

Flexible FDF Connectivity Mechanism

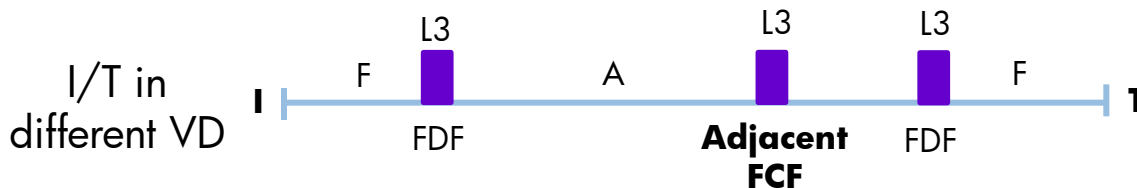
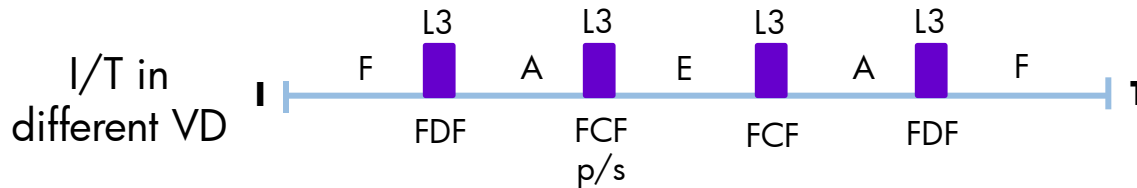
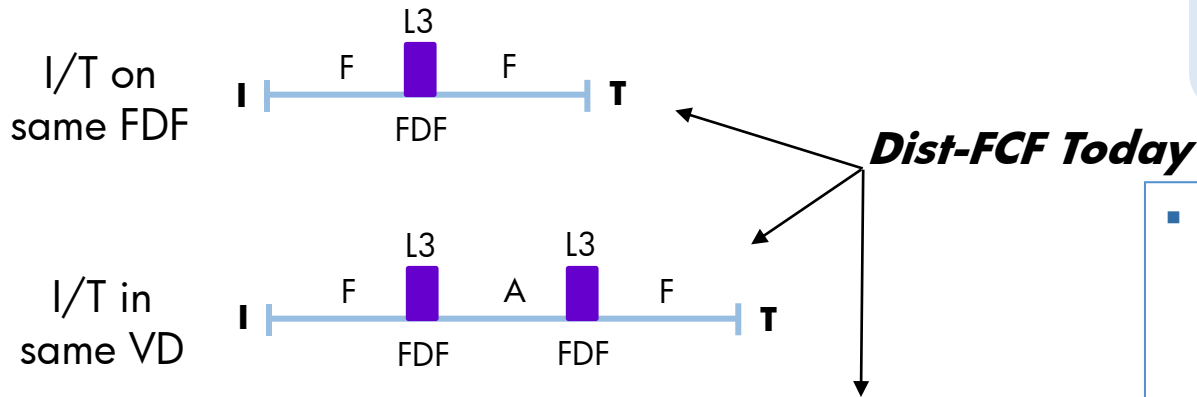
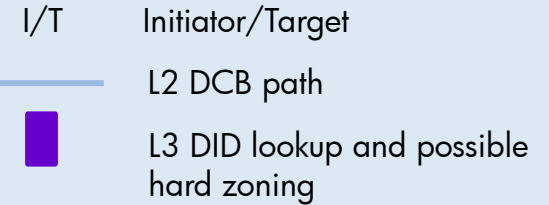
Siamack Ayandeh
Chief Architect (Datacenter)
HP Networking

October 17th, 2011
Siamack@HP.Com

Puttagunta, Krishna (Babu)
HP Storage Division



Desired Goal



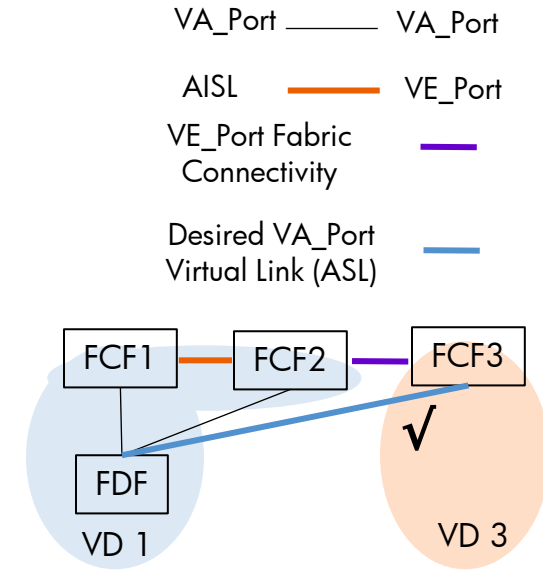
- Given the feedback in the Albuquerque meeting this more limited goal:
 - Minimizes change to existing drafts and standard's time schedule
 - Minimizes any additional complexity by re-using existing FIP ELP and VA_Port protocols
 - While enhancing the architecture

Flexible FDF Connectivity

Goal of Flexible FDF Connectivity is to allow an FDF to forward to an adjacent (egress) FCF with minimal change to existing drafts

Flexible FDF Connectivity Mechanisms -- summary

- Allow a VA_Port virtual link *for forwarding only* between an FDF and an **adjacent FCF** (blue link in diagram)
- Requires:
 - Addition of an adjacent FDF set to an FCF (FCF3 in diagram, please see next 2 slides)
 - Allow an FCF to setup VA_Port virtual links to this list of FDFs using FIP ELP (e.g. allow ELP with C=D=0 or add an additional bit) indicating that FCF is an adjacent FCF
 - Re use the VA_Port protocol, FDRN, FDUN, and add one new VA_Port SW_ILS (Please see slides 7 and 8)



FCF to FDF VA_Port virtual link setup procedures -- 11-026v3 recap

- FDF VA_Port_virtual_link MACs are discovered by FCFs through FIP discovery
- FCFs initiate the FIP ELP exchange (and reject any initiation from FDFs):
 - “A VA_Port/VE_Port capable FCF-MAC shall reject a received FIP ELP Request with the FDF/FCDF flag set to one with Reason Code ‘Protocol Error’ and Reason Code Explanation ‘Invalid Request’ “
 - “the FIP ELP Exchange shall be initiated by the VA_Port/VE_Port capable FCF-MAC and the VA_Port capable FDF-MAC shall reply, irrespective of the value of the Switch_Name field in the ELP payload”
- When **operational** a **FCF-MAC** shall instantiate VA_Port virtual links with VA_Port capable FDF-MACs belonging to FDFs that are part of the Distributed FCF’s FDF Set and discovered by FIP discovery
 - FIP ELP REQ from c/FCF to FDF with C=1 & D=0
 - FIP ELP ACC from FDF to c/FCF with C=0 & D=1
- So virtual domain configuration i.e. controlling FCF Set and FDF Set are only defined in the c/FCF
 - and communicated to an FDF using VA_Port protocol DFMD after VA_Port virtual link setup between c/FCF and an FDF

Adjacent FCF to FDF VA_Port virtual link setup procedures -- new

- We follow the same principle of operation as in 11-026v3 i.e. an adjacent FCF initiates an ELP exchange with an FDF based on its configuration of **adjacent FDF set**
- An adjacent FCF can be identified in FIP header either by:
 - Addition of an “adjacent FCF” bit in FIP header
 - Or setting of combination of C and D bits e.g. C=D=0
- **Fabric connectivity** of “adjacent FCF”, FCF2 to FCF3 virtual link, is a pre-requisite for successful completion of FIP_ELP and setup of the VA_Port virtual link
- Open to consideration of any additional protection say a complimentary **adjacent FCF set** per FDF as an option
- Following 11-026v3 text needs to be added/edited:

Add: section 1.1.5.1 Adjacent FDF Set: The Switch-Names of the FDFs that are part of an “adjacent FDF set”

Section 1.2 A distributed FCF is defined ... add c) the Switch_Names of the Adjacent FDF Set

Section 1.6.1 An FDF does not establish VE_Port to VE_Port Virtual Links, therefore an FDF-MAC shall reject a received FIP ELP Request with both Controlling FCF/Switch flag and FDF/FCDF flag set to zero (i.e., a FIP ELP Request coming from an FCF that is not a Controlling FCF) with Reason Code ‘Protocol Error’ and Reason Code Explanation ‘Invalid Request’.

An FDF-MAC shall also reject a received FIP ELP Request coming from a Controlling FCF other than the Controlling FCFs that define its Distributed FCF, with Reason Code ‘Logical Error’ and Reason Code Explanation ‘Not Authorized’.

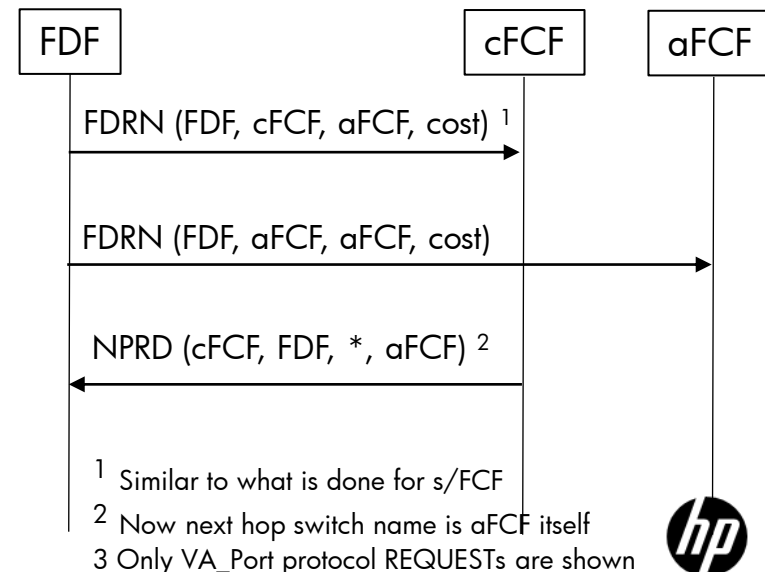
FDF to Adjacent FCF Route Setup – using 11-225v1

- The FCDF Reachability Notification (FDRN) SW_ILS is used by an FCDF to communicate to the Primary Controlling Switch that it has instantiated an ASL with another FCDF or with the Secondary Controlling Switch
- The same construct can be used for an FDF to notify its primary FCF that it has an VA_Port virtual link with an **adjacent FCF** (which is in a different domain)
- The c/FCF can use the adjacent_FCF switch name to determine its domain_ID and the FDRN Link Cost to calculate routes for the FDF's in its domain
- This information is then advertised in NPRD “Reachable Domain_ID” fields
- This SW_ILS (as well as the corresponding FDUN) is re-used between the FDF and its adjacent FCF

Table 6 - FDRN Request payload

Item	Size (bytes)
SW_ILS Code = XX00 0003h	4
Originating FCDF Switch_Name	8
Destination Controlling <i>or Adjacent Switch</i> Switch_Name	8
Reachable FCDF, Controlling Switch, <i>or Adjacent Switch</i> Switch_Name	8
Reserved	2
Link Cost	2

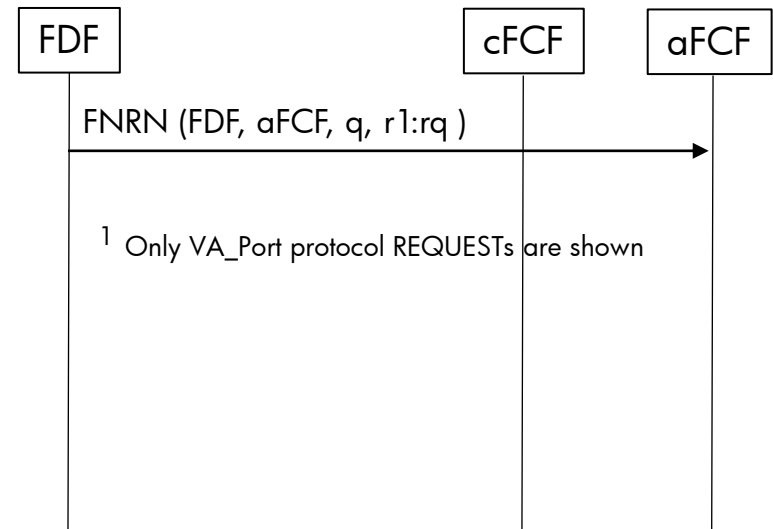
Edits to 11-225v1 are in blue
Similarly for FCDF Unreachability Notification (FDUN)



Adjacent FCF to FDF Route Setup

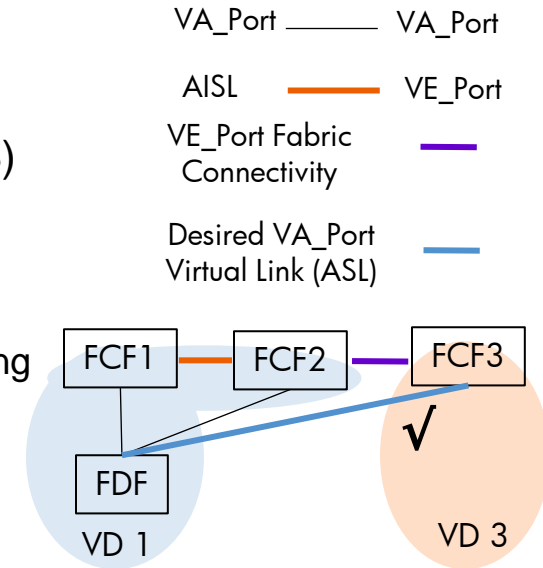
- From routing perspective an adjacent FCF has a:
 - VA_Port virtual link to an adjacent domain FDF identified by a Switch_Name and
 - Possibly multiple VE_ports to FDF's domain_ID (i.e. via the c/s FCF) and
 - Possibly multiple path to the c/s FCF domain_ID (i.e. via other FCFs)
- N_Port_ID ranges terminated by an FDF should be communicated to the adjacent FCF using an FDF N_Port_ID range notification (FNRN) using a new SW_ILS

Item	Size (bytes)
SW_ILS Code = XX00 000xh	4
Originating FCDF Switch_Name	8
Destination Adjacent Switch Switch_Name	8
Number of N_Port_ID Ranges (q)	4
N_Port_ID Range #1	4
N_Port_ID Range #2	4
N_Port_ID Range #q	4



Adjacent VA_Port virtual link maintenance

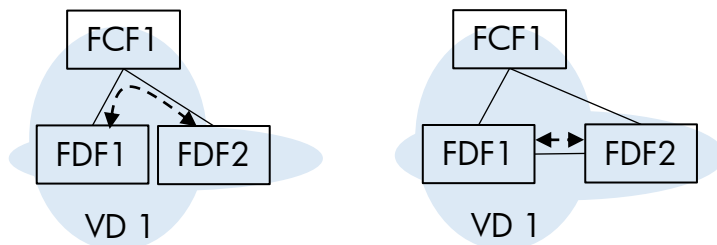
- **Fabric connectivity** of “adjacent FCF”, over **FCF2 to FCF3 virtual link**, is a pre-requisite for successful completion of FIP_ELP and setup of the VA_Port virtual link (from FDF to FCF3)
- If FCF3 loses fabric connectivity over its VE_Port virtual links or E_Ports then:
 - Routing tables would be updated in FCF1 and FCF3 for the corresponding domains via FSPF; this clears the forwarding path from FCF3 to FDF
 - And via a VA_Port protocol NPRD in FDF
 - FCF3 would de-instantiate the VA_Port virtual link using a FIP Clear Virtual Link
- Remaining slides identify the gains to BB6 model due to this additional connectivity



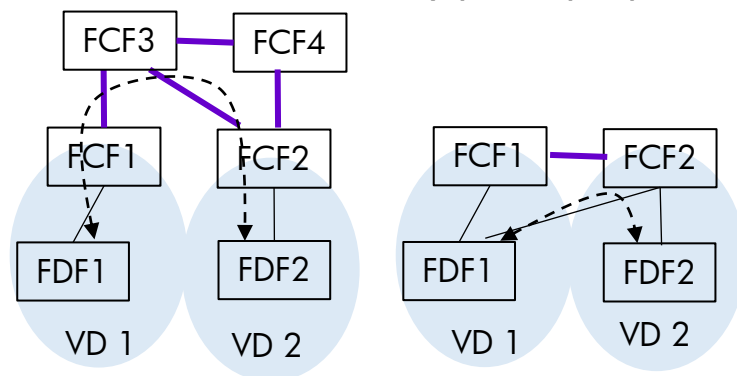
Gains – Flatter Topology

- Topology becomes more flexible leaving the options open for a topologically flatter network
 - This saves on cost
 - Future proofs to meet demands of new emerging rack designs
- Fewer L3 hops reduce latency and remove potential choke points; FCF by-pass is a goal of BB6

Intra domain FCF by-pass (already part of BB6)



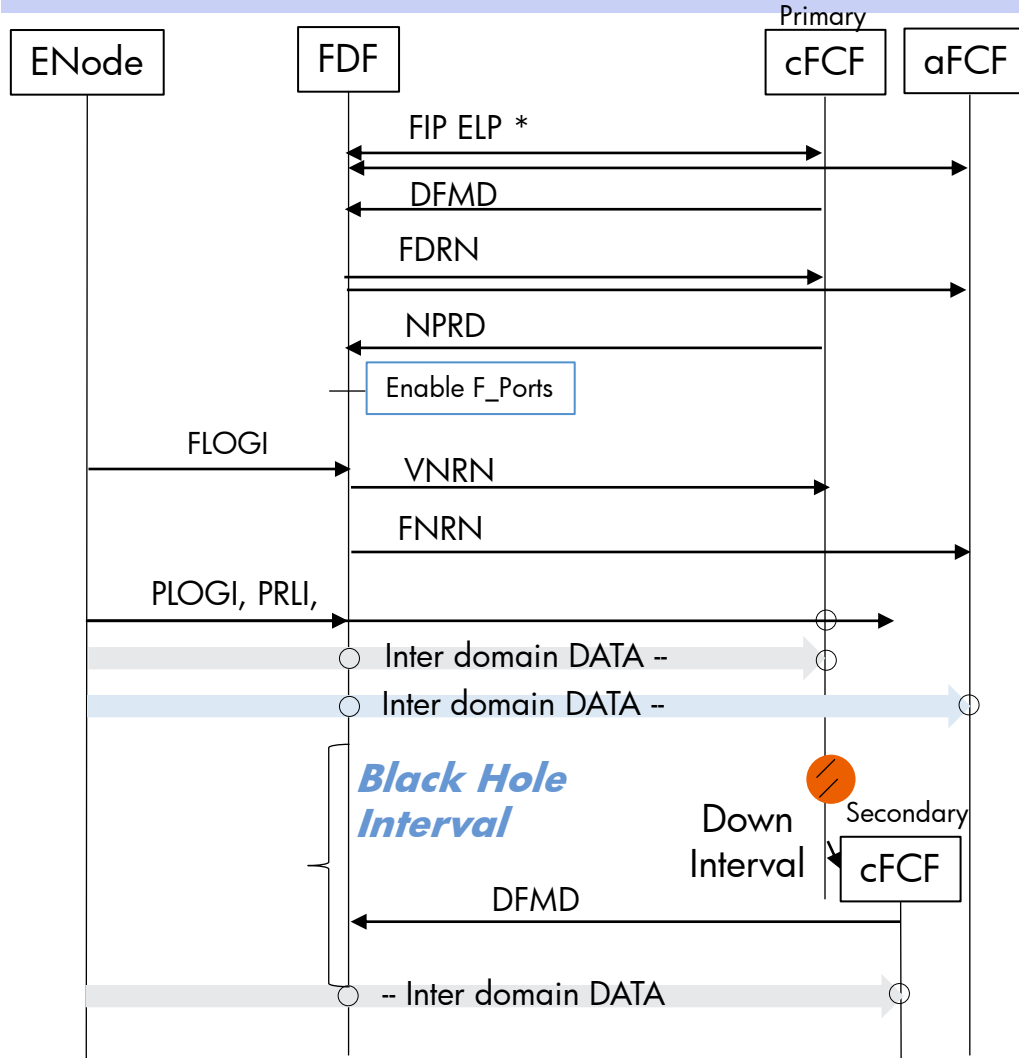
Inter domain FCF by-pass (proposed)



Gains – Control over load balancing independent of HA

- Gain control over load balancing as presence of a secondary FCF is optional, so as a minimum an alternate data path is desirable in the absence of s/FCF for load balancing
- Why is secondary FCF optional?
 - Please see contribution 11-xyz “High Availability Models of Operation in Fibre Channel”
- Hence having an alternative path from an FDF to adjacent FCFs for forwarding only, without sharing connection state, is desirable
 - This allows for load balancing to/from FDFs

Gains -- Load balancing reduces lost traffic in fail over event



- During a switch over from primary to secondary FCF data traffic is black holed for duration of Down_Interval+
- Load balancing using adjacent FCFs reduces the volume of lost data traffic and reduces impact on applications

* Only some protocol REQUESTs are shown

Gains -- Summary

- Topology becomes more flexible leaving the options open for a topologically flatter network
 - This can save on cost
 - Future proofs the standard to meet requirements of new emerging rack designs over the next 3 to 5 years
- Fewer L3 hops reduce latency and remove potential choke points; FCF by-pass is a goal of BB6
- Gain control over load balancing as presence of a secondary FCF is optional, so as a minimum an alternate data path is desirable in the absence of s/FCF for load balancing
- Load balancing reduces lost traffic to black holes during fail over hence improving HA operations
- BB6 group requirements are met
 - Technical: virtual domain preservation, FCF by-pass, use of trusted fabric/FDF for forwarding based on L3 at the edge where hardware based zoning is applied, and no change to adapters
 - Schedule and effort: Impact is kept to a minimum by re-using existing protocols

Thank You

