draft-sajassi-l2vpn-pbb-evpn-02.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)

July 28th, 2011

IETF Quebec City

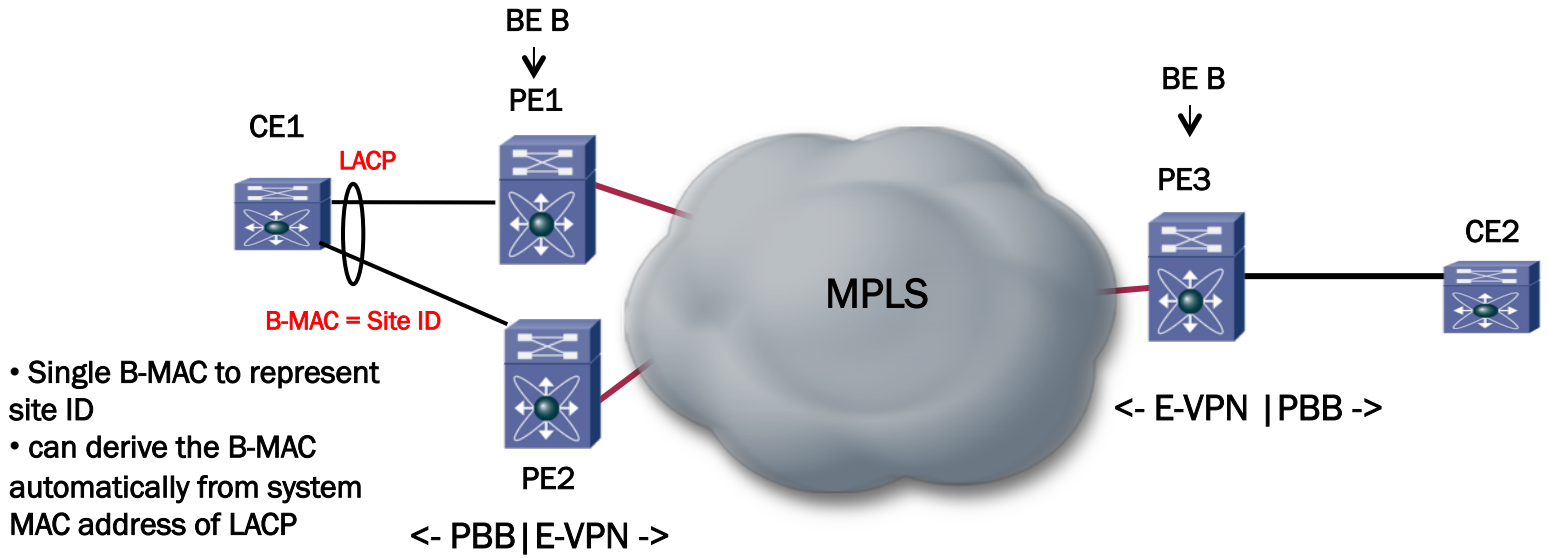
Additional Requirements (PBB-EVPN)

1. MAC Advertisement Route Scalability
   - To support millions of C-MAC addresses (million of VMs)
2. C-MAC Mobility with MAC sub-netting
   - Support C-MAC address mobility while retaining the scalability benefits of MAC sub-netting
3. C-MAC Address Learning and Confinement
   - MES nodes to maintain C-MAC addresses in their RIB & FIB for active flows ONLY

Additional Requirements (PBB-EVPN)

4. Interworking with TRILL & 802.1aq/.1Qbp networks and C-MAC Transparency
   - To avoid learning of C-MACs by DC WAN Edge PE
5. Per Site Policy
   - To support connectivity policy rules at the granularity of a site (or segment)
6. Avoiding C-MAC flushing
   - To avoid C-MAC flushing upon link, port, or node failure for multi-homed devices
7. Avoid transient loop for known unicast when doing egress MAC lookup

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

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

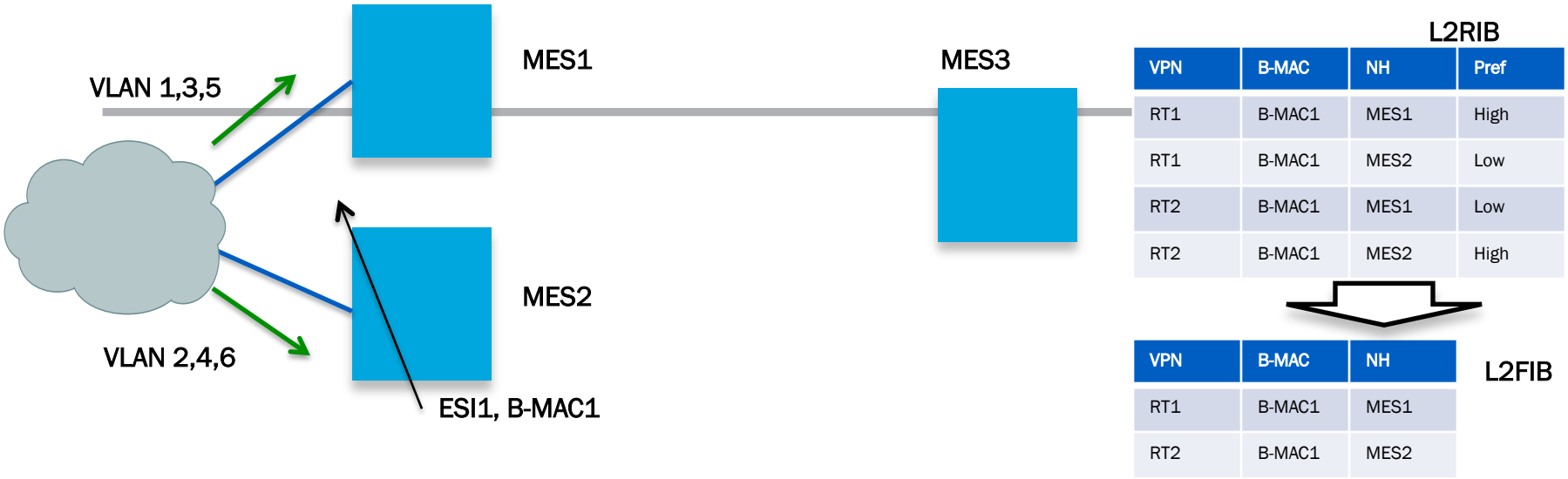
Advantages – Cont.

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

Network Multi-Homing: Active/Active Per VLAN Load-Balancing



VPN	B-MAC	NH	Pref
RT1	B-MAC1	MES1	High
RT1	B-MAC1	MES2	Low
RT2	B-MAC1	MES1	Low
RT2	B-MAC1	MES2	High

VPN	B-MAC	NH
RT1	B-MAC1	MES1
RT2	B-MAC1	MES2

- Each MES advertises two MAC routes per Ethernet Segment: one with high Local Pref for active VLANs, and another with low Local Pref for standby VLANs.
  - Use MES RD with LSBit reset (Standby) or set (Active) to differentiate the prefixes.
- Remote MES installs the route with higher local pref into FIB for associated B-MAC.
- MES1 MAC Routes:
  - Route 1: RD11, B-MAC1, RT1, RT3, RT5, Local Pref = High
  - Route 2: RD12, B-MAC1, RT2, RT4, RT6, Local Pref = Low
- MES2 MAC Routes:
  - Route 1: RD22, B-MAC1, RT1, RT3, RT5, Local Pref = Low
  - Route 2: RD21, B-MAC1, RT2, RT4, RT6, Local = High

Future Considerations

- ARP suppression
- Anything else ?