

draft-ietf-l2vpn-pbb-evpn-01.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)

March 26th, 2012

IETF Paris

Purpose

- To inform TRILL WG that there is already an L2VPN WG draft to address DC interconnection over MPLS/IP networks for TRILL islands
- L2VPN WG was re-chartered about a year ago to work on this EVPN/PBB-EVPN solution

History

- PBB-EVPN is part of EVPN family solution
- EVPN rev 00 was published in Oct 2010
 - It was a merged of R-VPLS and MAC-VPN drafts
 - R-VPLS and MAC-VPN were published in Mar 2010
- PBB-EVPN rev 00 was published in Mar 2011
- PBB-EVPN rev 03 describes how TRILL islands can be interconnected independently over WAN
 - It was published in Oct 2011
 - It was presented at L2VPN WG in Nov 2011

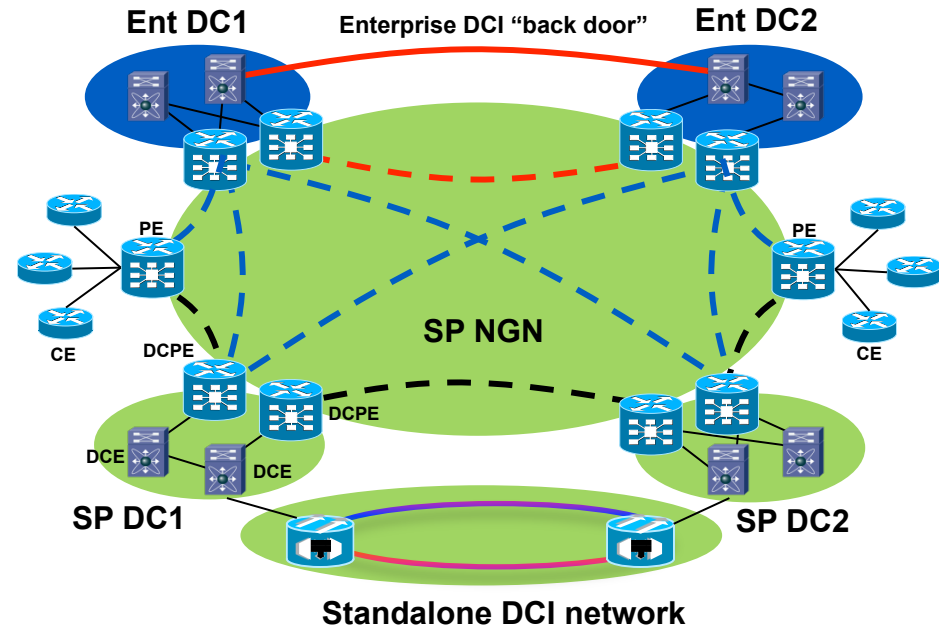
Scope of Work

- L2VPN WG was re-chartered to address
 - Requirements driven by cloud computing services and data centers as they apply to Layer-2 VPN services (including DC Interconnection over MPLS/IP networks)

DC Interconnect

- DCI Requirements:

- Multi-homing
- Scale (MAC-addresses, Number of Service Instances)
- Load balancing
- Optimal Forwarding
- Multicast optimization
- Multi-tenancy



- Interconnection models:

- Enterprise to Enterprise (E2E)
- Enterprise to Service Provider (E2SP)
- Service Provider to Service Provider (SP2SP)

Evolving Requirements for L2VPN

1. All-active Redundancy

- Flow Based Load Balancing
- Flow Based Multi-pathing
- Geo-redundancy and Flexible Redundancy Grouping

2. Simplified Provisioning and Operation

- Core Auto-Discovery
- Access Multi-homing Auto-Discovery
- New Service Interfaces

3. Optimal Multicast with LSM

- P2MP Trees
- MP2MP Trees

4. Fast Convergence

- Link/Port/Node Failure
- MAC Mobility

5. Scalable for SP virtual private cloud service:

- Support O(10 Million) MAC Addresses per DC
- Confinement of C-MAC Learning

6. Seamless interworking between TRILL / 802.1aq / 802.1Qbp and MST / RSTP

- Guarantee C-MAC Transparency on PE

7. Fast Convergence

- Avoiding C-MAC Flushing

Underline: Addressed by VPLS

 Addressed by E-VPN & PBB-EVPN

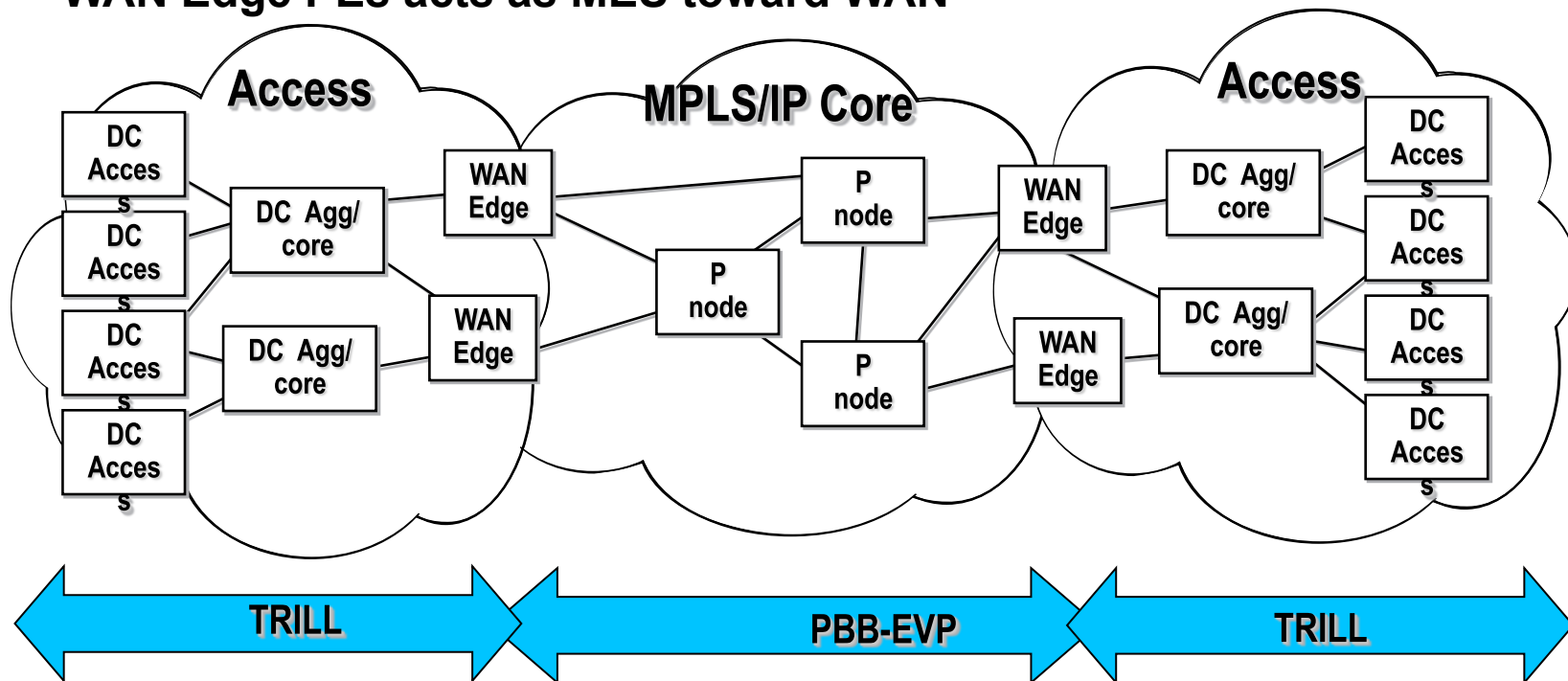
 Addressed by PBB-EVPN

Requirements wrt TRILL DCI

- Connecting TRILL islands **independently**
- Support Multi-tenancy for TRILL network provider
 - Allow multiple TRILL network providers share the same common MPLS/IP WAN by providing virtual private network for each TRILL network provider
 - Single-tenancy can be treated as a degenerate case of multi-tenancy
- Support optimum unicast & mcast forwarding for L2VPN service over both MPLS and IP networks
- Support active/active load balancing when TRILL network is multi-homed to PBB-EVPN
 - Active/active load balancing for per VLAN and per flow for both unicast and mcast (E-VPN and PBB-EVPN have extensive DF election mechanism)

Connecting TRILL Islands Independently

- No need to terminate TRILL encapsulation at WAN Edge PEs
- WAN Edge PEs act as Edge Rbridge wrt control plane and transit Rbridge wrt data-plane
- WAN Edge PEs acts as MES toward WAN



Advantages – Cont.

(copied from IETF 82 preso)

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/TRILL 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.

(copied from IETF 82 preso)

5. Per Site Policy

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

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

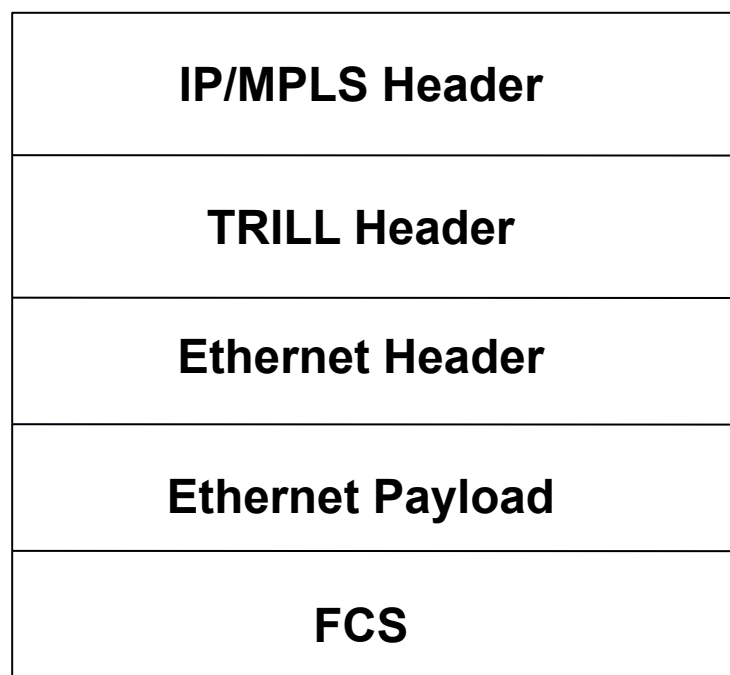
- Since B-MAC SA is always transmitted with every frame, checking of every frame against its source MAC SA for known unicast frame is already provided by PBB

BGP Encoding – Cont.

(copied from IETF 82 preso)

- 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 (copied from IETF 82 preso)



- **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

(copied from IETF 82 preso)

- 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 (copied from IETF 82 preso)

- 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.

Conclusion

- L2VPN for Cloud and DC applications (including DCI) over MPLS/IP has already been re-chartered into L2VPN WG
- DCI solution over MPLS/IP (using MP-BGP) seems to be out of the scope of the current TRILL charter
- Authors of PBB-EVPN WG draft think that this draft addresses the DC Interconnection for TRILL islands
- If there is anything missing, we'll be glad to work on it and address it