

Textual revisions to the High Availability Protocol

Siamack Ayandeh
Chief Architect (Datacenter)
HP Networking
December 10th, 2012
Siamack@HP.Com



Outline

12-434v3 identified the multiple failure scenarios and agreed resolutions for the HA protocol revisions

Following is the outline of the resulting *edits* to 12-035v3

1. Define the condition “Primary Controlling Switch is not anymore available”
2. Add clarifying text to state R2 (election in case of NULL ASIL set)
3. Add a “Distributed Switch Switch-name” to table-41 in 12-036v2
4. Revise state P2 for the case of N_Port ID conflict
5. Delete transition P2:S1 as it no longer exists (was labeled P2:S2 in error in figure 3)
6. Revise state S1 to wait for synchronization or in case primary is not anymore available; instantiate a new domain
7. Case of partitioned virtual domain is already resolved with above revisions

All new text is in RED

Condition “Primary FCF is not anymore available”: is defined as:

Down Timer expires as primary is not responsive, yet links are up

- ASIL set is not NULL
- And the ASL links in the virtual domain are up

Or if the primary fails, and the links attached to the primary go down

- The event is reported within the Down Timer expiry interval through FSPF i.e. secondary can tell that primary is isolated or down
- And the Down timer expires

In both scenarios the secondary can take over and become the primary

All new text is in RED

- **“State R2:Election. In this state a Controlling Switch determines if it operates as Primary or Secondary.**
- If the AISL Set is NULL, then the Controlling Switch exits this state **without election and moves to state P1. In state P1 each controlling FCF will use a unique virtual domain switch name to obtain a domain-ID using an RDI. Two virtual domains may therefore form. The FDF set will be partitioned between these two virtual domains based on connectivity or timing. Once the AISL set is not NULL, the two virtual domains may merge using ERP.**
- If the AISL Set is not NULL, then an ERP Exchange is performed...”

All new text is in RED



Section 1.2.2.2 of 12-036v2 edits

- Since ERP is used to potentially merge two virtual domains each with its own virtual domain switch-name; the following needs to be added to section 1.2.2.2 of 12-036v2.

Item	Size (byte)
Tag Value = 0011h	4
Length = variable	4
Distributed Switch Switch-name	8
Originating Controlling Switch Priority	4
Number of Allocated N_Port_ID Ranges (q)	4
Allocated N_Port_ID Range #1	4
Allocated N_Port_ID Range #2	4
Allocated N_Port_ID Range #q	4

All new text is in RED

- **Transition P2:P2: Occurs following the ERP Exchange performed when the AISL Set went from NULL to not-NULL if:**
 - a) there is no allocated N_Port_IDs conflict between the two Controlling Switches and this Switch is the one selected to remain Primary; or
 - b) there is an allocated N_Port_IDs conflict between the two Controlling Switches **and the controlling FCF(s) are the only two switches in a fabric**. In this case the AISL shall be isolated. **Each controlling FCF would continue as primary for its (portion) of the virtual domain.**

- **Transition P2:S2: Occurs following the ERP Exchange performed when the AISL Set went from NULL to not-NULL if there is no allocated N_Port_IDs conflict between the two Controlling Switches and this Switch is the one selected to become Secondary.**

P2:S1

- Delete this transition from figure 3.

- **Entry to this state is from R2:S1 or S2:S1 (when AISL set is NULL)**
- **State S1:Secondary Initialization. In this state a Controlling Switch waits for at least an AISL to be available.** When an AISL is available it performs the operations to become the Secondary Controlling Switch of the Distributed Switch. The Controlling Switch has to synchronize its state with the ~~one of the~~ Primary Controlling Switch. To this end the Controlling Switch:
 - 1) Requests to the Primary the FCDF topology through the GTFS (Get FCDF Topology State) SW_ILS;
 - 2) Requests to the Primary the Virtual Domain_IDs and N_Port_IDs Allocation state in the Distributed Switch through the GFNS (Get FDCF N_Port_IDs State) SW_ILS;
 - 3) Obtains the information associated with each N_Port_ID in the Name Server through the GE_ID CT Request; and
 - 4) Communicates the achieved state synchronization to the Primary through the SSA (Secondary Synchronization Achieved) SW_ILS.
- While in this state the Controlling Switch:
 - a) processes FDUN, FDRN, and NPZD Requests coming over the AISL from the Primary Controlling Switch, if at least an AISL is available;
 - b) sends RHello Requests every RHello_Interval over each of its AISLs and over each ASL through which the Primary is reachable; and
 - c) resets the Down_Timer to Down_Interval everytime an RHello Request is received over at least one AISL or ASL.

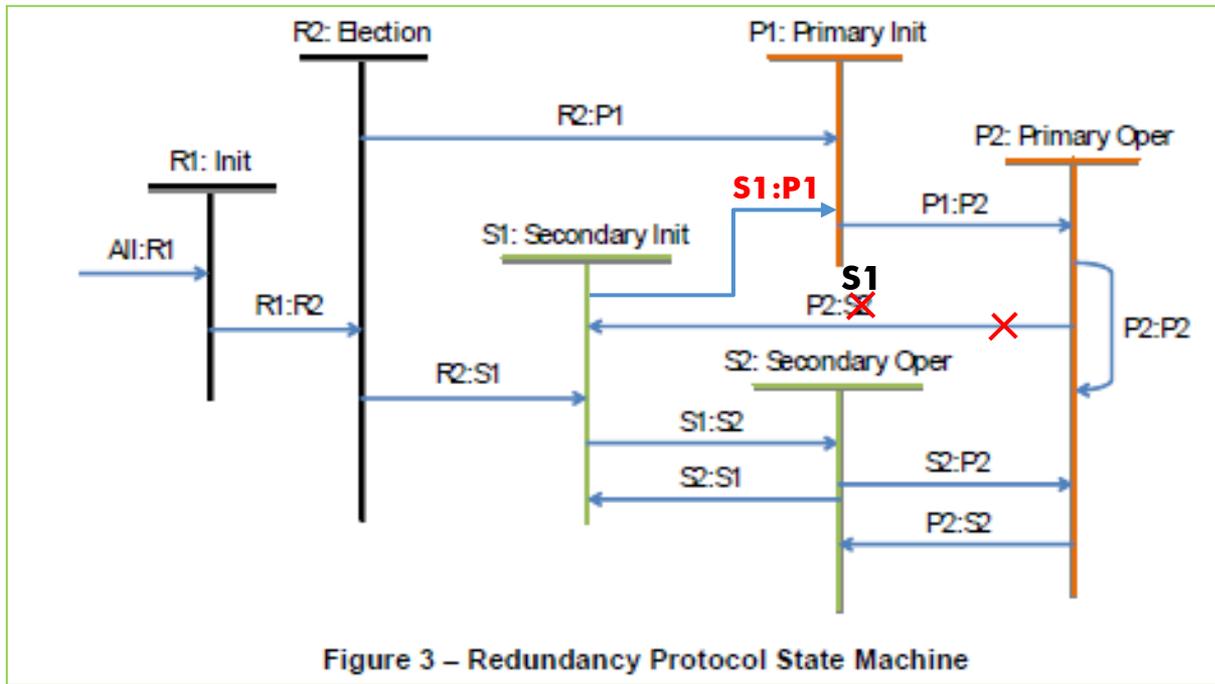
If the primary is not any more available the (secondary) c/FCF will transition to state P1

1.3.2.1 Partitioned virtual domain– what should the secondary FCF do?

Given that the secondary can detect a partitioned domain through:

- AISL Set is NULL (moves to state S1)
- No ASL(s) to the primary
- And that the Down_timer expires (moves to state P1)
- Based on the principle of operation that a secondary should not operate as a secondary without AISL connectivity, to the primary
- Secondary will suspend being a c/FCF for the virtual domain in question and allows the connections through the associated FCF(s) to time out
- And it moves to state P1 to instantiate a new virtual domain

Page 8: Figure 3: Revised Redundancy Protocol State Machine



Thank You

