

TRILL DataCenter/Campus/PBB Inter-connect over IP core with BGP

Bhargav Bhikkaji

Balaji Venkat Venkataswami

Ramasubramani Mahadevan

Shivakumar Sundaram

Narayana Swamy Perumal

[draft-balaji-trill-over-ip-multi-level-05.txt](#)

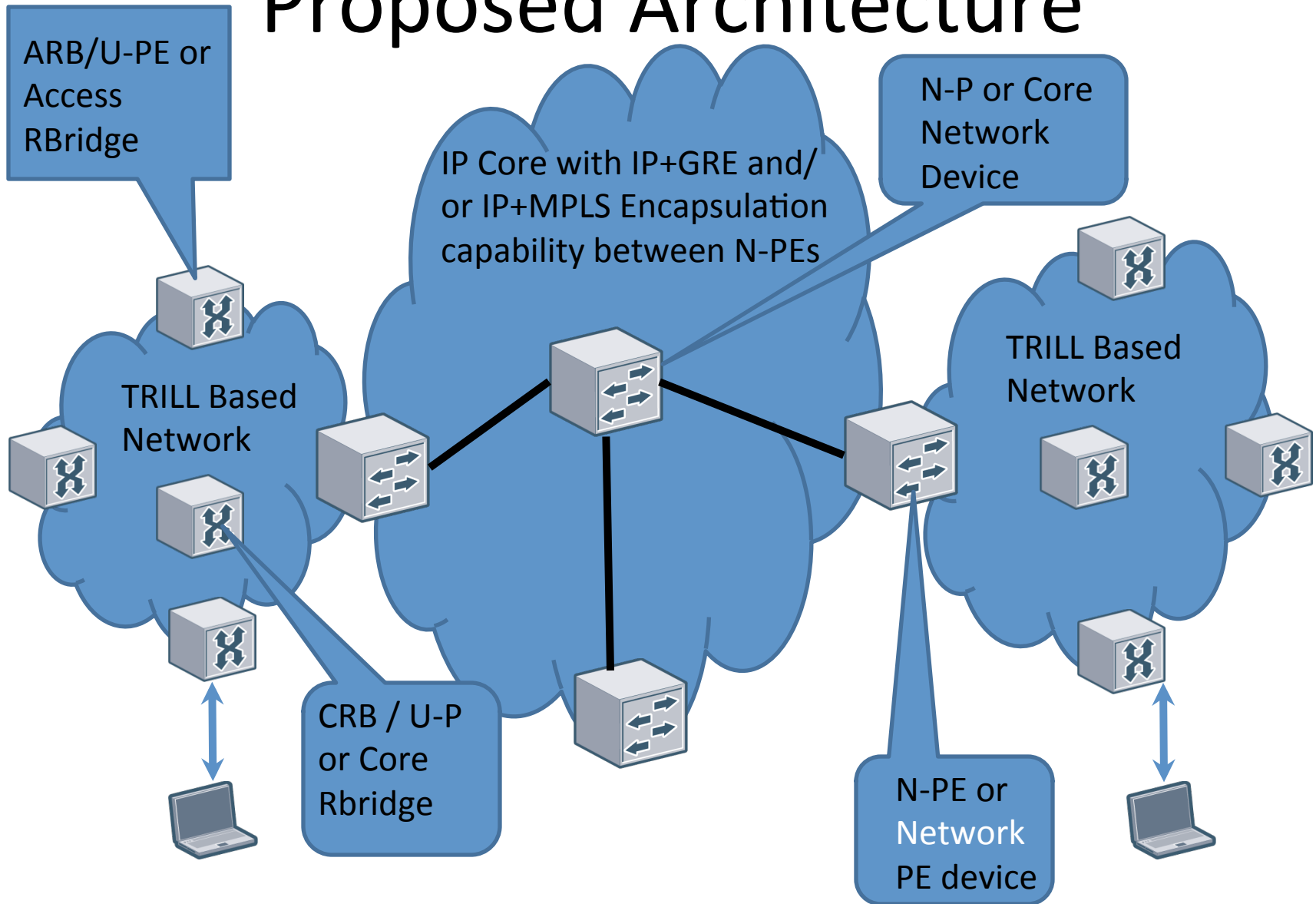
DELL-Force10

Introduction and Problem Statement

Interconnecting Disparate DataCenters

- Providing Layer 2 extension capabilities amongst different independently managed data centers running TRILL.
- Provide a solution agnostic to the technology used in the service provider network
- Provide auto-configured tunnels instead of pre-configured ones in the transport network.
- Allow for additional facilities and features like TE and QoS offered to be applied for these interconnected TRILL sites
- Such an interconnect should scale independently of the number of MAC routes in each of the interconnected sites.
- Allow a mechanism to facilitate inter-working of TRILL with other technologies such as STP, VXLAN, NVGRE.
- The problem statement also includes Campus networks deploying TRILL across multiple sites.

Proposed Architecture



Legend for Architecture

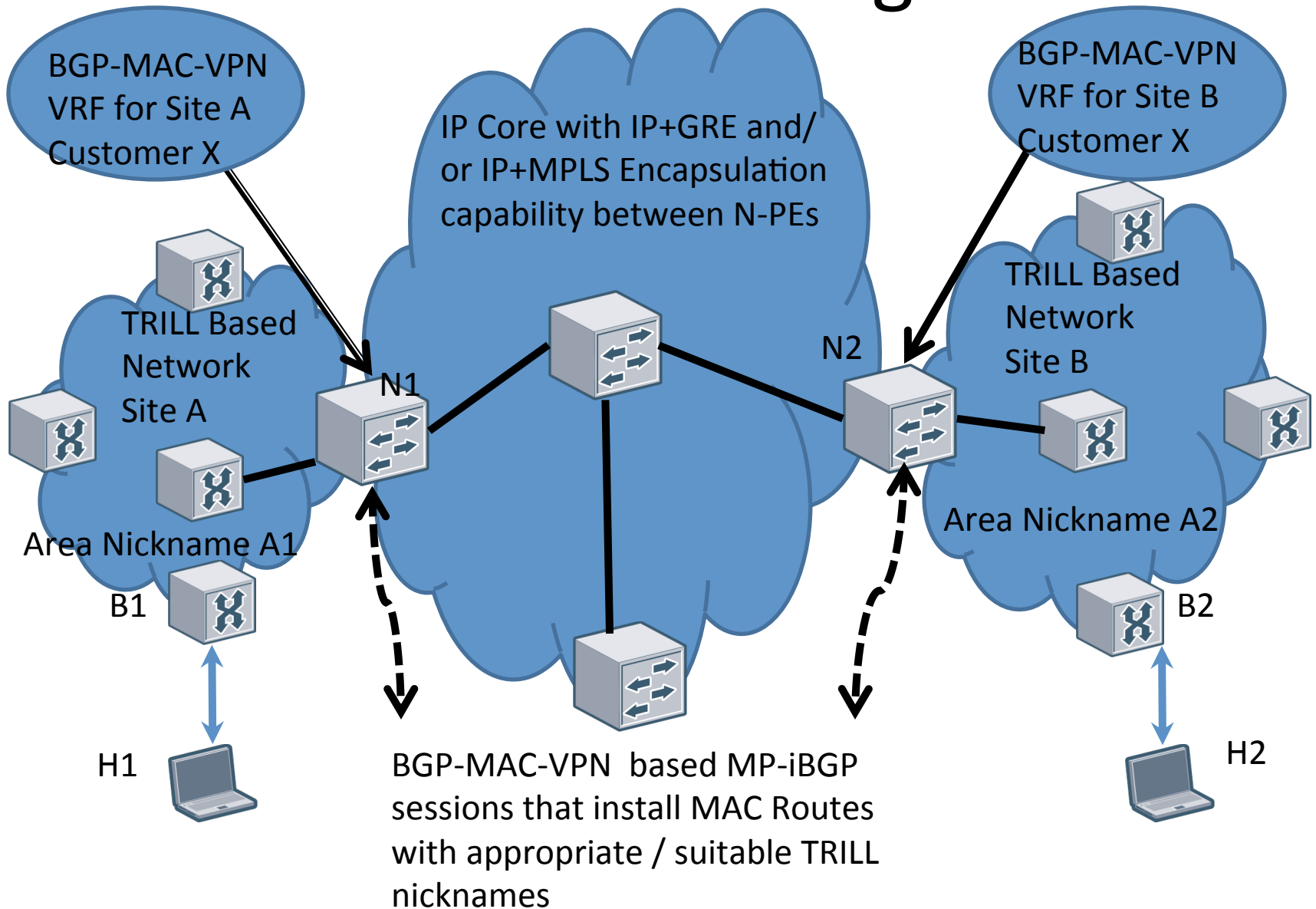
- **U-PE / ARB** : User-near PE device or Access Rbridge. U-PEs are edge devices in the TRILL site or tier-2 site.
- **U-Ps / CRB** : Core Rbridges in the TRILL sites. They do not have any hosts connected to them.
- **N-PE** : Network Transport PE device. This is a device with RBridge capabilities in the non-WAN facing side. On the WAN facing side it is a Layer 3 device supporting IP+GRE and/or IP+MPLS. On the non-WAN facing side it has support for VRFs one for each TRILL site that it connects to. It runs BGP to convey the BGP-MAC-VPN VRF routes to its peer N-PEs.
 - It may be implemented as a single device or across multiple devices.
- **N-P** : Network Transport WAN core device. This device is IP and/or IP +MPLS WAN core device that is part of the ISP / ISPs that provide the transport network that connect the disparate TRILL networks together.

Format of the L2-Forwarding-Table with regard to TRILL nicknames as next-hops

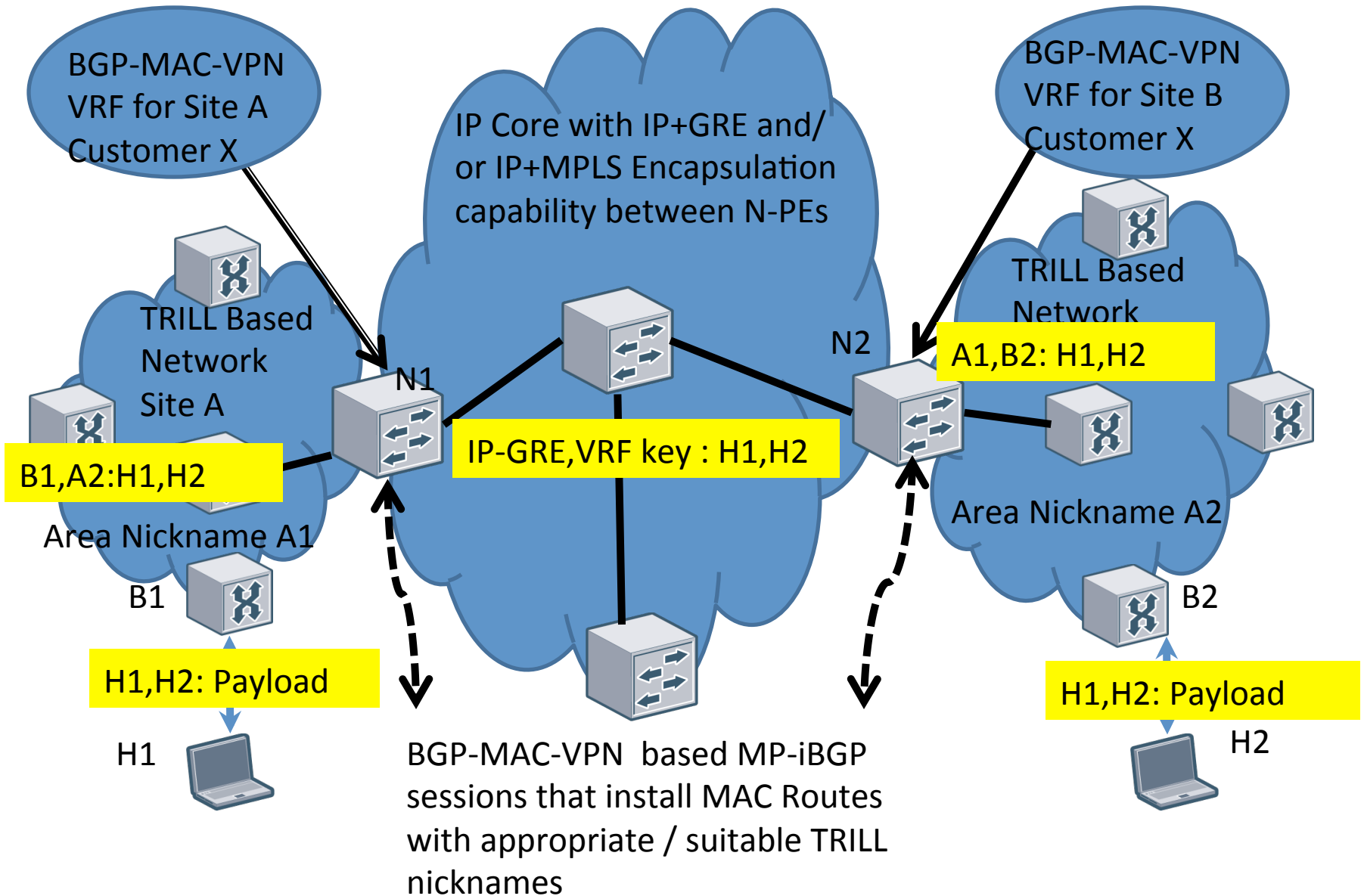
FORMAT of FORWARDING TABLE ENTRIES on an ACCESS RBRIDGE (ARB)

MAC Route / MAC Address to be reached	Next-hop Rbridge's Nickname
00:0f:bc:ef:d9:a6 [Local]	Local ARB Nickname : <16 bit Nickname>
00:ac:d9:ef:a5:bc [Non-Local]	Non-Local ARB AREA Nickname : <16 bit AREA Nickname>
...

Control Plane Exchanges for DC



Data Plane for DC



Uniqueness and Advantages

- o Uses existing protocols such as IS-IS for Layer 2 and BGP to achieve this.
- o Uses BGP-MAC-VPNs for transporting MAC-updates of customer devices between edge devices only.
- o Employs a hierarchical MAC-route hiding from the core Rbridges of the site. Employs a hierarchical VPN like solution to avoid routing state of sites within the transport core.
- o Multi-tenancy through the IP+GRE or IP+MPLS core is possible when N-PEs at the edge of the L3 core place various customer sites using the VPN VRF mechanism. This is otherwise not possible in traditional networks and using other mechanisms suggested in recent drafts.
- o The VPN mechanism also provides ability to use overlapping MAC address spaces within distinct customer sites interconnected using this proposal.
- o Mac Moves can be detected if source learning / Gratuitous ARP combined with the BGP-MAC-VPN update triggers a change in the concerned VRF tables.
- o PBB like functionality supported where P-VLAN and Customer VLAN are different spaces.
- o Uses regular BGP supporting MAC-VPN features, between transport core edge devices and the Tier-2 customer edge devices.
- o When new TRILL sites are added then no re-election in the Level 1 area is needed. Only the Pseudo-interface of the N-PE has to be added to the mix with the transport of the election PDUs being done across the transport network core.