

Multiple FCF Redundancy Proposal – Issues & Questions

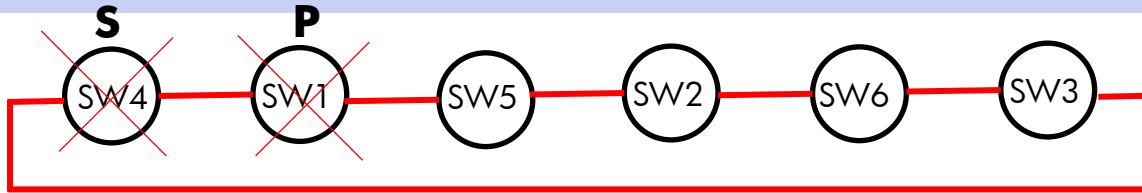
Siamack Ayandeh
Chief Architect (Datacenter)
HP Networking
December 14th, 2012
Siamack@HP.Com



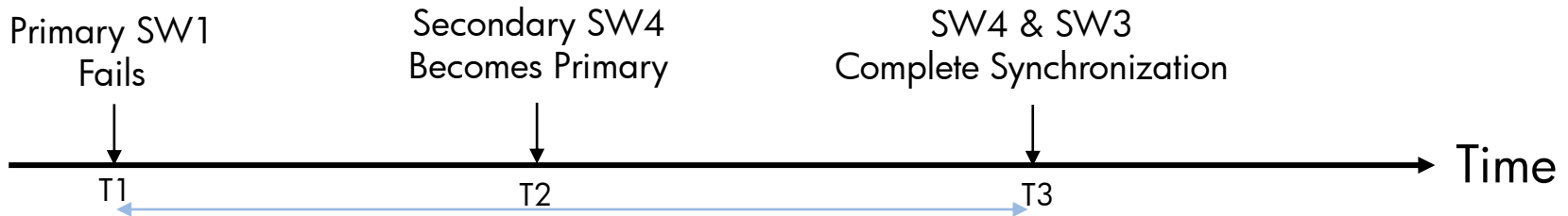
Purpose of HA

- Purpose of a distributed switch HA is to minimize disruption to active connections when a primary c/FCF fails

Does HA improve with more than two c/FCF(s) in 12-312v2?



- Despite having more than two FCF(s) in its HA model, 12-312v2 still is exposed to double failures
 - And that's due to the way HA protocol is defined, Once SW1 fails, the secondary becomes the primary, &
 - Finds a new FCF (say SW3) to synchronize its state to become the secondary FCF

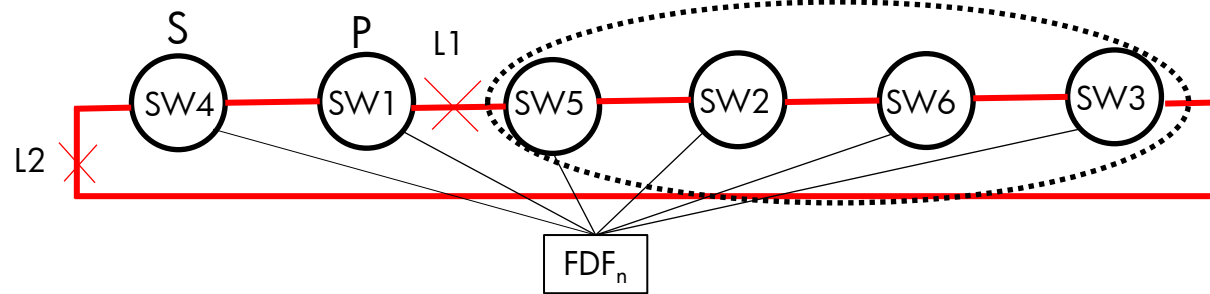


Window of time system is exposed to double failures

- If primary and secondary fail within the time it takes to synchronize SW3 as a secondary then the distributed switch loses all its connections
 - The distributed switch is therefore still exposed to double failures

The above system's cFCF availability is slightly better than dual redundancy

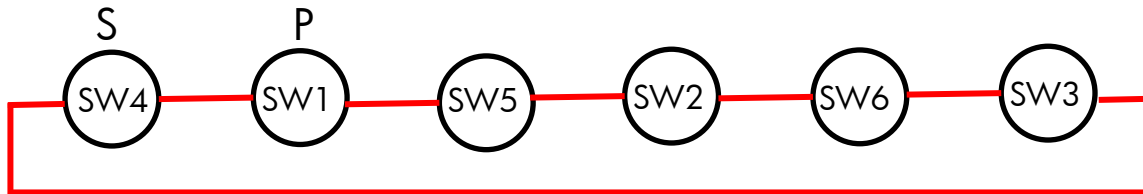
Also need redundant links to improve over dual redundancy



- Double failure of AISL(s) leads to isolation of SW5/2/6/3
- Therefore AISL redundancy has to keep pace with the switch redundancy to keep additional c/FCF(s) useful
 - A fully meshed AISL configuration of course adds the corresponding port costs
- Even if SW5/2/6/3 form a new virtual switch; FDF_n can not be part of it ! [see 12-434v3]
- *For a fully meshed system, cFCF availability is slightly better than dual redundancy & less than triple redundancy irrespective of SW5/2/6*

Beyond first two, HA is not improving much with increasing number of FCF(s)

What is SW6 HA objective?



- Range of HA Objectives:
 - Single device HA = 2x9(s)
 - Dual redundancy HA = 5x9(s) ← *Industry Norm*
 - Triple redundancy HA = 15x9(s)
- Despite having more than two FCF(s), 12-312v2 availability is slightly better than dual redundancy
 - And both schemes carry a number of single points of failure within a distributed switch
- *WG may want to design towards an HA objective* as opposed to number of c/FCF(s) per virtual domain
- 12-035vx offers quadruple c/FCF redundancy or better:
 - Implementations of FCF(s) in datacenters will likely be redundant to begin with (x2),
 - HA (12-035vx) adds dual redundancy (x2) x2,
 - And there is the traditional air gap fabric redundancy ((x2) x2) x2 for x8 times redundancy

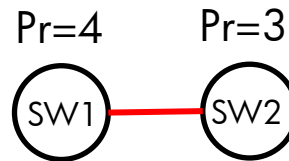
Additional Protocol Questions


- Questions re 12-312v2 & 12-070v4

Primary selection

Section 1.2.3

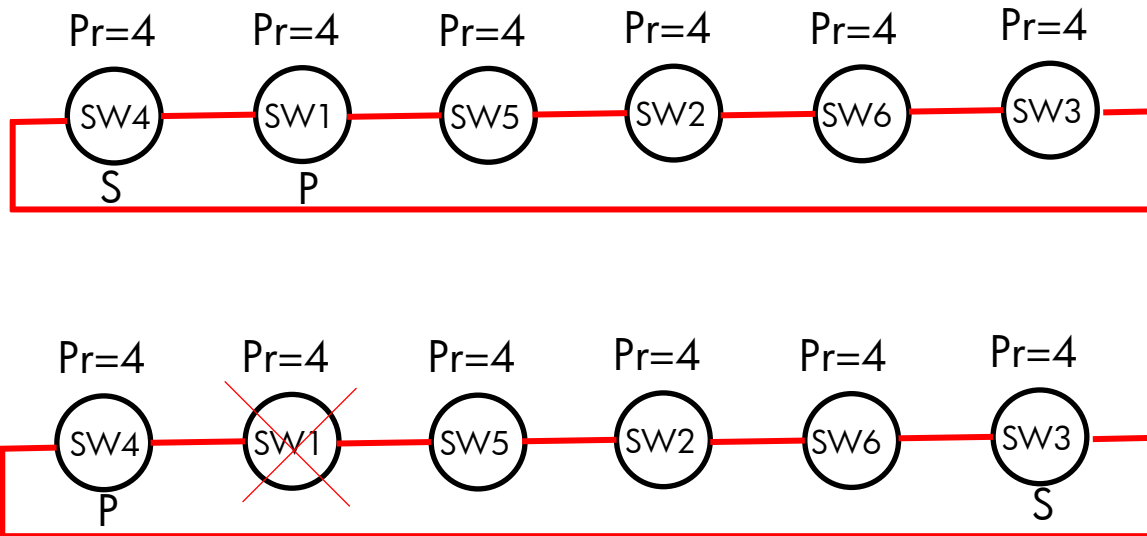
- a) Current_Primary: set to the lowest known `cSwitch_Priority`||Primary Controlling Switch
`Switch_Name`; and



Which switch will be the current primary? I take it is SW2 
Need a better description of the rule

How can an operator make sense of P/S selection




The primary chooses the secondary from the list of direct attached switches with the lowest known **Controlling Switch Switch_Name**



How can an operator looking at this system make sense of who is primary and who is secondary and why?

SW4 becomes the primary (and SW3 secondary) in violation of the selection rule in section 1.2.3

Build Distributed Switch (BDS) -- questions

- What triggers BDS?
- “The Primary Controlling Switch selection process is initiated by a Controlling Switch transmitting a Build Distributed Switch (BDS) DS_ILS (see 1.3.2.1) *on all Controlling Switch AISLs.*”
 - What’s meant by all AISL(s)? 
 - Is a switch responsible for re transmitting a BDS? 
- Is there some logic or a state machine for processing BDS? 

Thank You

