

draft-sajassi-l2vpn-pbb-evpn-03.txt

Ali Sajassi (Cisco), Nabil Bitar (Verizon), Aldrin
Issac (Bloomberg), Samer Salam (Cisco), Sami
Boutros (Cisco), Florin Balus (ALU), Wim
Henderickx (ALU), Clarence Filsfils (Cisco),
Dennis Cai (Cisco), Lizhong Jin (ZTE)

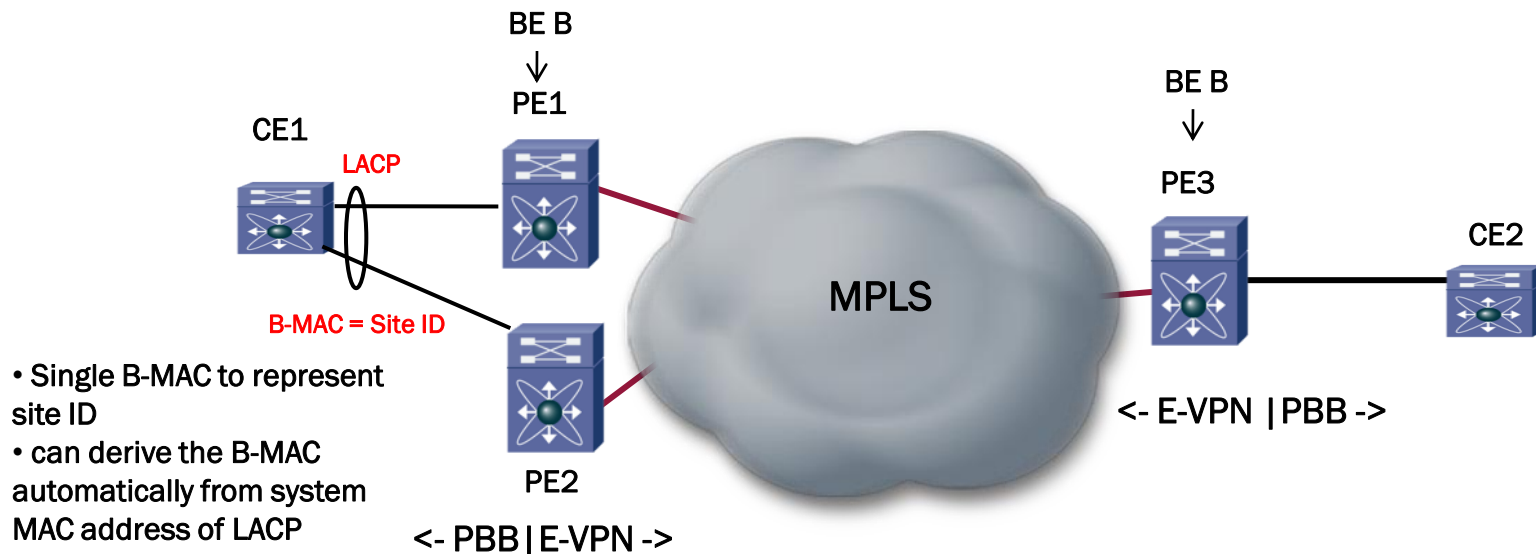
November 14th, 2011

IETF Taipei

History

- Rev 01
 - Review of E-VPN requirements
 - Additional PBB-EVPN requirements
 - Solution Overview & BGP Encoding
 - Advantages of PBB-EVPN
- Rev 02
 - Added MAC Mobility Extended Community
- Rev 03
 - Added new BGP route for TRILL
 - Added ARP suppression

Solution Overview



- Advertise local B-MAC addresses in BGP to all other PEs that have at least one VPN in common just like E-VPN
- Build a forwarding table from remote BGP advertisements just like E-VPN (e.g., association of B-MAC to MPLS labels)
- PEs perform PBB functionality just like PBB-VPLS
 - C-MAC learning for traffic received from ACs and C-MAC/B-MAC association for traffic received from core

Overview of Advantages

1. MAC Advertisement Route Scalability

- A single B-MAC represents a multi-homed site
- A single B-MAC can represent all single-homed sites
- A single C-MAC represents a single VM
- => several order of magnitude difference between C-MAC & B-MAC

2. C-MAC Mobility with MAC sub-netting

- Typically C-MACs are not managed and thus can not be sub-netted
- B-MACs on the other hand are always managed and can easily be sub-netted
- Even when C-MACs are sub-netted, VM mobility contradicts the effect of sub-netting

Advantages – Cont.

3. C-MAC Address Learning and Confinement

- With C-MAC learning in control plane, C-MACs are always in RIBs and maybe also in FIBs
- With C-MAC learning in data plane, C-MACs are never in RIBs and they are only present in FIBs for active flow.

4. Interworking with TRILL & 802.1aq/.1bp networks and C-MAC Transparency

- PBB encapsulation enables end-to-end tunneling of C-MAC addresses for the access networks thus avoiding termination and learning by DC WAN Edge PE

Advantages – Cont.

5. Per Site Policy

- Since B-MAC addresses are per site, BGP policy per MAC gives us very nice set of per-site policy

6. Avoiding C-MAC flushing

- Since B-MAC represent a site, a link, port, or node failure doesn't change the B-MAC address – it only changes number of next hop for that B-MAC

7. Avoid transient loop for known unicast when doing egress MAC lookup

- Since B-MAC SA is always transmitted with every frame, checking of every frame against its source MAC SA for known unicast frame is already provided by PBB

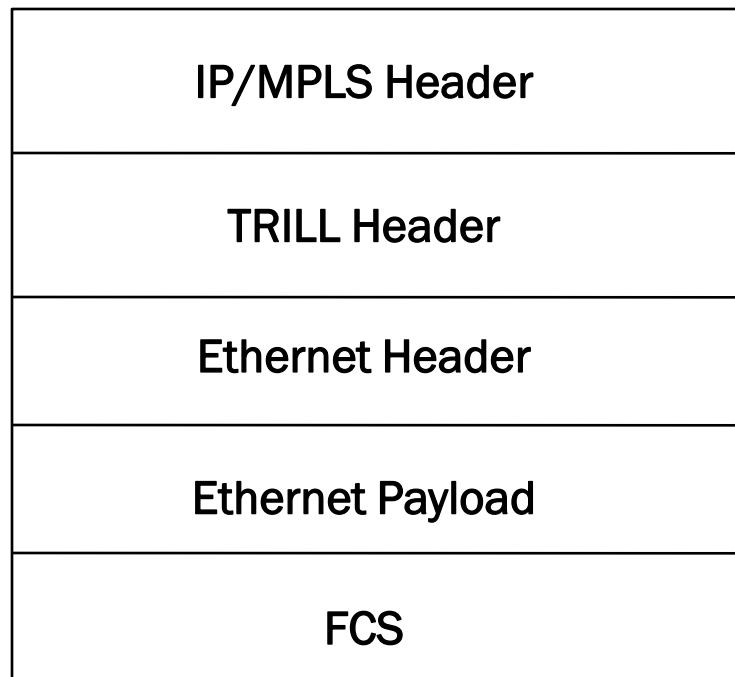
BGP Encoding

- Ethernet A-D is not needed and it is not used
 - Avoids different modes of operation associated with this route
 - Simplifies operation and provisioning
- MAC Mobility Extended Community (delta to rev 01)
 - It is a transitive extended community
 - When advertised with a B-MAC route, it signals all the C-MAC forwarding tables associated with the I-SIDs corresponding to the RTs should be flushed

BGP Encoding – Cont.

- TRILL Nickname Advertisement route
 - For supporting interconnection of TRILL islands over PBB-EVPN and maintain “independence” for each island
 - Similar to Ethernet Route – e.g., just replacing MAC address field with TRILL Rbridge Nickname field

Encapsulation of TRILL over MPLS



- Very Efficient Encapsulation
- Could have done with Ethernet encapsulation but it would have added additional 16-bytes of overhead plus it would have required exchange of TRILL next hop MAC addresses or use of well known MAC addresses

Operation for TRILL/802.1Qbp over MPLS

- TRILL Nicknames or .1aq/.1Qbp B-MAC addresses are exchanged among different IS-IS islands using BGP – BGP Provides “independence” among TRILL/.1Qbp islands – e.g., each island IS-IS can be run independently from all others
- It is assumed that TRILL Nicknames or .1aq/.1Qbp B-MAC addresses are globally unique in the network – e.g., site-id:Rbridge-id or site-id:mac
- Imposition/disposition operation for TRILL frames is similar to B-MAC frames except the MPLS label is associated with TRILL nickname instead of B-MAC

ARP Suppression

- Similar to E-VPN in operational principle
- Difference is that E-VPN advertises MAC/IP binding in control plane; whereas, PBB-EVPN uses the data-plane for this purpose
- PBB-EVPN MES nodes snoop ARP request or responses on the ACs or received over MPLS core
- Then they build a cache of MAC/IP binding from these messages and use this cache to respond to subsequent ARP messages received over local ACs and targeting hosts on remote MESes.

Next Step

- Authors think this draft is ready for the WG call