BROCADE

FC-SW-6
DISTRIBUTED SWITCH
>2 CONTROLLING SWITCHES

David Peterson

Howard Johnson

Office of the CTO

T11/14-123v0

Distributed Switch >2 cSwitches

Distributed Switch Operation

- Startup Process (i.e., set the base for further discussion)
- One Controlling Switch (review)
- Two Controlling Switches (review)
- Three Controlling Switches
  - Supports mesh topology
  - Supports non-mesh topology
  - RHello(s) between Primary and Secondary only
    - No RHello(s) to/from additional Controlling Switch(es)
  - Differences
    - Primary cSwitch selection (ERP vs BDS/EDSP)
    - Brocade – push dSwitch info via Distributed Switch Sync (DSS), VSD, DSD
    - Cisco – pull dSwitch info via Get FCDF Topology State (GFTS) and Get FDCF N_Port_IDs State (GFNS)

Distributed Switch

Startup Process (1)

- Perform Fabric initialization and configuration as normal SW-6 Switch
- Determine Controlling Switch role (i.e., Primary, Secondary, or Additional, via redundancy protocol)
- Controlling Switches instantiate A_Port link(s) with FCDF(s) and FCDFs send FDRNs to PcSwitch
  - FCDF(s) send FDRN(s) to PcSwitch
  - PcSwitch sends received FDRN info to ScSwitch and AcSwitch(es), if present

Distributed Switch

Startup Process (2)

- PcSwitch sends DSMD to FCDF(s)
- PcSwitch sends NPRD to all FCDF(s)
  - *NPRD is not sent to ScSwitch or AcSwitch
- FCDF(s) now able to initiate ELP to other FCDF(s)

Distributed Switch

Startup Process (3)

- FCDF(s) now able accept FLOGI(s) from end device(s)
  - FCDF(s) send VNRN request to PcSwitch
    - PcSwitch allocates N_Port_ID, recomputes Zoning ACLs, generates RSCNs, updates FC Name Server
    - PcSwitch sends appropriate NPZD(s) to ScSwitch and AcSwitch(es), if any, and each FCDF
    - After receiving NPZD DS_ACC from ScSwitch and FCDF that sent VNRN, PcSwitch sends VNRN DS_ACC with FLOGI LS_ACC to FCDF that sent VNRN
  - FCDF sends LS_ACC for FLOGI(s)

One cSwitch

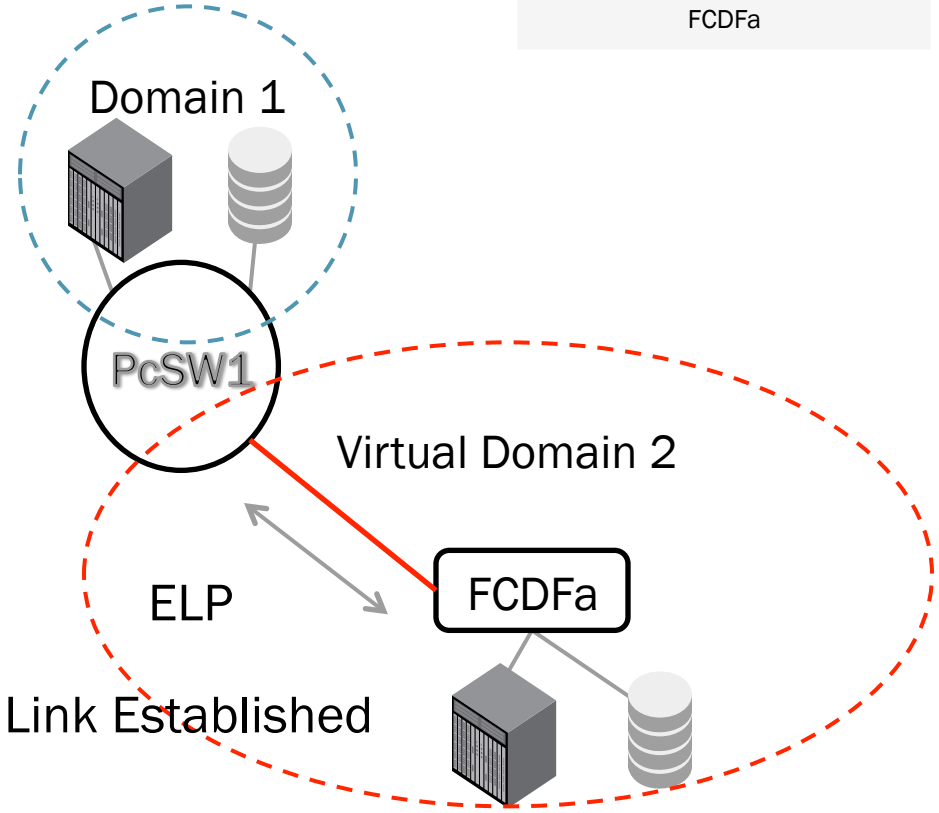
Initialization Sequence (Single FCDF)

Distributed Switch

A_Port link instantiation

Distributed Switch Membership Set	
Controlling Switch Set	
cSW1	Primary
FCDF Set	
FCDFa	

- ELP
  - Initiated by PcSW1
    - Based on Distributed Switch Membership Set (see text in 17.1 - Overview)
  - Responded by FCDFa
- A_Port Link Established
  - Upon PcSW1 receiving SW_ACC
- End Device Status at FCDFa
  - FLOGI LS_RJT "Logical Busy"

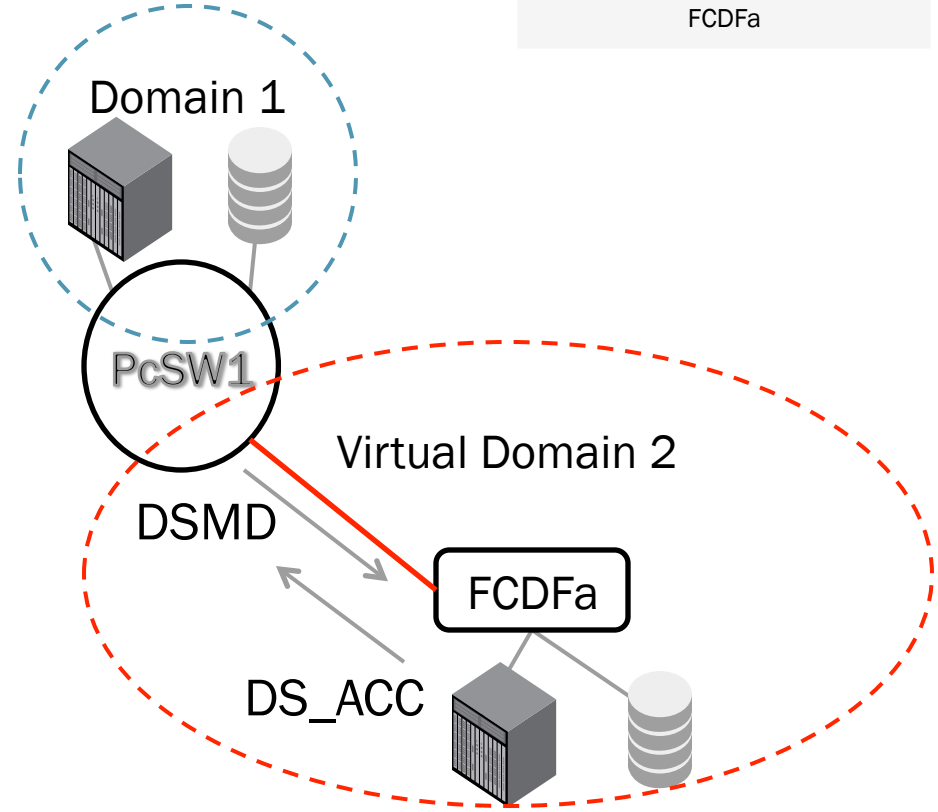


Distributed Switch

Distribute Switch Membership Distribution (DSMD)

Distributed Switch Membership Set	
Controlling Switch Set	
cSW1	Primary
FCDF Set	
FCDFa	

- DSMD request
  - PcSW1 sends Distributed Switch Membership Set to FCDFa
- DS_ACC
  - FCDFa sends DS_ACC to PcSW1
    - FCDF descriptor - Number of Ports & RNID Specific Node-Identification Data
  - FCDFa can report links to other cSWs
    - Responds with SW_ACC to a ELP from another cSW in the Distributed Switch Membership Set
    - Sends FDRN to PcSW1
      - Following SW_ACC to ELP
  - FCDFa can now initiate ELP with downstream FCDFs
    - Reports links to other FCDFs via FDRNs to PcSW1 following SW_ACC to ELP
- End Device Status
  - LS_RJT "Logical Busy"

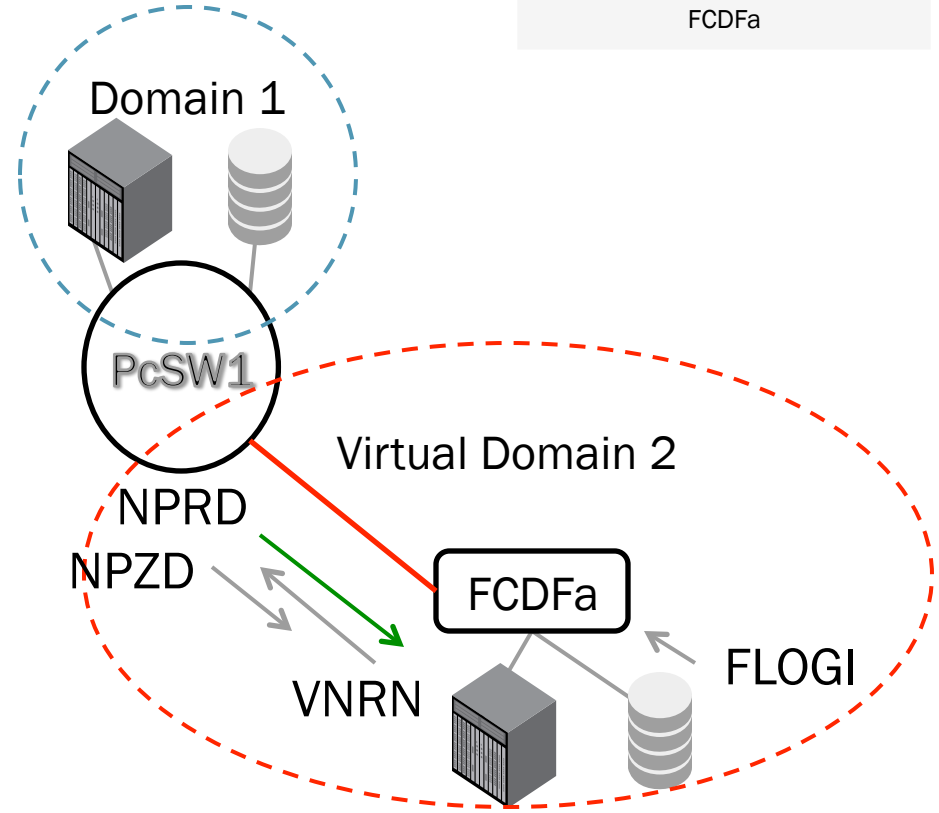


Distributed Switch

N_Port_ID Route Distribution (NPRD) & FLOGI

- PcSW1 sends NPRD request
  - Describes the Virtual Domain routes to FCDFa
- FCDFa can now accept FLOGIs
  - Sends VNRN to PcSW1 upon receiving FLOGI requests
    - PcSwitch allocates N_Port_ID, recomputes Zoning ACLs, generates RSCNs, updates FC Name Server
    - PcSwitch sends appropriate NPZD(s) to ScSwitch, if present, and each FCDF
    - After receiving NPZD DS_ACC from ScSwitch, if present, and FCDF that sent VNRN, PcSwitch sends VNRN DS_ACC with FLOGI LS_ACC to FCDF that sent VNRN
  - FCDF sends LS_ACC for FLOGI(s)
- End Device Status at FCDFa
  - Active following NPZD DS_ACC from FCDFa

Distributed Switch Membership Set	
Controlling Switch Set	
cSW1	Primary
FCDF Set	
FCDFa	



Two cSwitches

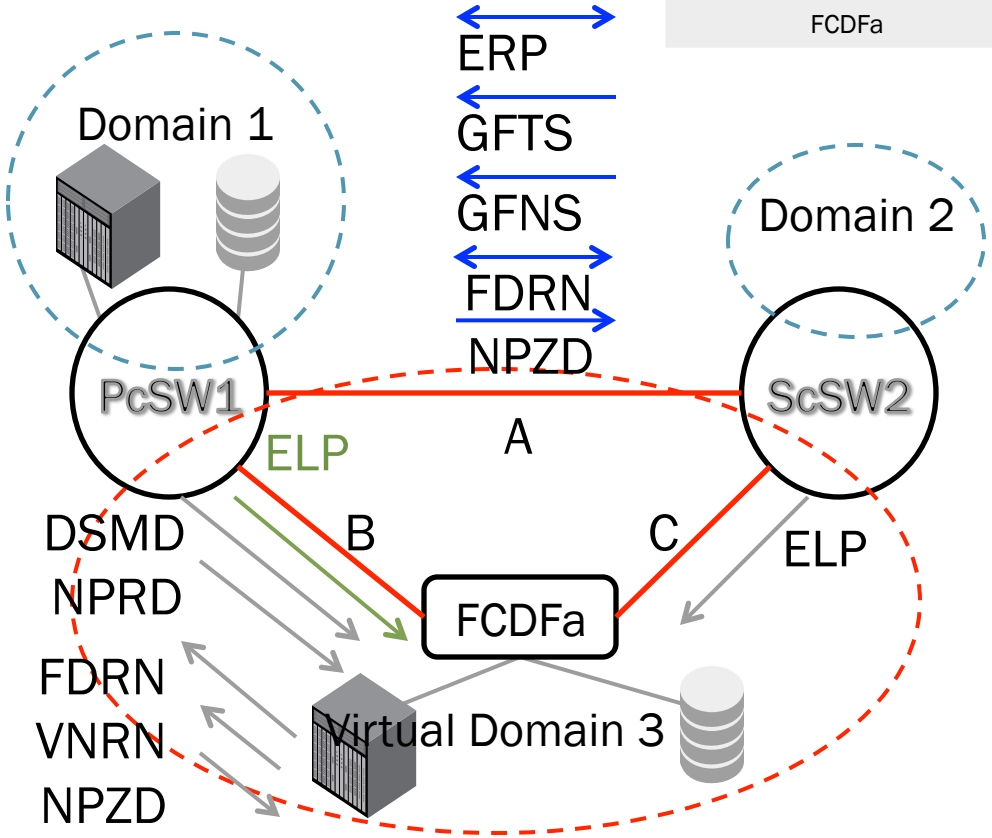
Initialization Sequence (Single FCDF)

Distributed Switch

2 Controlling Switch (current)

- Link A – PcSW1-ScSW2 Exchanges
  - PsSW1 and ScSW2 exchange ERP
  - ScSW1 sends GFTS/GFNS
  - PcSW1 notifies ScSW2 that FCDFa link B exists (FDRN)
  - ScSW2 notifies PcSW1 that FCDFa link C exists (FDRN)
  - PcSW1 sends NPZD per VNRN processing
- Link B – PcSW1-FCDFa Exchanges
  - PcSW1 sends ELP to FCDFa
    - FCDFa replies with SW_ACC
  - PcSW1 establishes membership (DSMD)
  - PcSW1 establishes routes (NPRD)
  - *FCDFa notifies PcSW1 that ScSW2 link C exists (FDRN)
  - FCDFa facilitates device login (VNRN, NPZD)
- Link C – ScSW2-FCDFa Exchanges
  - ScSW2 sends ELP to FCDFa
    - FCDFa replies with SW_ACC

Distributed Switch Membership Set	
Controlling Switch Set	
cSW1	Primary
cSW2	Secondary
FCDF Set	
FCDFa	

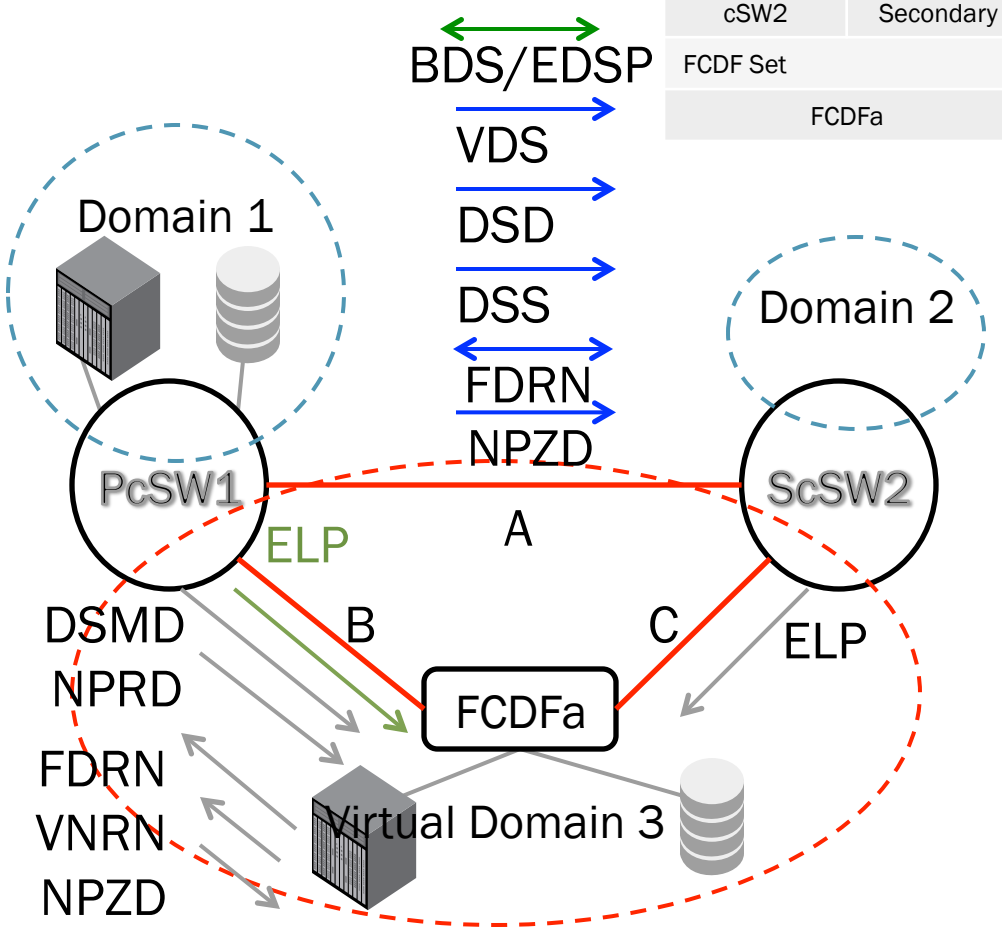


Distributed Switch

>2 Controlling Switch

- Link A – PcSW1-ScSW2 Exchanges
  - Any cSW initiates Primary cSW selection (BDS/ EDSP)
  - PcSw1 sends VDS, DSD, DSS to ScSW2
  - PcSW1 notifies ScSW2 that FCDFa link B exists (FDRN)
  - ScSW2 notifies PcSW1 that FCDFa link C exists (FDRN)
  - PcSW1 sends NPZD per VNRN processing
- Link B – PcSW1-FCDFa Exchanges
  - PcSW1 sends ELP to FCDFa
    - FCDFa replies with SW_ACC
  - PcSW1 establishes membership (DSMD)
  - PcSW1 establishes routes (NPRD)
  - *FCDFa notifies PcSW1 that ScSW2 link C exists (FDRN)
  - FCDFa facilitates device login (VNRN, NPZD)
- Link C – ScSW2-FCDFa Exchanges
  - ScSW2 sends ELP to FCDFa
    - FCDFa replies with SW_ACC

Distributed Switch Membership Set	
Controlling Switch Set	
cSW1	Primary
cSW2	Secondary
FCDF Set	
FCDFa	

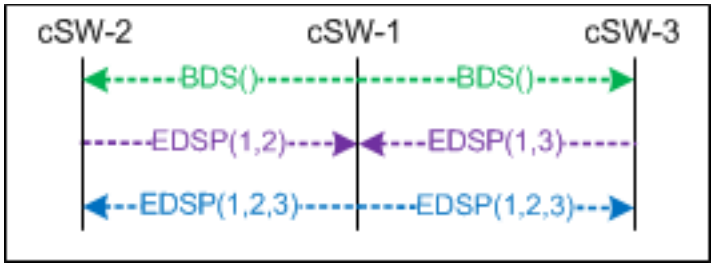


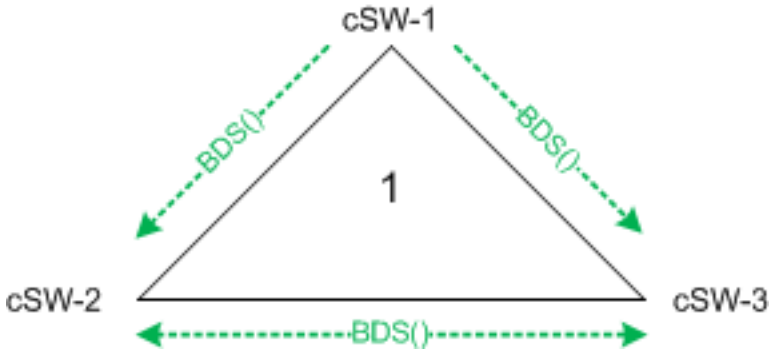
Three cSwitches

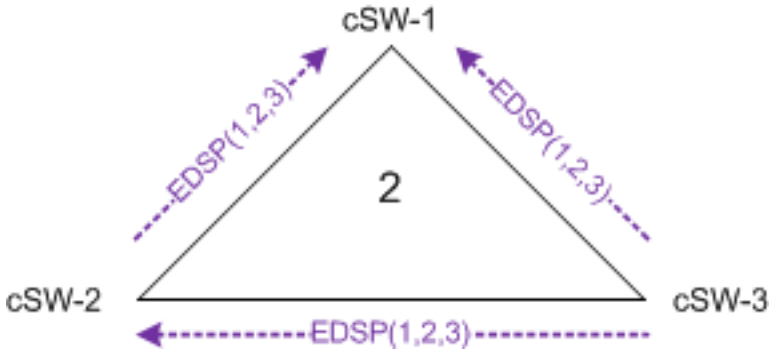
Initialization Sequence (Single FCDF)

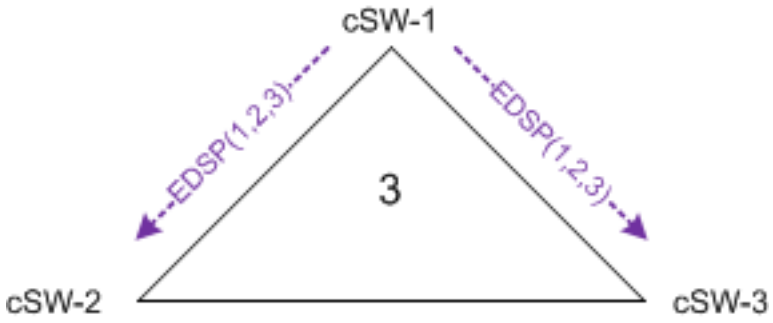
Primary cSwitch Selection

Run Primary cSwitch selection process (using well known methods)





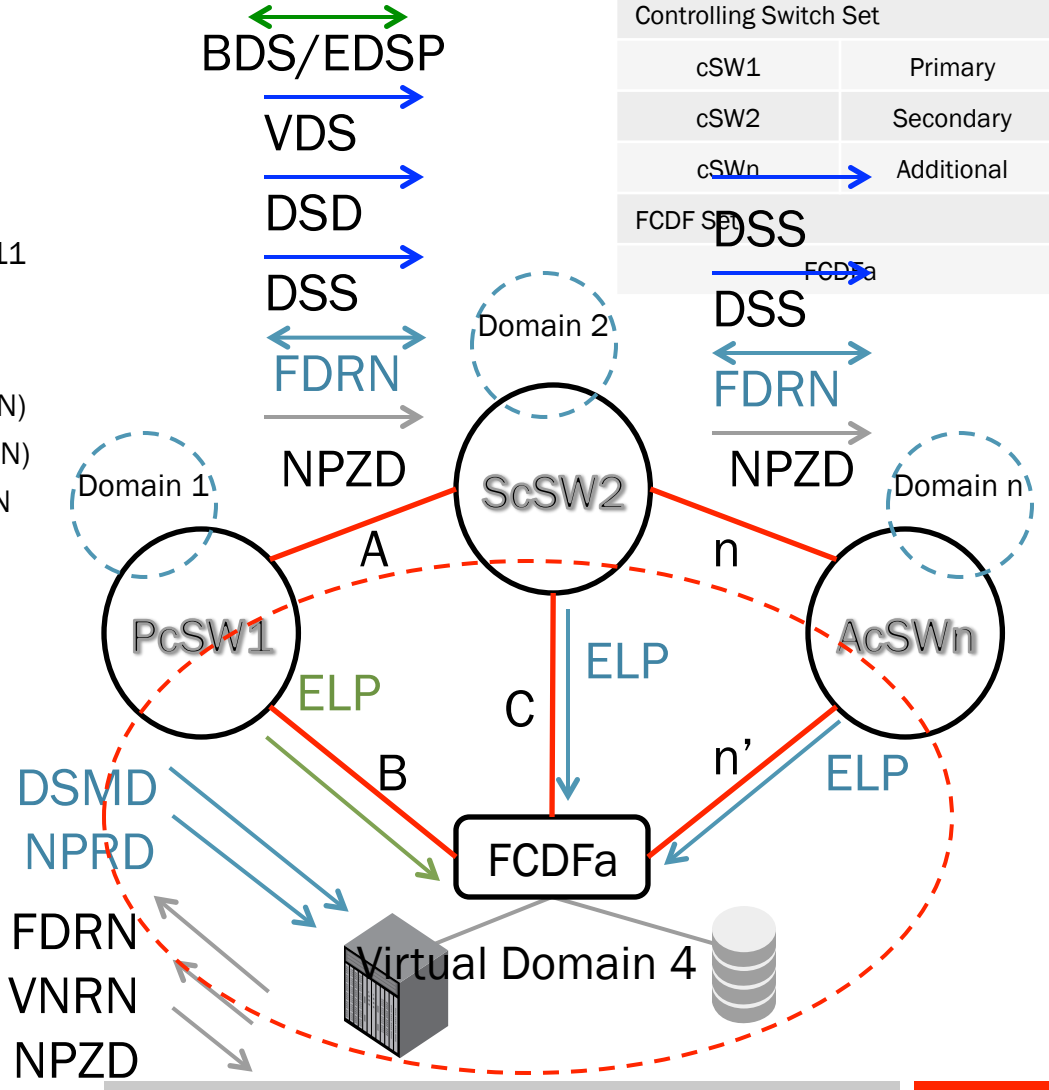




Distributed Switch

>2 Controlling Switch

- Link A and Link n – PcSW1-ScSW2-AcSWn Exchanges
  - Any cSW initiates Primary cSW selection – see slide 11 for link A Exchanges
  - PcSW1 notifies ScSW2 and AcSWn that link B exists (FDRN)
  - ScSW2 notifies PcSW1 that FCDFa link C exists (FDRN)
  - AcSWn notifies PcSW1 that FCDFa link n' exists (FDRN)
  - PcSW1 sends NPZDs to ScSW2 and AcSWn per VNRN processing
- Link B – PcSW1-FCDFa Exchanges
  - PcSW1 sends ELP to FCDFa
    - FCDFa replies with SW_ACC
  - PcFCF1 establishes membership (DSMD)
  - PcFCF1 establishes routes (NPRD)
  - FCDFa notifies PcSW1 that link C and link n' exist (FDRNs)
  - FCDFa facilitates device login (VNRN, NPZD)
- Link C – ScSW2-FCDFa Exchanges
  - ScSW2 sends ELP to FCDFa
    - FCDFa replies with SW_ACC
- Link n' – AcSWn-FCDFa Exchanges
  - AcSWn sends ELP to FCDFa
    - FCDFa replies with SW_ACC



Failure Cases

Primary cSwitch, Secondary cSwitch, Additional cSwitch

Primary cSwitch

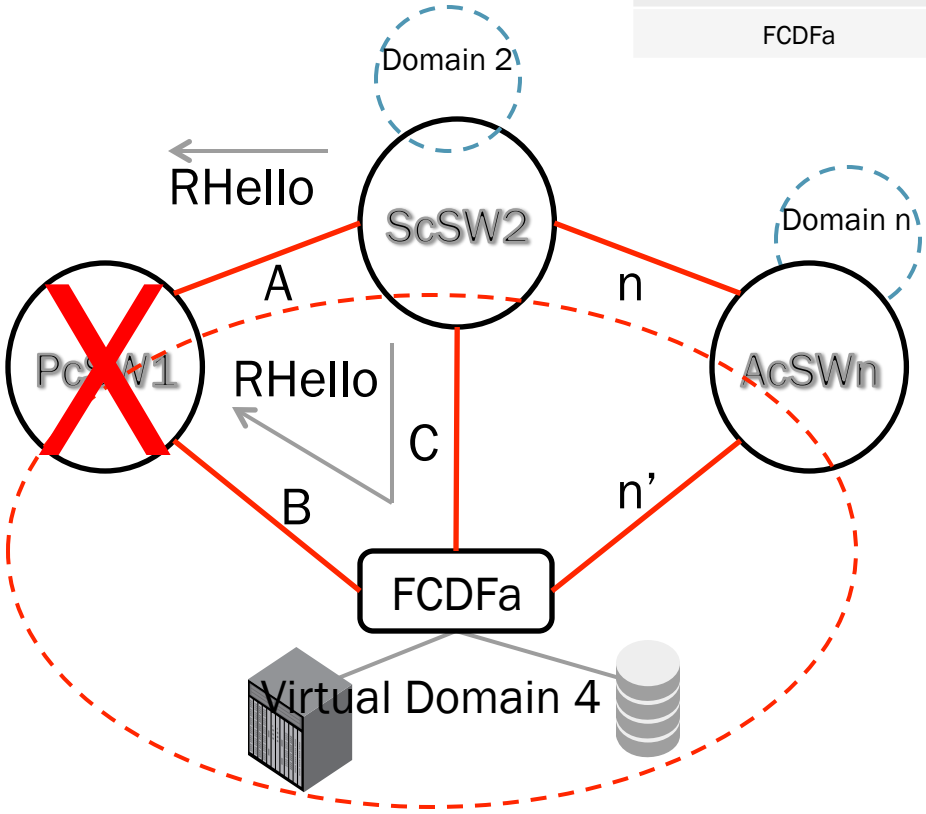
Failure

Distributed Switch Membership Set	
Controlling Switch Set	
cSW1	Primary
cSW2	Secondary
cSWn	Additional
FCDF Set	
FCDFa	

Distributed Switch

>2 Controlling Switch – Non-mesh – PcSW1 Fails

- ScSW2 detects PcSW1 failure via Rhello(s)/physical layer and transitions to Primary State/Role
- Is Rhello also sent via link n/ n'?

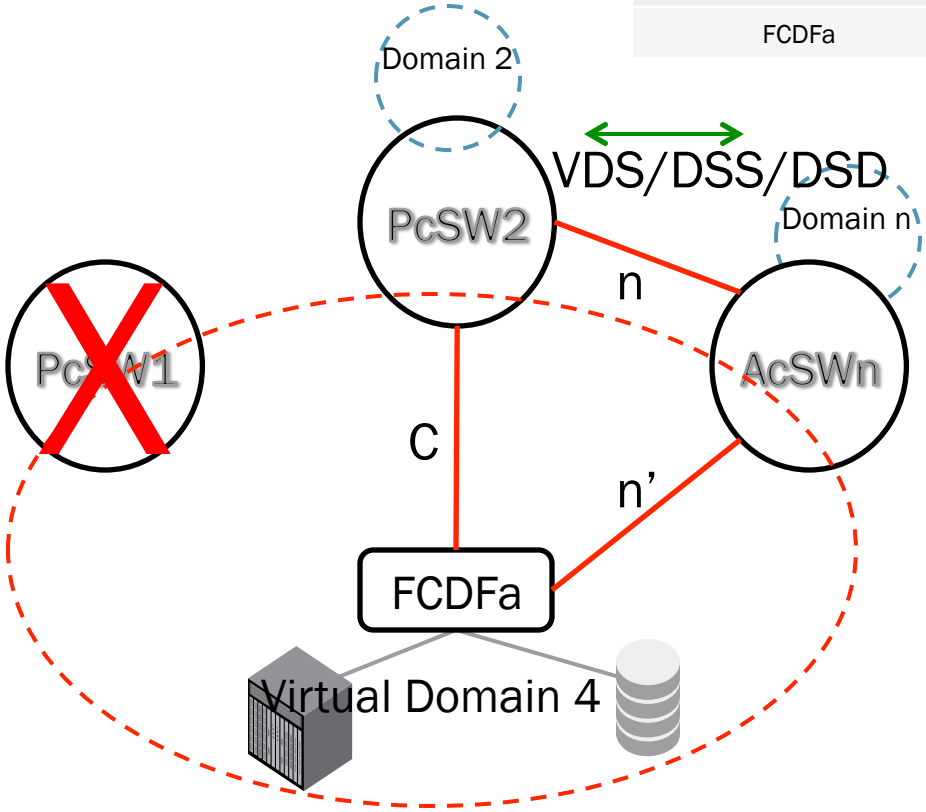


Distributed Switch

>2 Controlling Switch – Non-mesh – PcSW1 Fails

- Link n – PcSW1-AcSWn Exchanges
  - PcSW2 sends VDS, DSS, and DSD to AcSWn and AcSWn becomes Secondary

Distributed Switch Membership Set	
Controlling Switch Set	
cSW1	Primary
cSW2	Secondary
cSWn	Additional
FCDF Set	
FCDFa	

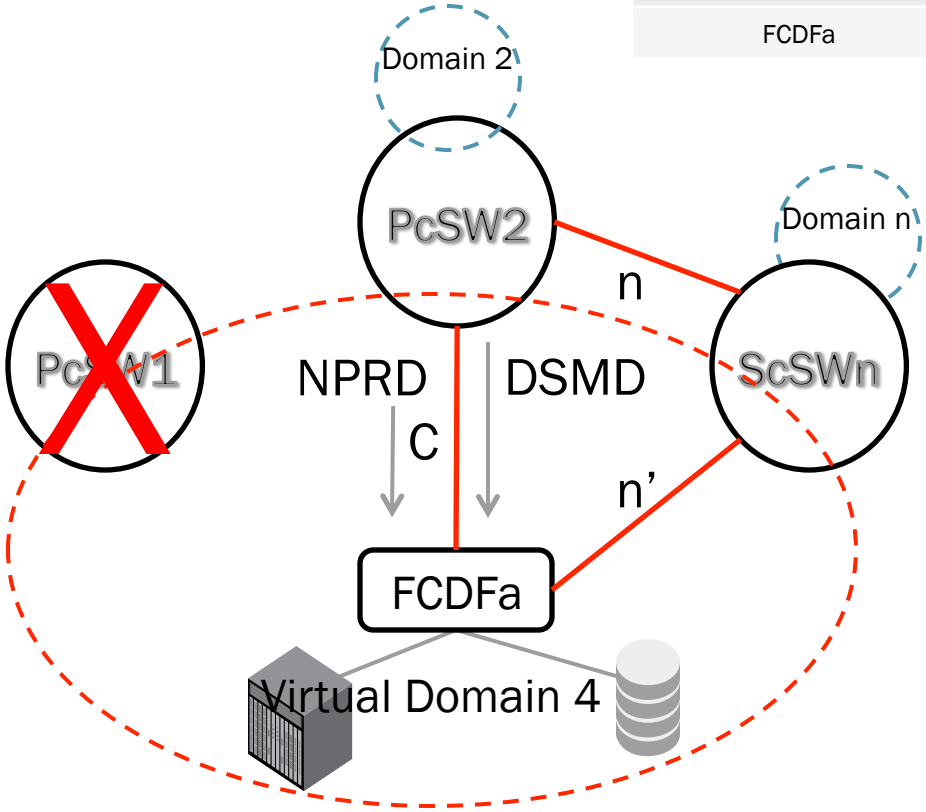


Distributed Switch Membership Set	
Controlling Switch Set	
cSW1	Primary
cSW2	Secondary
cSWn	Additional
FCDF Set	
FCDFa	

Distributed Switch

>2 Controlling Switch – Non-mesh – PcSW1 Fails

- Link n – PcSW2-ScSWn Exchanges
  - PcSW2 sends NPZDs to ScSWn per VNRN processing
- Link C – PcSW2-FCDFa Exchanges
  - PcSW2 establishes membership (DSMD)
  - PcSW2 establishes routes (NPRD)
  - FCDFa facilitates device login (VNRN, NPZD)
- Link n' – ScSWn-FCDFa Exchanges
  - None



Secondary cSwitch

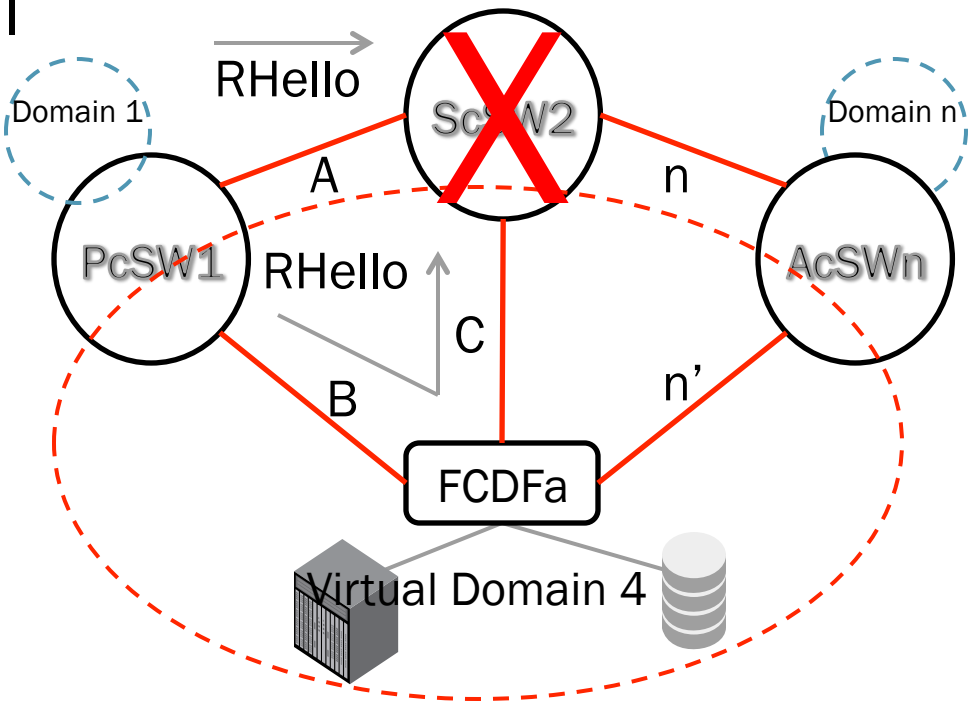
Failure

Distributed Switch Membership Set	
Controlling Switch Set	
cSW1	Primary
cSW2	Secondary
cSWn	Additional
FCDF Set	
FCDFa	

Distributed Switch

>2 Controlling Switch – Non-mesh – ScSW2 Fails

- PcSW1 detects ScSW2 failure via Rhello(s)/physical layer

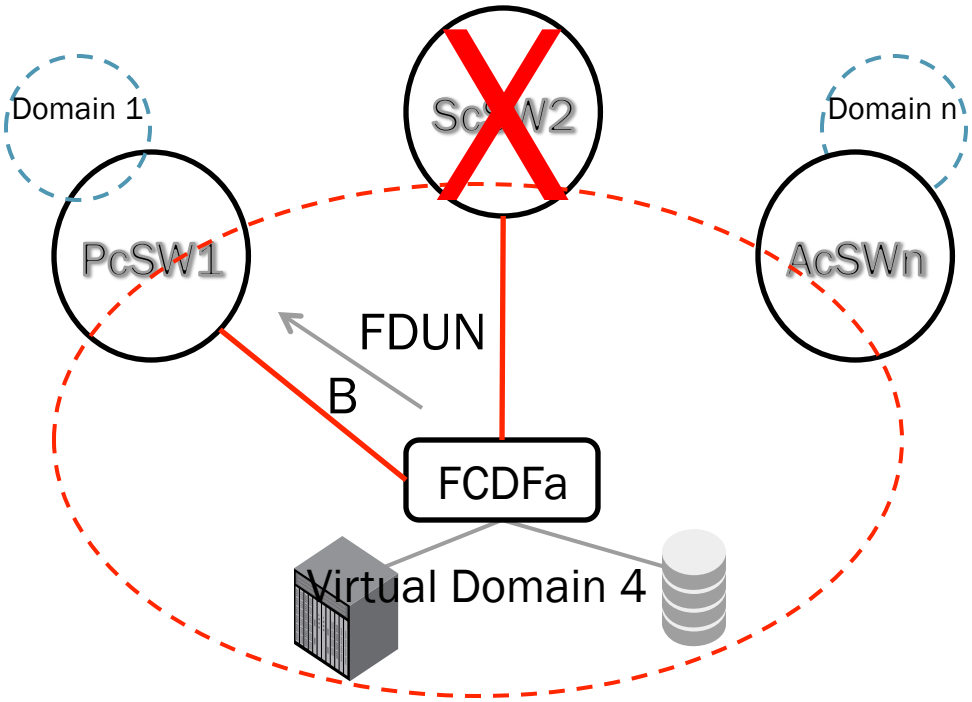


Distributed Switch Membership Set	
Controlling Switch Set	
cSW1	Primary
cSW2	Secondary
cSWn	Additional
FCDF Set	
FCDFa	

Distributed Switch

>2 Controlling Switch – Non-mesh – ScSW2 Fails

- PcSW1 remains primary
- AcSWn segments from Distributed Switch and link n' is de-instantiated
- Link B - PcSW2-FCDFa Exchanges
  - FCDFa sends FDUN to PcSW1

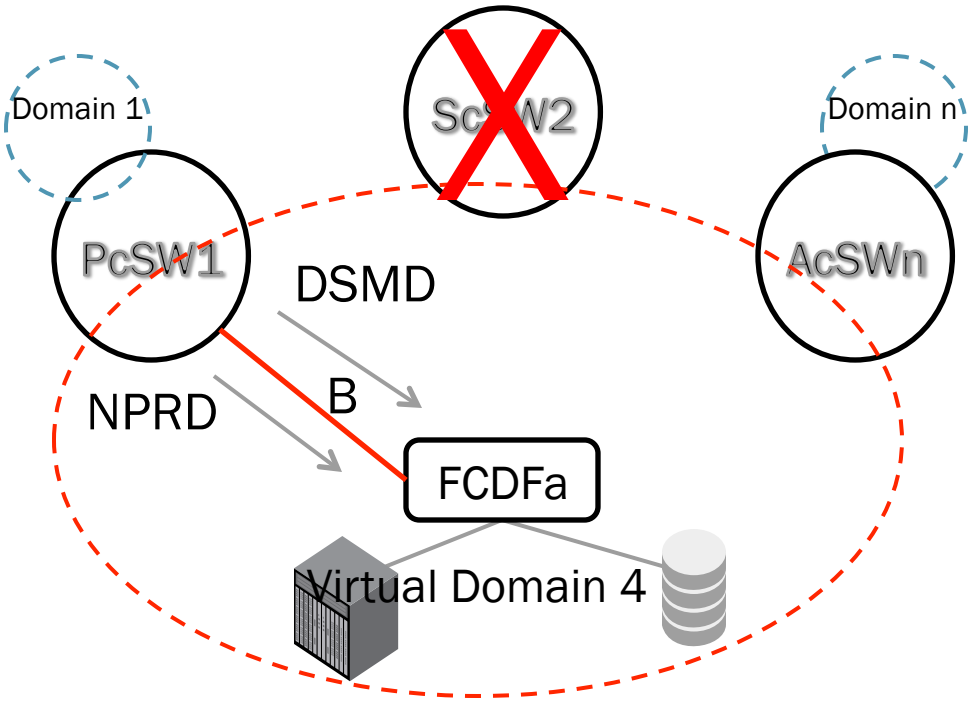


Distributed Switch Membership Set	
Controlling Switch Set	
cSW1	Primary
cSW2	Secondary
cSWn	Additional
FCDF Set	
FCDFa	

Distributed Switch

>2 Controlling Switch – Non-mesh – ScSW2 Fails

- Link B - PcSW2-FCDFa Exchanges
  - PcSW1 sends DSMD and NPRD to FCDFa



Additional cSwitch

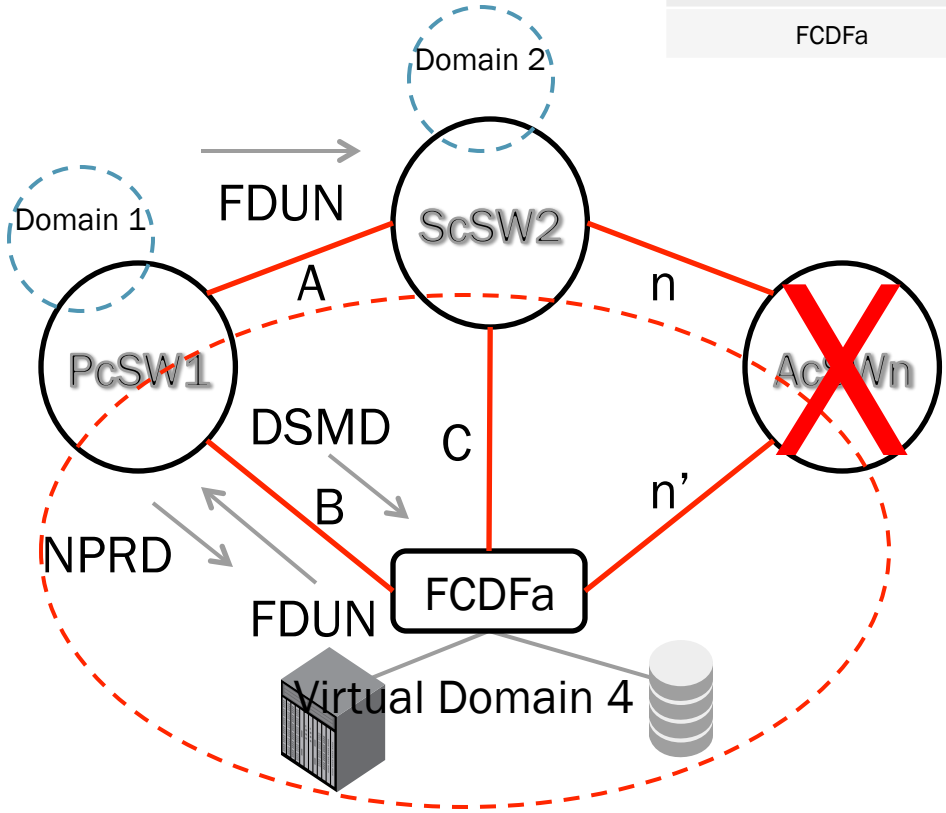
Failure

Distributed Switch Membership Set	
Controlling Switch Set	
cSW1	Primary
cSW2	Secondary
cSWn	Additional
FCDF Set	
FCDFa	

Distributed Switch

>2 Controlling Switch – Non-mesh - AcSWn Fails

- Link A – PcSW1-ScSW2 Exchanges
  - ScSW2 and PcSW1 know link n does not exist via FSPF
  - PcSW1 notifies ScSW2 that FCDFa link n' does not exist (FDUN)
- Link B – PcSW1-FCDFa Exchanges
  - FCDFa notifies PcSW2 that link n' does not exist (FDUN)
  - PcSW1 sends DSMD to FCDFa
  - PcSW1 sends NPRD to FCDFa
- Link C – ScSW2-FCDFa Exchanges
  - None



BROCADE

This slide intentionally left blank

Thank You

Reference

Questions

- Are Additional Controlling Switch(es) part of the Controlling Switch Set?
  - Some folk mentioned an AcSwitch functions as an FCDF
  - In my mind an AcSwitch functions as a Controlling Switch not operating in the Primary or Secondary state
    - As such an AcSwitch needs the same information as a ScSwitch to route received frames received from the Virtual Domain

Membership Set descriptor >2 cSwitch

17.8.3.12 Membership Set descriptor

The format of the Membership Set descriptor is shown in table 228

Table 228 – Membership Set descriptor format

Item	Size (Bytes)
Tag Value = 000Bh	4
Length = variable	4
Fabric_Name	8
Virtual Domain Switch_Name	8
Primary Controlling Switch Switch_Name	8
Secondary Controlling Switch Switch_Name	8
Number of Additional Controlling Switches (m)	4
Additional Controlling Switch Switch_Name #1	8
Additional Controlling Switch Switch_Name #2	8
...	
Additional Controlling Switch Switch_Name #m	8
Number of FCDFs (n)	4
FCDF Switch_Name #1	8
FCDF Switch_Name #2	8
...	
FCDF Switch_Name #n	8