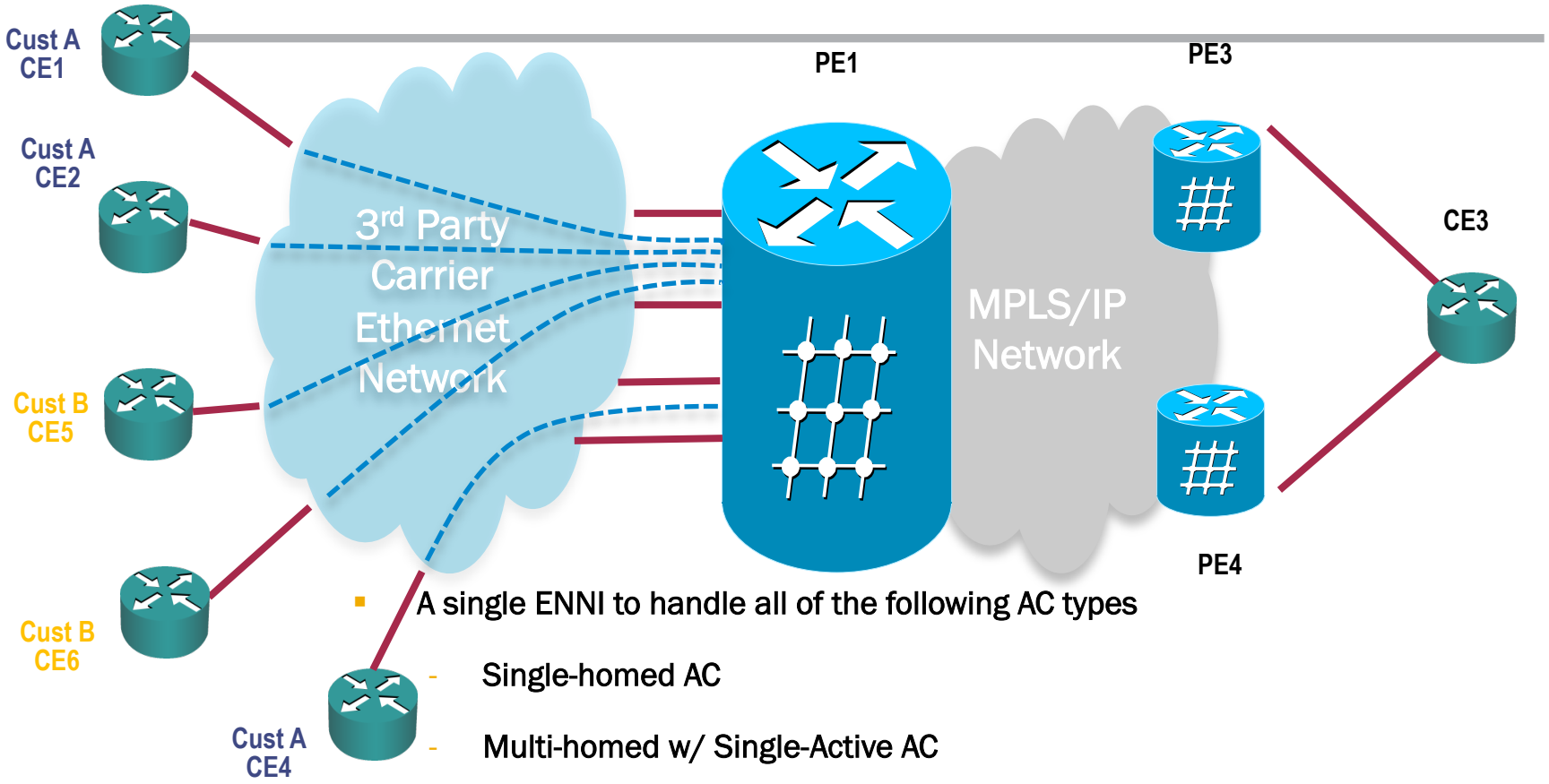
draft-sajassi-bess-evpn-virtual-eth-segment-00.txt

A. Sajassi (Cisco), P. Brissette (Cisco), R. Schell (Verizon), J. Drake (Juniper), T. Singh (Juniper)

IETF 91, November 2014

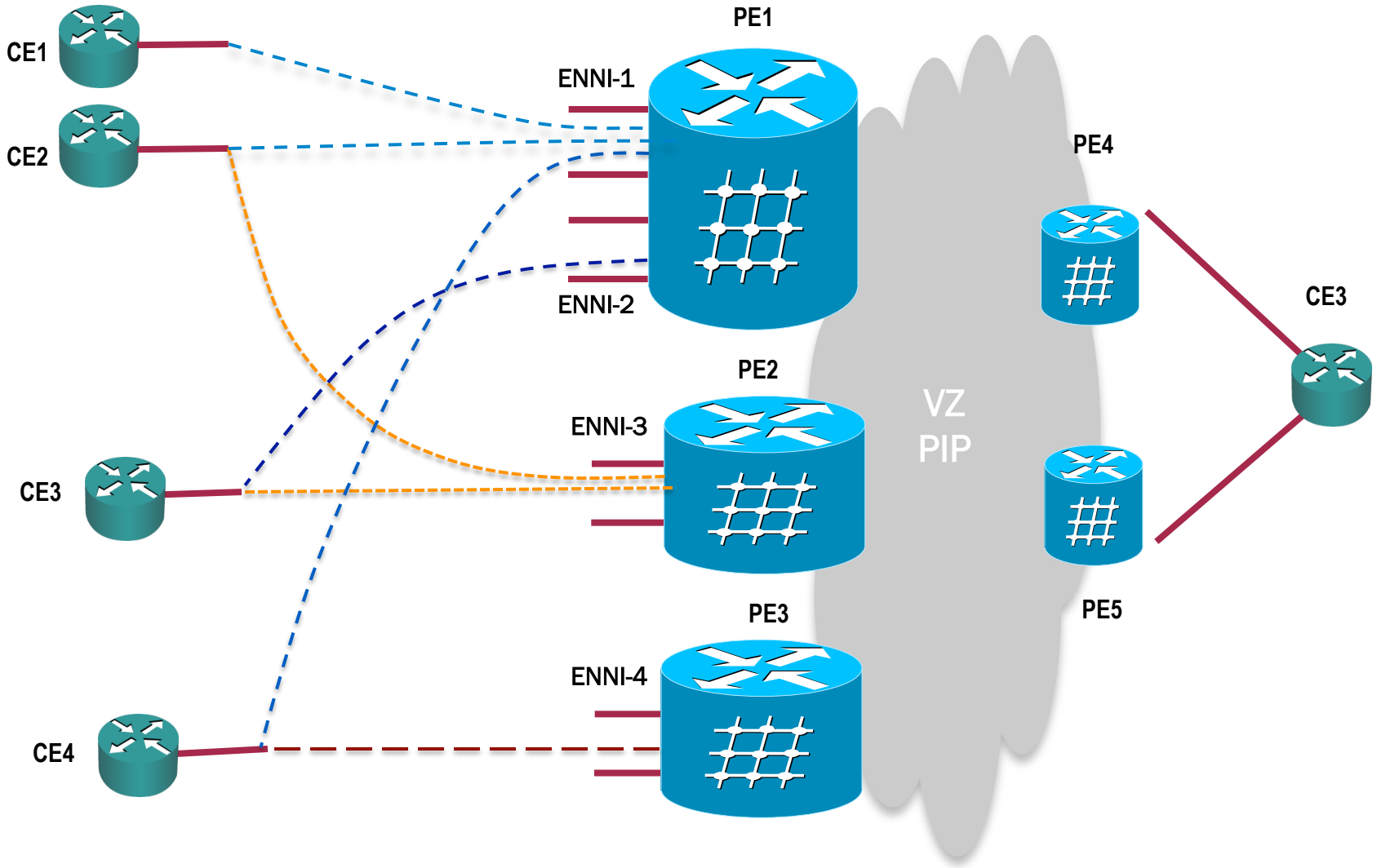
Hawaii

General Requirements



- A single ENNI to handle all of the following AC types
  - Single-homed AC
  - Multi-homed w/ Single-Active AC
  - Multi-homed w/ All-Active AC (optional)
- A single ENNI can carry many ACs of the same type – e.g., several thousands single-homed ACs, + several hundreds of multi-homed ACs
- Switching must be supported among active ACs belonging to the same service instance on the same ENNI

MHD/MHN (both SA/AA) and SH on same ENNI



Characteristics

- Different ACs are represented by different VLANs.
  - So that they can take different paths within .1Q network for redundancy
- Relationship between ACs and physical ports (ENNIs) is arbitrary
  - Cannot assume all VLANs are available on a set of ENNI ports (as is the case with current ES definition)

General Requirements – Cont.

- An AC can carry one or more VLANs
  - An AC carries typically a single VLAN (VLAN-based or VLAN bundle)
  - In general case, an AC can carry several VLANs (VLAN-aware bundle)
- A redundancy group consists of two or more ACs on different PEs (w/ one AC per PE) – e.g., an ESI corresponds to this AC set
  - Different redundancy groups can be spread across different ENNIs – e.g., ENNI ports may not be common for all redundancy groups

General Requirements – Cont.

- For PBB-EVPN, I-SID mapping is still done as before
  - A single VLAN can map to an I-SID
  - Several VLANs can map to an I-SID
- For PBB-EVPN, DF election is still based on <ESI, I-SID> as before
  - For I-SID-a, ESI-x is active on PE1 but ESI-x is blocked on PE2

Provisioning Requirements

- For A-A & S-A MH scenarios, ES in here corresponds to a set of ACs (as opposed to physical ports) – e.g., two or more ACs is associated with an ESI.
- For S-A MH, there is no additional requirements for BMAC allocations beyond what is in the PBB-EVPN draft
- For A-A MHD/MHN, there will be one B-MAC per ESI as before (optional)
- Minimize introduction of new parameters (and thus the overhead of their provisioning) for Virtual ES as much as possible

Fast Convergence Requirement

- There should be a mechanism equivalent to mass-withdraw such that upon an ENNI failure, only a single message is need to the remote PEs in the redundancy group to activate fail-over procedures as opposed to tens of thousands of BGP messages

OAM Requirements

- Need to do monitoring on per AC basis and the ability to switch from AC-a on PE1 to AC-b on PE2 upon AC-a failure

Failure/Recovery Requirements

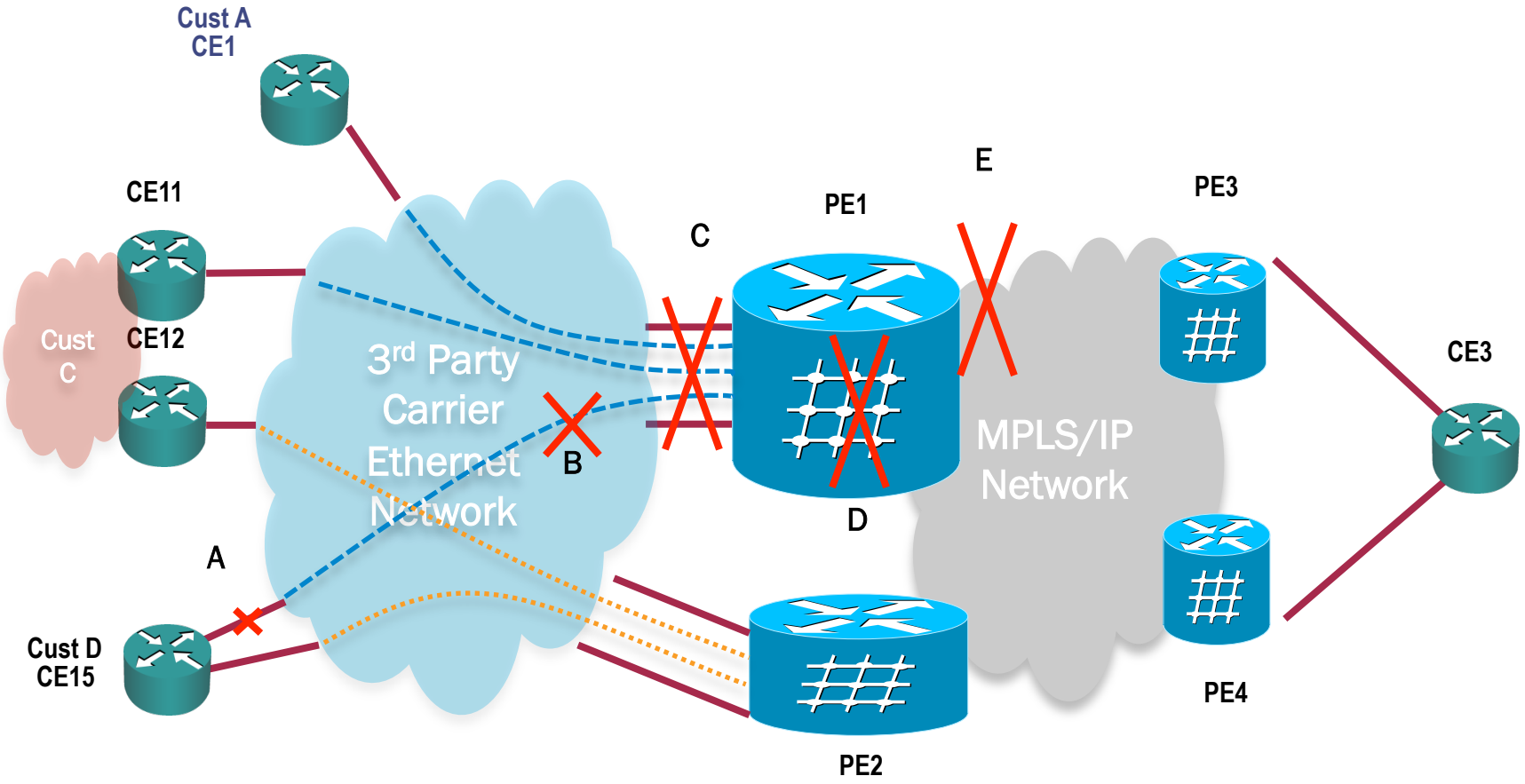
- Failure & failure recovery of a SH AC MUST NOT impact any other ACs (within its own EVI) or other EVIs – no flushing what so ever
- Failure & failure recovery of A/A MH AC MUST NOT impact any other ACs (within its own EVI) or other EVIs – no flushing what so ever
- Failure & failure recovery of S/A MH AC MUST only impact the service instance associated with that AC – e.g., flushing SHALL be limited to that single service instance (I-SID)
- Failure & failure recovery of S/A MH AC MAY only impact C-MACs associated with MHD/MHNs for that service instance – e.g., flushing SHALL be limited to single service instance (I-SID) and only CMACs for S-A MHD/MHNs. In other words, CMAC for SHD/SHN SHALL NOT be impacted.

Solution for PBB-EVPN

- Existing PBB-EVPN solution (per IETF draft) is leveraged with the following simple modifications:
  - Associate ACs with ESI (instead of physical ports)
- BMAC assignment remains the same as before (to meet the MUST requirements)
  - One shared BMAC for both SHD/SHN and S-A MHD/MHN
  - One individual BMAC per A-A customer MHD/MHN
- For S-A/A-A scenarios, DF election is done as before per (ESI, I-SID)

=> Existing PBB-EVPN solution can be leveraged entirely as long as ES is model as {set of ACs/EFPs) instead of (set of physical ports)

Failure Handling



Failure Handling

- A & B: CFM running on PE detects AC failure and advertises the shared B-MAC with the MAC mobility Ext. Comm. and a list of affected I-SIDs (typically one!)
  - In this approach flushing is limited to the CMACs associated with the shared BMAC for the affected I-SID(s)
  - If it is needed to confine the flooding further, then we can use two shared BMACs – one for SH and another for S-A MH. In this case, the flooding will be limited to C-MACs associated with S-A MHDs/MHNs for the affected I-SID
- C: If ENNI port fails, then the corresponding PE advertises the shared BMAC with MAC mobility Ext. Comm. but without any I-SID list
  - Or send multiple messages to cover all the I-SIDs
- D: This results in shared BMAC to be withdrawn and effectively giving the result for “C”
- E: Same as “D”

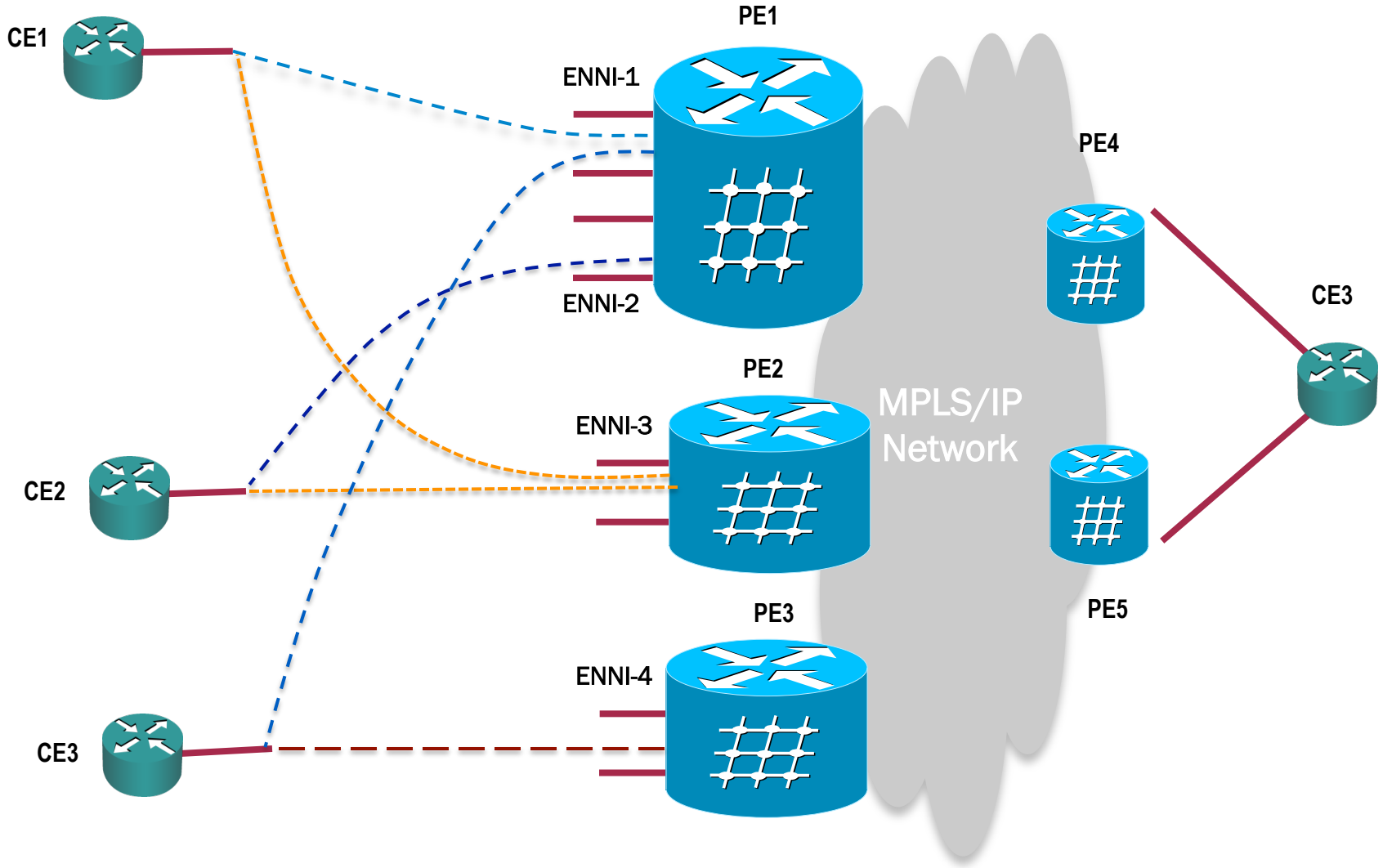
Solution with Fast Convergence

- Upon ENNI failure, we need to do two things:
  a) DF election for all impacted ESIs
  b) Flushing of BDs for the impacted I-SIDs
- Fast convergence of (b) is automatically achieved when shared BMAC is advertised with MAC mobility Extend Community (either without any I-SID list or with a list of I-SIDs)
- How do we do fast convergence for (a) ???
  - One possible way is to color all vES's for the same ENNI with the same color. Then when the ENNI fails, the corresponding PE withdraws that color

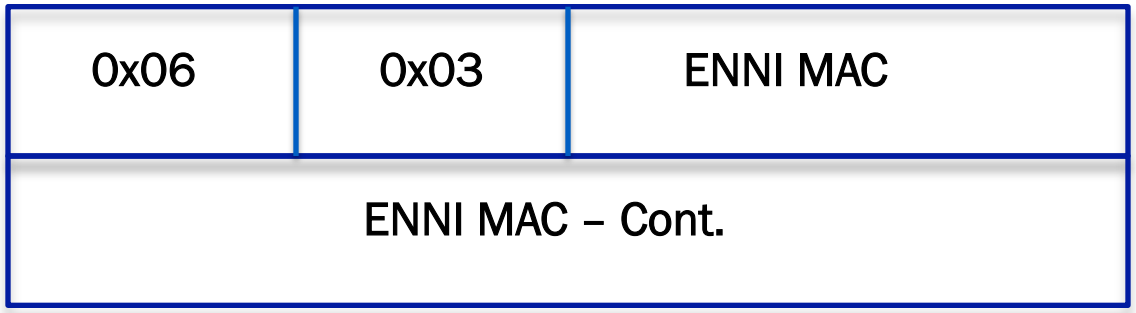
Fast Convergence for (a) – Cont.

- Procedure
  1. When a vES is configured, the PE color the vES with this MAC and advertises the ES routes of that ENNI with this color (this MAC is ONLY used in Control Plane)
  2. All other PEs (in the group) take note of this color
  3. Upon an ENNI failure, the PE sends the flush message as before but with this new Ext. Comm.
  4. Other PEs will use this info to
     a. Flush their impacted CMACs
     b. initiate DF election across all affected vES's
  5. PEs also upon receiving ES withdrawals, will clean up their tables

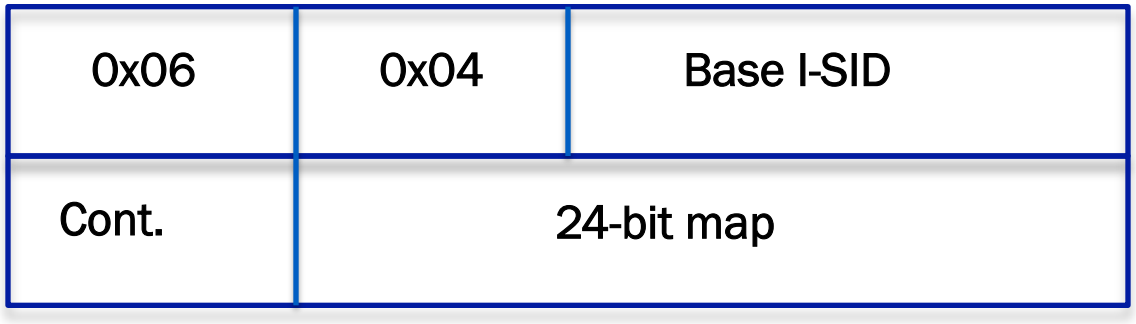
Fast Convergence Upon ENNI Failure



New EVPN BGP Extended Community



New EVPN BGP Extended Community



- 24-bit map represents the next 24 I-SID after the base
- For example based I-SID of 10025 with 24-bit map of zero means, only a single I-SID of 10025.
- I-SID of 10025 with bit map of 0x000001 means there are two I-SIDs, 10025 and 10026