White Paper

Ethernet VPN (EVPN) and Provider Backbone Bridging-EVPN: Next Generation Solutions for MPLS-based Ethernet Services

Introduction and Application Note

Last Updated: 5/2014

Ethernet VPN (EVPN) and Provider Backbone Bridging EVPN (PBB-EVPN) are next generation solutions that provide Ethernet multipoint services over MPLS networks. EVPN is different compared to existing Virtual Private LAN Service (VPLS) offerings due to its use of control-plane based MAC learning over the core. EVPN has been designed from the ground up to handle sophisticated access redundancy scenarios, per-flow load balancing, and operational simplicity. PBB-EVPN inherits all of the benefits of EVPN, while combining PBB (IEEE 802.1ah) and EVPN functions in a single node. This allows PBB-EVPN to simplify control-plane operation in the core, provide faster convergence and enhance scalability, when compared to EVPN. EVPN and PBB-EVPN applications include Data Center Interconnect (DCI) and carrier Ethernet E-LAN services.

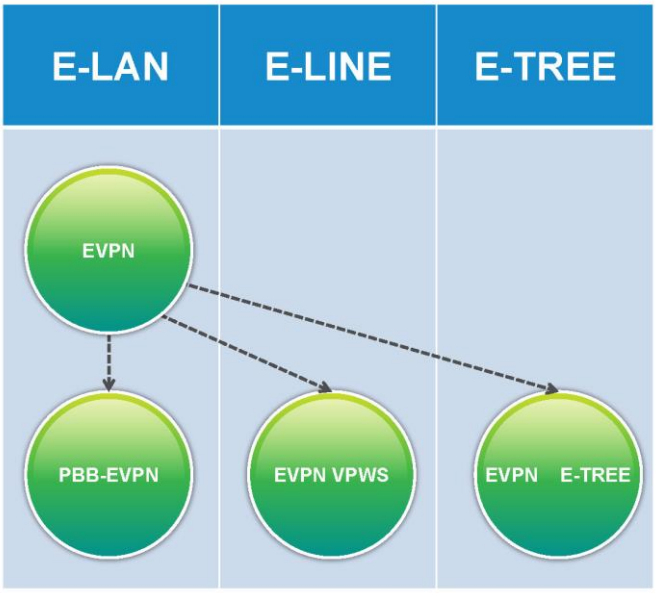
EVPN and PBB-EVPN Overview

EVPN and PBB-EVPN are members of the family of EVPN technologies, which encompasses next generation Ethernet L2VPN solutions that use Border Gateway Protocol (BGP) as control-plane for MAC address signaling / learning over the core as well as for access topology and VPN endpoint discovery. The introduction of EVPN marks a significant milestone for the industry, as it aligns the well-understood technical and operational principles of IP VPNs to Ethernet services. Operators are now able to leverage their experience and the scalability characteristics inherent to IP VPNs for their Ethernet offerings.

As depicted in Figure 1 below, the EVPN family includes various solutions that address Ethernet multipoint (E-LAN), Ethernet point-to-point (E-LINE) and Ethernet rooted-multipoint (E-TREE) service types. This whitepaper focuses on the E-LAN solutions. Altogether, the EVPN family enables all types of Ethernet services under a common architecture. These solutions are currently under standardization by the IETF L2VPN Working Group[1].

[1] http://datatracker.ietf.org/wg/l2vpn/

Figure 1. The EVPN Technology Family



Addressing DCI Requirements

MPLS Layer 2 VPNs based on pseudowires (PWs)[2] have been widely deployed in service provider and enterprise networks. L2VPN applications range from Ethernet business services, to fixed / mobile convergence and enterprise campus layer 2 transport.

Recently Data Center Interconnect (DCI) has become a leading application for Ethernet multipoint L2VPNs. As customers deploy virtualization to consolidate servers and become more flexible in their data centers, they demand higher agility, lower cost and resource optimization among DC sites. With that in mind, a DCI solution must support the following:

- Cloud bursting
- Disaster recovery and business continuity
- Workload (Virtual Machine (VM)) and data (storage) mobility

VM mobility, storage clustering and other data center services require nodes and servers in the same layer 2 network to be extended across data centers over the WAN.

Overcoming VPLS Limitations

For scenarios requiring Ethernet multipoint connectivity, MPLS deployments leverage Virtual Private LAN Services (VPLS)[3]. A VPLS service is built with a full-mesh of Ethernet pseudowires among MPLS Provider Edge (PE) routers that are part of a given layer-2 broadcast domain. A VPLS PE emulates an Ethernet bridge by performing data-plane MAC learning (also known as “flood and learn” operation) against traffic arriving over the pseudowires and local interfaces.

[2] IETF RFCs 3916, 3985, 4664 introduced the requirements and architecture for pseudowire-based L2VPNs
[3] IETF RFC 4761 and 4762

In the case of DCI, several requirements cannot be addressed solely by VPLS. First, in order to keep the data center always ON, and to utilize the resources and links as efficiently as possible, data centers need per-flow load balancing between DC switches and DCI routers. Second, highly virtualized multi-tenant service provider and large enterprise data centers require solutions that can address a high number of VLANs and MAC addresses. Finally, fast convergence is a must in order to minimize downtime and packet loss due to any network topology changes.

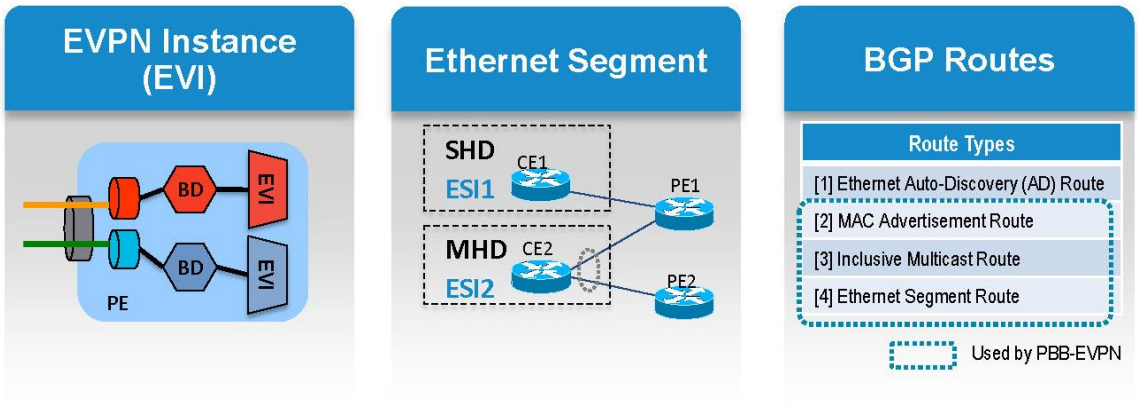
Existing VPLS PE dual-homing solutions[4] provide limited load balancing support, including: active / standby or active / active per-VLAN scenarios. If the traffic is heavier on one particular VLAN, administrative intervention would be required to manually re-assign VLANs to individual PEs. In cases of per-VLAN load balancing, manual administration is required to compensate for the lack of access auto-discovery and automatic service carving mechanisms. In general, VPLS was not designed to address the challenges associated with dual-homing and per-flow load balancing.

In addition, VPLS scalability, in terms of number of PEs, is limited by the maximum number of pseudowires that an implementation allows on a given VPLS Virtual Forwarding Instance (VFI) (this is typically in the low hundreds).

EVPN and PBB-EVPN Concepts

As depicted in Figure 2 below, there are three fundamental building blocks for this technology.

Figure 2. EVPN and PBB-EVPN Concepts



EVPN Instance (EVI)

An EVI represents a VPN on a PE router. It serves the same role of an IP VPN Routing and Forwarding (VRF), and EVIs are assigned import / export Route Targets (RTs). Depending on the service multiplexing behaviors at the User to Network Interface (UNI), all traffic on a port (all-to-one bundling), or traffic on a VLAN (one-to-one mapping), or traffic on a list / range of VLANs (selective bundling) can be mapped to a Bridge Domain (BD). This BD is then associated to an EVI for forwarding towards the MPLS core.

Ethernet Segment (ES)

Ethernet Segments are used to represent “site” connections. They are associated with the access-facing interfaces (physical or logical (Ethernet bundles)) that they represent, and are assigned a unique identifier (referred as Ethernet Segment Identifier (ESI)). Furthermore, a site can be connected to one or more PEs attached to a

[4] draft-ietf-l2vpn-vpls-multihoming; draft-ietf-pwe3-iccp

common MPLS core in order to provide access redundancy. With this, the solutions can support PE geo-redundancy and multi-homing (beyond dual-homing) requirements. Lastly, a site can be used to connect to a single device or to an access network.

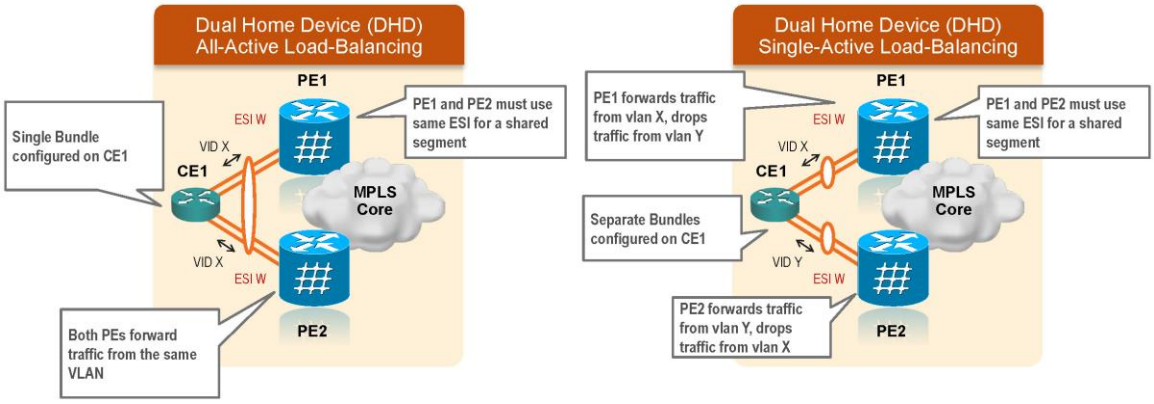
The following scenarios are available:

- Single-Homed Device (SHD)
- Multi-Homed Device (MHD)
- Single-Homed Network (SHN)
- Multi-Homed Network (MHN)

For MHDs (see Figure 3 below), EVPN and PBB-EVPN define two load balancing modes of operation:

- **All-active load balancing:** Supports multi-homed devices with per-flow load balancing. This mode allows the access device to connect via a "single" Ethernet bundle to multiple PEs and to send / receive traffic of the "same" VLAN from all the PEs in the same Ethernet segment.
- **Single-active load balancing:** Supports multi-homed devices with per-vlan load balancing. In this mode, the access device connects via "separate" Ethernet bundles to multiple PEs. PE routers in turn automatically perform service carving in order to divide VLAN forwarding responsibilities across the PEs in the Ethernet segment. The access device learns via the data-plane which Ethernet bundle to use for a given VLAN.

Figure 3. All-Active and Single-Active Load Balancing



EVPN BGP Routes and Extended Communities

EVPN / PBB-EVPN PEs signal and learn MAC addresses over the core via BGP[5]. For this purpose, a new address-family[6] in Multi-Protocol BGP (MP-BGP) and new BGP extended communities allow PE routers to advertise and learn prefixes that identify MAC addresses and Ethernet segments over the network. This contrasts with existing VPLS solutions, which rely solely on data-plane learning. Control-plane learning brings a number of benefits that allow EVPN and PBB-EVPN to address the VPLS shortcomings listed previously, including support for multi-homing with per-flow load balancing. Finally, by using BGP as a common VPN control-plane, providers can now leverage their operational experience and scalability characteristics inherent to IP VPNs, for their Ethernet offerings.

[5] On the access side, local MAC addresses continue to be learned based on data-plane
[6] A common AF has been assigned to EVPN and PBB-EVPN (AFI 25 (L2VPN) and SAFI 70 (EVPN))

A total of four new BGP routes have been defined in order to collectively achieve the following functions:

- MAC address reachability
- MAC mass withdrawal
- Access split-horizon filtering
- Aliasing
- VPN endpoint discovery
- Redundancy Group discovery
- Designated forwarder election

When used in conjunction with an MPLS data-plane, these BGP routes also signal the MPLS labels associated with MAC addresses and Ethernet segments. This separates the new technologies from existing L2VPN solutions. EVPN and PBB-EVPN no longer require the signaling and maintenance of pseudowires. EVPN MPLS labels represent Multipoint-to-Point (MP2P) Label Switched Paths (LSPs). Similar to IP VPNs, an EVPN MPLS label is signaled for use by the rest of the PEs in the VPN. There is no longer a requirement to signal separate point-to-point pseudowire Virtual Circuit (VC) labels for each remote PE. This brings the scale of L2VPNs to the same level of IP VPNs.

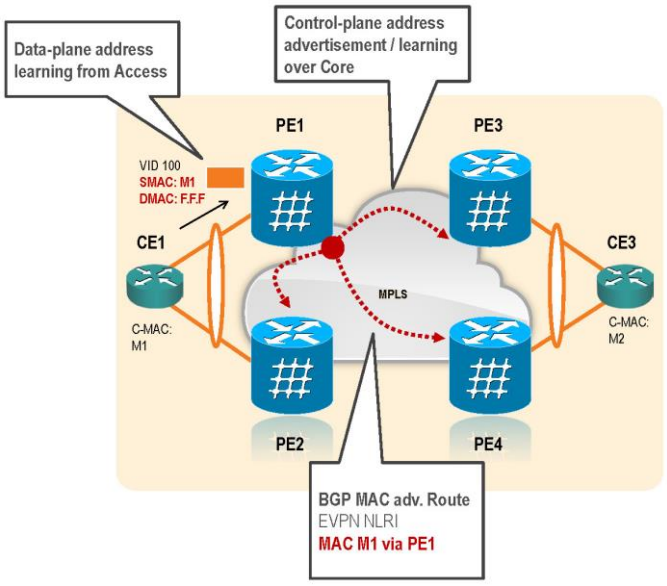
EVPN Operation

At startup, PEs exchange EVPN routes in order to advertise the following:

- **Ethernet segment reachability:** The PE discovers remote ESIs and their corresponding redundancy mode (all-active or single-active). In case of segment failures, PEs withdraw the routes used at this stage in order to trigger fast convergence by signaling a MAC mass withdrawal on remote PEs.
- **Redundancy Group membership:** PEs connected to the same Ethernet segment (multi-homing) automatically discover each other and elect a Designated Forwarder (DF) that is responsible for forwarding Broadcast, Unknown unicast and Multicast (BUM) traffic for a given EVI.
- **VPN membership:** The PE discovers all remote PE members of a given EVI. In the case of a multicast ingress replication model, this information is used to build the PE's flood list associated with an EVI.

After the initial EVPN route exchange, and when customer traffic arrives at the PE, EVPN MAC advertisement routes distribute reachability information over the core for each customer MAC address learned on local Ethernet segments (see Figure 4 below). Each EVPN MAC route announces the customer MAC address and the Ethernet segment associated with the port where the MAC was learned from, as well as a MPLS label. This EVPN MPLS label will be used later by remote PEs when sending traffic destined to the advertised MAC address.

Figure 4. The EVPN Operation



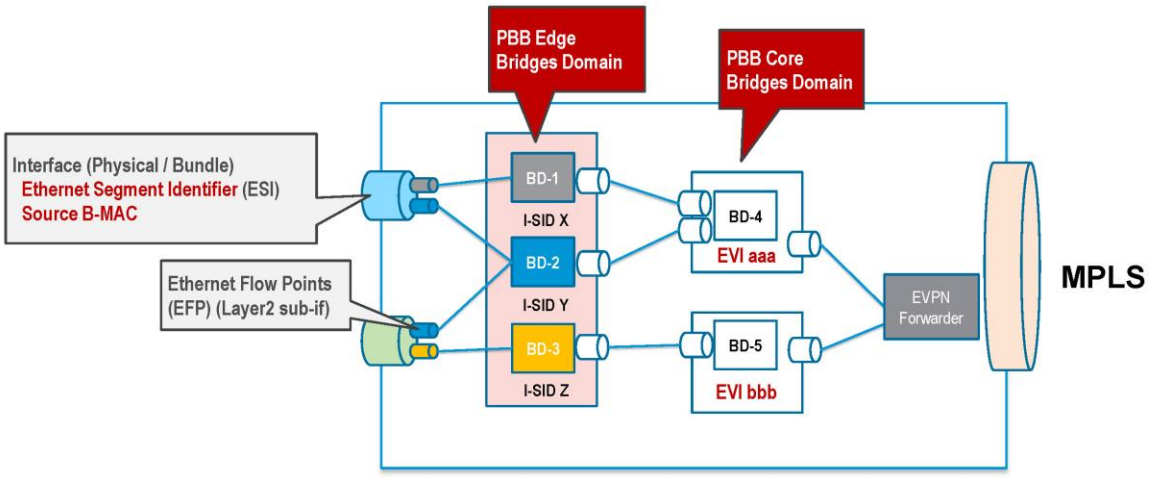
PBB-EVPN Operation

PBB (IEEE 802.1ah) defines a hierarchy of bridging devices (BEBs and BCBs[7]) that includes provisions to scale up the number of services that can be multiplexed into a single Backbone VLAN (B-VLAN) and to prevent backbone bridges from learning Customer MAC (C-MAC) addresses. This is accomplished with the addition of Ethernet headers that include a new set of source / destination Backbone MAC addresses (B-MAC) (referred to as MAC-in-MAC function) and a scalable 24-bit service instance identifier (I-SID[8]).

A PBB-EVPN PE combines the functions of a PBB BEB bridge and an EVPN PE, where PBB encapsulated traffic is mapped to MPLS LSPs using EVPN MPLS labels. On the access side, traffic from the UNI is mapped to a bridge domain associated with a PBB I-SID (referred to as “PBB Edge BD” in IOS-XR). One or more PBB Edge BDs are then associated with a “PBB Core BD” that is mapped to an EVI for traffic forwarding towards the MPLS core (see Figure 5 below).

[7] BEB = Backbone Edge Bridge / BCB = Backbone Core Bridge
[8] I-SID = Backbone Service Instance Identifier; 24-bit I-SID versus 12-bit VLAN ID (2^{24} = 16,777,216; 2^{12} = 4,096)

Figure 5. The PBB-EVPN Architecture

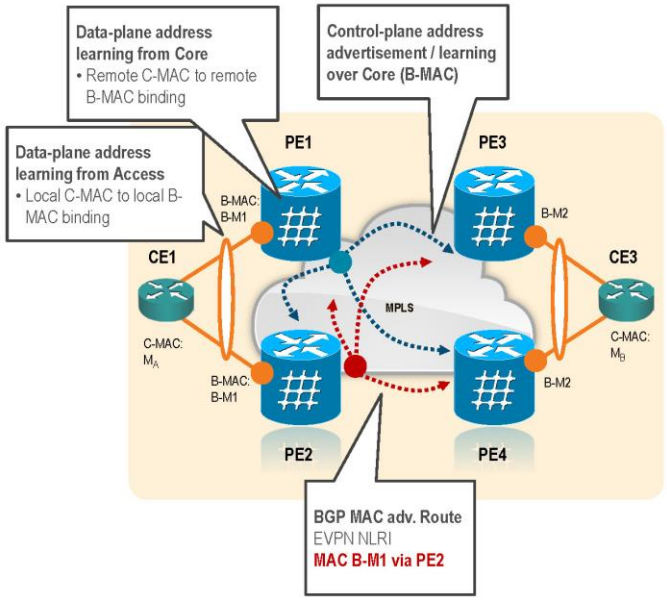


At startup, PBB-EVPN PEs exchange EVPN routes in order to advertise the following:

- **Redundancy group membership:** PEs connected to the same Ethernet segment (multi-homing) automatically discover each other and elect a DF that is responsible for forwarding BUM traffic for a given PBB I-SID.
- **VPN membership:** PE discovers all remote PE members of a given PBB I-SID. In the case of a multicast ingress replication model, this information is used to build the PE's flood list associated with a PBB I-SID.
- **Backbone MAC reachability:** Using EVPN MAC advertisement routes, a PE distributes reachability information for the B-MACs associated with local Ethernet segments. B-MACs are used to represent "sites" in PBB-EVPN.

After the initial EVPN route exchange, and when customer traffic arrives at the PE, source C-MACs are learned in the MAC table via data-plane learning, and PBB encapsulation is performed. Remote PEs across the core would then receive the PBB encapsulated traffic, and learn via data-plane the C-MACs, alongside the associated remote B-MACs, that must be used for traffic in the reverse direction (see Figure 6 below). At steady state, and with active conversations, the destination C-MAC of incoming traffic is inspected in the MAC table. If it is found, the PE would impose a PBB header, using the previously learnt destination B-MAC. Subsequently, the BGP routing table would indicate the remote PE, alongside the EVPN MPLS label associated with the destination B-MAC.

Figure 6. The PBB-EVPN Operation



As a result of its use of PBB, PBB-EVPN provides the following advantages over an EVPN solution:

- **Lower BGP control-plane scale requirements:** With the advertisement of B-MACs (instead of C-MACs) in BGP, PBB-EVPN dramatically reduces the amount of prefixes, and in turn the resources (CPU and memory) consumed by BGP. This benefits both PE and Route Reflector (RR) routers in all deployments, regardless of the C-MAC scale.
- **Simpler BGP control-plane operation and faster failure convergence:** PBB-EVPN provides a simpler control-plane operation in the core. For example, access link failures are handled with a single prefix withdrawal of the MAC route for the B-MAC assigned to the failed Ethernet Segment (regardless of the number of C-MACs learned on it). In EVPN, a failed link must be followed by the withdrawal of all the MAC routes for each C-MAC learned on the failed link.
- **Faster MAC move handling**: PBB-EVPN handles C-MAC moves (e.g. virtual machine moves) in the data-plane and it is completely transparent to BGP. PEs would simply re-learn in the data-plane the new B-MAC associated with the moved C-MAC. In contrast, EVPN handles MAC moves in the control-plane with re-advertisements of MAC routes. In scenarios where the MAC move is unintentional, e.g. as a consequence of an Ethernet loop, the looping traffic could translate in constant BGP updates.

EVPN and PBB-EVPN Applications

EVPN / PBB-EVPN standardization work at the IETF was started in order to address VPLS shortcomings for DCI applications, especially in the area of per-flow load balancing. These new technologies provide an extensible and flexible multi-homing VPN solution for intra-subnet connectivity among hosts / VMs over an MPLS / IP network.

EVPN / PBB-EVPN solutions are now also considered as the preferred choice for operators offering Ethernet LAN services to enterprise customers. This is especially true in greenfield deployments for providers who, for example, never deployed VPLS-based solutions.

For brownfield deployments that have existing VPLS networks, there are two solution categories to address a graceful introduction and migration to EVPN / PBB-EVPN:

- **Seamless Integration[9] :** A solution that allows co-existence of EVPN / PBB-EVPN and VPLS / PBB-VPLS technologies in the same E-LAN service. A capable PE is able to discern remote PE capabilities and establish proper protocol relationships in a mixed environment. This allows for the smooth insertion / upgrade of PEs in the network.
- **Interworking:** A solution where gateway nodes (PEs with dual capabilities) perform the interconnection of EVPN / PBB-EVPN and VPLS / PBB-VPLS networks. This allows for a smooth interconnection of new networks, without disrupting deployed ones.

For many service providers, Ethernet point-to-point (E-LINE) services are typically more prevalent than multi-point ones. A subset of the EVPN principles can be applied for E-LINE, and considering, no MAC leaning needs to be performed. Under standardization at the IETF, the EVPN VPWS draft[10] describes how EVPN can be used to support Virtual Private Wire Service (VPWS) in MPLS / IP networks. EVPN enables the following characteristics for VPWS: single-active as well as all-active multi-homing with flow-based load balancing, eliminates the need for single-segment and multi-segment PW signaling, and provides fast protection using data-plane prefix independent convergence, upon node or link failure. Operators therefore can consolidate point-to-point and multi-point services together under the same EVPN architecture.

Some operators have also expressed interest in deploying E-TREE services, a rooted-multipoint service type. In an E-TREE service, endpoints are labeled as either root or leaf sites. Root sites can communicate with all other sites. Leaf sites can communicate with root sites but not with other leaf sites. The EVPN E-TREE draft[11] describes how the functional requirements for E-TREE service can be met with EVPN and PBB-EVPN extensions, that include root / leaf indicators, alongside EVPN MAC advertisement routes, allowing PEs to perform respective filtering.

Lastly, there is significant work at the IETF on solutions that leverage EVPN control-plane for scenarios in which inter-subnet forwarding among hosts / VMs across different IP subnets is required, while maintaining the multi-homing capabilities of EVPN. This work is referred to as EVPN Integrated Routing and Bridging (IRB) draft[12].

Introducing PBB-EVPN in Cisco ASR 9000

Continuing its success and leadership in the market, the Cisco ASR 9000 series is the first router in the industry to support PBB-EVPN. This technology is currently available on the entire Cisco ASR 9000 series platform family, ranging from the smallest fixed model (ASR 9001-S) to the largest modular model (ASR 9922).

ASR 9000 routers running IOS-XR 4.3.2 / 5.1.1 or later releases with enhanced Ethernet line cards support PBB-EVPN.

[9] draft-sajassi-l2vpn-evpn-vpls-integration
[10] draft-boutros-l2vpn-evpn-vpws
[11] draft-sajassi-l2vpn-evpn-etree
[12] draft-sajassi-l2vpn-evpn-inter-subnet-forwarding

Figure 7. The Cisco ASR 9000 Series Platform Family

Conclusion

Cisco is leading the way in the standardization and implementation of next generation L2VPN solutions based on the Ethernet VPN (EVPN) solution family. EVPN is a solution for Ethernet multipoint services, with advanced multi-homing capabilities, using BGP for distributing MAC address reachability information over the MPLS network, while bringing the same operational and scale characteristics of IP VPNs to L2VPNs.

PBB-EVPN is a solution that combines the functionality of EVPN with Provider Backbone Bridging (PBB). PBB-EVPN provides significant advantages, including lower control-plane scale, simpler control-plane operation, and faster convergence, making it a superior solution for Data Center Interconnection (DCI) and E-LAN offerings.

Beyond DCI and E-LAN applications, the EVPN solution family provides a common foundation for all Ethernet service types; including E-LINE, E-TREE, as well as data center routing and bridging scenarios.

Lastly, the Cisco ASR 9000 raises the bar once again, by becoming the first router to deliver an implementation of PBB-EVPN to the market that is ready for live deployments.

For More Information

Read more about the Cisco ASR 9000 series, or contact your local account representative.
Read more about How to Implement PBB-EVPN.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Printed in USA C11-731864-00 05/14