draft-sajassi-bess-evpn-vpls-all-active-00.txt
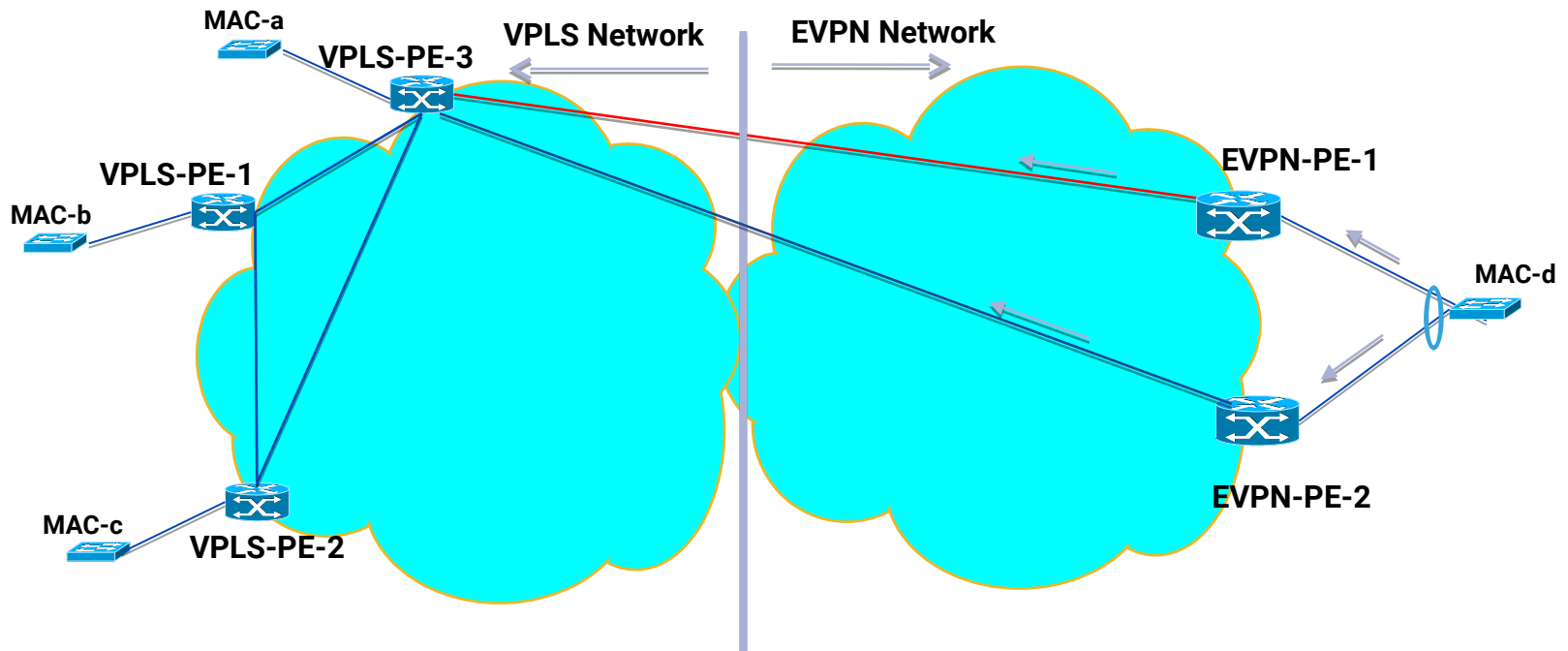
A. Sajassi (Cisco), S. Salam (Cisco), P. Brissette (Cisco), L. Jalil (Verizon)

IETF 99, July 2017
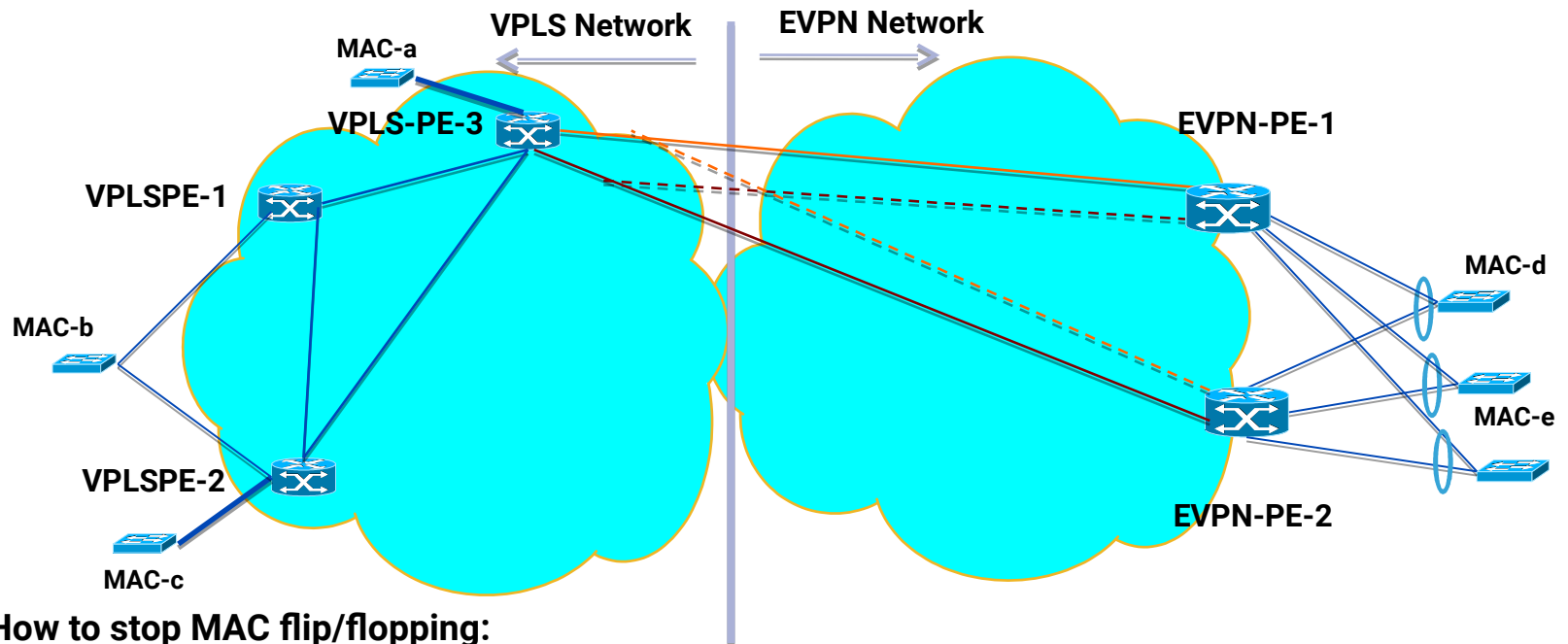
Prague

Issues w/ VPLS interop for All-Active MH

- **Traffic from MAC-d to MAC-a is load-balancing between red & blue PWs**
- **VPLS-PE3 learning is constantly switched between red & blue PWs**
- **Constant MAC flip/flopping causes high overhead on NPU**

How to Resolve the MAC Flip/Flopping Issue?

How to stop MAC flip/flopping:
- Regardless of which dual-homed EVPN PEs, traffic from dual-homed device arrives at, use the same PW toward destination VPLS PE
- If traffic from dual-homed device arrives at DF PE, then use the "Primary" PW from the DF PE
- If traffic from dual-homed device arrives at non-DF PE, then use Alias PW (same as "Primary" PW from DF PE)

What does it do to Load Balancing?
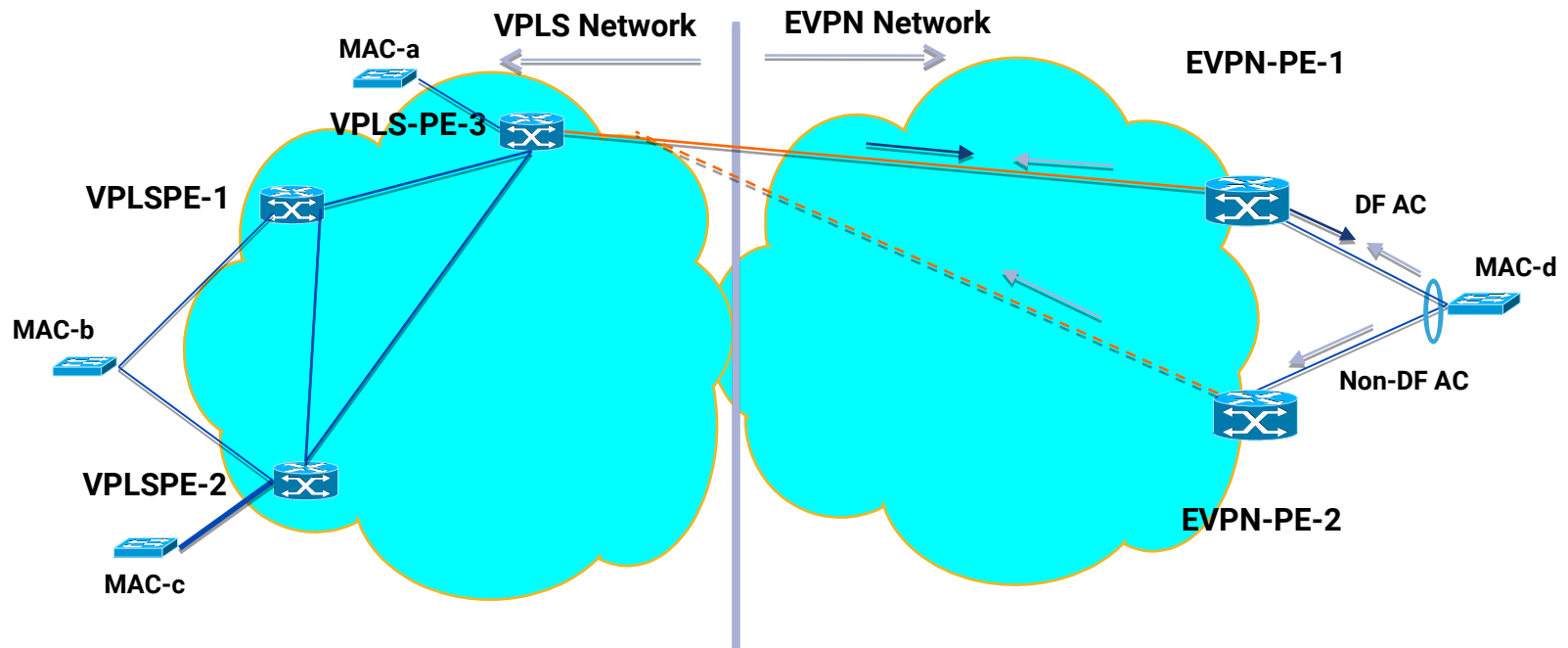
- From EVPN to VPLS direction: traffic "from" multi-homed device is load balanced on a per flow basis
- From VPLS to EVPN direction: traffic "to" multi-homed device is load balanced on a per (VLAN, site) basis – ie, traffic is sent to only one of the multi-homing EVPN PEs
- With this load-balancing scheme, we can devise a simple enough solution for All-Active multi-homing

What Control Plane Changes Are Needed by EVPN PEs?

1. Synch up VPLS PW labels (advertised by VPLS PEs) among EVPN PEs in the redundancy group

   ❑ This is easy for VPLS-BGP because from VPLS-BGP route advertisement, each multi-homing EVPN PE can derive the Alias PW (VE-IDs & label blocks are advertised in VPLS-BGP)

   ❑ Assuming DHD, each EVPN multi-homing PE keeps two PW labels toward a destination VPLS PE

      ➢ One label for its own (i.e., Primary PW)

      ➢ Another label for the other EVPN PE in the dual-homing (i.e., Alias PW)

2. Synch up MAC addresses learned over primary PW to the other Multi-Homing EVPN PE(s)
   ❑ This is already done in EVPN for Attachment Circuits. It can be easily extended to PWs (the way to extend it is by adding VE-ID or NH address of VPLS PE along with MPLS label in ESI field)

NOTE: no new EVPN routes or attributes should be needed !!

What Data Plane Changes Are Needed by EVPN PEs? Know Unicast

1. Each dual-homing EVPN PE performs DF election for a given Ethernet Segment (DHD) as before, and it is either a DF or not DF for that <ES, EVI>

2. If known-unicast traffic is received by the DF PE, then it uses its own PW label

3. If known unicast traffic is received by the non DF PE, then it uses the Alias PW label (PW for the other dual-homing EVPN PE)

❑ In other words, it uses DF / non-DF status of the incoming interface, to chose which of the two PWs to choose from

What Data Plane Changes Are Needed by EVPN PEs? BUM Traffic

- BUM traffic received over ACs:
  - If BUM traffic is received by the DF PE, then it uses its own replication list
  - If BUM traffic is received by the non-DF PE, then it uses the Alias replication list (replication list consisting of Alias PWs)
    - In other words, it uses DF / non-DF status of the incoming interface, to chose which of the two replication list to use

- BUM traffic received over VPLS PWs:
  - Received BUM traffic from VPLS PW should be flagged w/ BUM marking per EVPN MPLS ingress flooding mechanism
  - At the egress AC <ES, EVI>, all BUM VPLS traffic gets filtered on non-DF AC based on this BUM marking (just like EVPN traffic)
  - This feature already exists for EVPN !!

Failure Handling − Nothing new here!

- Link failure on DF EVPN PE
  - ○ DF EVPN PE sends mass withdraw per-ES and per-EVI to other EVPN PEs
    - ➢ Other EVPN PEs adjust their path list to only point to the non-DF PE
  - ○ DF EVPN PE sends VPLS flush messages to other VPLS PEs
  - ○ Non-DF EVPN PE becomes DF and it uses its own VPLS PW label for sending traffic toward VPLS PEs

- Link failure on non-DF EVPN PE
  - ○ Non-DF EVPN PE sends mass withdraw per-ES and per-EVI to other EVPN PEs
  - ○ Other EVPN PEs adjust their path list to only point to the DF PE
  - ○ Nothing to do wrt VPLS PEs

Next Steps

- Beef up the text for VPLS-LDP

- More discussions on the list