

Additional FIP Functions

T11/08-263v1, June 2008

Claudio DeSanti, Cisco

Joe Pelissier, Cisco

Landon Noll, Cisco

Silvano Gai, Cisco

Craig Carlson, QLogic

Ed McGlaughlin, QLogic

Alan Spalding, QLogic

Dave Peterson, Brocade

Bob Snively, Brocade

Bob Nixon, Emulex

Parag Bhide, Emulex

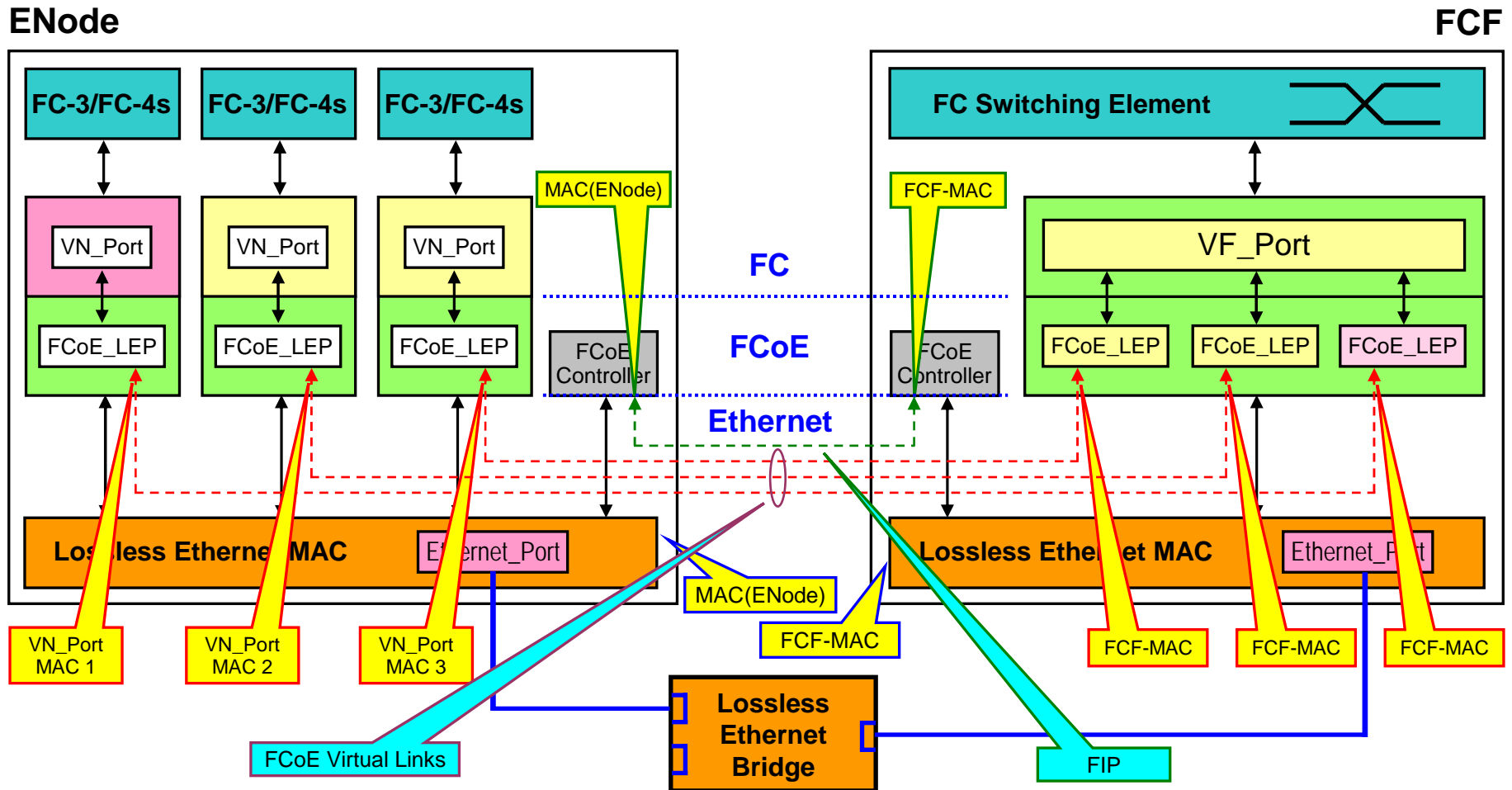
Pat Thaler, Broadcom

Uri Elzur, Broadcom

Agenda

- **FCoE “Soft State”**
- **FIP Clear Virtual Link**
- **FIP Discovery Summary**
- **FIP Descriptors Summary**
- **FIP Messages Summary**

FIP and FCoE Virtual Links



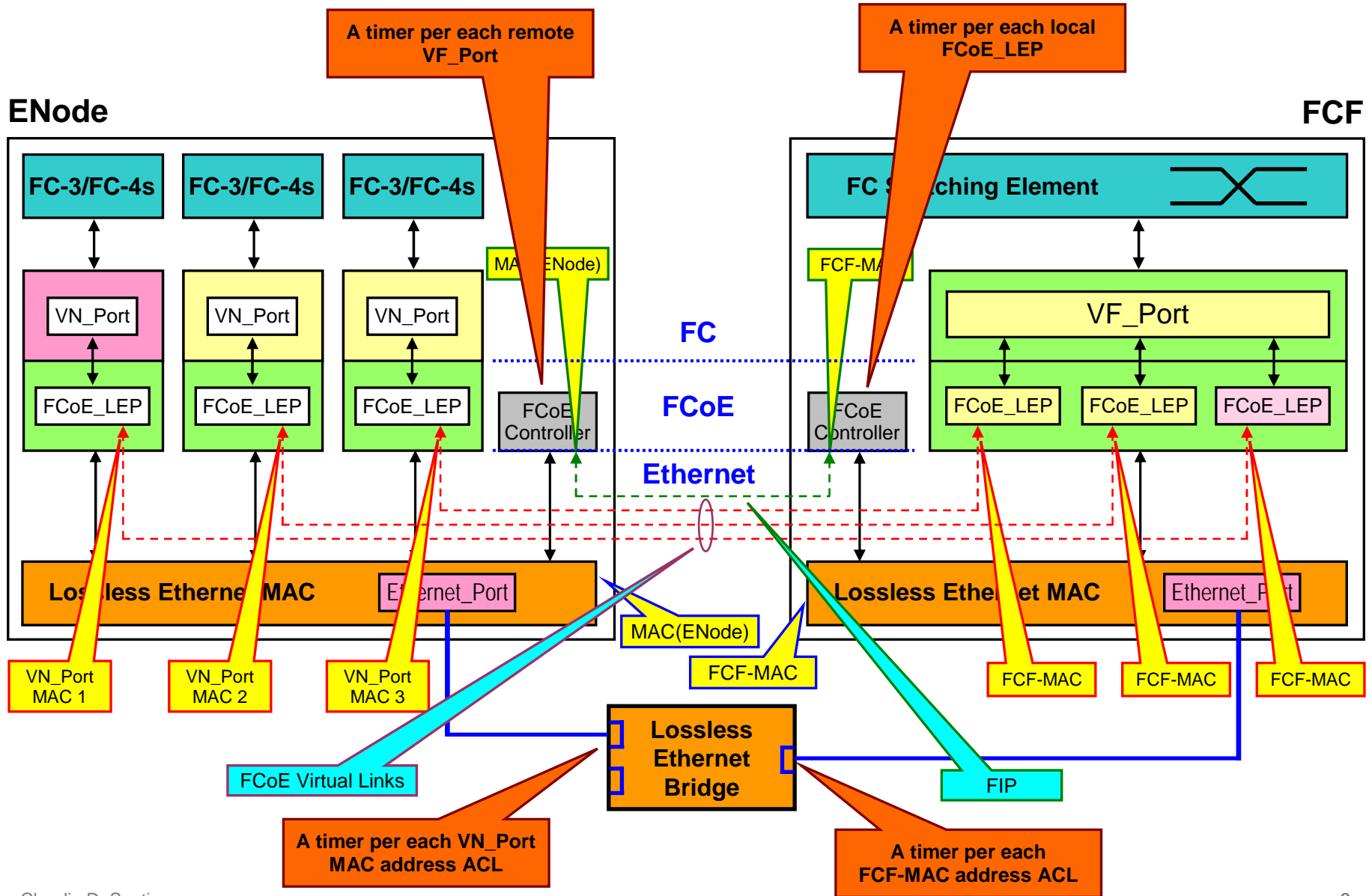
Issues with LKA

- **LKA is an ELS to be generated by VN_Ports when no FC traffic is present**
- **But it is not implemented**
 - Because current drivers/implementations do not have the capability of inspecting the FC traffic**
 - True also for F_Ports and E_Ports**
 - This capability is not easy to add**
- **Conclusion: a better solution is needed!**

Soft vs. Hard State

- **In native Fibre Channel, F_Ports, N_Ports and E_Ports have “hard state”**
 - i.e., state depending by the physical layer signaling “the link is up”
- **With FCoE “hard state” is no longer possible**
 - The physical layer cannot provide anymore this signaling
 - All state has to be “soft” (i.e., timed)
- **Existence of an “entity” is dependent on the reception of appropriate Keep Alive messages from a remote entity**
 - After some missing messages (e.g., 3), the entity should be de-instantiated
- **This is true for every entity (except the FCoE Controller)**
 - VN_Ports, VF_Ports, VE_Ports
 - Related ACLs in intermediate Lossless Ethernet bridges

FIP and FCoE Virtual Links



The FCoE Controller

- **The functional entity implementing the FIP protocol**

Performs the FIP Discovery protocol

Instantiates VN_Ports, VF_Ports, VE_Ports

On successful completion of FIP FLOGI/FDISC/ELP

De-instantiates VN_Ports, VF_Ports, VE_Ports

When needed

Monitors the “health” of VN_Ports, VF_Ports, VE_Ports and generates appropriate periodic “keep alive” messages on behalf of them

Maintains timers per each appropriate entity in order to de-instantiate it if some consecutive “keep alive” messages are not received

Keep Alive: ENode's MAC View

- **Per each FCF-MAC, there is a single VF_Port to which the set of ENode's VN_Ports instantiated through NPIV are logged in**

- **The MAC address of that VF_Port is the FCF-MAC**

Receiving a periodic FIP message from that FCF-MAC is a good enough indication of the health of that VF_Port

Periodic Advertisements already provide this function

After a certain number of missing Advertisements (e.g., 3) the VF_Port is considered unreachable and the associated VN_Ports de-instantiated (i.e., implicitly logged out)

- **Each VN_Port may be associated with a different MAC address**

Generating a periodic FIP message from each MAC address is needed to indicate the corresponding VN_Port is operational

every FKA_ADV_PERIOD (e.g., 30 seconds)

A periodic unicast and unidirectional FIP Keep Alive message

a message without reply

Keep Alive: VF_Port capable FCF's View

- **Per each ENode, there is a single VF_Port to which the set of ENode's VN_Ports instantiated through NPIV are logged in**
- **Each VN_Port may be associated with a different MAC address**

Receiving a periodic FIP message from each MAC address is a good enough indication of the health of that VN_Port

A periodic unicast and unidirectional FIP Keep Alive message

a message without reply

After a certain number of missing FIP Keep Alive messages (e.g., 3) the VN_Port is considered unreachable (i.e., implicitly logged out) and the associated FCoE_LEP de-instantiated

a FIP Clear Virtual Link removing that VN_Port should be generated

- **The MAC address of that VF_Port is the FCF-MAC**

Generating a periodic FIP message from that FCF-MAC is needed to indicate the VF_Port is operational

Periodic Advertisements already provide this function

Advertisements as Keep Alive messages

- **Using periodic multicast Advertisements rather than unicast Keep Alive messages simplifies FCFs processing**
 - If 100 ENodes are logged in with an FCF, send 1 message rather than 100
 - But does not allow to identify the case of a specific VF_Port that is no more operational
- **The FCF-MAC's FCoE Controller should generate a FIP Clear Virtual Link when a VF_Port monitored by it becomes not operational**
 - Removing all VN_Ports connected to that VF_Port
 - The FCoE Controller monitors all VF_Ports
- **A FIP Keep Alive erroneously received from a VN_Port once logged in and now logged out should be replied with a FIP Clear Virtual Link**
 - Removing that VN_Port
- **The period of multicast Advertisements and FIP Keep Alive messages should be the same**
 - FKA_ADV_PERIOD**
 - Provided to ENodes by FCFs in Advertisements**
 - Therefore it can be also snooped by intermediate Ethernet bridges

Keep Alive: Intermediate Bridge's View

- **Intermediate Lossless Ethernet bridges may instantiate ACLs by snooping FIP messages**

E.g., FIP FLOGI/FDISC Accepts

- **Each ACL should have an associated timer**

E.g., a timer per each VN_Port related ACL

- **Snooping periodic FIP Keep Alive messages allows to keep a VN_Port ACL in place**

After a certain number of missing FIP Keep Alive messages (e.g., 3) the VN_Port is considered unreachable (i.e., implicitly logged out) and the associated ACL de-instantiated

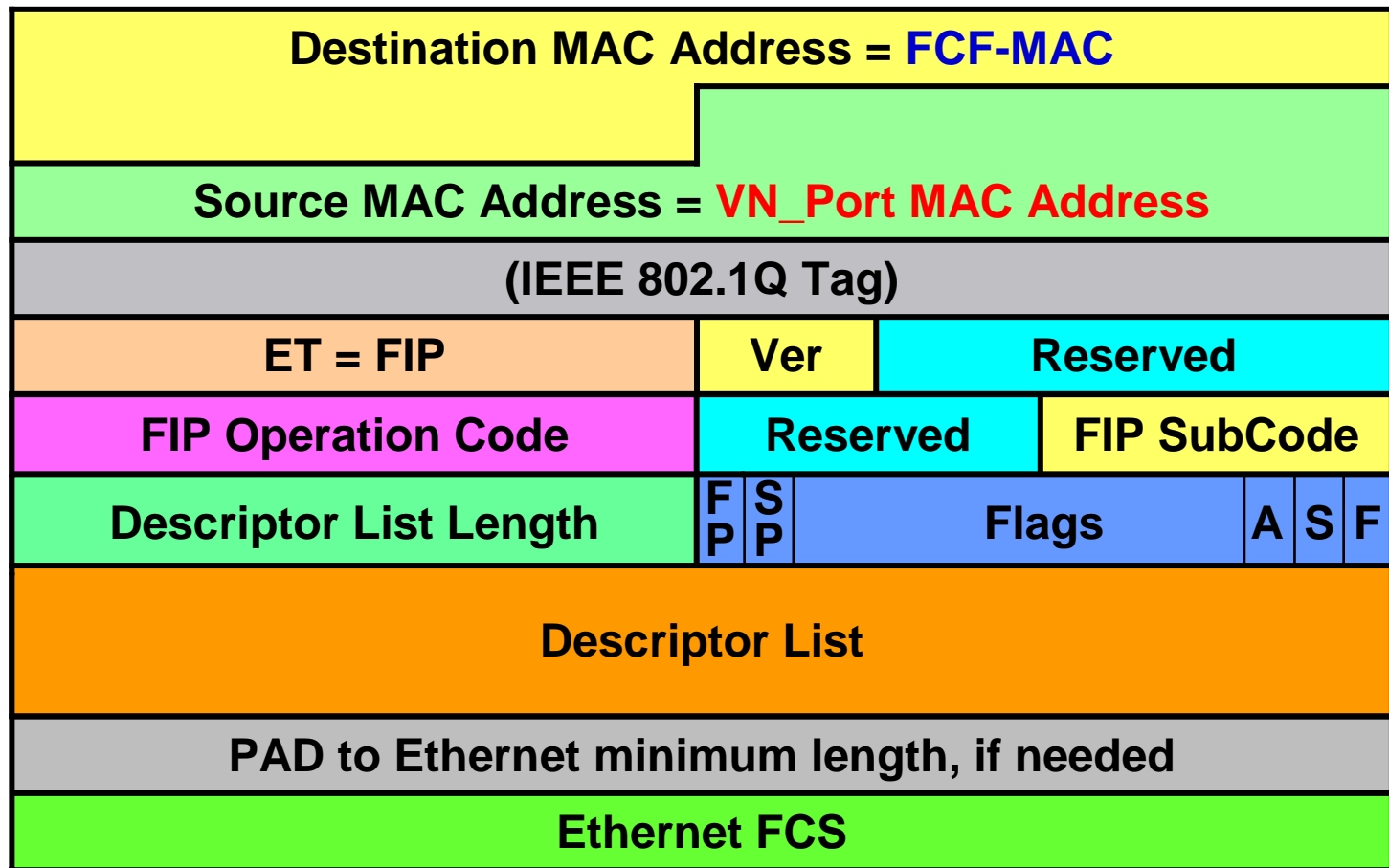
- **Snooping periodic Advertisements allows to keep a VF_Port capable FCF-MAC ACL in place**

After a certain number of missing Advertisements (e.g., 3) the VF_Port capable FCF-MAC is considered unreachable and the associated ACL de-instantiated

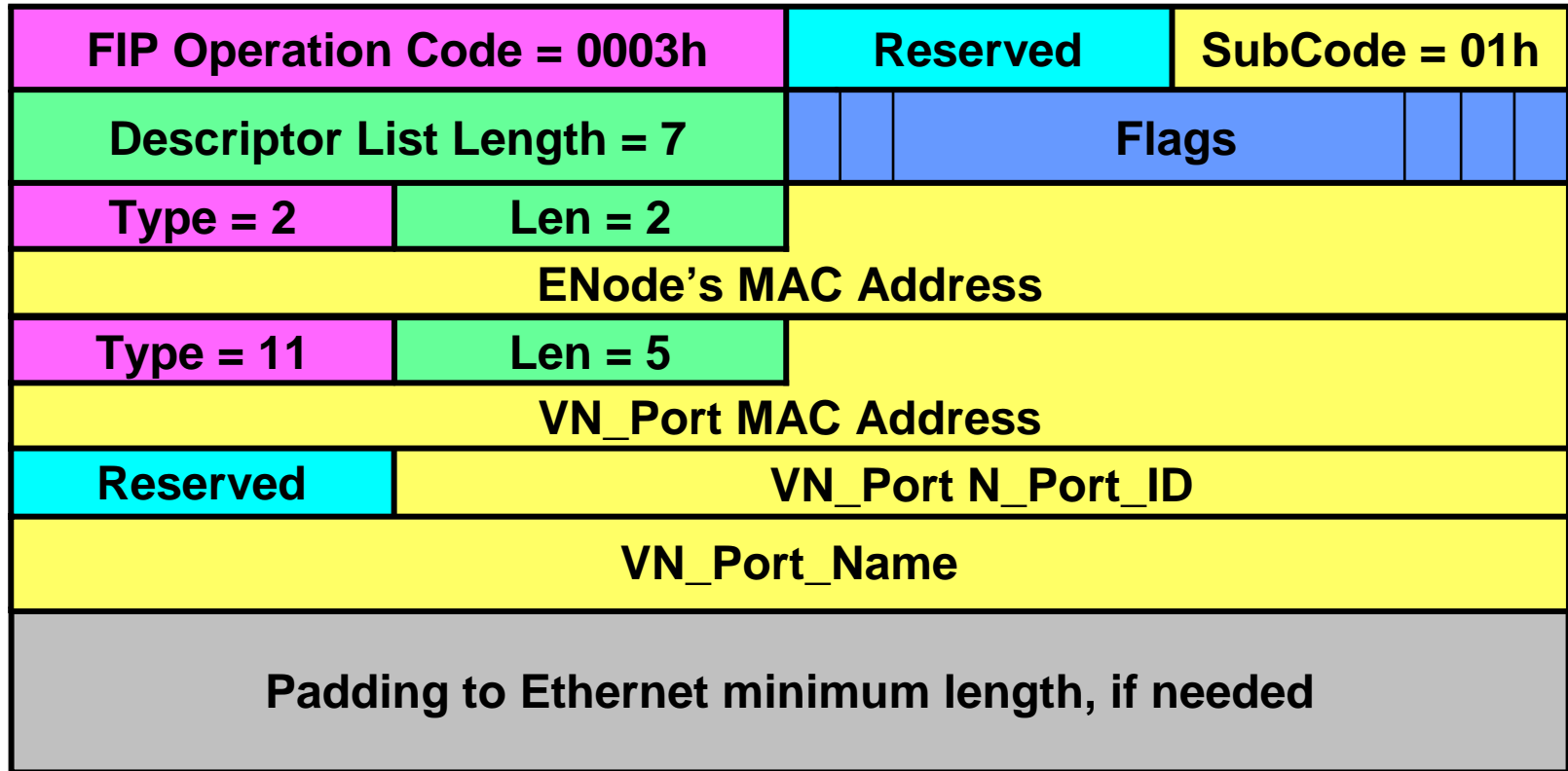
FIP Operation Codes and SubCodes

Operation Code	SubCode	Operation
0001h	01h	Discovery, Solicitation
	02h	Discovery, Advertisement
0002h	01h	FLOGI/FDISC/LOGO/ELP, Request
	02h	FLOGI/FDISC/LOGO/ELP, Reply
0003h	01h	FIP Keep Alive
	02h	FIP Clear Virtual Link
All others	All others	Reserved

FIP Keep Alive (1)



FIP Keep Alive (2)



A minimum length Ethernet frame
(low overhead)

Keep Alive: VE_Port capable FCF's View

- **VE_Port capable FCF-MACs are symmetric**

E.g., they generate periodic Advertisements

Receiving a periodic Advertisement from the other FCF-MAC address is a good enough indication of the health of that FCF-MAC

After a certain number of missing Advertisements (e.g., 3) the remote FCF-MAC is considered unreachable and the associated VE_Port de-instantiated

Keep Alive: Intermediate Bridge's View

- **Intermediate Lossless Ethernet bridges may instantiate ACLs by snooping FIP messages**
 - E.g., FIP ELP Accepts
- **Each ACL will have an associated timer**
 - E.g., a timer per each VE_Port capable FCF-MAC related ACL
- **Snooping periodic Advertisements allows to keep a VE_Port capable FCF-MAC ACL in place**
 - After a certain number of missing Advertisements (e.g., 3) the VE_Port capable FCF-MAC is considered unreachable and the associated ACL de-instantiated

Agenda

- FCoE “Soft State”
- FIP Clear Virtual Link
- FIP Discovery Summary
- FIP Descriptors Summary
- FIP Messages Summary

A FIP Clean-up Function

- **Native Fibre Channel ports may perform a link initialization**
 - On a N_Port to F_Port link this removes all N_Port_IDs (VN_Ports)
- **It is desirable to have something similar also on the virtualized FCoE environment**

A **FIP Clear Virtual Link** function

When invoked, it removes VN_Ports, VF_Port and Virtual Links between an ENode's MAC and a VF_Port capable FCF-MAC

When invoked, it removes VE_Ports and Virtual Links between two VE_Port capable FCF-MACs

- **An ENode has already this capability: send another FLOGI**
 - Equivalent to a LOGO of everything followed by the FLOGI
- **FCFs do not have this capability**
 - A Fabric LOGO removes a single VN_Port
 - Fabric LOGO is not currently implemented on FC Switches
- **Instead define FIP Clear Virtual Link for FCFs**
- **Fabric LOGO remains for ENodes**

FIP Clear Virtual Link

- **Unidirectional message (i.e., without reply) that can be generated only by an FCF**
- **Carries the FCF-MAC Address and the Switch_Name of the generating FCF**
- **When sent to an ENode's MAC, carries the list of VN_Ports to be removed**

For a “full” FIP Clear Virtual Link, all VN_Ports are listed

To remove some VN_Ports, only those ones are listed

Should not exceed the standard Ethernet size (i.e. 1500 bytes)

if the list of VN_Ports does not fit, multiple frames may be generated instead than one

- **When sent to an FCF-MAC, carries the destination FCF-MAC address**

The VE_Port address to be removed

FCF to ENode FIP Clear Virtual Link Payload

FIP Operation Code = 0003h		Reserved	SubCode = 02h	
Descriptor List Length = ...			Flags	
Type = 2	Len = 2	FCF-MAC Address		
Type = 4	Len = 3	Reserved		
Switch_Name				
Type = 11	Len = 5	VN_Port #1 MAC Address		
Reserved	VN_Port #1 N_Port_ID			
VN_Port_Name #1				
...				
Type = 11	Len = 5	VN_Port #n MAC Address		
Reserved	VN_Port #n N_Port_ID			
VN_Port_Name #n				

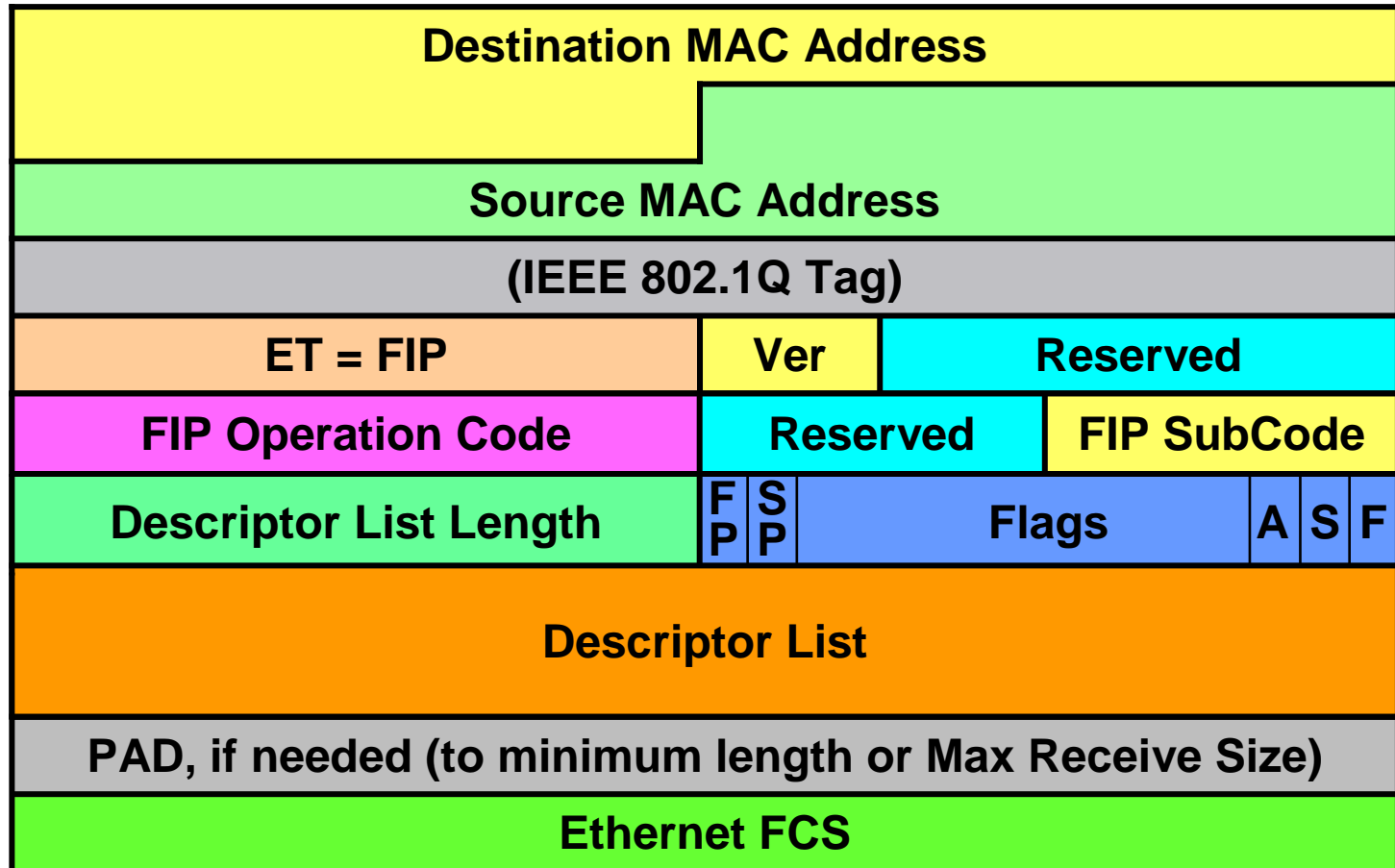
FCF to FCF FIP Clear Virtual Link Payload

FIP Operation Code = 0003h		Reserved	SubCode = 02h	
Descriptor List Length = 10		Flags		
Type = 2	Len = 2	Source FCF-MAC Address		
Source FCF-MAC Address				
Type = 4	Len = 3	Reserved		
Source Switch_Name				
Type = 11	Len = 5	Destination FCF-MAC Address		
Destination FCF-MAC Address				
Reserved	FFFFFFDh (Fabric Controller WKA)			
Destination VE_Port_Name				

Agenda

- FCoE “Soft State”
- FIP Clear Virtual Link
- **FIP Discovery Summary**
- FIP Descriptors Summary
- FIP Messages Summary

FIP Frame Format



FIP Discovery Protocol (1)

- **VF_Port capable FCF-MACs periodically send multicast Advertisements**
Every FKA_ADV_PERIOD (e.g., 30 seconds)
Not jumbo!
- **On receiving Advertisements, ENodes' FCoE Controllers create an entry per FCF-MAC in a FCF List**
By default ordered on the basis of the value carried in the Priority descriptor
Each entry has some flags:
 - ‘Max Receive Size verified’ (set to zero for entries created from multicast Advertisements, set to one when a unicast jumbo Advertisement is received)
 - ‘Available for Logins’ (set to one when the FCF is able to accept additional FLOGI/FDISC requests, set to zero otherwise)
- **ENodes' FCoE Controllers select a subset of the Available FCF-MACs for Login (the “FCF Login Set”)**
On the basis of a “local policy”
By default the one(s) with higher priority (i.e., lower priority value)

FIP Discovery Protocol (2)

- **Each FCF-MAC of the FCF Login Set needs to be verified for Max Receive Size support before performing the Login**
 - By sending a unicast Solicitation and receiving a unicast Advertisement**
 - Unicast Advertisements may carry updated information specific for the soliciting ENode**
- **The periodic reception of multicast Advertisements allows ENode's FCoE Controllers to continuously verify the FCF-MACs reachability**
- **The 'A' flag in FIP Advertisements carries the information "Available for Login"**
 - This flag is updated on the FCF list / FCF Login Set on reception of Advertisements**

FIP Discovery Protocol (3)

- **When an ENode's FCoE Controller becomes operational it can just wait to receive the Advertisements and operate as described in (1) and (2)**

This may take some time

- **Shortcut: the ENode's FCoE Controller sends a multicast Solicitation when it becomes operational**
- **FCF-MACs respond with Unicast Advertisements**
Jumbo padded
- **The ENode's FCoE Controller may then directly proceed to the Login phase**

FIP Discovery Protocol (4)

- **VE_Port capable FCF-MACs periodically send multicast Advertisements**

Every FKA_ADV_PERIOD (e.g., 30 seconds)

Not jumbo!

- **On receiving Advertisements, VE_Port capable FCF-MAC's FCoE Controllers create an entry per FCF-MAC in a FCF List**

By default ordered on the basis of the value carried in the Priority descriptor

Each entry has some flags:

'Max Receive Size verified' (set to zero for entries created from multicast Advertisements, set to one when a unicast jumbo Advertisement is received)

'Available for ELP' (set to one when the FCF is able to accept additional ELP requests, set to zero otherwise)

FIP Discovery Protocol (5)

- **Each FCF-MAC of the FCF List needs to be verified for Max Receive Size support before performing the ELP**

By sending a unicast Solicitation and receiving a unicast Advertisement

Unicast Advertisements may carry updated information specific for the soliciting FCF-MAC

- **The periodic reception of multicast Advertisements allows VE_Port capable FCF-MAC's FCoE Controllers to verify the reachability of other VE_Port capable FCF-MACs**
- **The 'A' flag in FIP Advertisements carries the information "Available for ELP"**

This flag is updated on the FCF list on reception of Advertisements

FIP Discovery Protocol (6)

- **When a VE_Port capable FCF-MAC's FCoE Controller becomes operational it can just wait to receive the Advertisements and operate as described in (4) and (5)**

This may take some time

- **Shortcut: the VE_Port capable FCF-MAC's FCoE Controller sends a multicast Solicitation when it becomes operational**
- **VE_Port capable FCF-MACs respond with Unicast Advertisements**

Jumbo padded

- **The VE_Port capable FCF-MAC's FCoE Controller may then directly proceed to the ELP phase**

Dealing with Multicasts and Periodicity

- **Periodic transmission of multicast messages as well as responses to multicast messages may create undesirable synchronizations**
- **The interval between subsequent transmissions of multicast Advertisements should be randomized**
 - When the clock expires, wait a random time
- **Advertisements sent in response to a multicast Solicitation should be also delayed by a random time**
- **The random time should be uniformly distributed between 0 and MAX_RANDOM_DELAY**
 - 100 milliseconds is the value for MAX_RANDOM_DELAY

Advertisements Generation Rules

- Advertisements are sent by FCF-MACs
- Advertisements may be Solicited (i.e., sent in response to a Solicitation) or unsolicited (i.e., sent periodically)
- Solicited Advertisements are unicast messages, have the Solicited bit set to one and are padded to the Max Receive Size specified in the Solicitation they are responding to
- Solicited Advertisements are sent to the MAC address carried in the MAC address descriptor contained in the Solicitation they are responding to
- Unsolicited Advertisements are multicast messages, have the Solicited bit set to zero and are not padded
- Unsolicited Advertisements are periodically sent:
 - to the multicast address 'All-ENode-MACs' by VF_Port capable FCF-MACs
 - to the multicast address 'All-FCF-MACs' by VE_Port capable FCF-MACs

Advertisements Processing Rules (1)

- **The FCoE Controllers of ENode's MACs and VE_Port capable FCF-MACs maintain an FCF list**
- **Each Advertisement creates an entry in the FCF list**
 - Solicited Advertisements (Jumbo) create entries with the 'Max Receive Size verified' flag set to one**
 - Unsolicited Advertisements create entries with the 'Max Receive Size verified' flag set to zero**
- **An FCF having multiple FCF-MACs may respond to a multicast Solicitation with multiple unicast Advertisements**
 - Having the same value in the Switch_Name descriptor**
- **The Priority descriptor indicates which FCF-MACs are preferred to establish Logins**
 - They have higher priority in the FCF List**
- **The MAC address carried in the MAC address descriptor is used to establish virtual links**

Advertisements Processing Rules (2)

- **The FC-MAP descriptor is used by an ENode's MAC supporting Fabric Provided MAC Addresses to configure such addresses**
- **The FC-MAP descriptor is used by a VE_Port capable FCF-MAC to detect anomalous network configurations**

Because the FC-MAP shall be the same for FCFs belonging to the same Fabric

An FCF ignores an Advertisement if the value carried in the FC-MAP descriptor is not equal to its configured FC-MAP

An exception may be signaled

- **The FC-MAP descriptor may be ignored if set to zero**
- **An ENode's MAC should detect an anomalous configuration if:**
 - different Advertisements carry different FC-MAP values**
 - different Advertisements carry different Fabric_Names**
- **The Fabric_Name descriptor may be ignored by VE_Port capable FCF-MACs**

Advertisement

FCF Capabilities:
 FP = 1 if FPMA supported
 SP = 1 if SPMA supported

A = 1b if FCF available for Login/ELP

F = 1b

FIP Operation Code = 0001h		Reserved	SubCode = 02h
Descriptor List Length = 13		FP	SP
		Flags	
Type = 1	Len = 1	Reserved	Priority
Type = 2	Len = 2		
FCF-MAC Address			
Type = 3	Len = 2	Reserved	
FC-MAP			
Type = 4	Len = 3	Reserved	
Switch_Name			
Type = 5	Len = 3	Reserved	
Fabric_Name			
Type = 12	Len = 2	Reserved	
FKA_ADV_PERIOD			
Padding to Max Receive Size of soliciting entity, if solicited (i.e., if S=1b), otherwise no padding			

Solicitations Generation Rules

- **Solicitations are sent:**

- By ENodes' MACs to solicit VF_Port capable FCF-MACs

- F=0b

- By VE_Port capable FCF-MACs to solicit other VE_Port capable FCF-MACs

- F=1b

- **Multicast Solicitations are sent to the multicast address 'All-FCF-MACs'**

Solicitations Processing Rules

- **The Switch_Name descriptor is used by a VE_Port capable FCF to detect Solicitations coming from its own FCF-MACs**

If the value carried in the Switch_Name descriptor is the FCF's own Switch_Name, the FCF does not respond

- **The FC-MAP descriptor is used by a VE_Port capable FCF-MAC to detect anomalous network configurations**

Because the FC-MAP shall be the same for FCFs belonging to the same network

An FCF ignores a Solicitation if the value carried in the FC-MAP descriptor is not equal to its configured FC-MAP

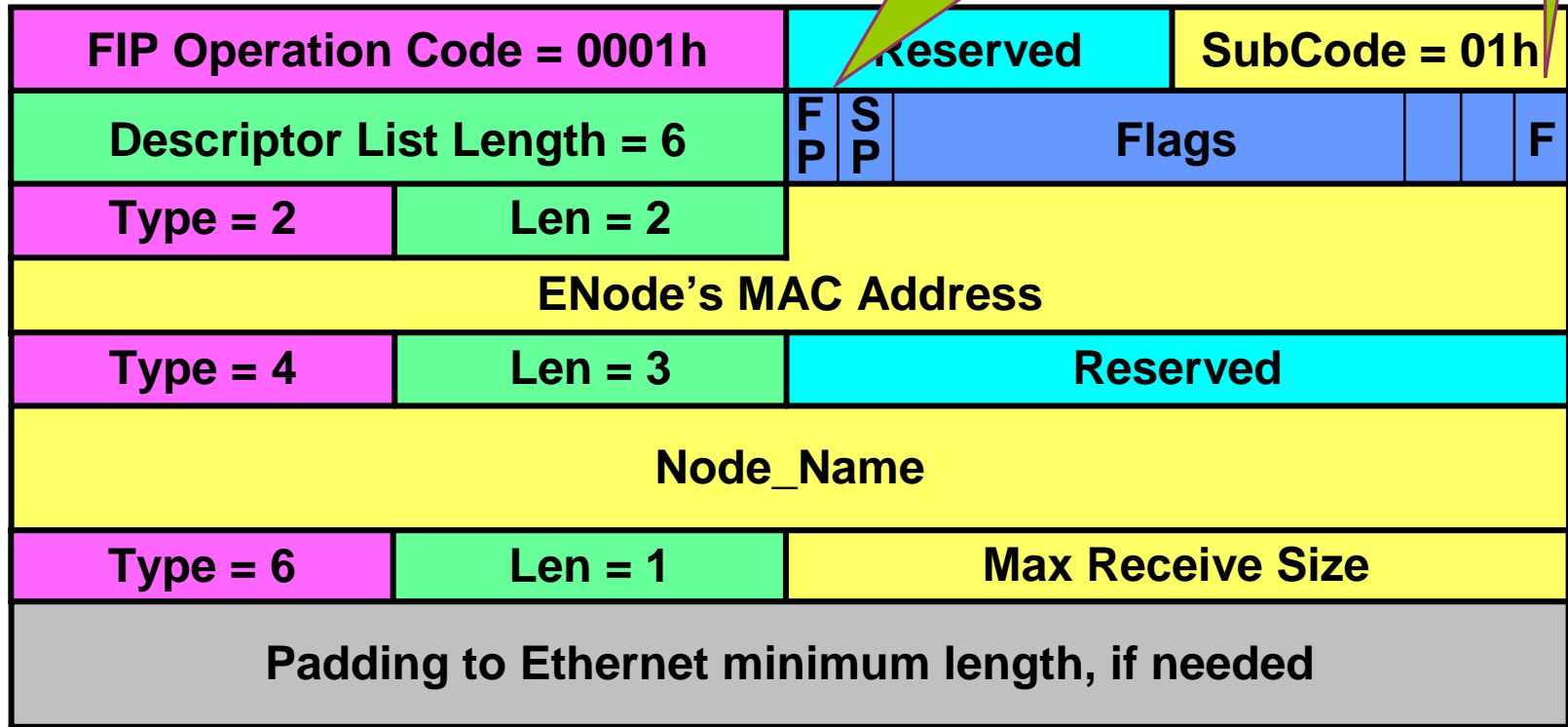
An exception may be signaled

- **The FC-MAP descriptor may be ignored if set to zero**

Solicitation from ENode

ENode Capabilities:
 FP = 1 if FPMA supported
 SP = 1 is SPMA supported

F = 0b

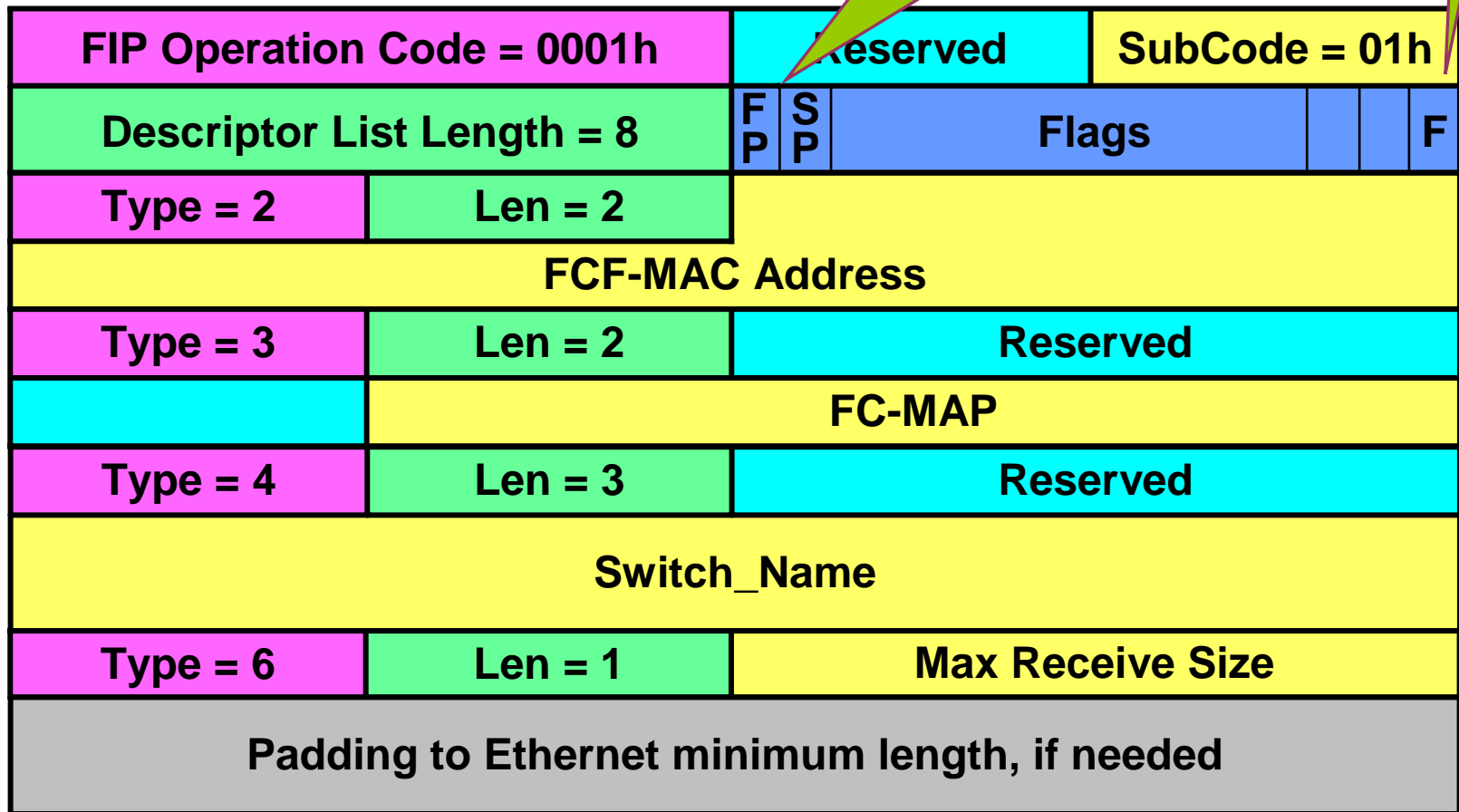


Solicits VF_Port capable FCF-MACs

Solicitation from FCF

FCF Capabilities:
 FP = 1 if FPMA supported
 SP = 1 is SPMA supported

F = 1b

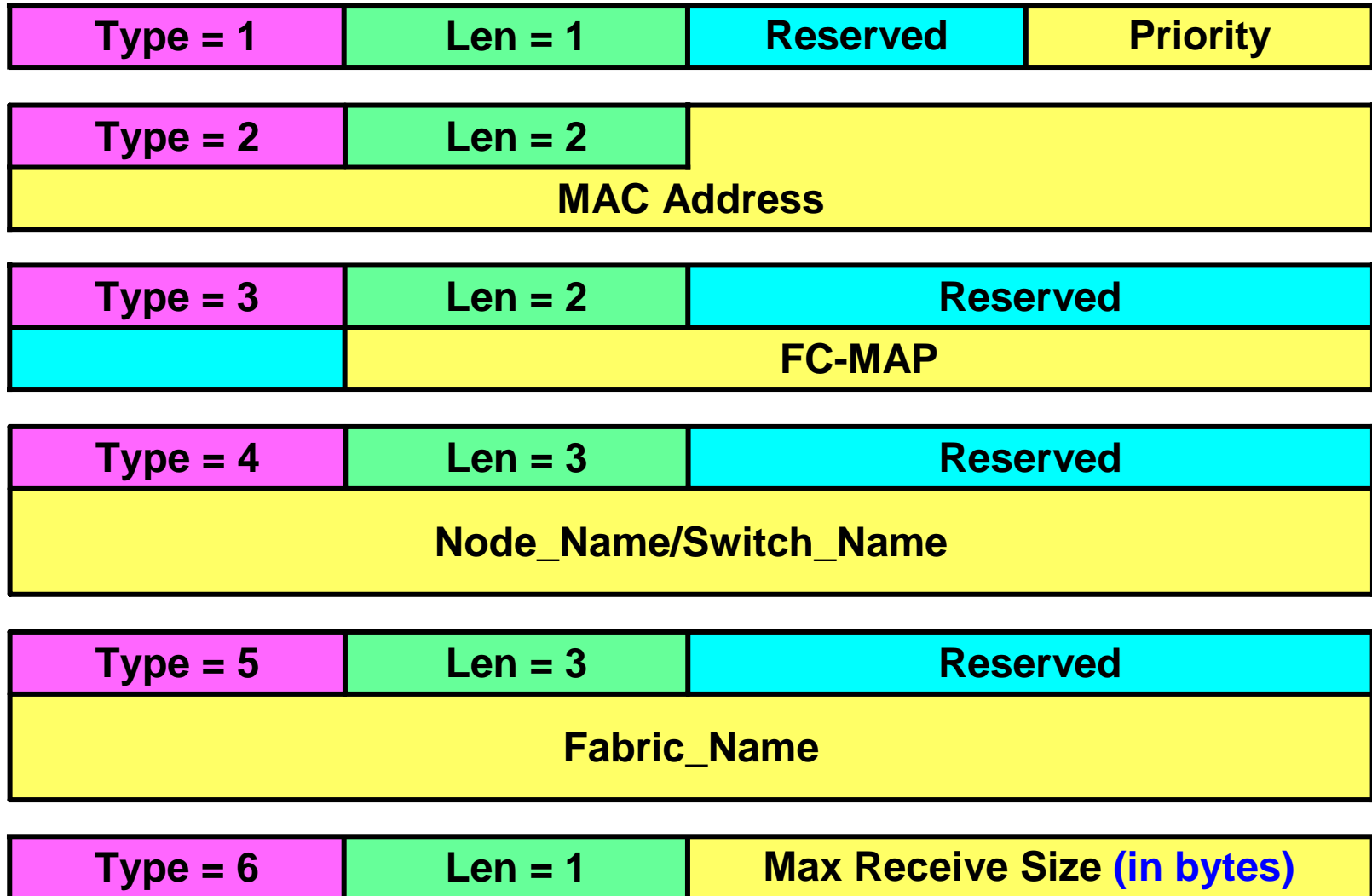


Solicits VE_Port capable FCF-MACs

Agenda

- FCoE “Soft State”
- FIP Clear Virtual Link
- FIP Discovery Summary
- **FIP Descriptors Summary**
- FIP Messages Summary

FIP Descriptors (1)



Lengths are measured in 32-bit words

FIP Descriptors (2)

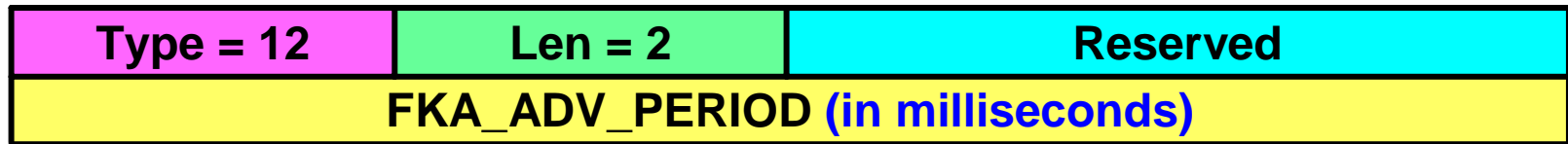
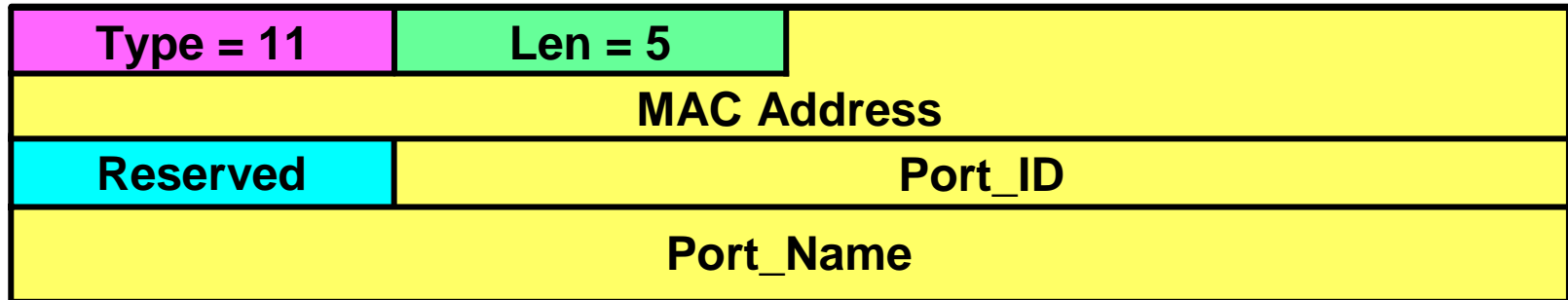
Type = 7	Len = 36/9	Reserved
FLOGI Request, FLOGI LS_ACC/LS_RJT (no CRC, SOF, nor EOF)		

Type = 8	Len = 36/9	Reserved
NPIV FDISC Request, FDISC LS_ACC/LS_RJT (no CRC, SOF, nor EOF)		

Type = 9	Len = 11/8/9	Reserved
Fabric LOGO Request, LOGO LS_ACC/LS_RJT (no CRC, SOF, nor EOF)		

Type = 10	Len = 33/9	Reserved
ELP Request, ELP SW_ACC/SW_RJT (no CRC, SOF, nor EOF)		

FIP Descriptors (3)



- **FKA_ADV_PERIOD**

- Should be equal among all FCFs

- Default value: 8000 milliseconds

- Special value: 0

- If set to zero, no Keep Alives are requested

Do we need anything else?

Compatibility of Descriptors

- **At every iteration of the FIP definition we defined some new descriptors**

It is quite likely this may happen also in the future

We need also to allow some experimentation

Deployments are just starting

- **How to ensure backward compatibility?**

Split the Descriptor Type space

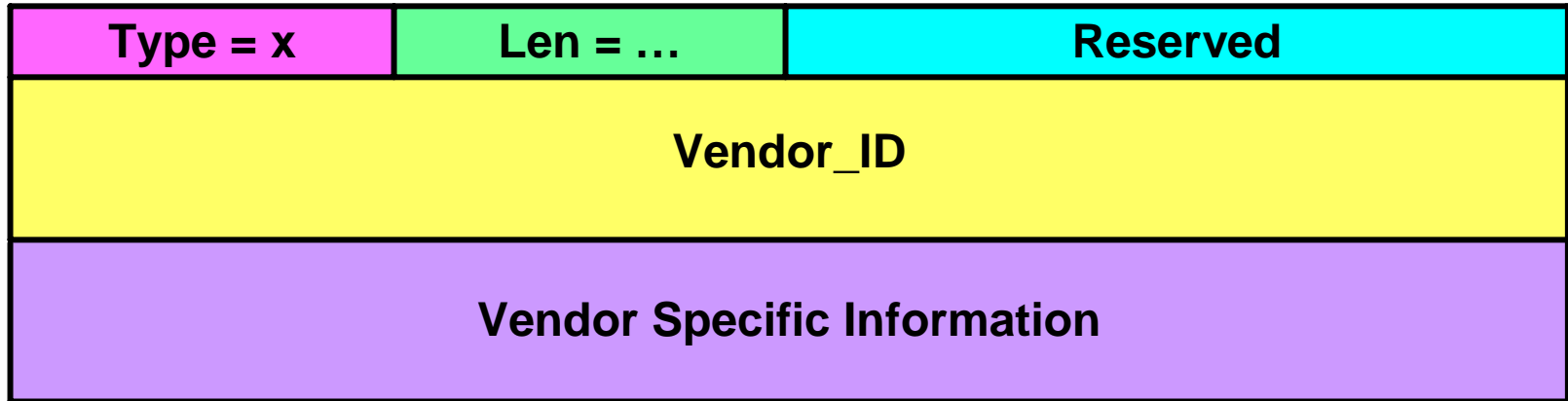
0-127: Critical. An implementation receiving a FIP message carrying an unknown critical descriptor shall ignore the entire message

128-255: Non-critical. An implementation receiving a FIP message carrying an unknown non-critical descriptor shall ignore the unknown descriptor and process the message

Descriptor Types

Criticality	Type	Name
Critical	0	Reserved
	1	Priority
	2	MAC Address
	3	FC-MAP
	4	Node_Name/Switch_Name
	5	Fabric_Name
	6	Max Receive Size
	7	FLOGI
	8	NPIV FDISC
	9	Fabric LOGO
	10	ELP
	11	VN_Port Identification
	12	FKA_ADV_Period
	13	Vendor_ID
14 – 127	Reserved	
Non-Critical	128 – 240	Reserved
	240 – 254	Vendor Specific
	255	Reserved

Vendor Specific Descriptors



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

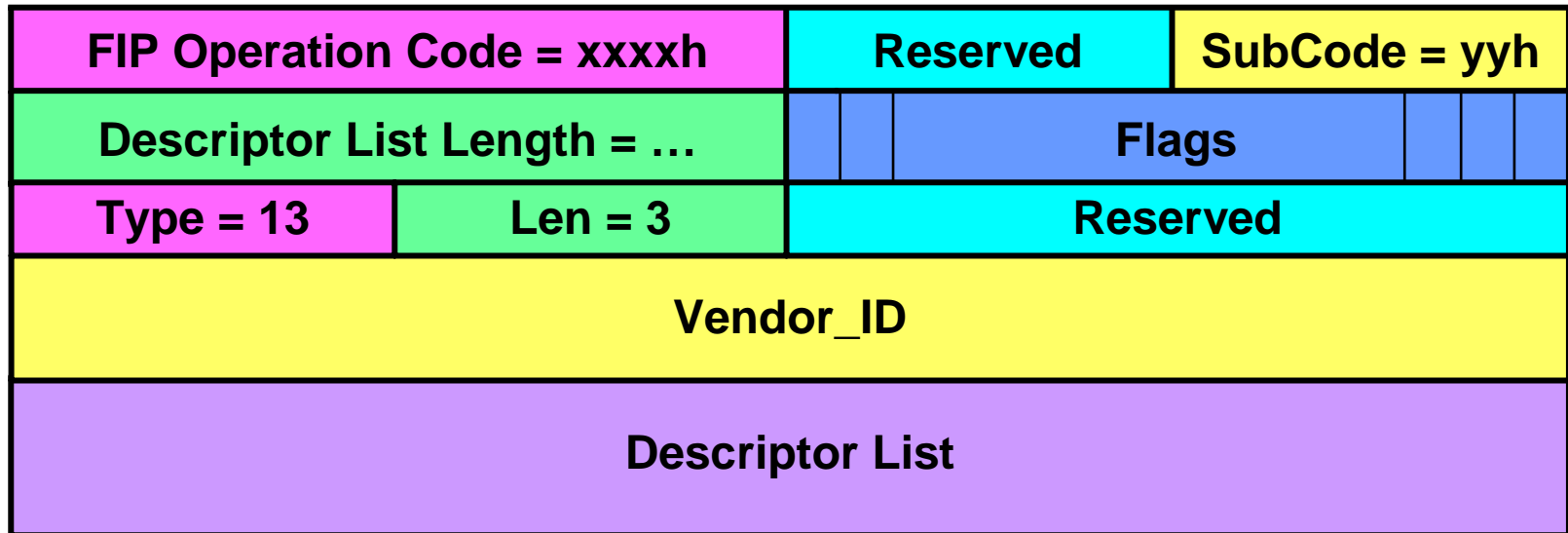
Agenda

- FCoE “Soft State”
- FIP Clear Virtual Link
- FIP Discovery Summary
- FIP Descriptors Summary
- FIP Messages Summary

Vendor Specific 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
0003h	01h	FIP Keep Alive
	02h	FIP Clear Virtual Link
FFF8h .. FFFEh	00h .. FFh	Vendor Specific
All others	All others	Reserved

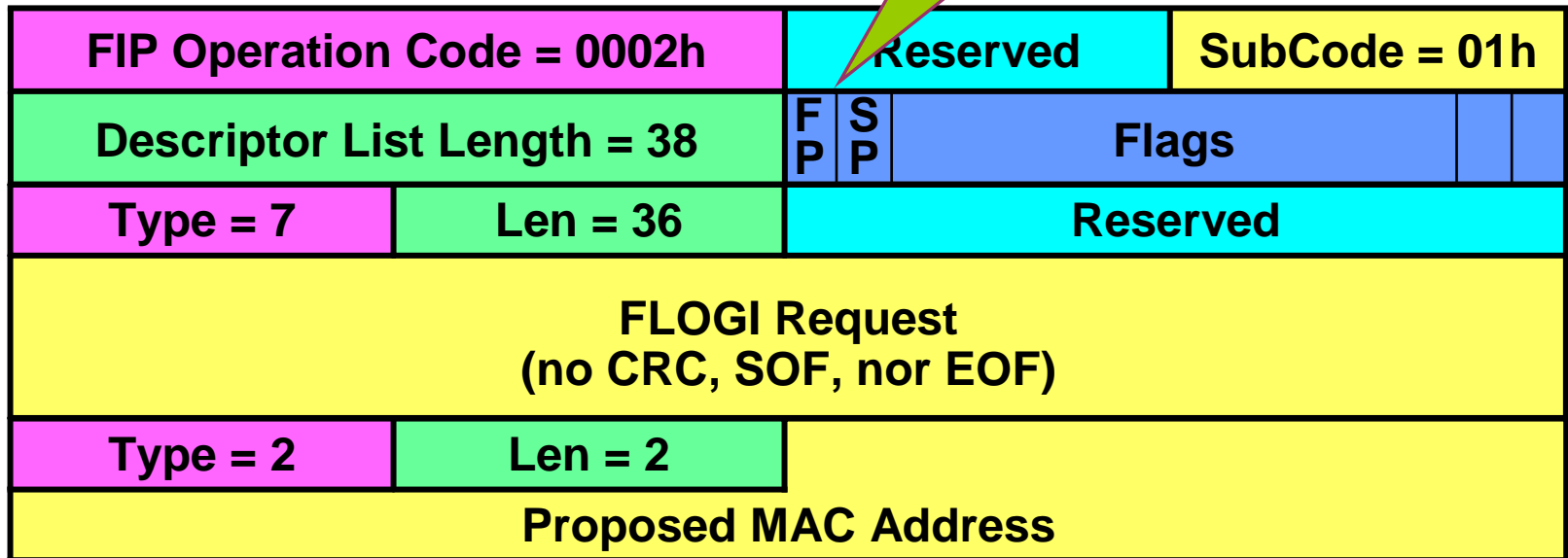
Vendor Specific FIP Messages



- 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

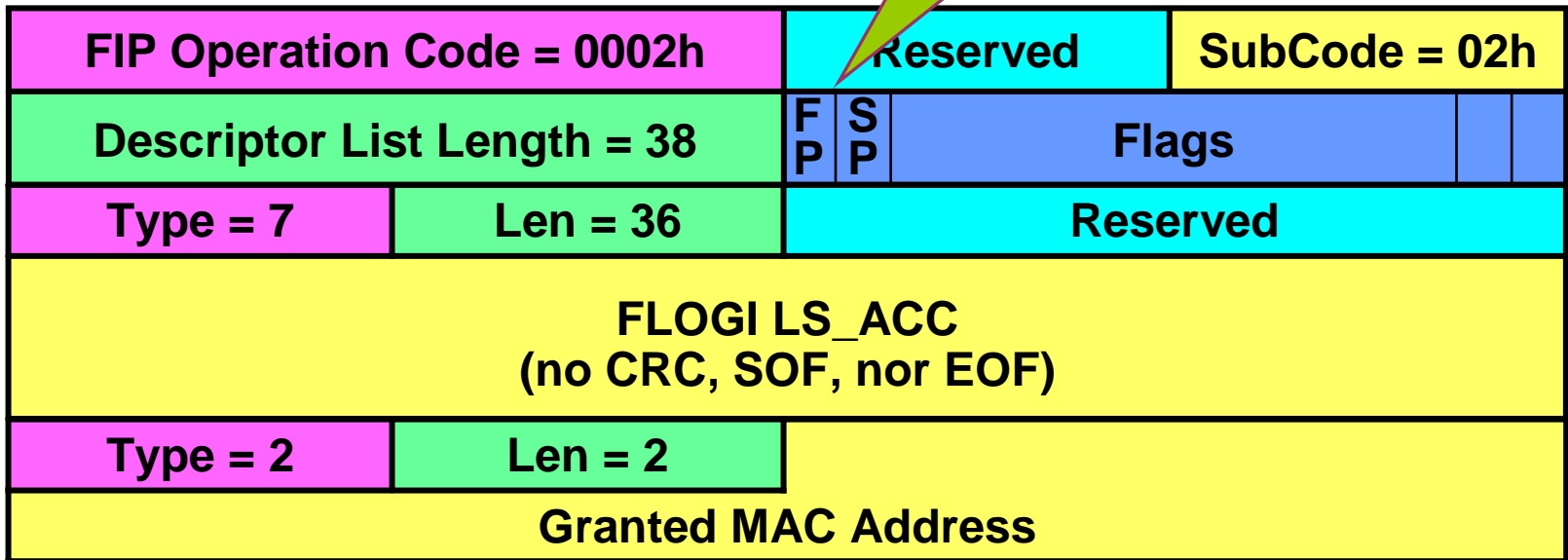
FIP FLOGI Request (ENode → FCF)

FP = 1 if FPMA requested, or
SP = 1 if SPMA requested

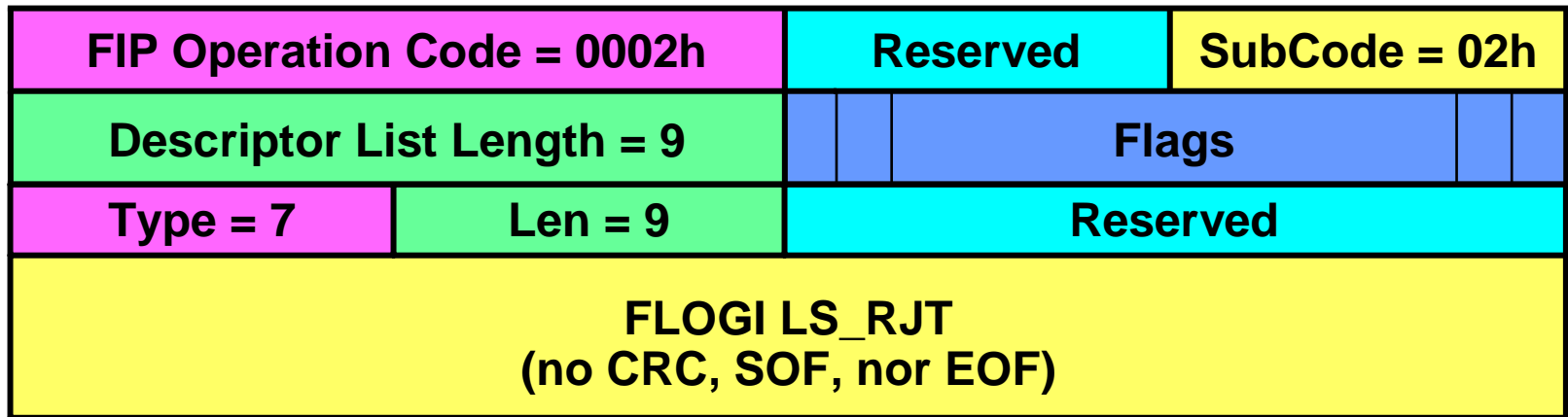


FIP FLOGI Accept (FCF → ENode)

FP = 1 if FPMA granted, or
SP = 1 if SPMA granted

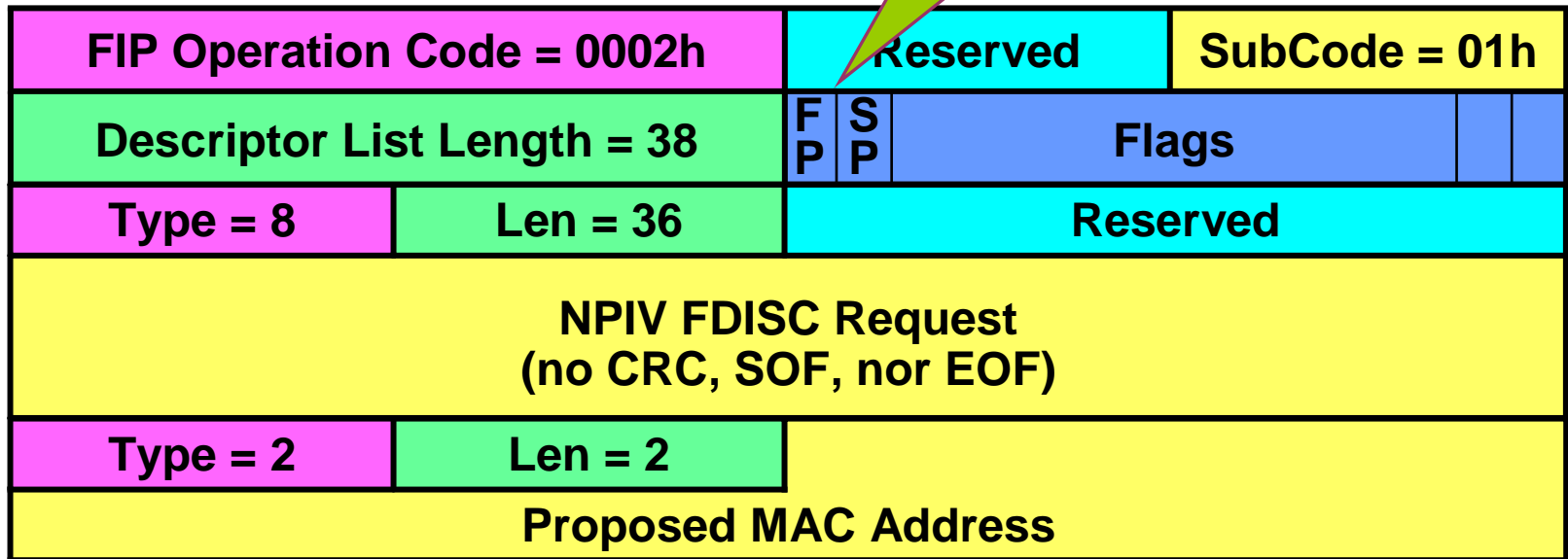


FIP FLOGI Reject (FCF → ENode)



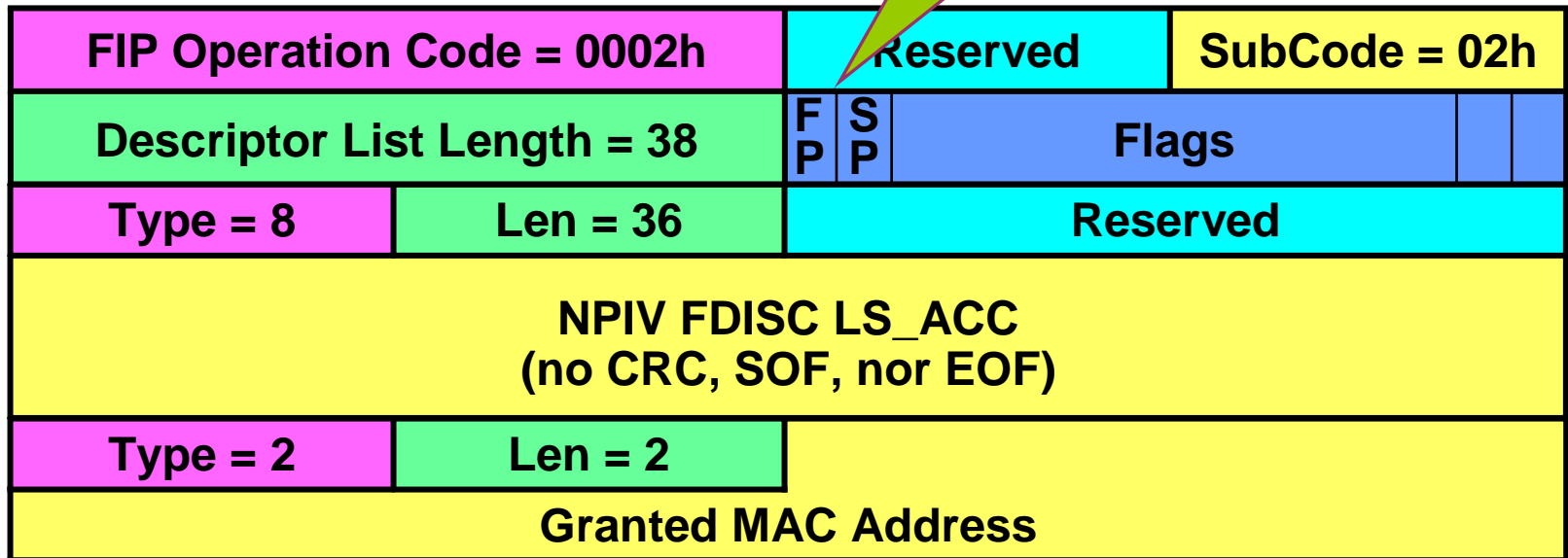
FIP NPIV FDISC Request (ENode → FCF)

FP = 1 if FPMA requested, or
SP = 1 if SPMA requested

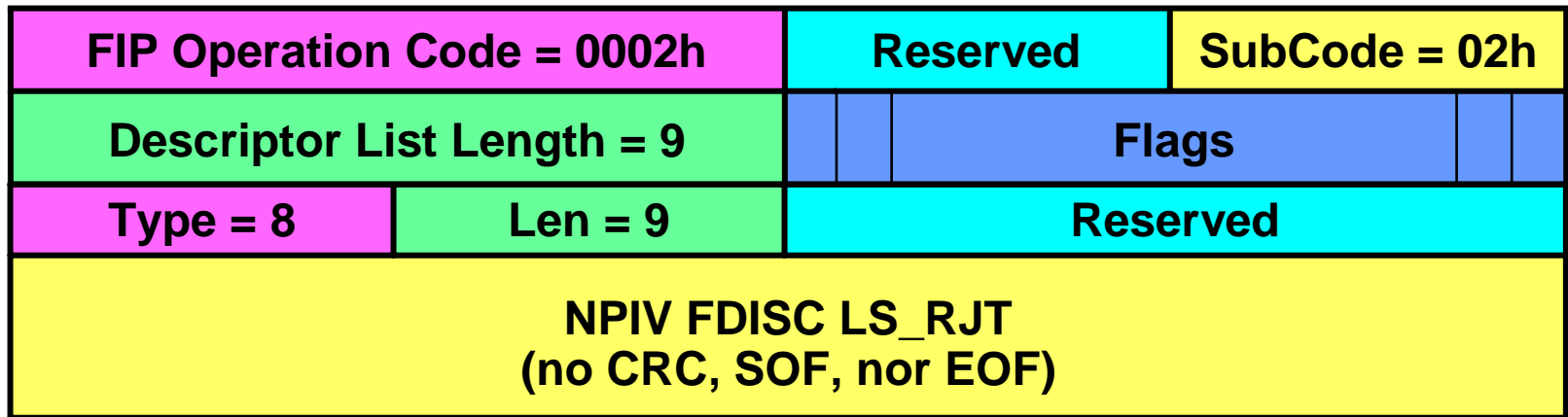


FIP NPIV FDISC Accept (FCF → ENode)

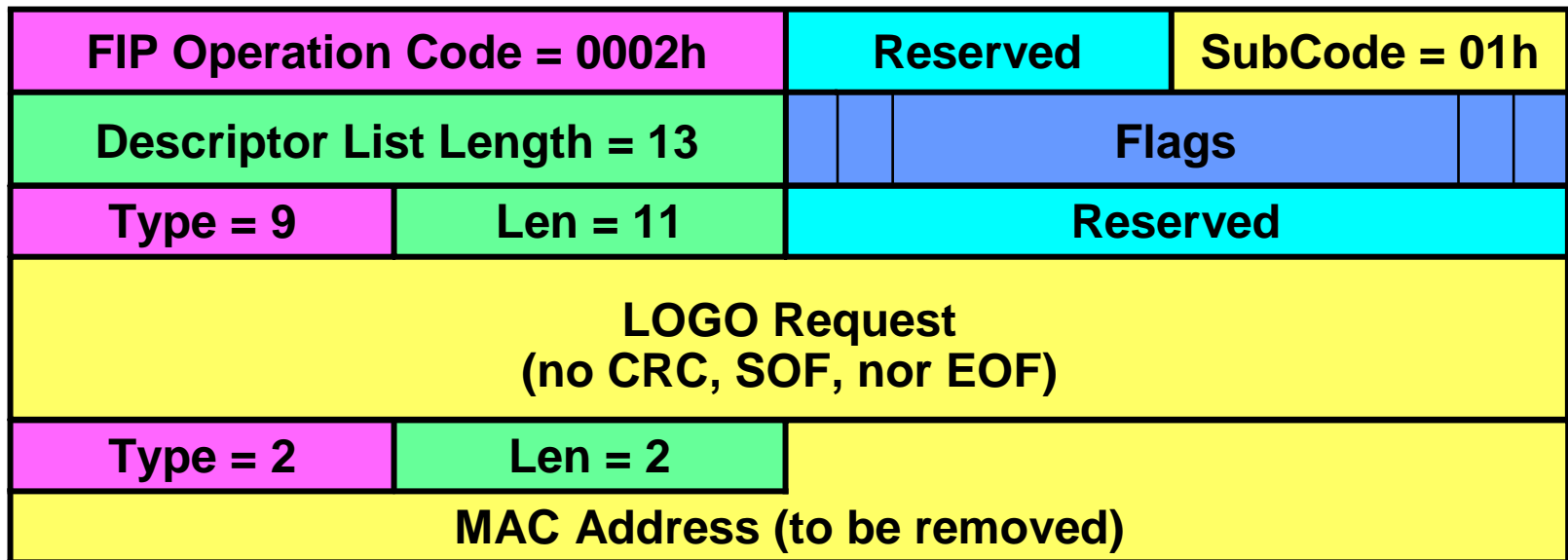
FP = 1 if FPMA granted, or
SP = 1 if SPMA granted



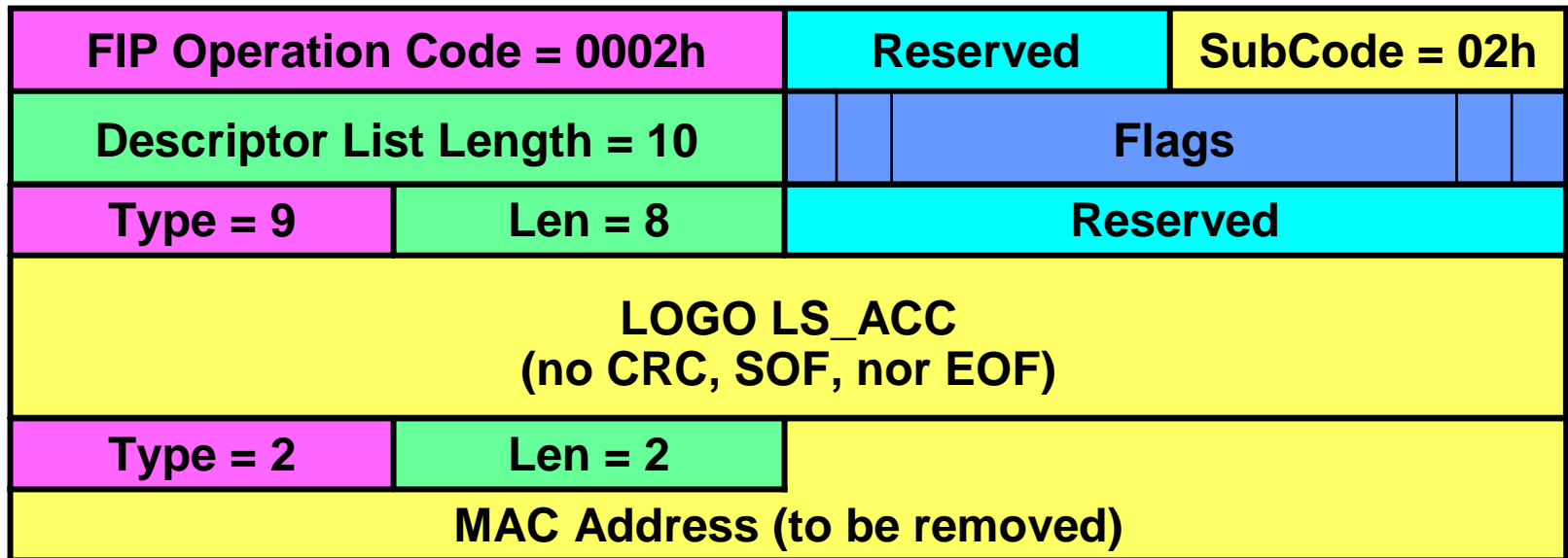
FIP NPIV FDISC Reject (FCF → ENode)



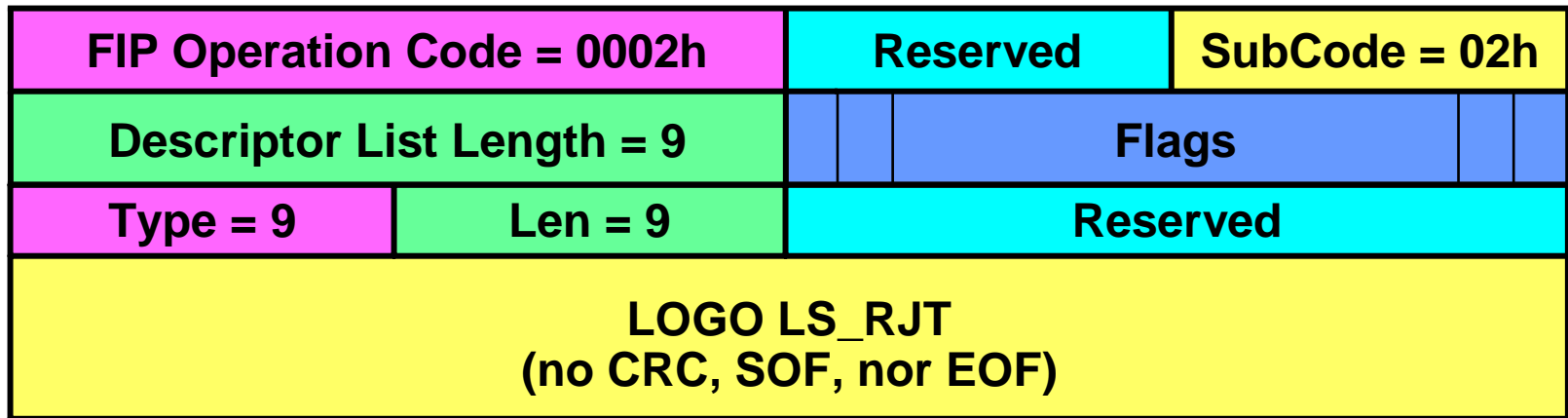
FIP Fabric LOGO Request (ENode → FCF)



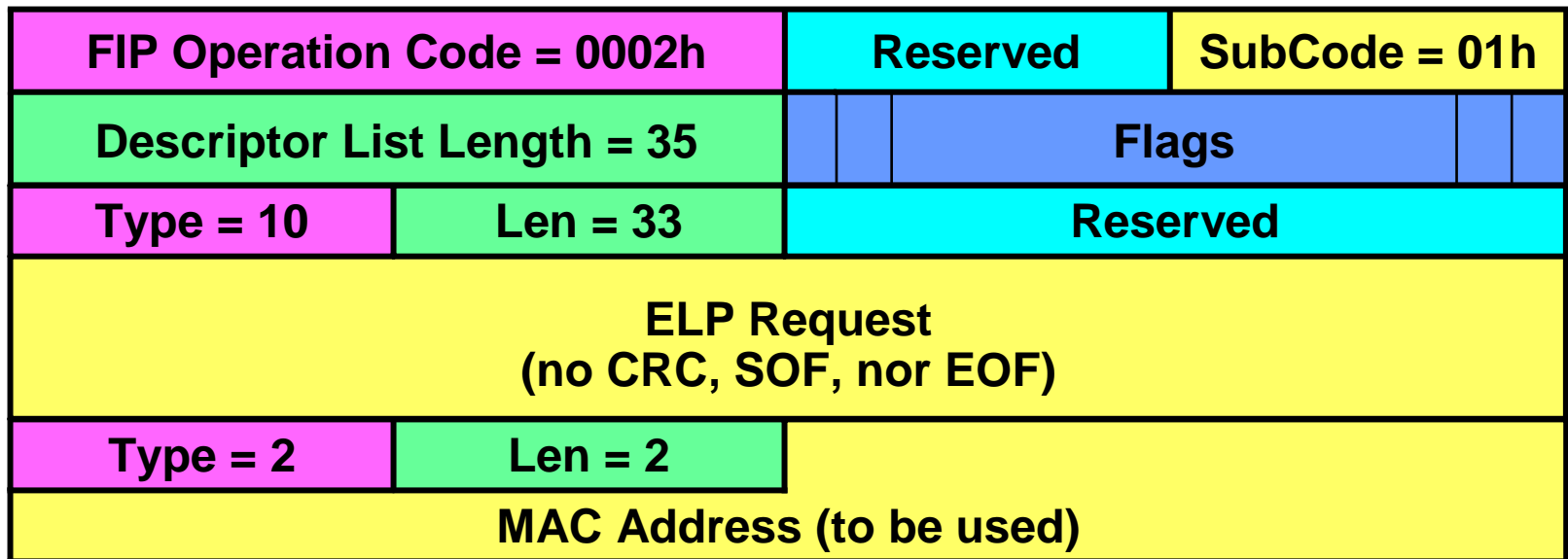
FIP Fabric LOGO Accept (ENode ← FCF)



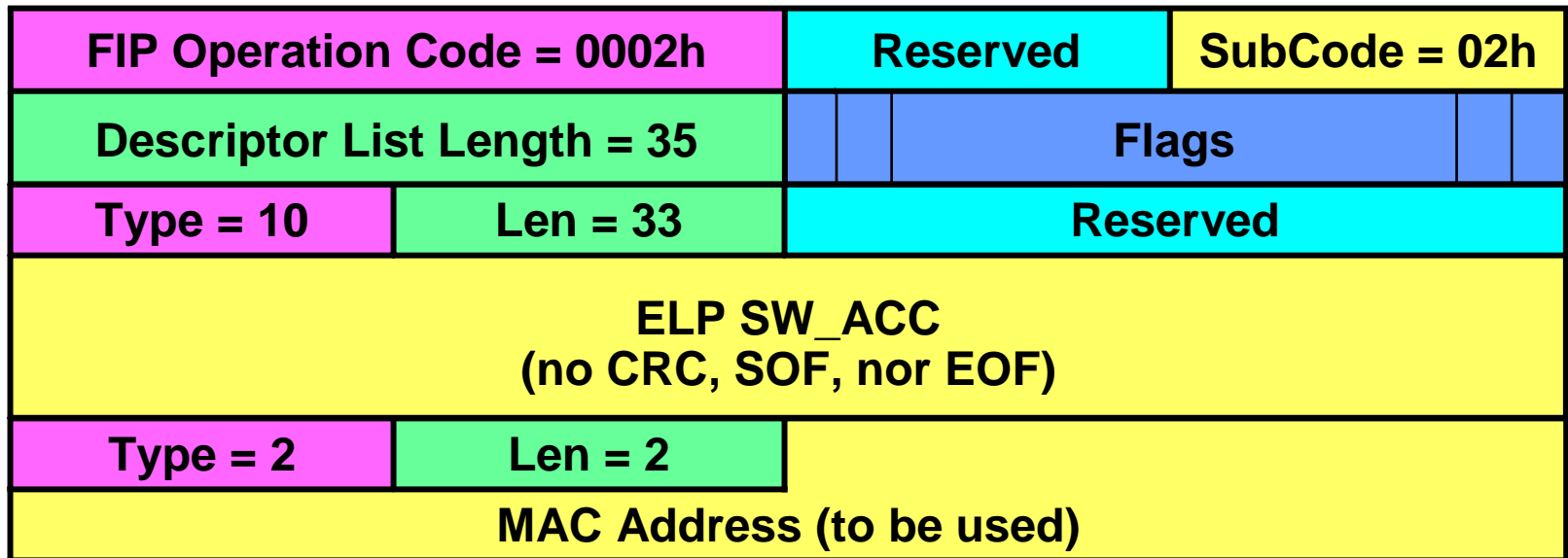
FIP Fabric LOGO Reject (ENode ← FCF)



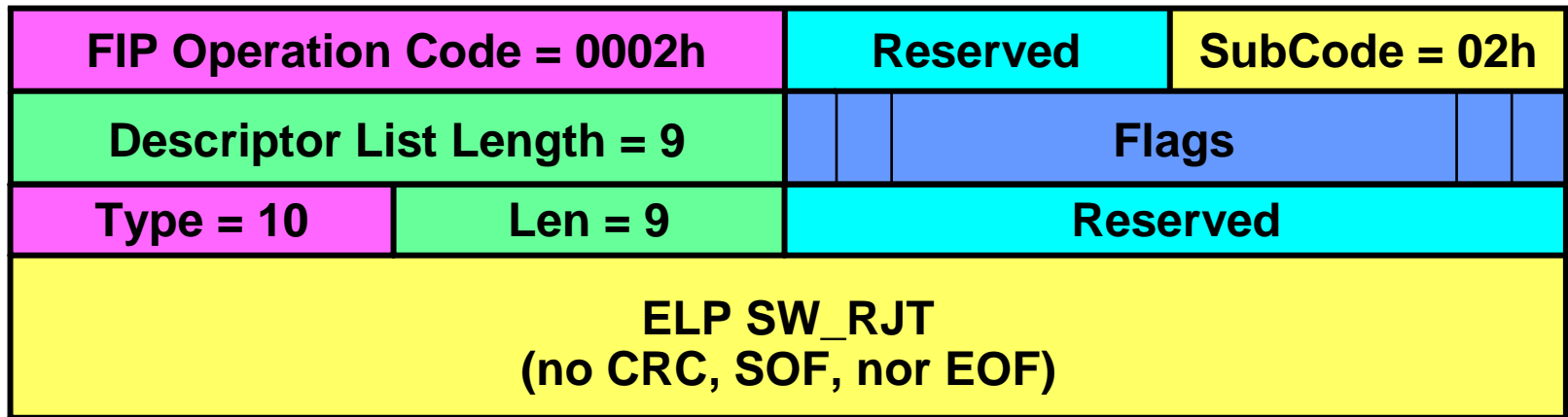
FIP ELP Request (FCF \leftrightarrow FCF)



FIP ELP Accept (FCF \leftrightarrow FCF)



FIP ELP Reject (FCF \leftrightarrow FCF)



Motion

- **To accept the FIP protocol specified in this document as the FIP protocol to be specified in FC-BB-5**

Thank You