BESS Workgroup A. Sajassi (Editor)
INTERNET-DRAFT S. Salam
Intended Status: Standard Track Cisco
N. Del Regno
Verizon
J. Rabadan
Nokia

Expires: October 15, 2018 April 15, 2018

(PBB-)EVPN Seamless Integration with (PBB-)VPLS
draft-ietf-bess-evpn-vpls-seamless-integ-03

Abstract

This draft specifies procedures for backward compatibility of the (PBB-)EVPN solution with (PBB-)VPLS and provides mechanisms for seamless integration of the two technologies in the same MPLS/IP network on a per-VPN-instance basis. Implementation of this draft enables service providers to introduce (PBB-)EVPN PEs in their brownfield deployments of (PBB-)VPLS networks.

Status of this Memo

Copyright and License Notice

Table of Contents

1 Introduction

VPLS and PBB-VPLS are widely-deployed L2VPN technologies. Many Service Providers (SPs) who are looking at adopting EVPN and PBB-EVPN want to preserve their investment in the (PBB-)VPLS networks. Hence, they require procedures by which (PBB-)EVPN technology can be introduced into their brownfield (PBB-)VPLS networks without requiring any upgrades (software or hardware) to these networks. This document specifies procedures for the seamless integration of the two technologies in the same MPLS/IP network.

```
                         VPLS PE
                          +---+
                          |PE1|
                          +---+
                            /
       EVPN/VPLS PE  +---------------+   EVPN/VPLS PE
            +---+    |               |   +---+
            |PE4|----|    MPLS/IP    |---|PE5|
            +---+    |     Core      |   +---+
                     |               |
                     +---------------+
                       /        \
                    +---+     +---+
                    |PE2|     |PE3|
                    +---+     +---+
                  VPLS PE     VPLS PE
```

Figure 1: Seamless Integration of (PBB-)EVPN PEs & (PBB-)VPLS

Section 2 provides the details of the requirements. Section 3 specifies procedures for the seamless integration of VPLS and EVPN networks. Section 4 specifies procedures for the seamless integration of PBB-VPLS and PBB-EVPN networks. Section 5 discusses the solution advantages.

It should be noted that the scenarios for PBB-VPLS integration with EVPN and VPLS integration with PBB-EVPN are not covered in this document because there haven't been any requirements from service providers for these scenarios. The reason for that is that deployments which employ PBB-VPLS typically require PBB encapsulation for various reasons. Hence, it is expected that for those deployments the evolution path would be from PBB-VPLS towards PBB-EVPN. Furthermore, the evolution path from VPLS is expected to be towards EVPN.

1.1. Specification of Requirements

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

1.2. Terms and Abbreviations

B-MAC: Backbone MAC

B-VID: Backbone VLAN ID

Broadcast Domain: In a bridged network, the broadcast domain corresponds to a Virtual LAN (VLAN), where a VLAN is typically represented by a single VLAN ID (VID) but can be represented by several VIDs where Shared VLAN Learning (SVL) is used per [IEEE.802.1ah].

Bridge Table: An instantiation of a broadcast domain on a MAC-VRF

RIB: Routing Information Base - An instantiation of a routing table on a MAC-VRF

FIB: Forwarding Information Base - An instantiation of a forwarding table on a MAC-VRF

CE: A Customer Edge device, e.g., a host, router, or switch.

EVI: An EVPN Instance spanning the Provider Edge (PE) devices participating in that EVPN.

MAC-VRF: A Virtual Routing and Forwarding table for Media Access Control (MAC) addresses on an EVPN PE.

MAC address: Media Access Control address

ES: When a customer site (device or network) is connected to one or more PEs via a set of Ethernet links, then that set of links is referred to as an "Ethernet Segment".

ESI: An Ethernet Segment Identifier is a unique non-zero identifier that identifies an ES

Ethernet Tag: An Ethernet Tag identifies a particular broadcast domain, e.g., a VLAN. An EVPN instance consists of one or more broadcast domains

MHD: Multi-Homed Device

MHN: Multi-Homed Network

P2MP: Point-to-Multipoint

PBB: Provider Backbone Bridge

PE: Provider Edge device

VSI: Virtual Switch Instance

VPLS: Virtual Private LAN Service

Single-Active Redundancy Mode: When only a single PE, among all the PEs attached to an Ethernet segment, is allowed to forward traffic to/from that Ethernet segment for a given VLAN, then the Ethernet segment is defined to be operating in Single-Active redundancy mode.

All-Active Redundancy Mode: When all PEs attached to an Ethernet segment are allowed to forward known unicast traffic to/from that Ethernet segment for a given VLAN, then the Ethernet segment is defined to be operating in All-Active redundancy mode.

(PBB-)EVPN: refers to both, PBB-EVPN and EVPN. This document uses this abbreviation when a given description applies to both technologies.

(PBB-)VPLS: refers to both, PBB-VPLS and VPLS. As for EVPN, this abbreviation is used when the text applies to both technologies.

VPLS A-D: refers to Virtual Private LAN Services with BGP-based Auto Discovery as in [RFC6074].

PW: Pseudowire.

2. Requirements

Following are the key requirements for backward compatibility between (PBB-)EVPN and (PBB-)VPLS:

1. The solution MUST allow for staged migration towards (PBB-)EVPN on a site-by-site basis per VPN instance - e.g., new EVPN sites to be provisioned on (PBB-)EVPN PEs.

2. The solution MUST require no changes to existing VPLS or PBB-VPLS PEs, not even a software upgrade.

3. The solution MUST allow for the coexistence of PEs running (PBB-)EVPN and (PBB-)VPLS for the same VPN instance and single-homed segments.

4. The solution MUST support single-active redundancy of multi-homed networks and multi-homed devices for (PBB-)EVPN PEs.

5. In case of single-active redundancy, the participant VPN instances MAY span across both (PBB-)EVPN PEs and (PBB-)VPLS PEs as long as the MHD or MHN is connected to (PBB-)EVPN PEs. In case of an ES link failure, the (PBB-)EVPN PEs will send a BGP mass-withdraw to the EVPN peers OR MAC advertisement with MAC Mobility extended community for PBB-EVPN AND follow existing VPLS MAC Flush procedures with the VPLS peers.

6. The support of All-Active redundancy mode across both (PBB-)EVPN PEs and (PBB-)VPLS PEs is outside the scope of this document.

These requirements collectively allow for the seamless insertion of the (PBB-)EVPN technology into brown-field (PBB-)VPLS deployments.

3 VPLS Integration with EVPN

In order to support seamless integration with VPLS PEs, this document requires that VPLS PEs support VPLS A-D per [RFC6074] and EVPN PEs support both BGP EVPN routes per [RFC7432] and VPLS A-D per [RFC6074]. All the logic for this seamless integration SHALL reside on the EVPN PEs. If a VPLS instance is setup without the use of VPLS A-D, it is still possible (but cumbersome) for EVPN PEs to integrate into that VPLS instance by manually configuring PWs to all the VPLS PEs in that instance (i.e., the integration is no longer seamless).

3.1 Capability Discovery

The EVPN PEs MUST advertise both the BGP VPLS A-D route as well as the BGP EVPN Inclusive Multicast Ethernet Tag (IMET) route for a given VPN instance. The VPLS PEs only advertise the BGP VPLS A-D route, per current standard procedures specified in [RFC4761], [RFC4762] and [RFC6074]. The operator may decide to use the same Route Target (RT) to identify a VPN on both EVPN and VPLS networks. In this case, when a VPLS PE receives the EVPN IMET route, it MUST ignore it on the basis that it belongs to an unknown SAFI. However, the operator may choose to use two RTs - one to identify the VPN on VPLS network and another for EVPN network and employ RT-constrained [RFC4684] in order to prevent BGP EVPN routes from reaching the VPLS PEs.

When a EVPN PE receives both a VPLS A-D route as well as an EVPN IMET route from a given remote PE for the same VPN instance, it MUST give preference to the EVPN route for the purpose of discovery. This ensures that, at the end of the route exchanges, all EVPN capable PEs discover other EVPN capable PEs in addition to the VPLS-only PEs for that VPN instance. Furthermore, all the VPLS-only PEs would discover the EVPN PEs as if they were standard VPLS PEs. In other words, when the discovery phase is complete, the EVPN PEs would have discovered all the PEs in the VPN instance along with their associated capability: EVPN or VPLS-only. Whereas the VPLS PEs would have discovered all the PEs in the VPN instance as if they were all VPLS-only PEs.

3.2 Forwarding Setup and Unicast Operation

The procedures for forwarding state setup and unicast operation on the VPLS PE are per [RFC8077], [RFC4761], [RFC4762].

The procedures for forwarding state setup and unicast operation on the EVPN PE are as follows:

- The EVPN PE MUST establish a PW to a remote PE from which it has received only a VPLS A-D route for the corresponding VPN instance, and MUST set up the label stack corresponding to the PW FEC. For seamless integration between EVPN and VPLS PEs, the PW that is setup between a pair of VPLS and EVPN PEs is between the VSI of the VPLS PE and the MAC-VRF of the EVPN PE.

- The EVPN PE must set up the label stack corresponding to the MP2P VPN unicast FEC to any remote PE that has advertised EVPN IMET route.

- If a EVPN PE receives a VPLS A-D route followed by an EVPN IMET route from the same PE and a PW is already setup to that PE, then the EVPN MUST bring that PW operationally down.

- If a EVPN PE receives an EVPN IMET route followed by a VPLS A-D route from the same PE, then the EVPN PE will setup the PW but MUST keep it operationally down.

- In case VPLS AD is not used in some VPLS PEs, the EVPN PEs need to be provisioned manually with PWs to those remote VPLS PEs for each VPN instance. In that case, if a EVPN PE receives an EVPN IMET route from a PE to which a PW exists, the PW will be brought operationally down.

When the EVPN PE receives traffic over the VPLS PWs, it learns the associated C-MAC addresses in the data-plane. The C-MAC addresses

learned over these PWs MUST be injected into the bridge table of the associated MAC-VRF on that EVPN PE. The learned C-MAC addresses MAY also be injected into the RIB/FIB tables of the associated MAC-VRF on that EVPN PE. For seamless integration between EVPN and VPLS PEs, since these PWs belong to the same split-horizon group as the MP2P EVPN service tunnels, then the C-MAC addresses learned and associated to the PWs will NOT be advertised in the control plane to any remote EVPN PEs. This is because every EVPN PE can send and receive traffic directly to/from every VPLS PE belonging to the same VPN instance.

The C-MAC addresses learned over local Attachment Circuits (ACs) by an EVPN PE are learned in data-plane. For EVPN PEs, these C-MAC addresses MUST be injected into the corresponding MAC-VRF and advertised in the control-plane using BGP EVPN routes. Furthermore, the C-MAC addresses learned in the control plane via the BGP EVPN routes sent by remote EVPN PEs, are injected into the corresponding MAC-VRF table.

3.3 MAC Mobility

In EVPN, host addresses (C-MAC addresses) can move around among EVPN PEs or even between EVPN and VPLS PEs.

When a C-MAC address moves from an EVPN PE to a VPLS PE, then as soon as BUM traffic is initiated from that MAC address, it is flooded to all other PEs (both VPLS and EVPN PEs) and the receiving PEs update their MAC tables (VSI or MAC-VRF). The EVPN PEs do not advertise the C-MAC address learned over PW to each other because every EVPN PE learns it directly over its associated PW to that VPLS PE. If only known-unicast traffic is initiated from the moved C-MAC address toward a known C-MAC, then this can result in black-holing of traffic destined to the C-MAC that has moved until there is a BUM traffic originated with the moved C-MAC address as the source MAC address (e.g., as a result of MAC age-out timer expires) but this is the typical behavior of VPLS PEs.

When a C-MAC address moves from a VPLS PE to an EVPN PE, then as soon as BUM or known-unicast traffic is initiated from that C-MAC address, the C-MAC is learned and advertised in BGP to other EVPN PEs and MAC mobility procedure is exercised among EVPN PEs. For BUM traffic, both EVPN and VPLS PEs learn the new location of the moved C-MAC address; however, if there is only known-unicast traffic, then only EVPN PEs learn the new location of the C-MAC that has moved but not VPLS PEs. This can result in black-holing of traffic sent from VPLS PEs destined to the C-MAC that has moved until there is a BUM traffic originated with the moved C-MAC address as the source MAC address (e.g., as a result of MAC age-out timer expires) but this is the typical behavior of VPLS PEs.

3.4 Multicast Operation

3.4.1 Ingress Replication

The procedures for multicast operation on the VPLS PE, using ingress replication, are per [RFC4761], [RFC4762], and [RFC7080].

The procedures for multicast operation on the EVPN PE, for ingress replication, are as follows:

- The EVPN PE builds a replication sub-list to all the remote EVPN PEs per EVPN instance as the result of the exchange of the EVPN IMET routes per [RFC7432]. This will be referred to as sub-list A. It comprises MP2P service tunnels used for delivering EVPN BUM traffic [RFC7432].

- The EVPN PE builds a replication sub-list per VPLS instance to all the remote VPLS PEs. This will be referred to as sub-list B. It comprises PWs from the EVPN PE in question to all the remote VPLS PEs in the same VPLS instance.

The replication list, maintained per VPN instance, on a given EVPN PE will be the union of sub-list A and sub-list B. Note that the PE must enable split-horizon over all the entries in the replication list, across both PWs and MP2P service tunnels.

3.4.2 P2MP Tunnel

The procedures for multicast operation on the EVPN PEs using P2MP tunnels are outside of the scope of this document.

4 PBB-VPLS Integration with PBB-EVPN

In order to support seamless integration between PBB-VPLS and PBB-EVPN PEs, this document requires that PBB-VPLS PEs support VPLS A-D per [RFC6074] and PBB-EVPN PEs support both BGP EVPN routes per [RFC7432] and VPLS A-D per [RFC6074]. All the logic for this seamless integration SHALL reside on the PBB-EVPN PEs.

4.1 Capability Discovery

The procedures for capability discovery are per Section 3.1 above.

4.2 Forwarding Setup and Unicast Operation

The procedures for forwarding state setup and unicast operation on

the PBB-VPLS PE are per [RFC8077] and [RFC7080].

The procedures for forwarding state setup and unicast operation on the PBB-EVPN PE are similar to that of section 3.2 except for the following:

- For seamless integration between EVPN and VPLS PEs, the PW that is setup between a pair of PBB-VPLS and PBB-EVPN PEs, is between B-components of PBB-EVPN PE and PBB-VPLS PE per section 4 of [RFC7041].

- When the PBB-EVPN PE receives traffic over the PBB-VPLS PWs, it learns the associated B-MAC addresses in the data-plane. The B-MAC addresses learned over these PWs MUST be injected into the bridge table of the associated MAC-VRF on that PBB-EVPN PE. The learned B-MAC addresses MAY also be injected into the RIB/FIB tables of the associated the MAC-VRF on that BPP-EVPN PE. For seamless integration between PBB-EVPN and PBB-VPLS PEs, since these PWs belongs to the same split-horizon group as the MP2P EVPN service tunnels, then the B-MAC addresses learned and associated to the PWs will NOT be advertised in the control plane to any remote PBB-EVPN PEs. This is because every PBB-EVPN PE can send and receive traffic directly to/from every PBB-VPLS PE belonging to the same VPN instance.

- The C-MAC addresses learned over local Attachment Circuits (ACs) by an PBB-EVPN PE are learned in data-plane. For PBB-EVPN PEs, these C-MAC addresses are learned in I-component of PBB-EVPN PEs and they are not advertised in the control-plane per [RFC7623].

- The B-MAC addresses learned in the control plane via the BGP EVPN routes sent by remote PBB-EVPN PEs, are injected into the corresponding MAC-VRF table.

4.3 MAC Mobility

In PBB-EVPN, a given B-MAC address can be learnt either over the BGP control-plane from a remote PBB-EVPN PE, or in the data-plane over a PW from a remote PBB-VPLS PE. There is no mobility associated with B-MAC addresses in this context. Hence, when the same B-MAC address shows up behind both a remote PBB-VPLS PE as well as a PBB-EVPN PE, the local PE can deduce that it is an anomaly and notify the operator.

4.4 Multicast Operation

4.4.1 Ingress Replication

The procedures for multicast operation on the PBB-VPLS PE, using ingress replication, are per [RFC7041] and [RFC7080].

The procedures for multicast operation on the PBB-EVPN PE, for ingress replication, are as follows:

- The PBB-EVPN PE builds a replication sub-list per I-SID to all the remote PBB-EVPN PEs in a given VPN instance as a result of the exchange of the EVPN IMET routes, as described in [RFC7623]. This will be referred to as sub-list A. It comprises MP2P service tunnels used for delivering PBB-EVPN BUM traffic.

- The PBB-EVPN PE builds a replication sub-list per VPN instance to all the remote PBB-VPLS PEs. This will be referred to as sub-list B. It comprises PWs from the PBB-EVPN PE in question to all the remote PBB-VPLS PEs in the same VPN instance.

- The PBB-EVPN PE may further prune sub-list B, on a per I-SID basis, if [MMRP] is run over the PBB-VPLS network. This will be referred to as sub-list C. This list comprises a pruned set of the PWs in the sub-list B.

The replication list maintained per I-SID on a given PBB-EVPN PE will be the union of sub-list A and sub-list B if [MMRP] is NOT used, and the union of sub-list A and sub-list C if [MMRP] is used. Note that the PE must enable split-horizon over all the entries in the replication list, across both pseudowires and MP2P service tunnels.

4.4.2 P2MP Tunnel - Inclusive Tree

The procedures for multicast operation on the PBB-EVPN PEs using P2MP tunnels are outside of the scope of this document.

5 Solution Advantages

The solution for seamless integration of (PBB-)EVPN with (PBB-)VPLS has the following advantages:

- When ingress replication is used for multi-destination traffic delivery, the solution reduces the scope of [MMRP] (which is a soft-state protocol) to only that of existing VPLS PEs, and uses the more robust BGP-based mechanism for multicast pruning among new EVPN PEs.

- It is completely backward compatible.

- New PEs can leverage the extensive multi-homing mechanisms and provisioning simplifications of PBB-EVPN:

a. Auto-sensing of MHN / MHD
b. Auto-discovery of redundancy group
c. Auto-provisioning in DF election and VLAN carving

6 Security Considerations

No new security considerations beyond those for VPLS and EVPN.

7 IANA Considerations

This document has no actions for IANA.

8 References

8.1 Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March <https://www.rfc-editor.org/info/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8077] Martini, et al., "Pseudowire Setup and Maintenance using the Label Distribution Protocol", RFC 8077, February 2017.

[RFC7432] Sajassi et al., "BGP MPLS Based Ethernet VPN", RFC 7432, February, 2015.

[RFC7623] Sajassi et al., "Provider Backbone Bridging Combined with Ethernet VPN (PBB-EVPN)", RFC 7623, September, 2015.

[RFC4761] Kompella, K., Ed., and Y. Rekhter, Ed., "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC 4761, January 2007, <http://www.rfc-editor.org/info/rfc4761>.

[RFC4762] Lasserre, M., Ed., and V. Kompella, Ed., "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", RFC 4762, January 2007, <http://www.rfc-editor.org/info/rfc4762>.

[RFC6074] Rosen et al., "Provisioning, Auto-Discovery, and Signaling

in Layer 2 Virtual Private Networks (L2VPNs)", RFC 6074,
January 2011.

8.2 Informative References

[MMRP] Clause 10 of "IEEE Standard for Local and metropolitan area networks - Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks", IEEE Std 802.1Q, 2013.

[RFC7041] Balus et al., "Extensions to VPLS PE model for Provider Backbone Bridging", RFC 7041, November 2013.

[RFC7080] Sajassi et al., "VPLS Interoperability with Provider Backbone Bridges", RFC 7080, December, 2013.

[IEEE.802.1ah] IEEE, "IEEE Standard for Local and metropolitan area networks - Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks", Clauses 25 and 26, IEEE Std 802.1Q, DOI 10.1109/IEEESTD.2011.6009146.

[RFC4684] Marques et al., "Constrained Route Distribution for Border Gateway Protocol/MultiProtocol Label Switching (BGP/MPLS) Internet Protocol (IP) Virtual Private Networks (VPNs)", RFC 4684, November, 2006.

Authors' Addresses

Ali Sajassi
Cisco
Email: sajassi@cisco.com

Samer Salam
Cisco
Email: ssalam@cisco.com

Nick Del Regno
Verizon
Email: nick.delregno@verizon.com

Jorge Rabadan

Nokia
Email: jorge.rabadan@nokia.com