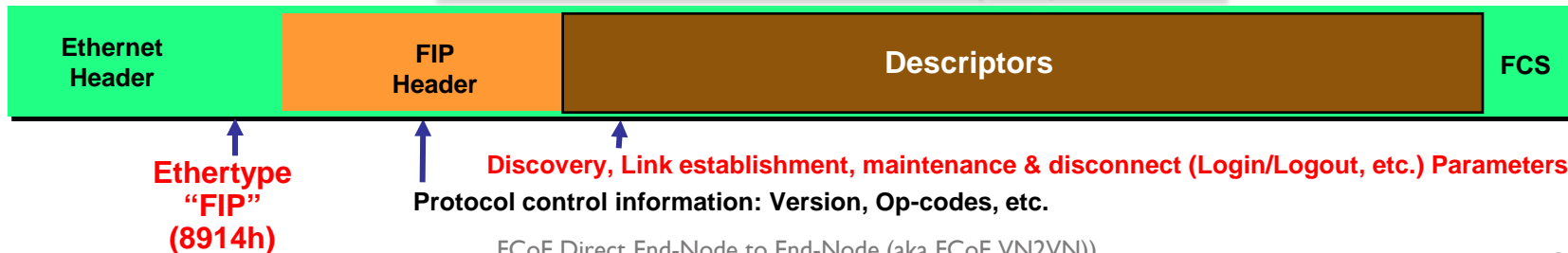
FC Encapsulation Into Ethernet Frames (2 FCoE Related Packet types)



Fibre Channel over Ethernet (FCoE) Packets



FCoE Initialization Protocol (FIP) Packets





➤ Discovery Phase

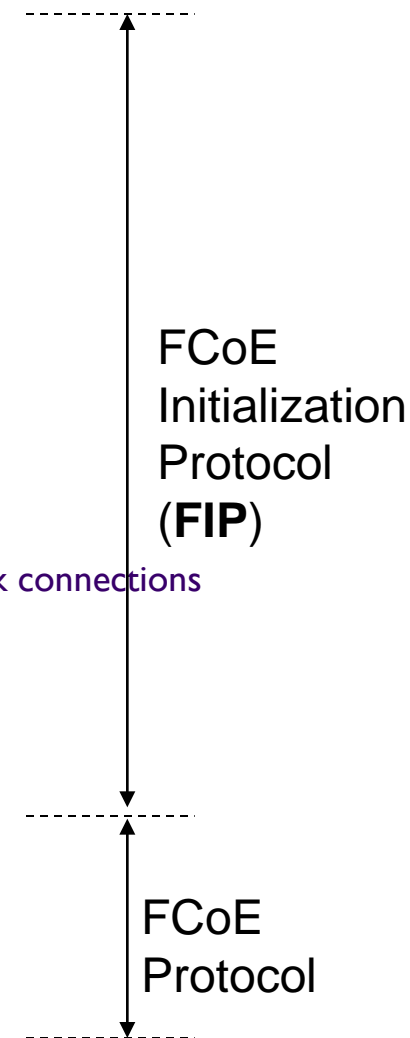
- ◆ FCoE (w/o VN2VN)
 - › FCFs Discover each other, & form a Fabric
 - › ENodes Discover FCFs & Potential VN_Port ← → VF_Port pairing
- ◆ FCoE VN2VN
 - › VN2VN capable ENodes Discover each other

➤ Login Phase

- ◆ FCoE (w/o VN2VN)
 - › ENodes chose among discovered FCFs' Ports for Virtual Link connections
- ◆ FCoE VN2VN
 - › VN2VN capable ENodes chose among discovered VN2VN Ports for Virtual Link connections
- ◆ Both Use: FLOGI, FLOGI ACC, LOGO, etc ...

➤ End-to-End path control & Data Transfer Phase

- ◆ PLOGI/PRLI
- ◆ All other FC protocol frames (FC4 ULPs. etc.)



FC's Encapsulation in Ethernet (FCoE)

Word	31-24	23-16	15-8	7-0
0	Destination MAC Address (6 Bytes)			
1				
2	Source MAC Address (6 Bytes)			
3	ET=FCoE (16 bits)		Ver (4b)	Reserved (12 bits)
4	Reserved			
5	Reserved			
6	Reserved			SOF (8 bits)
7	Encapsulated FC Frame FC Frame = Minimum 28 Bytes (7 Words) → Maximum 2180 Bytes (545 Words) (including FC-CRC)			
...				
n				
n+1	EOF (8 bits)	Reserved		
n+2	Ethernet FCS			

Optional IEEE 802.1q
4 Byte Tag goes here →

This field varies
In size →

Ethernet frame
Size is 64 Bytes to
2220Bytes →

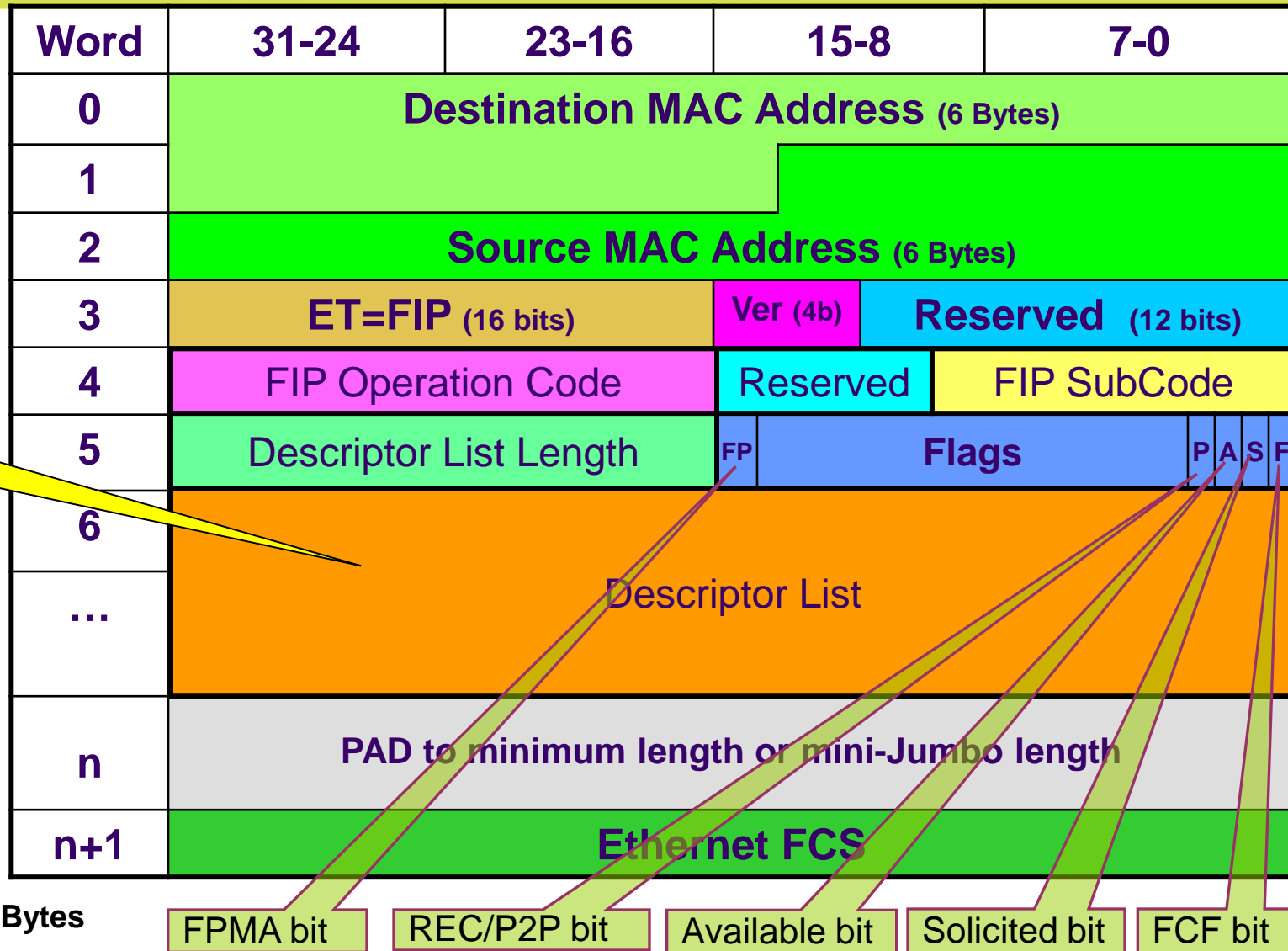
FIP Operation Format

Optional IEEE 802.1q
4 Byte
Tag goes here

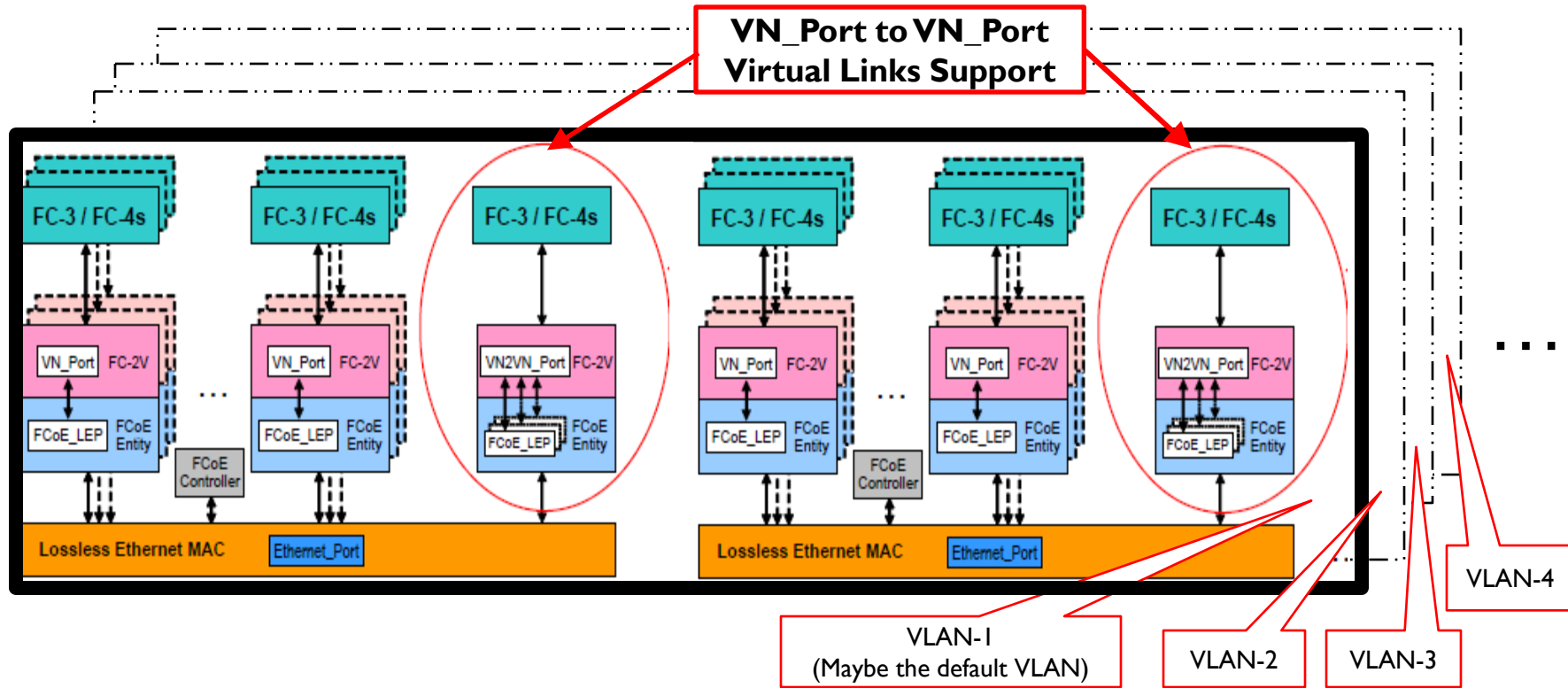
See Appendix
Below for
Descriptor list items

Descriptor list
varies
In size

Ethernet frame
size
Is 64Bytes to 2220Bytes



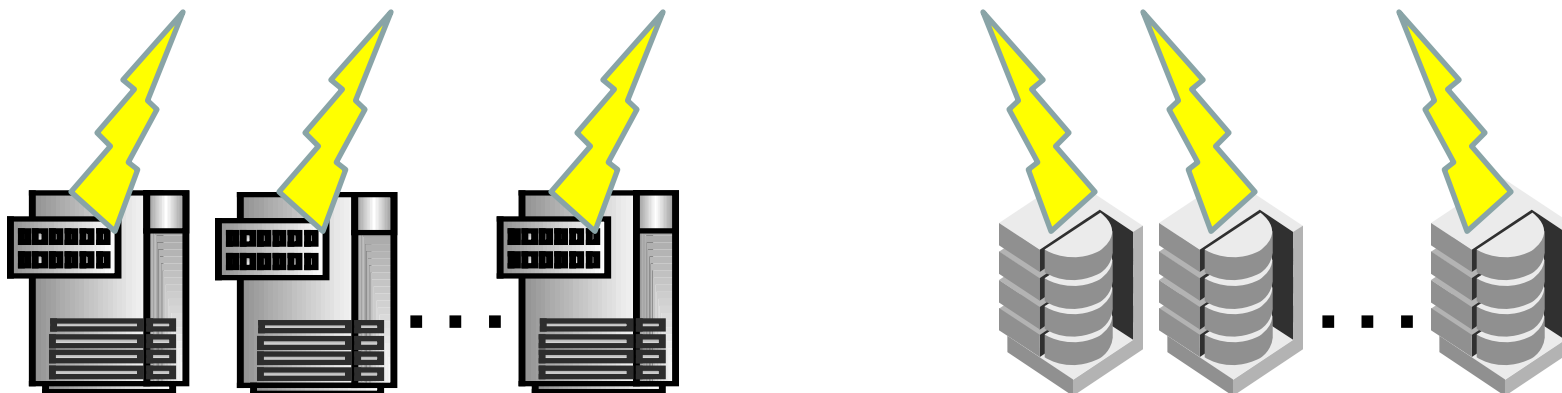
ENode (HBA) Model



Each ENode (HBA/CNA) may have multiple Physical Ethernet Ports
Each Physical Port may have multiple Logical VN_Ports

Each instantiation's N_Port_ID & MAC Address is independent of the others
There can be duplicates (if they are in different VLANs)

Discovery and Link Instantiation (FIP -- FCoE Initiation Protocol)

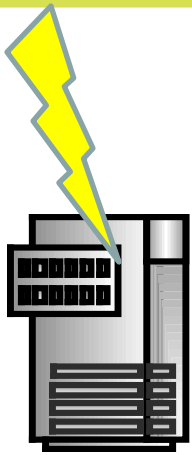


1. As each VN2VN enabled FCoE End-Node starts-up it will Randomly Generate its own:
 - FC LU-ID (N_Port_ID) and Ethernet MAC Address
2. Then Each VN2VN End-Node, on behalf of its VN_Port, Multicasts a PROB with the Generated MAC Address & LU-ID
 - And listens for conflict responses
3. If Address/ID Conflict message is received, the process will repeat (at step1) until no Conflict messages are received
4. If no Conflicts are received, the End-Node instantiates the VN_Port and Multicasts its CLAIM to the MAC Address & LU-ID (N_Port_ID) and announces its “Capabilities” (FC-4 Features)

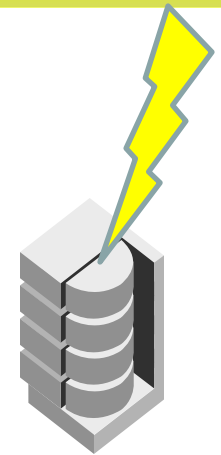


PROBE Responses

Neighbor Table



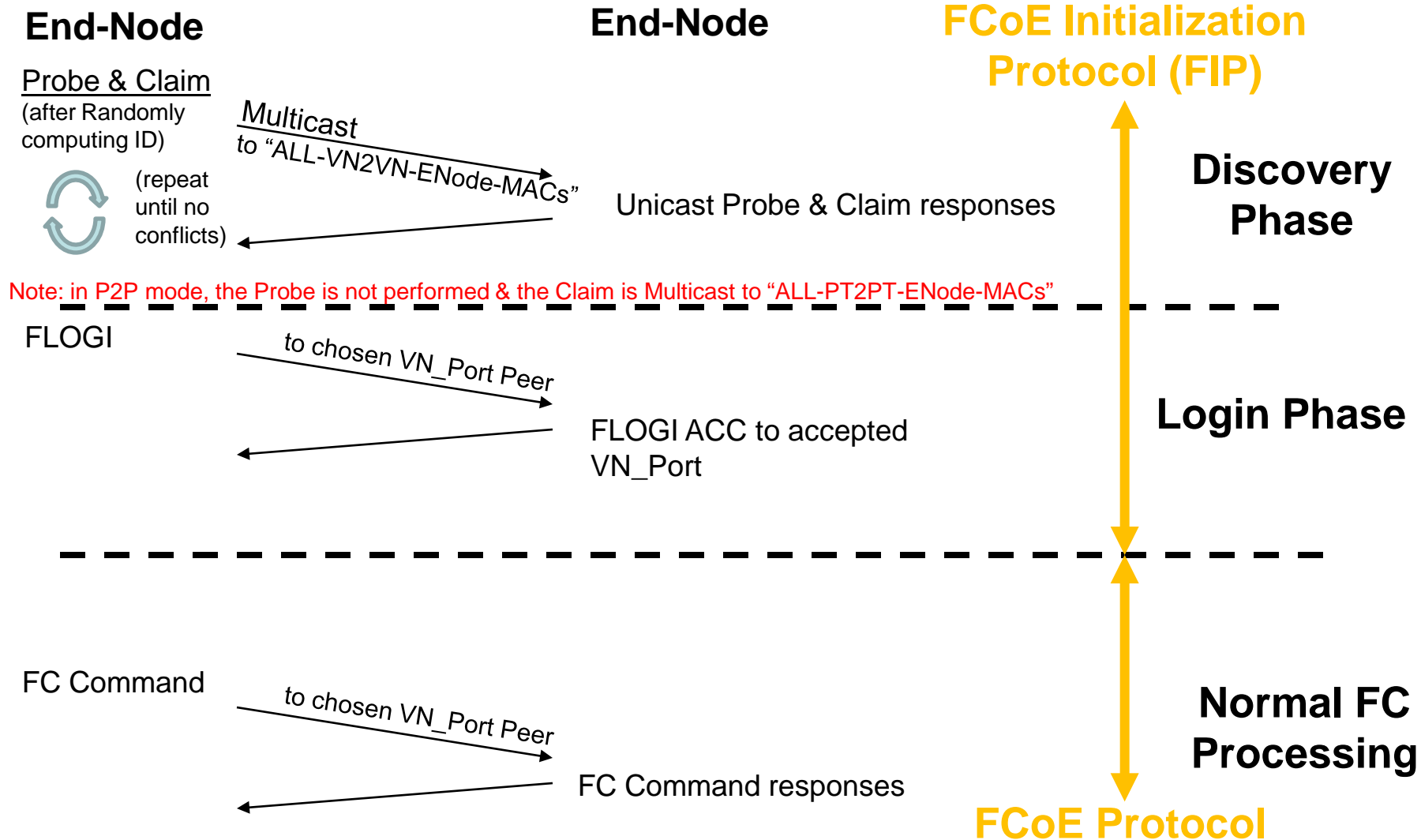
Name	MAC Address	LU-ID	Capabilities (FC-4 Features)



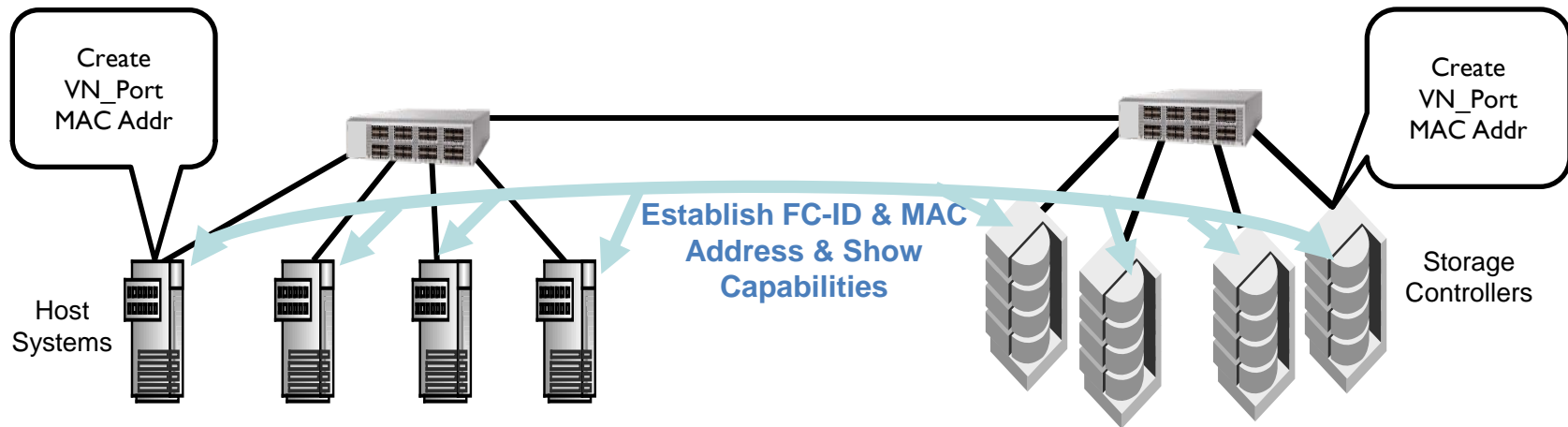
4. Each VN2VN enabled FCoE End-Node receiving a CLAIM will respond with its own information & record the received CLAIMed information into a Neighbor Table – Including:
 - N_Port_Name, MAC Address & LU-ID and
 - Capabilities of the CLAIMing VN_Port (Initiator/Target, etc.)

5. Upon receiving CLAIM response messages, the CLAIMing End-Node will record the received CLAIM response information into its own Neighbor Table – Including:
 - N_Port_Name, MAC Address & LU-ID and
 - Capabilities of the responding VN_Port (Initiator/Target, etc.)

VN2VN Initial Login Flow Ladder



Summary of Discovery Process (for VN₂VN)



Each VN₂VN Capable End-Node will (for each of its VN_Ports):

- Randomly chooses an identity (FC-ID & MAC Address)
- Multicast (& Respond with) FIP messages (called "Probe, Probe Response, Claim, & Claim Response")

which:

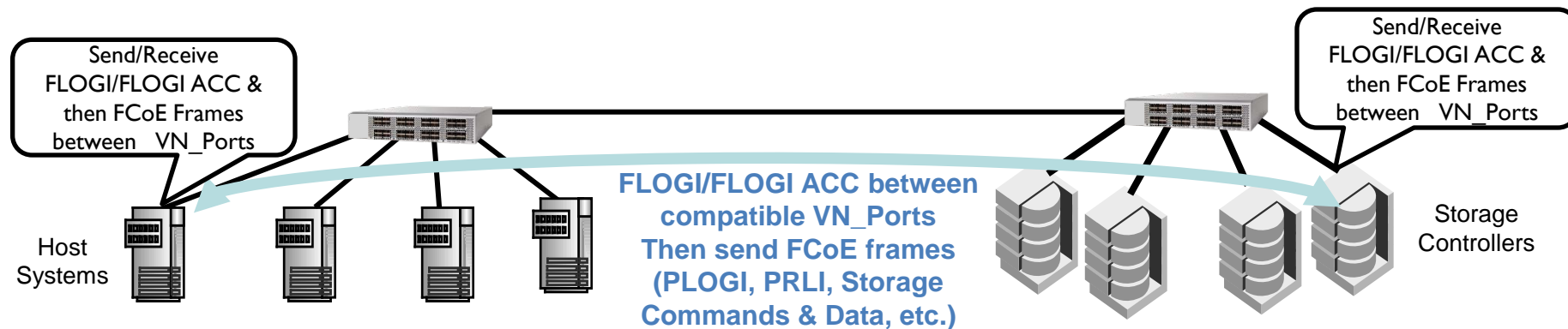
- Insure ID uniqueness
(repeat the Random ID creation and the Probe & Claim as needed)
- Permit population of the neighbor tables

This will announce the VN_Ports' identities and the VN_Ports' capabilities to other ENodes

- The capabilities will be used for choosing a peer VN_Port

Note: FC-ID & MAC Address should be saved, if possible, for next Reboot

Summary of Link Instantiation (for VN₂VN)



After IDs and Potential Partners (VN_Ports) are identified within the Level 2 Ethernet:

- The Initiators chose their Targets and FLOGI & FLOGI ACC FIP frames are exchanged
- Then PLOGI/PLOGI ACC & PRLI/PRLI ACC FCoE Frames are exchanged
- Thus Instantiating the VN₂VN Logical Link

→ Then FCoE frames will be exchanged directly between the VN_Ports ←

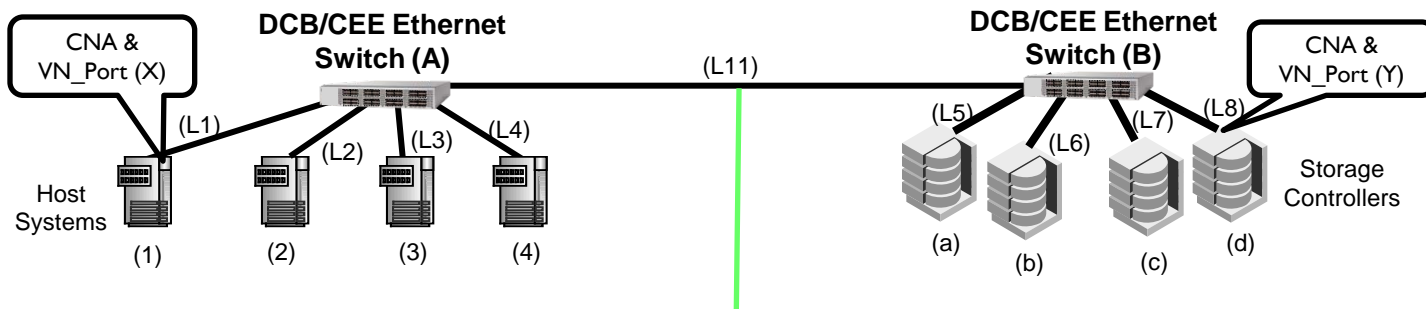


After Link Instantiation all VN_Port's IDs will be Periodically Beaconsed (Multicast)

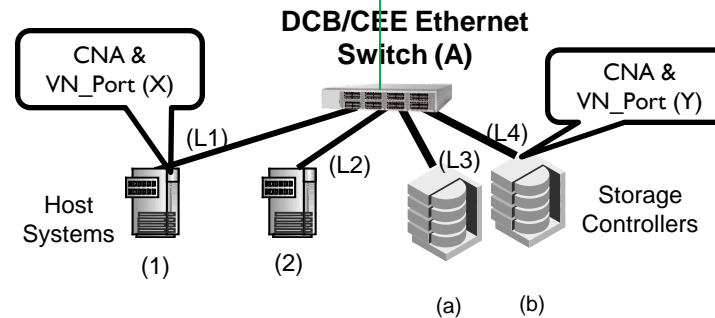
(Beaconing permits detection of link loss (via time-outs) & new or incorrect LAN joins)

Topologies

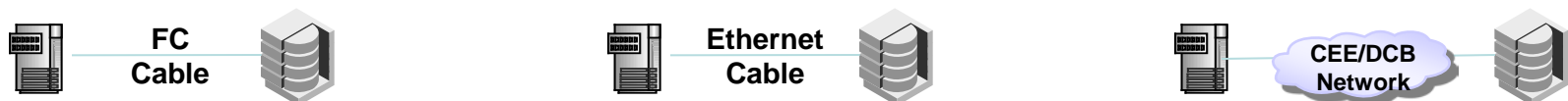
Direct VN_Port to VN-Port (VN2VN)



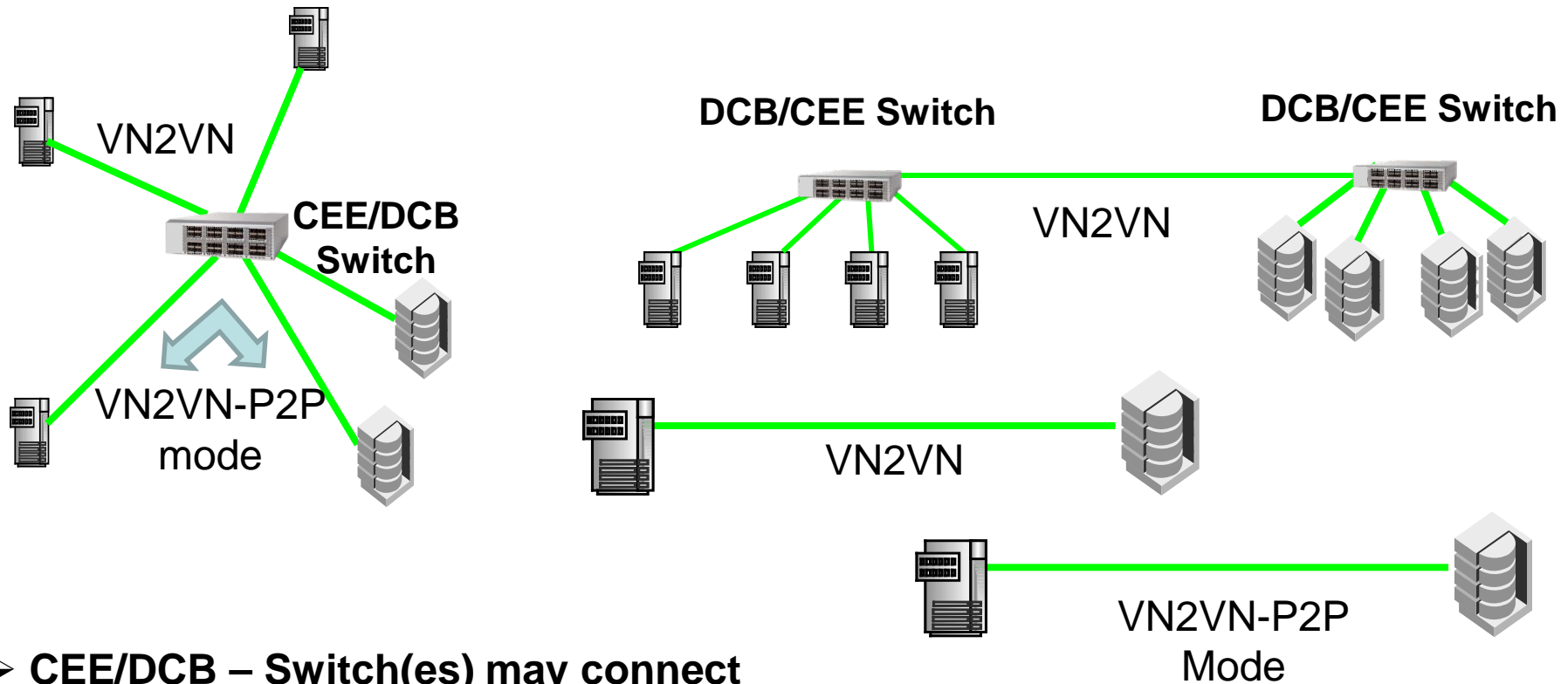
A CNA to CNA FCoE path between these Switch ports is now also possible even without an FCF using Direct VN2VN mode



It is now possible to connect End-to-End as shown below



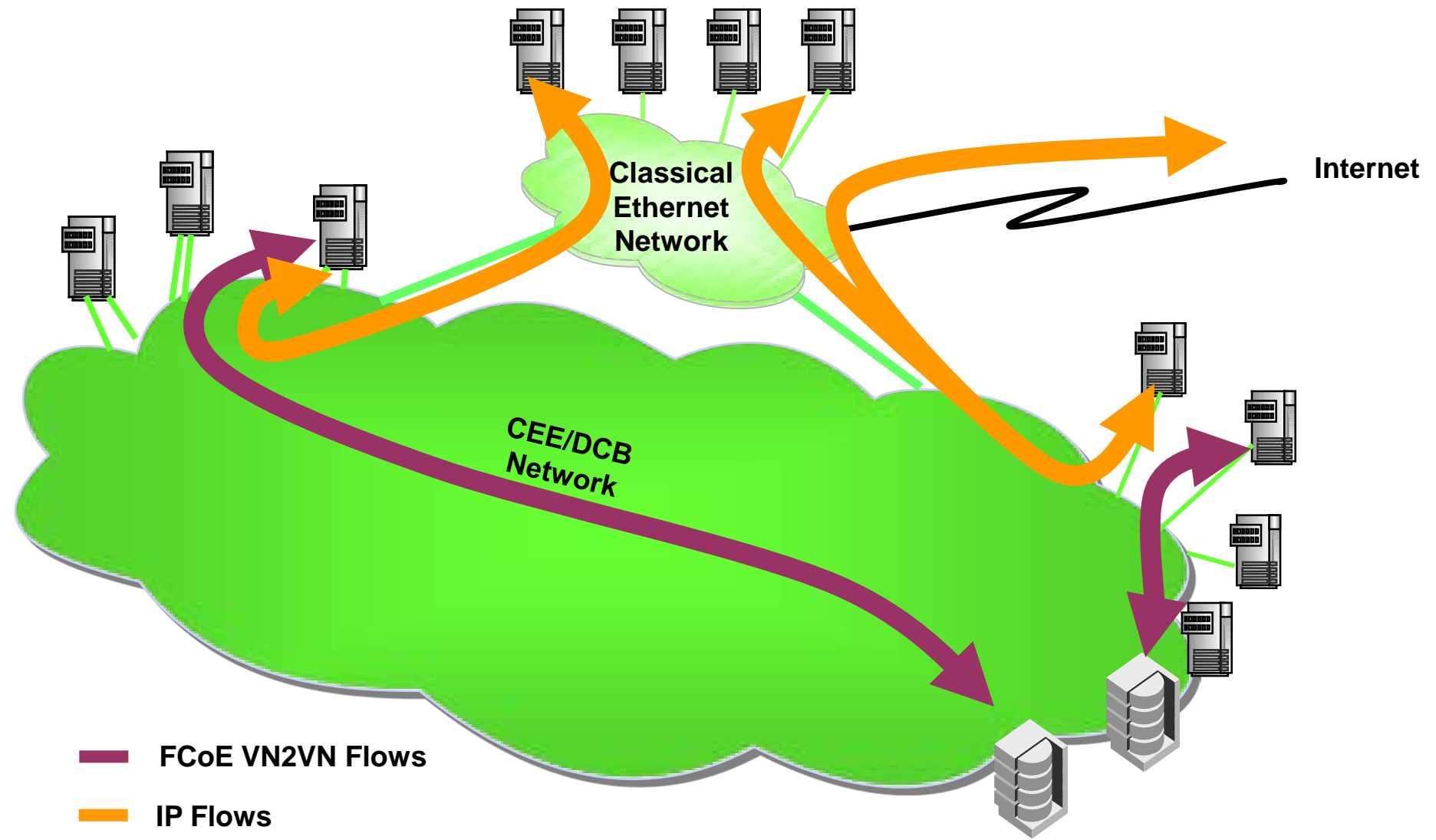
VN2VN Interconnect Including P2P



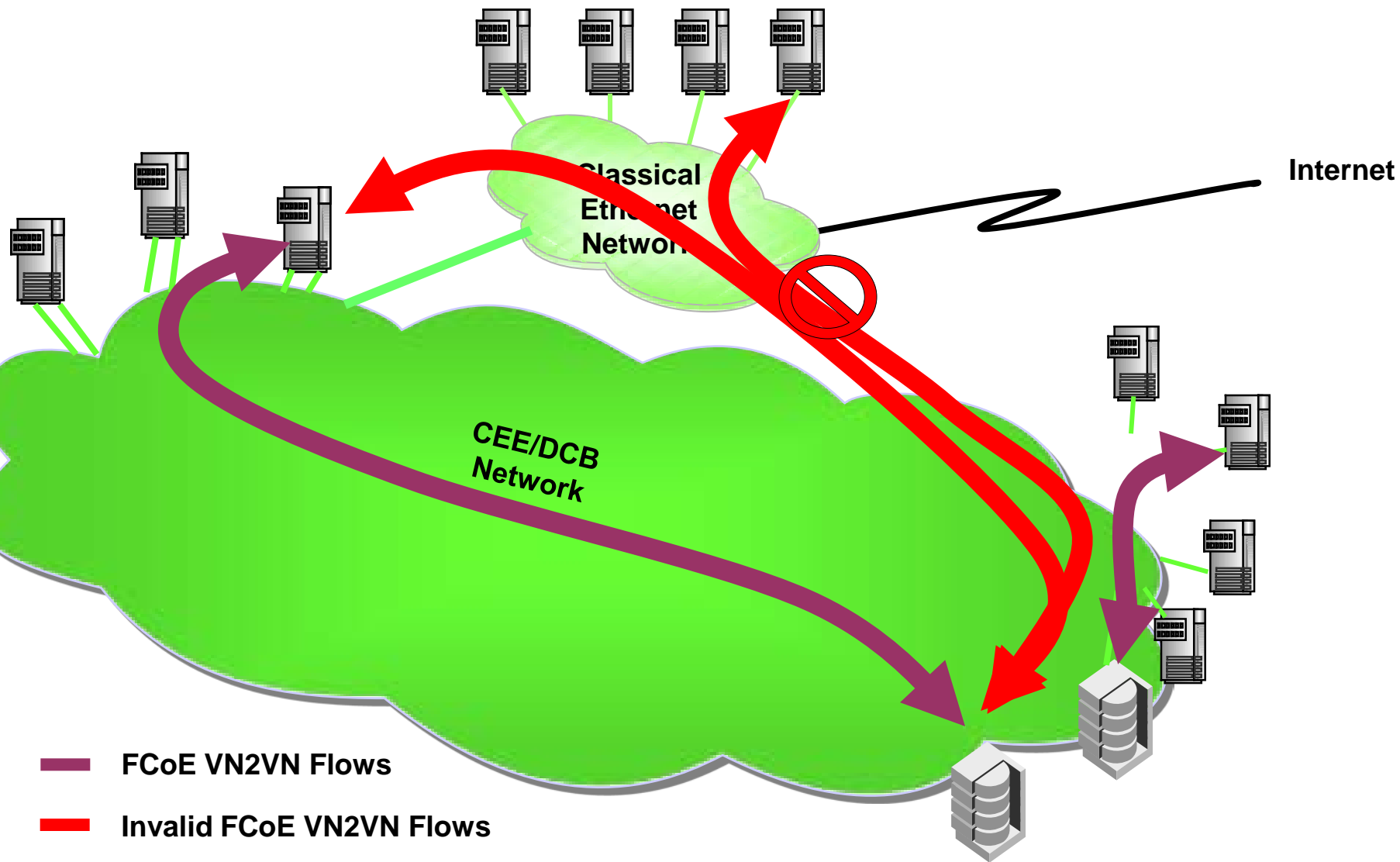
- **CEE/DCB – Switch(es) may connect**
 - A number of **VN2VN** capable **VN_Ports** together
 - Pairs of some configured **VN2VN-P2P ENodes**
(Requires physically/logically configuring the Switch)
- **A single Wire may connect**
 - **VN2VN ENodes (without P2P capability/configuration)**
 - **A pair of VN2VN-P2P ENodes**

Scenarios

Scenario 1: FCoE & IP Flows



Scenario 2: FCoE Right & Wrong



Summary

- ◆ T11.3's FC-BB-6 Ad-Hoc Working Group accepted VN2VN specification for inclusion in the next published standard (due 2012)
 - ◆ Vendors may produce products before that
- ◆ FCoE VN2VN is a simple, efficient mechanism for encapsulating Fibre Channel in Ethernet frames on a Lossless Ethernet type Network
 - ◆ Not a traditional Ethernet Interface or fabric
 - ◆ A New Network – A lossless Ethernet Network
 - ◆ Defined in the IEEE 802.1 standards working group (and called DCB)
 - ◆ FC protocols frames will just be inserted into these Ethernet frames
- ◆ Specification permits the installation to evolve from simple VN2VN into Full FCoE Fabrics (and even real FC Fabrics)
 - ◆ Can start small with simple Networks (even Point-to-Point)
 - › Perhaps with all software Initiators
 - ◆ May grow into larger FCoE fabrics in an evolutionally way
 - ◆ Full FCoE Fabrics and VN2VN networks can co-exist as installations grow
 - ◆ When upgrading you can continue to use the same physical components
- ◆ Now FC protocols can span the small, medium and Enterprise size networks
 - ◆ Only one protocol type (FC) is needed within the Data Center even if using different wire types
 - ◆ Different size companies can merge and their Data Center will easily merge also

- Please send any questions or comments on this presentation to **SNIA**: tracktutorials@snia.org

Many thanks to the following individuals
for their contributions to this tutorial.

SNIA Education Committee

Claudio DeSanti Silvano Gai Joe Pelissier John Hufferd

- For additional information refer to
<http://www.t11.org/fcoe>

Thank You!

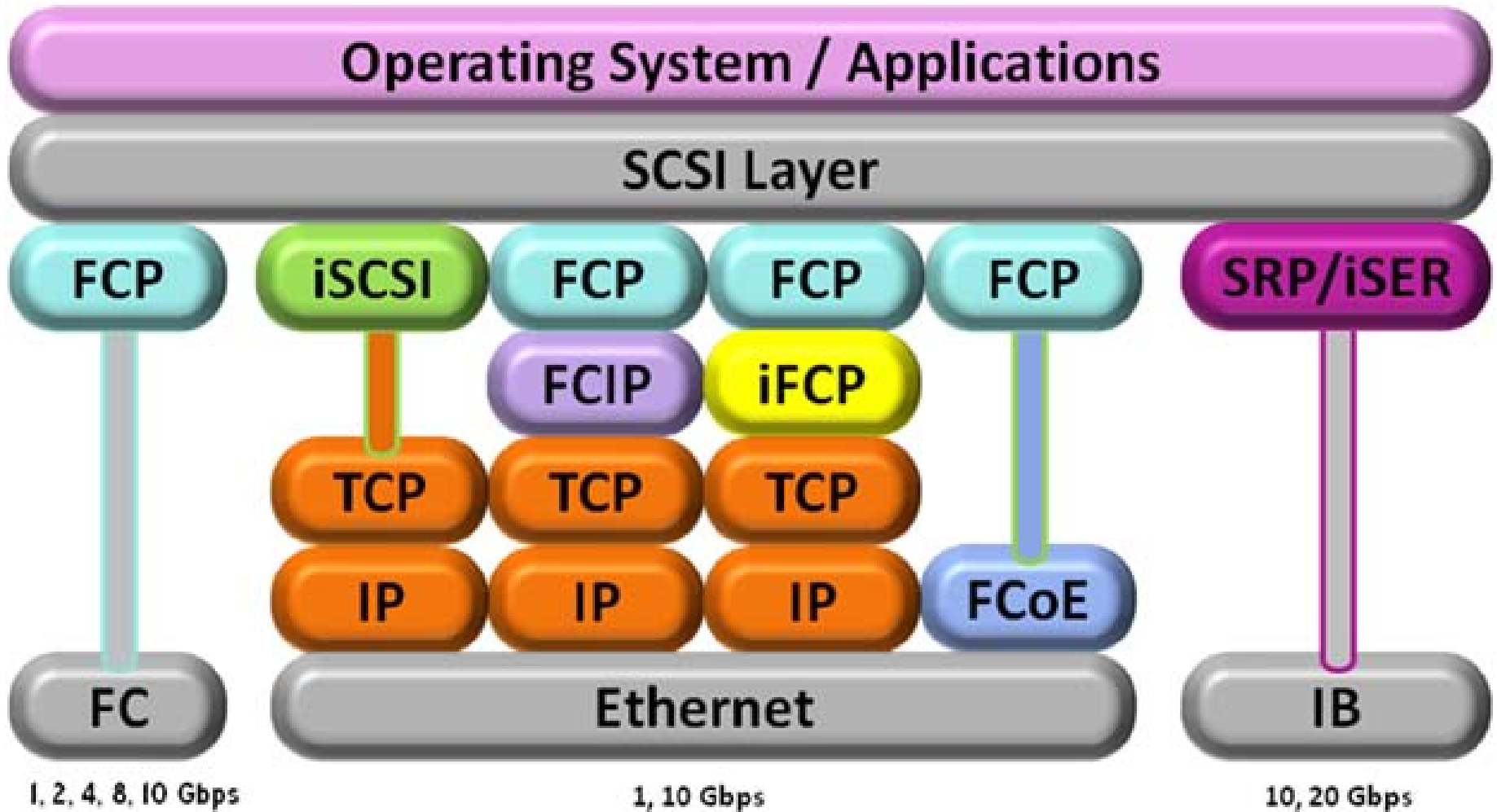


Education

- FCoE Relation to ISO Layers
- FIP Descriptors
- FIP Operation layouts



FCoE Relation to ISO Layers



FIP Descriptors



FIP Operation Codes

Operation Code	Subcode	Operation
0001h	01h	Discovery, Solicitation
	02h	Discovery, Advertisement
0002h	01h	FLOGI/FDISC/LOGO/ELP, Request
	02h	FLOGI/FDISC/LOGO/ELP, Reply, Reject
0003h	01h	FIP Keep Alive
	02h	FIP Clear Virtual Link
0004h	01h	FIP VLAN Request
	02h	FIP VLAN Notification
0005h	01h	N_Port_ID Probe Request
	02h	N_Port_ID Probe Reply
	03h	N_Port_ID Claim Notification
	04h	N_Port_ID Claim Response
	05h	N_Port_ID Beacon
FFF8h .. FFFEh	00h .. FFh	Vendor Specific
All others	All others	Reserved

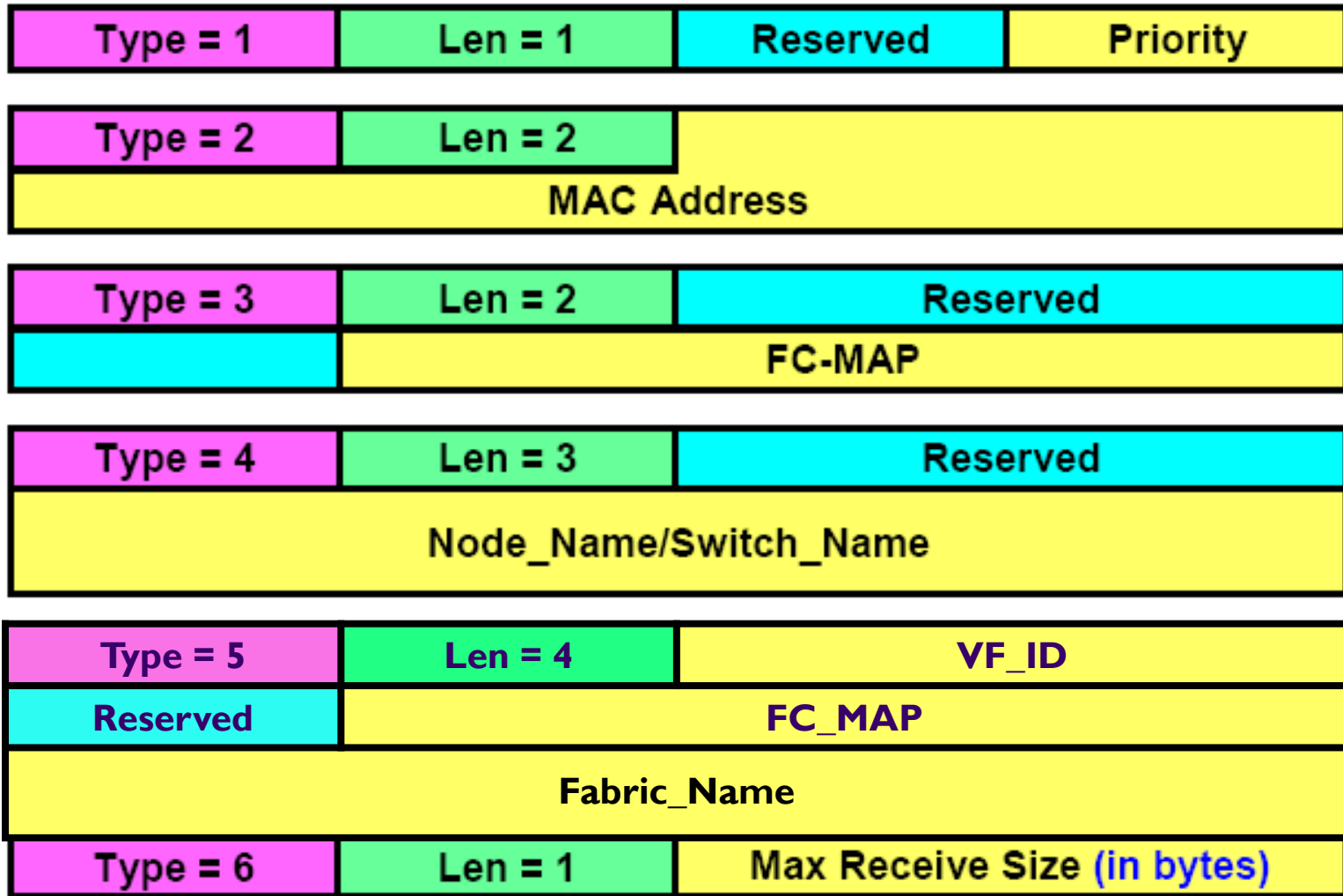


Type and Values of FIP Descriptors

Type	Value
00	Reserved
01	Priority
02	MAC Address
03	FC-MAP
04	Node_Name / Switch_Name
05	Fabric
06	Max FCoE Size
07	FLOGI
08	NPIV FDISK

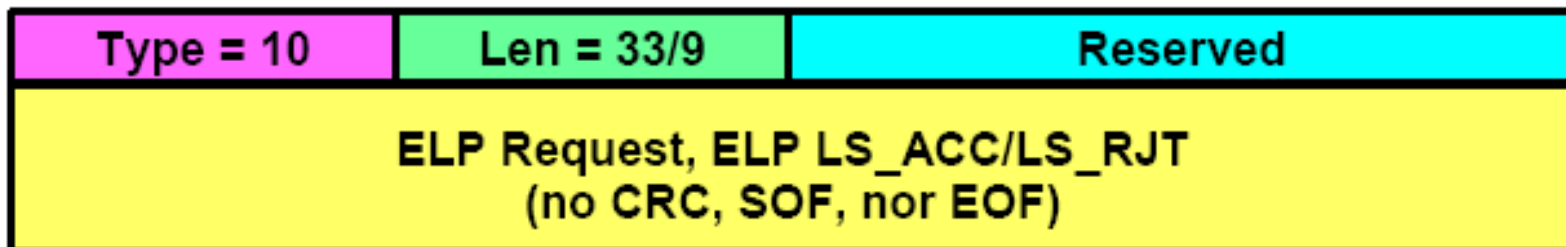
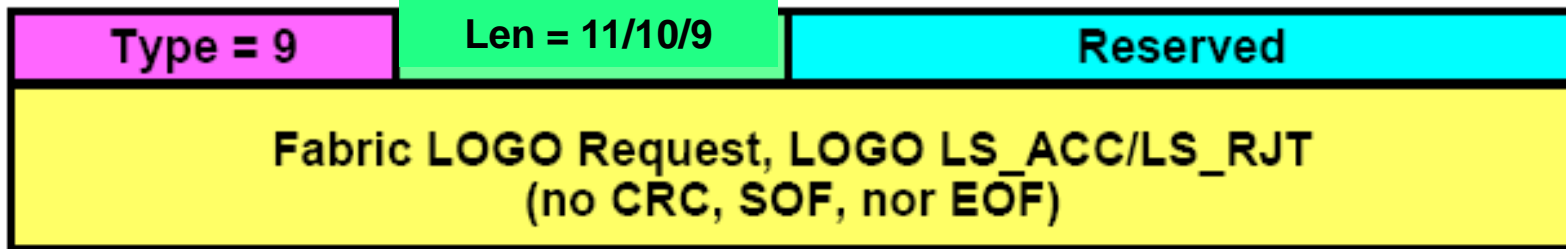
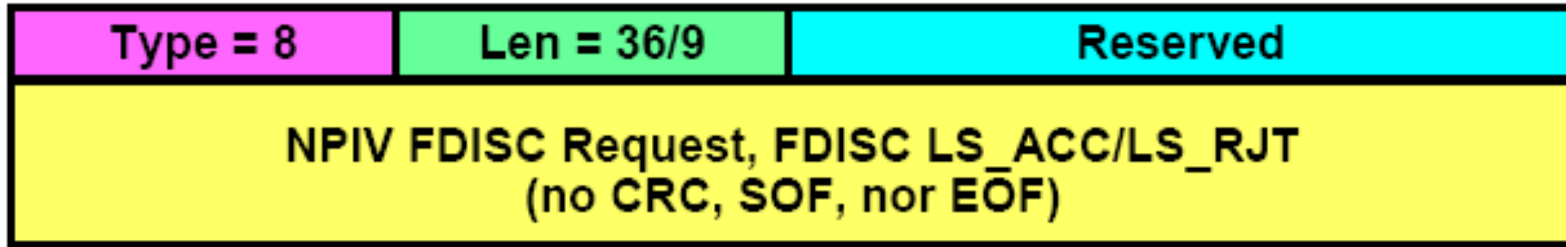
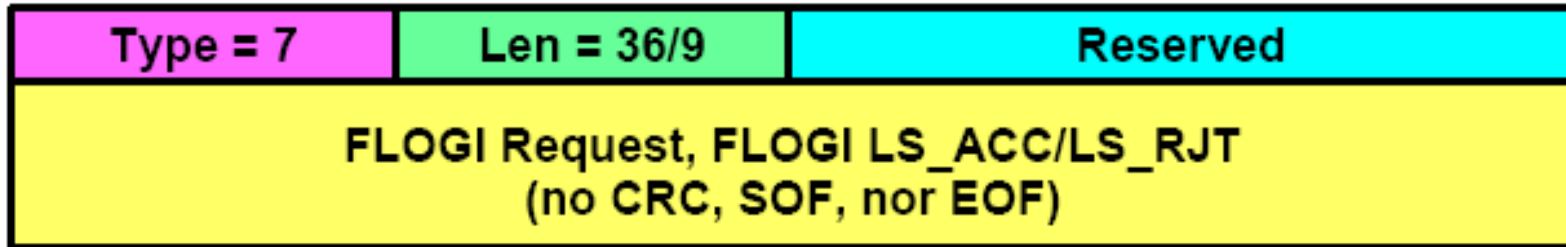
Type	Value
09	Fabric Log Out
10 (0Ah)	ELP
11 (0Bh)	VN_Port ID
12 (0Ch)	FKA_ADV_Period
13 (0Dh)	Vendor ID
14 (0Eh)	VLAN
15 (0Fh)	FC-4 Attributes
16 (10h) – 240 (F0h)	Reserved (128-240 are Non Critical)
241 (F1h) – 254 (FEh)	Vendor Specific
255 (FFh)	Reserved

FIP Descriptors (1)



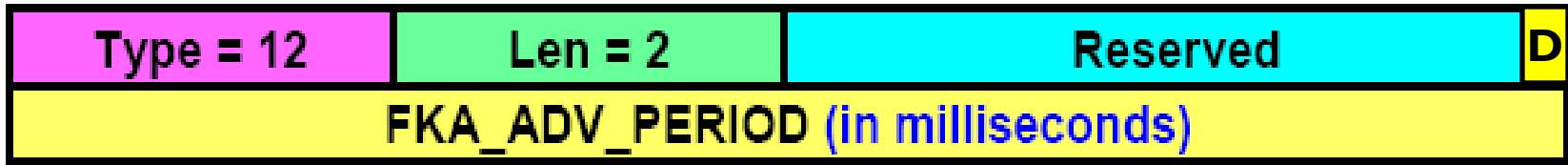
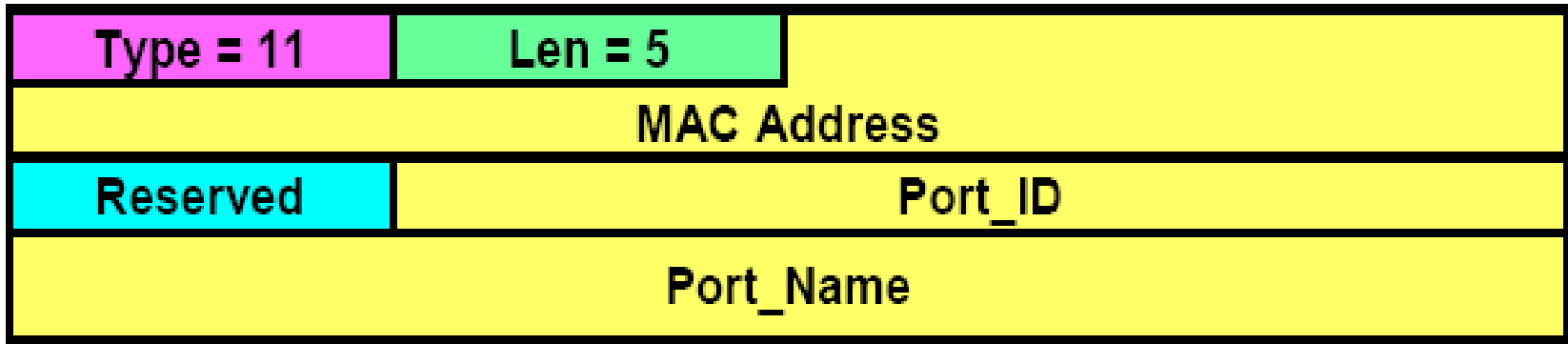
Lengths are measured in 32-bit words

FIP Descriptors (2)

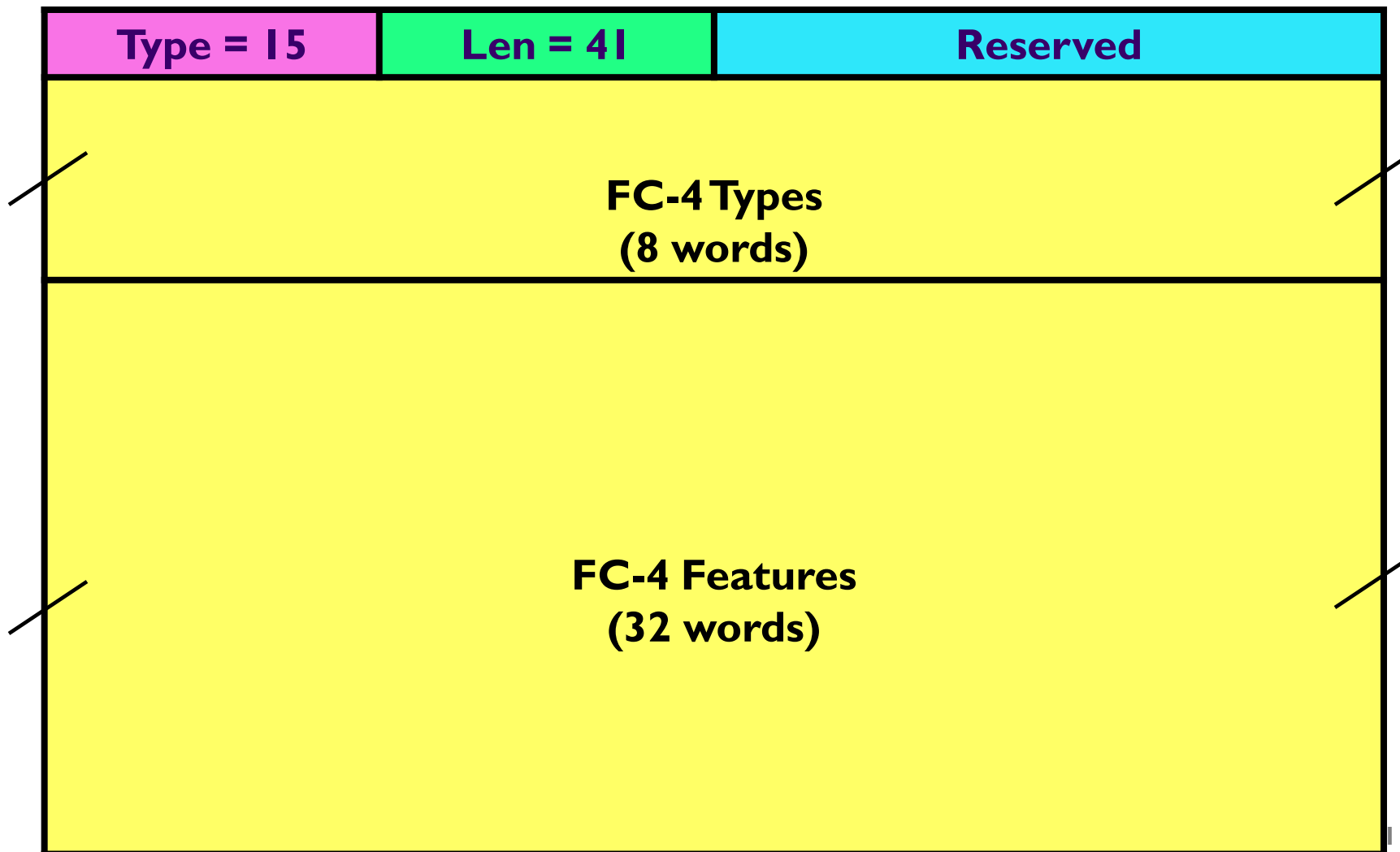




FIP Descriptors (3)

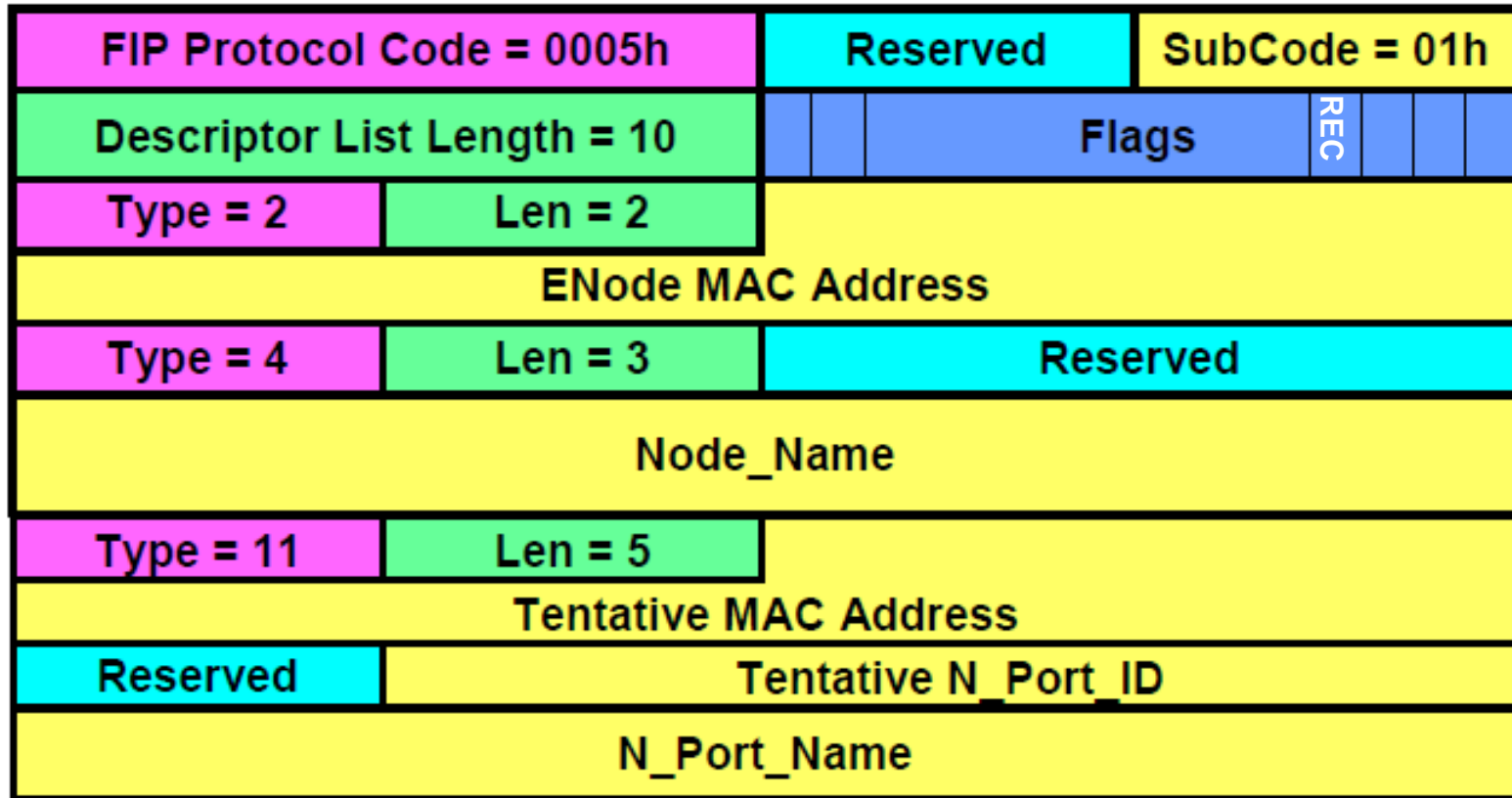


FIP Descriptor (4) – VN2VN



FIP Operation Layouts for VN2VN

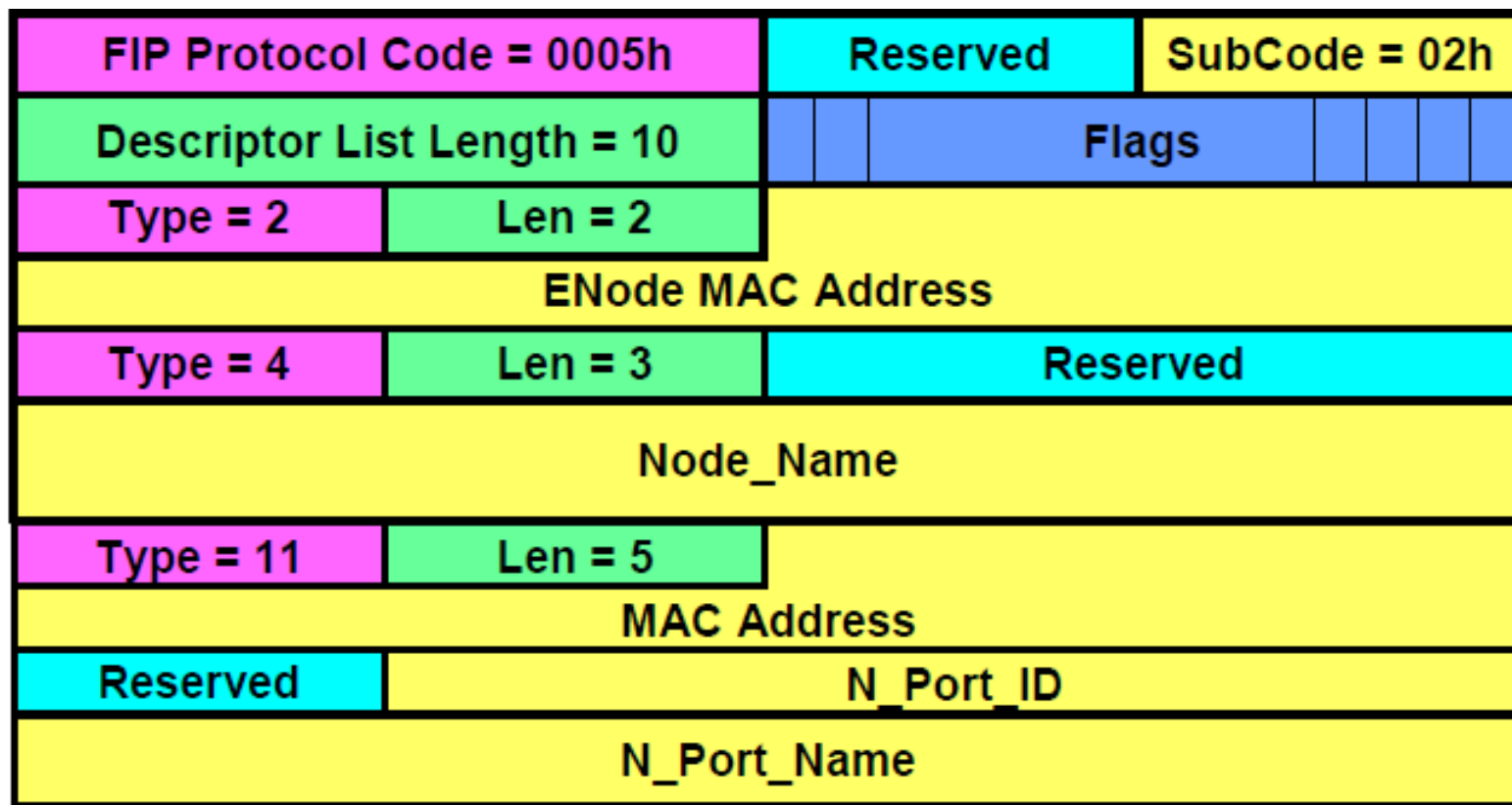
N_Port_ID Probe Request



Note: the REC flag is to signify that the VN_Port_ID & MAC Address has been previously saved/recorded

- SA = ENode MAC Address
- DA = All-VN2VN-ENode-MACs or All-PT2PT-ENode-MACs

N_Port_ID Probe Reply



- SA = Source ENode MAC Address
- DA = Destination ENode MAC Address

N_Port_ID Claim Notification

FIP Protocol Code = 0005h		Reserved	SubCode = 03h	
Descriptor List Length = 52		Flags		p2p
Type = 2	Len = 2	ENode MAC Address		
Type = 4	Len = 3	Reserved		
Node_Name				
Type = 11	Len = 5	VN Port MAC Address		
Reserved	N Port ID			
N_Port_Name				
Type = 15	Len = 41	Reserved		
FC-4 Types				
FC-4 Features				
Type = 6	Len = 1	Max FCoE Size		

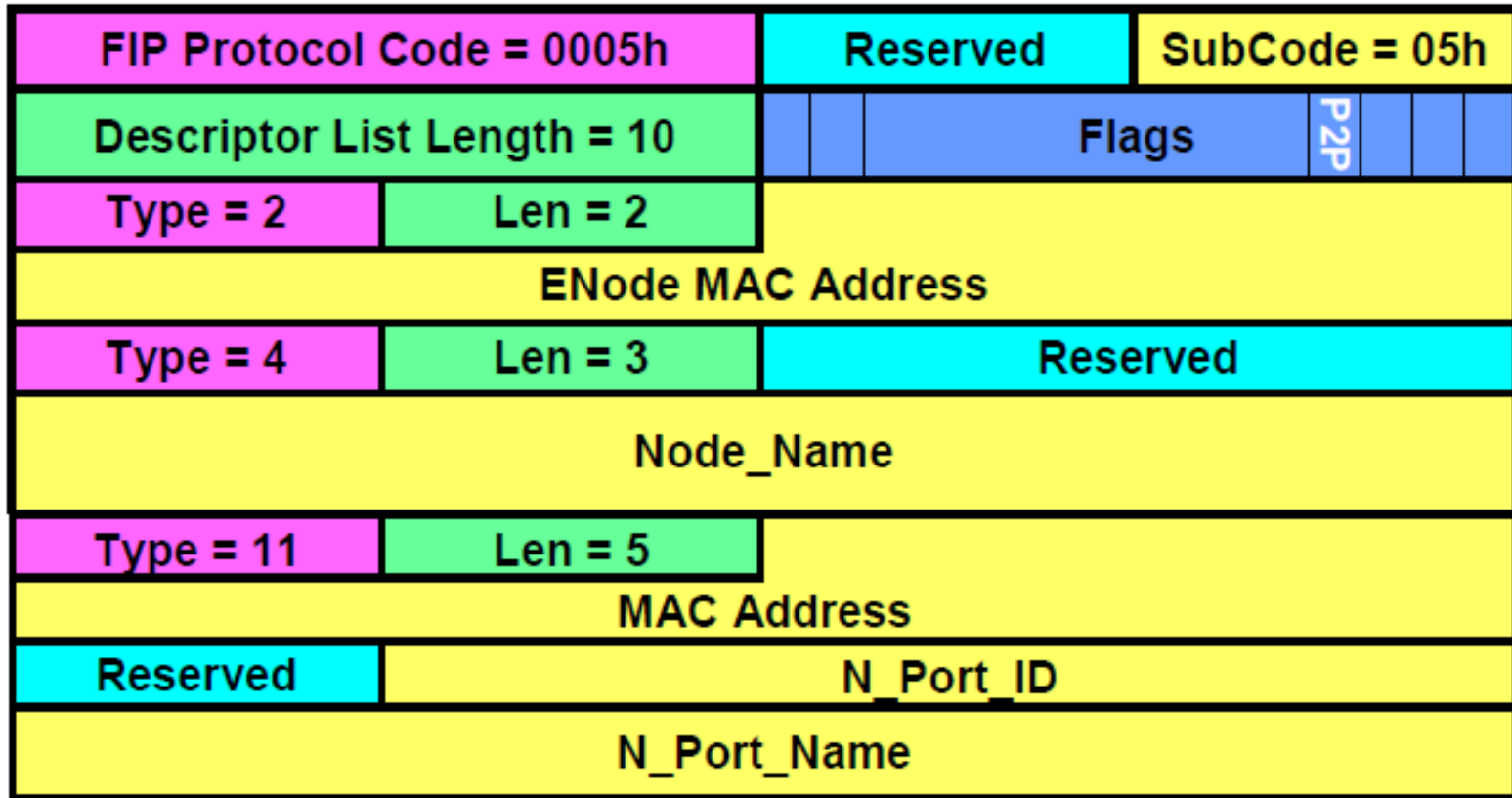
- SA = ENode MAC Address
- DA = All-VN2VN-ENode-MACs or All-PT2PT-ENode-MACs

N_Port_ID Claim Response

FIP Protocol Code = 0005h		Reserved	SubCode = 04h	
Descriptor List Length = 52		Flags		p2p
Type = 2	Len = 2	ENode MAC Address		
Type = 4	Len = 3	Reserved		
Node_Name				
Type = 11	Len = 5	VN Port MAC Address		
Reserved	N_Port_ID			
N_Port_Name				
Type = 15	Len = 41	Reserved		
FC-4 Types				
FC-4 Features				
Type = 6	Len = 1	Max FCoE Size		
FIP Pad to Max FCoE Size				

- SA = Source ENode MAC Address
- DA = Destination ENode MAC Address

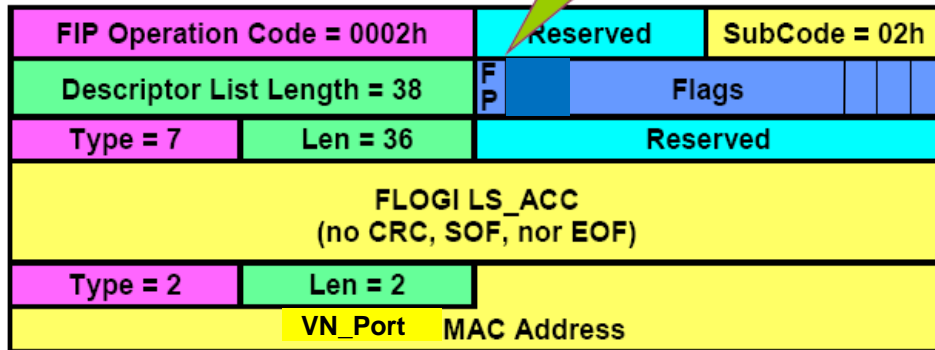
N_Port_ID Beacon



- SA = VN_Port MAC Address (i.e., VN2VN-FC-MAP || LUID)
- DA = All-VN2VN-ENode-MACs or All-PT2PT-ENode-MACs

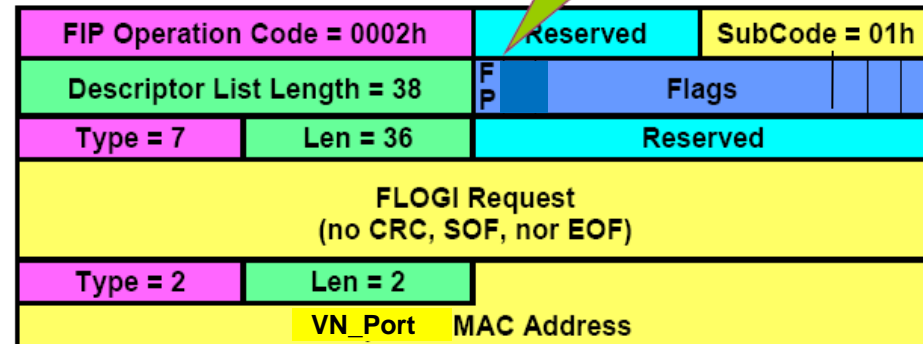
FIP FLOGI (VN2VN/PT2PT)

FP = 1



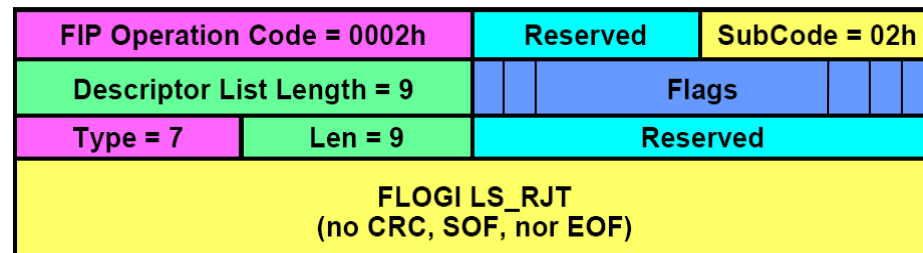
Accept (ENODE → ENODE)

FP = 1



Request (ENODE → ENODE)

Reject (ENODE → ENODE)



FIP Fabric LOGO (VN2VN/PT2PT)

FIP Operation Code = 0002h	Reserved	SubCode = 01h
Descriptor List Length = 13	Flags	
Type = 9	Len = 11	Reserved
LOGO Request (no CRC, SOF, nor EOF)		
Type = 2	Len = 2	
MAC Address (to be removed)		

Request (ENODE → ENODE)

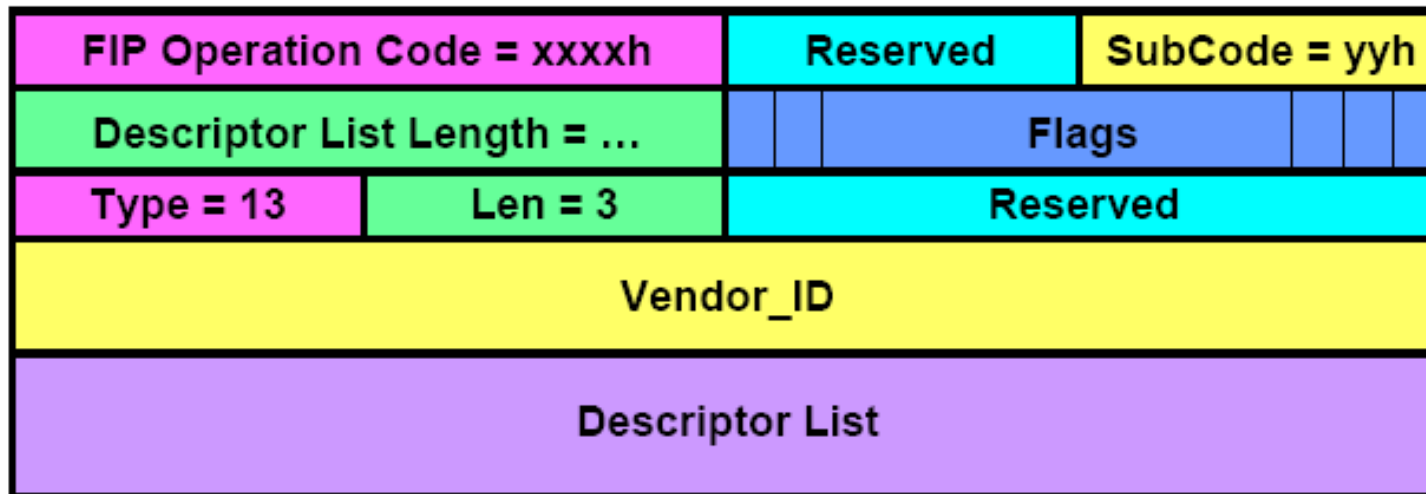
FIP Operation Code = 0002h	Reserved	SubCode = 02h
Descriptor List Length = 10	Flags	
Type = 9	Len = 8	Reserved
LOGO LS_ACC (no CRC, SOF, nor EOF)		
Type = 2	Len = 2	
MAC Address (to be removed)		

Accept (ENODE → ENODE)

Reject (ENODE → ENODE)

FIP Operation Code = 0002h	Reserved	SubCode = 02h
Descriptor List Length = 9	Flags	
Type = 9	Len = 9	Reserved
LOGO LS_RJT (no CRC, SOF, nor EOF)		

Vendor Specific FIP Message



- An unknown Vendor Specific message shall be ignored
Normal operation will be the result
- A device is never required to generate a Vendor Specific message for normal operation

Vendor Specific Operation Code = FFF8 -- FFFE with any SubCode
Vendor Specific Type (in Descriptor List) = F1h – FEh
Reserved Types = 0Fh – F0h, & FFh