draft-sajassi-bess-evpn-ip-aliasing-00.txt

A. Sajassi (Cisco), G. Badoni (Cisco), P. Warade (Cisco), S. Pasupula (Cisco)

IETF 99, July 2017

Prague

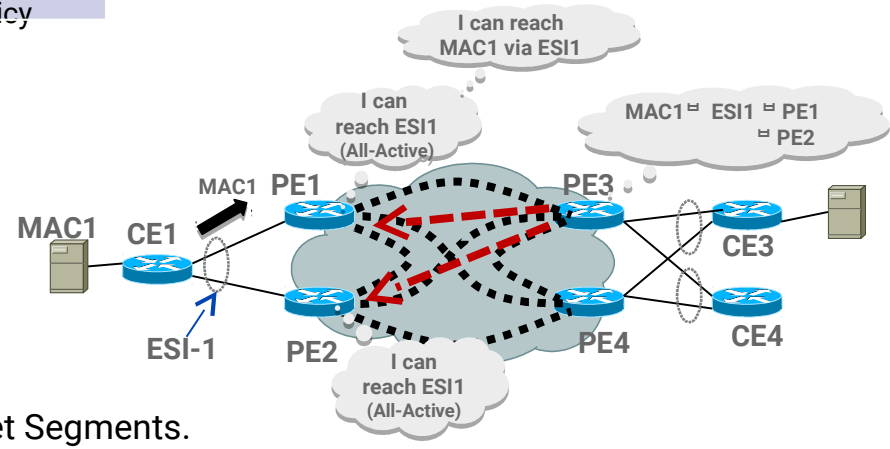
Objectives

- To improve load-balancing for IP routing in symmetric IRB by extending EVPN aliasing procedures to IP forwarding for symmetric IRB
  - Symmetric IRB only maintains ARP entries for its locally connected hosts. Thus, it only performs L3 forwarding for remote hosts on ingress PEs;
  - Whereas, asymmetric IRB maintains ARP entries for remote hosts and performs both L3+L2 forwarding for remote hosts. Thus it leverages L2 aliasing procedures for its load-balancing

Aliasing Procedure in EVPN (for MAC addresses)

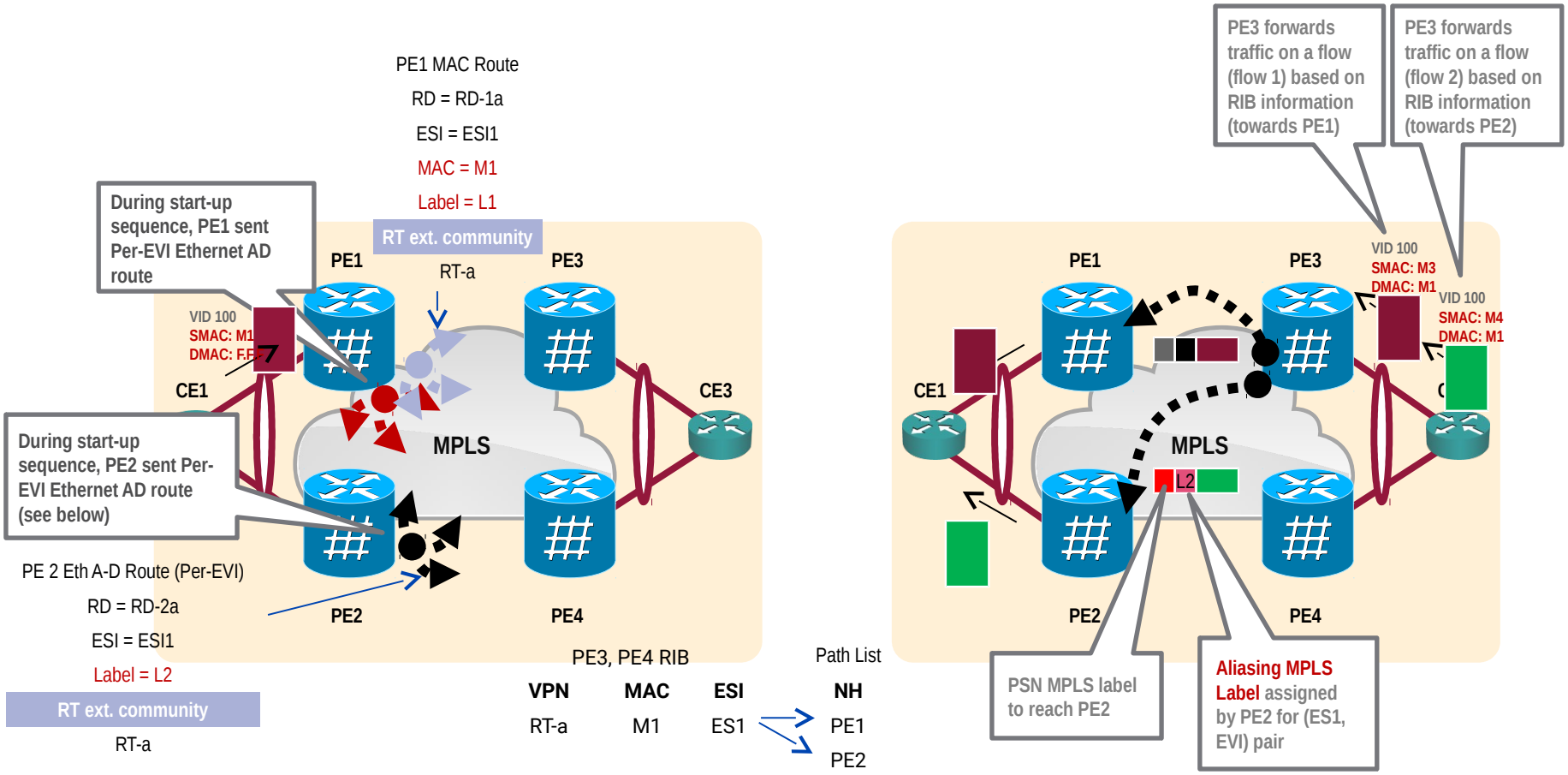
Route Type	Usage	Benefits
Ethernet A-D Route (Type 1)	• Advertising Split-Horizon Label • Aliasing • Mass Withdraw of addresses • SH/AA MH Indication	• Loop avoidance – even transient • Efficient load • Fast convergence • balancing • Per-site policy

Challenge:
How to load-balance traffic towards a multi-homed device across multiple PEs when MAC addresses are learnt by only a single PE?



- PEs advertise in BGP the ESIs of local multi-homed Ethernet Segments.
  - All-Active Redundancy Mode indicated
- When PE learns MAC address on its AC, it advertises the MAC in BGP along with the ESI of the Ethernet Segment from which the MAC was learnt.
- Remote PEs can load-balance traffic to a given MAC address across all PEs advertising the same ESI.

L2 Unicast Forwarding and Aliasing



Aliasing for IP addresses

- Basically repeat RFC 7432 aliasing procedure for L3 EVI (instead of L2 EVI)
- Besides advertising Ethernet AD per EVI route for L2 EVI (MAC-VRF), also advertise it for L3 EVI (IP-VRF)
- This Eth AD per EVI is advertised with RT corresponding to the L3 EVI (IP-VRF)
- Remote PE uses the alias route, to build next hop adjacencies for that ES/EVI
- Both MAC/IP route and Eth AD/EVI route needs to be validated by Eth AD/ES route as before

Mass withdraw upon failure

- Just as before, when an access link (or node) failure happens, the Eth AD per ES route corresponding to the failed ES is withdrawn
- This time around this route withdraw is sent with not only RTs for L2 EVI but also RTs for L3 EVI
- The receiving PEs use this route to trim down their next hop lists for both MAC addresses (using L2 RTs) and IP addresses (using L3 RTs)

Next Step

- Discussions on the Mailing list