BESS
Internet-Draft
Intended status: Standards Track
Expires: May 4, 2020

W. Lin
Juniper Networks, Inc.
B. Wen
V. Kozak
Comcast
J. Rabadan
Nokia

November 1, 2019

EVPN and BGP-based L2VPN Seamless Integration
draft-evpn-bgp-based-l2vpn-seamless-integration-00

Abstract

This document presents a seamless integration solution for BGP-based Layer-2 VPN (L2VPN) and EVPN to provide point-to-point Virtual Private Wire Service (VPWS). In addition, this document also extends the existing seamless integration for multipoint Ethernet VPN service with all-active multihoming support. The specified solution allows the coexistence of EVPN and L2VPN services under the same point-to-point or multipoint VPN instance. By using this seamless integration solution, a service provider can introduce EVPN into their existing L2VPN network or migrate from an existing L2VPN based network to EVPN.

Requirements Language

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.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 4, 2020.

Copyright Notice

Table of Contents

1. Introduction

[RFC6624] specifies a point-to-point L2VPN solution by using BGP auto-discovery and signaling. This BGP-based L2VPN service may offer point-to-point service using different types of L2 encapsulation, such as Ethernet, frame relay, etc., and with single home or active-standby redundancy.

EVPN VPWS leverages the latest EVPN technology and brings extra functions to Layer 2 point-to-point Ethernet service, such as all-active redundancy, load balancing and mass withdrawal. All-active redundancy also makes it easier to achieve fast convergence on an access link or node failure.

When expanding an existing L2VPN network with Ethernet encapsulation, service provider may want to deploy EVPN VPWS to provide additional Layer 2 point to point Ethernet services, and at the same time some of the customer traffic may still need to be terminated on the existing L2VPN PEs within the service provider network.

This document describes a seamless-integration solution that allows the co-existence of point-to-point Ethernet services using BGP-based L2VPN procedure per [RFC6624] or EVPN VPWS procedure per [RFC8214] under the same VPN network and over the same MPLS/IP network. Service providers may also use the seamless integration solution for migration a traditional L2VPN network to EVPN VPWS based network.

For the multipoint Ethernet VPN service, [RFC8560] specifies a seamless integration solution for VPLS and EVPN with single home and

single-active redundancy support. This document extends the seamless integration solution defined in [RFC8560] with all-active multihoming support for PEs that can support both VPLS per [RFC4761] and EVPN procedures. In the extended solution, VPLS [RFC4761] procedure is used to establish PWs to the rest of VPLS PEs in the same VPN network. Support for using VPLS [RFC4762] procedure to set up PWs to the rest of VPLS PEs is outside the scope of this document.

In this document, section 5 and 6 describe the requirements and operation model for the seamless integration solution for point-to-point Ethernet VPN. Section 6 covers the solution and procedure in more detail.

The extended seamless integration solution for multipoint Ethernet VPN is covered in Section 7.

2. Terminology

AC: Attachment Circuit. In EVPN VPWS, an attachment circuit for EVPN is also referred to as an Ethernet Segment (ES).

L2: Layer 2

VPWS: Virtual Private Wire Service

Point to point: P2P

P2P Ethernet Service: a point-to-point L2 service where the hand-off between a Provide Edge (PE) and a Customer Edge (CE) is based on L2 Ethernet. In this document a P2P Ethernet service is established based on control plane procedure specified in this document or EVPN VPWS [RFC8214] or BGP based L2VPN [RFC6624]. Forwarding is based on using an MPLS label as the demultiplexer.

L2VPN PE: a PE supports L2VPN services based on the procedures specified in [RFC6624]

EVPN VPWS PE: a PE supports EVPN VPWS based on the procedures specified in [RFC8214]. In this document an EVPN VPWS PE may also be referred to as an EVPN PE for short. An EVPN PE may or may not support seamless integration solution specified in this document.

BGP VPLS PE: a PE supports VPLS procedure and multipoint Ethernet VPN service defined in [RFC4761].

Composite PE: In the context of a point-to-point Ethernet VPN, a composite PE is a PE that can provide seamless integration solution specified in this document based on both L2VPN procedure per [RFC6624] and EVPN VPWS procedure per [RFC8214] under the same VPN instance. In the context of a multipoint Ethernet VPN, a composite PE is a PE that can provide seamless integration solution based on [RFC8560] as well as the extended procedure specified in this document under section 7.

L2VPN Route: a BGP NLRI used for auto-discovery and signaling for L2VPN per [RFC6624]. [RFC6624], in turns, uses BGP VPLS NRLI defined in [RFC4761] for L2VPN. Through out this document, the terms "L2VPN A-D route" and "L2VN route" are used exchangeable.

BGP-VPLS route: a BGP NLRI used for auto-discovery and signaling for BGP-based VPLS per [RFC4761].

EVPN E-AD per EVI Route: an EVPN Ethernet A-D per EVI route used for auto-discovery and signaling for EVPN VPWS per [RFC8214].

This document does not distinguish between "all-active" and "active-active" and they are used interchangeably. The same applies to "single-active" and "active-standby".

This document also uses the terms "P2P Ethernet service" and "VPWS" interchangeably. For simplicity, this document may refer to a P2P Ethernet service as a P2P service for short.

This document also makes frequent use of the terminologies specified in [RFC4761], [RFC6624], [RFC7432] and [RFC8214]

3. L2VPN PE, EVPN PE and Composite PE

There are three types of PEs defined in this seamless integration solution: L2VPN PE, EVPN PE and composite PE. Under a given Layer 2 Ethernet VPN, the type of PE is categorized by the technology it is provisioned for. For instance, a PE that is provisioned to use L2VPN and EVPN on the same VPN service is considered a composite PE. A L2VPN PE that provides BGP-VPLS service per [RFC4761] is also

referred to as BGP-VPLS or VPLS PE for short.

Also in this document in the context of a given Layer 2 Ethernet VPN, an EVPN PE is a PE that is provisioned to provide only the EVPN solution per [RFC8214], or [RFC7432] or both, but not seamless integration solution. It is irrelevant whether an EVPN PE is capable to support seamless integration solution.

For example, for a non-L2VPN PE, a network administrator may know a-priori that the PE does not need to establish any P2P Ethernet service that involves L2VPN PE under a given Layer 2 Ethernet VPN instance. In this case, the PE can be provisioned to act only as an EVPN PE for that VPN even though it is capable of providing seamless integration procedure. If such a prior knowledge is unavailable, then a PE SHALL be provisioned to act as a composite PE if it is capable of. Otherwise, it is unable to establish a P2P Ethernet service with a L2VPN PE.

The term "homogeneous PEs" refers to PEs that are of the same types. Unless explicitly specified in this specification, a PE's type applies to a given Layer 2 Ethernet VPN instance. A PE may act as an EVPN PE for one VPN, but as a composite PE for another VPN.

4. Requirements

The seamless integration solution for point-to-point Ethernet VPN meets the following requirements:

- It must allow L2VPN, EVPN and composite PEs to participate in the same Layer 2 Ethernet VPN instance.

- The composite PE, the PE that supports the seamless integration solution, must be backward compatible to support both EVPN VPWS and L2VPN when Ethernet is used as the hand-off between the PE and CE. The composite PE must support the establishment of a layer 2 P2P Ethernet service with a L2VPN PE or an EVPN PE.

- No change should be required for any exiting L2VPN PEs beyond what

are already specified in [RFC6624].

- The seamless integration solution must support a CE single homed to PEs of different types: L2VPN, EVPN and composite PEs.

- The seamless solution must support active-standby, also known as single-active, redundancy for L2VPN PEs or EVPN PEs or composite PEs, as long as PEs connecting to the same multihomed CE are of the same type.

- Composite PEs provisioned for all-active multihoming for their multithemed CE(s) MUST work with L2VPN PE(s) working in single home or active-standby multihoming.

- The solution SHALL support control word forwarding procedure defined in [RFC4448].

- The solution SHALL support staged migration to EVPN VPWS network when all L2VPN PEs are upgraded to support EVPN VPWS.

The requirements for the seamless integration solution for multipoint Ethernet VPN are specified in [RFC8560] and they are also reiterated in section 7.

5. Model of Operation for Seamless Integration of Point-to-point Ethernet VPN

5.1. Point-to-point Ethernet VPN

In the seamless integration solution described in this document, PEs participating in a VPN offer point-to-point Layer 2 connections between different customer sites, and Ethernet is used as the Layer 2 hand-off between a PE and a CE. Under the seamless integration

solution, two different techniques can be used to establish P2P Ethernet services under the same VPN: some P2P Ethernet services may use the technique specified per [RFC6624], while others may use the technique specified per [RFC8214]. [RFC6624] uses the terminology of "Layer 2 VPN (L2VPN)". [RFC8214] uses the terminology of "Ethernet VPN (EVPN)". In this document, we refer to a VPN that is capable of offering Layer 2 Ethernet services by using both L2VPN and EVPN VPWS technologies as a point-to-point Ethernet VPN.

5.2. Operation Model for Seamless Integration

A PE participating in a point-to-point Ethernet VPN offers P2P Ethernet services with different remote PEs. By nature of point-to-point service, there is no requirement for full mesh among all the PEs participating in the same point-to-point Ethernet VPN instance.

The seamless integration solution allows the coexistence of composite PE, L2VPN PE and EVPN PE under the same VPN instance. It allows the establishment of P2P Ethernet services over the same MPLS/IP core: (a) between two homogenous PEs, or (b) between a composite PE and a L2VPN PE, or (c) between a composite PE and a EVPN PE.

A composite PE can establish a P2P Ethernet service with a L2VPN PE and different a P2P service with an EVPN PE. It is the sole responsibility of a composite PE to seamless integrate with L2VPN PEs and EVPN PEs.

There will be no P2P service between an EVPN PE and a L2VPN PE in the same L2 Ethernet VPN as an EVPN PE is provisioned only to provide the procedure/function per EVPN VPWS.

5.3. Seamless Integration for Single Homing or Multihoming

L2VPN offers single home as well as active-standby multihoming support, but not active-active multihoming support. Under the seamless integration solution, a composite PE can integrate with L2VPN PE(s) working in:

Case 1: single home

Case 2: active-standby multihoming with its peer L2VPN PE(s)

A composite PE supports seamless integration with EVPN PE(s) working in:

Case 1: single home

Case 2: single-active multihoming with its peer EVPN PE(s)

Case 3: all-active multihoming with its peer EVPN PE(s)

While providing seamless integration solution, a composite PE may provide single home support as well as single-active or all-active multihoming support support to its locally attached CE.

For single-active multihoming, there are two options that a multihomed CE may connect to a redundant set of composite PEs:

1. Through a LAG interface while the composite PEs working in a port-active for single-active multihoming, and the DF or non-DF role on the composite PE is elected on a per port basis.

2. Through a separate interface to each composite PE working in single-active multihoming, and the DF or non-DF role on the composite PE is elected on a per access interface basis.

5.4. Control Plane Overview

In the seamless integration solution, a L2VPN PE continues to use the standard procedure per [RFC6624] without any change or additional new procedure. An EVPN PE also continues to use procedure per [RFC8214] without any change or additional new procedure.

A composite PE follows the seamless integration procedure defined in this document.

A composite PE uses EVPN VPWS procedure per [RFC8214] to establish a P2P Ethernet service with an EVPN PE.

5.5. Data Plane Overview

Regardless of the type of a PE, data traffic continues to be carried over a MPLS/IP tunnel from an ingress PE to an egress PE. At the egress PE, an MPLS label is used as the demultiplexer to identify the attachment circuit for a P2P Ethernet service.

6. Seamless Integration Solution for Point-to-point Ethernet VPN

It is the sole responsibility of a composite PE to provide seamless integration solution with a L2VPN PE. So the focus of the solution is the composite PE. This section and its sub-sections follow specify the solution and procedures a composite PE provides.

6.1. Local ID and Remote ID

Similar to other PEs, a composite PE is provisioned for the VPN it participates through Route Target(s) and a Route Distinguisher (RD). For each P2P Ethernet service, the PE involved is provisioned with a pair of local and remote IDs. The local ID identifies an local attachment circuit associated with a P2P service, while the remote ID identifies an attachment circuit attached to a remote PE.

For a given P2P Ethernet service, a local ID for a PE is the remote ID for its corresponding remote PE. It is required that that both PEs involved in a P2P Ethernet service must have a matching pair of local/remote IDs correspondingly. In the BGP signaling procedure for auto-discovery, only local ID is signaled in the control plane, but not remote ID.

In L2VPN, the ID used to identify an attachment circuit associated with a P2P service is referred to as a VE ID or site ID which is a 16-bit integer. A valid VE-ID for L2VPN is in the range of 1 to 0xFFFE.

In EVPN VPWS, the ID used to identify an attachment associated with a P2P service is referred as an EVPN VPWS service instance identifier which is a 24-bit integer. A valid service instance identifier for EVPN VPWS is in the range of 1 to 0xFFFFFF.

A p2p Ethernet service using L2VPN procedure MUST keep its local/ remote ID within the range of 0x1 to 0xFFFE.

6.2. Composite PE Control Plane Procedure

This section and the sub-sections under it cover the control plane procedure of a composite PE to interact with other types of PEs.

6.2.1. Auto-Discovery

All three types of PEs defined in this document continue to use MP-BGP for auto-discovery. An auto-discovery procedure involves two parts: A PE needs to identify itself for other PEs to discover it, and a PE needs to auto discover other PEs. Auto-discovery is only meaningful to PEs participating in the same VPN.

A composite PE needs to identify itself and discover other PE(s) participating in the same point-to-point Ethernet VPN. If a composite PE does not know a-priori the type of remote PE for a given P2P Ethernet service it tries to establish, a composite PE MUST

participate in both L2VPN and EVPN auto-discovery procedures per

[RFC6624] and [RFC8214] except in the cases specified in section 6.2.5.

Similar to a L2VPN or EVPN PE, a composite PE uses Route Target community to identify itself as a part of a point-to-point Ethernet VPN instance. A composite PE announces itself through both BGP L2VPN A-D route and EVPN E-AD per EVI route, and with the RT(s) belong to the VPN it participates. A network operator may choose to use different RT(s) to identify L2VPN PEs and EVPN PEs participating in the same VPN. In this case, A composite PE needs to be provisioned with RTs used by L2VPN PEs and EVPN PEs.

A composite PE discovers other L2VPN PEs by processing L2VPN A-D routes that have route target(s) matching its import RT(s). At the same time, a composite PE discovers other EVPN or composite PEs by processing EVPN E-AD per EVI routes that have the RT(s) matching its import RT(s).

At the end of discovery procedure, a L2VPN PE discovers all L2VPN PEs and all composite PEs participating in the same VPN. However a L2VPN cannot distinguish a L2VPN from a composite PE. From a point of L2VPN PE, all composite PEs are L2VPN PEs.

Also at the end of discovery procedure, an EVPN PE discovers all EVPN PEs and all composite PEs participating in the same VPN. Similarly, an EVPN PE cannot distinguish an EVPN PE from a composite PE. From a point of EVPN PE, all composite PEs are EVPN PEs.

6.2.2. Control Plane Signaling

In the seamless integration solution, a composite PE relies on MP-BGP signaling to exchange information for each of its P2P Ethernet service. A composite PE uses the procedures defined in [RFC6624] and [RFC8214] for control plane signaling, and by default it originates both a L2VPN route and an EVPN E-AD per EVI route for each of its P2P Ethernet service. Note that these are the same routes used for auto-discovery.

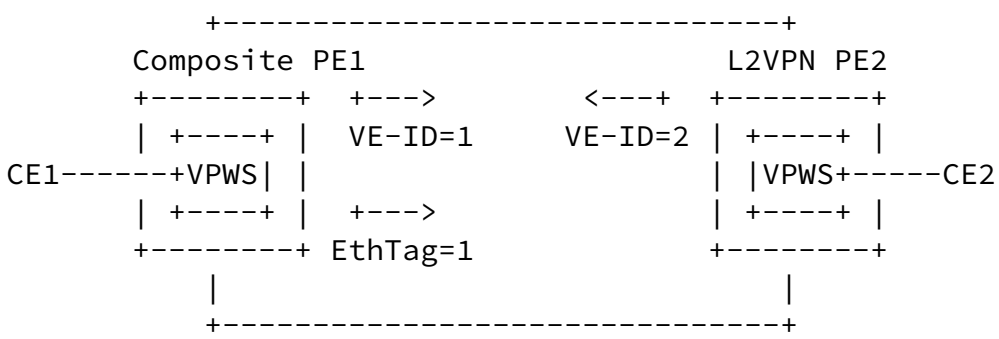

Figure 1: BGP-L2VPN and EVPN-VPWS integration

As it is shown in Figure 1 above, PE1 is provisioned to be a composite PE. PE1 originates both L2VPN A-D route with local VE-ID (1) as well as EVPN E-AD per EVI route with local Ethernet Tag ID (1) in their corresponding NLRIs. PE2 is a L2VPN PE and it originates a L2VPN A-D route with a local VE-ID (2) in its NLRI. A p2p Ethernet service is established between PE1 and PE2 based the L2VPN procedure per [RFC4761] when both PEs have the matching local/remote VE-IDs.

A composite PE may be optimized to originate either L2VPN route or EVPN E-AD per EVI route, but not both based on its provisioning model. Please see section 6.2.6 for detail.

If a CE is multihomed to composite PEs in multihoming mode, each composite PE also originates an EVPN E-AD per ES route and EVPN Ethernet segment per [RFC8214].

6.2.3. Status of an Attachment Circuit

A composite PE uses status vector TLV to notify other L2VPN PEs

through its L2VPN route the status of its attachment circuit per . A composite PE updates the corresponding L2VPN route with an updated status vector when there is a status change in its attachment circuit.

A composite PE withdraws its corresponding EVPN E-AD route per procedure defined in [RFC8214] when its locally attached Ethernet segment goes down.

6.2.4. Layer 2 Extended Community

A composite PE uses L2VPN info extended community for L2VPN per [RFC6624]. It shall support L2 encapsulation of type 4 and type 5.

A composite PE uses EVPN Layer 2 attribute extended community specified in [RFC8214] for EVPN, and it attaches the Layer 2 extended community in the EVPN A-D route it originates.

6.2.5. Port-active Multihoming and DF election

For the seamless migration, it is desirable that a multihomed CE uses a LAG interface to connect to a redundant set of composite PEs, such that when L2VPN PE involved in a Layer 2 P2P Ethernet service is migrated to support EVPN-VPWS, there is no need to touch the multihomed CE device if at that stage the redundant set of composite PEs are changed to provide all-active multihoming.

In addition, if the LACP protocol is running for the interface and while in single-active scenario, it is recommended a non-DF composite PE sends out-of-sync state for the interface instead of operational down. To that end, each composite PE is required to play a DF or non-DF role on a per port basis instead of per VLAN or per (ES, VLAN) basis.

To support multihomed CE connecting to the composite PEs working in a single-active multihoming scenario through a LAG interface, each composite PE must support port-active load-balancing, similar as it is specified in section 3 of [EVPN-MH-PA] except that a composite PE must also provide L2VPN functionality per [RFC6624].

Please note that per port DF/non-DF role can be achieved by using one of the standard based DF election algorithms, as long as the algorithm can be easily carried out on a per port basis, such as the preference based DF election when both the ESI and preference are configured on a per port basis.

Supporting port based single-active multihoming on the composite PEs with its multihomed CE using LAG interface does not change the control plane signaling, and it is oblivious to L2VPN PE. Since we cannot change the behavior of a L2VPN PE, a composite PE will continue to signal the preference for L2VPN on a per access interface basis through the Layer 2 extended community alongside its corresponding L2VPN A-D route. A L2VPN PE continues to carry the DF election based on its normal L2VPN process.

6.2.6. Optimization

With the simplest provisioning model, if a composite PE does not know a-priori whether the remote PE for a given P2P service is a L2VPN PE or an EVPN PE, the composite needs to participate in the auto-discovery and signaling procedures for both L2VPN and EVPN. This works well as it allows a composite to establish a P2P service with

different types of PEs composite PE, and to switch from using a L2VPN PW to EVPN VPWS dynamically during the migration process.

The simples provisioning model may not be optimal though, in that a composite PE originates twice as many A-D routes as they are required to establish the number of P2P services it is provisioned to. Therefore in some scenario, it is desirable that a composite PE be optimized to perform either L2VPN or EVPN VPWS procedure for a given P2P service, but not both.

For a composite PE, if a Service Provider has the prior knowledge about the types of remote PEs for some or all of its P2P Ethernet services, reducing the number of routes a composite PE originates can be achieved through the configuration. Based on the configuration, a composite may advertise EVPN route but not L2VPN A-D route for a P2P Ethernet service, or vice versa. It is up to the Service Provider to decide based on the network requirement.

6.3. Composite PE Forwarding Procedure

A composite PE supports forwarding procedure defined in [RFC6624] and [RFC8214].

When a packet arrives at an ingress composite PE, the composite PE adds a VPN service label based on the AC that packet arrives at, and it encapsulates the packet and sends it through a tunnel to the egress PE.

- A composite PE will not forward customer traffic to the L2VPN PE playing a non-DF role

- If a composite PE detects that two or more EVPN PEs are attached to the same ES and they are working in all-active mode, it will load balance the traffic among the EVPN PEs.

- If a composite PE detects that two or more EVPN PEs are attached to the same ES and they are working in single-active mode, it will only forward the traffic to the EVPN PE playing a DF role.

- If a set of composites PEs work in all-active multihoming mode for the same multihomed CE, then regardless of DF or Non-DF role each composite PE plays, it must forward the packet received from its multihomed CE to the remote L2VPN DF PE.

- If a composite PE receives both L2VPN and EVPN A-D routes from a remote PE for the same p2p Ethernet service, the composite should install forwarding routes in a make-before-break fashion:

a. For the traffic coming from the remote PE to its local access interface direction, to achieve a fast failover, the composite may install forwarding routes based on both L2VPN and EVPN A-D routes. However, to save system resource in a scaled setup, the composite may choose to install only the forwarding route for the EVPN A-D route and it should do so before it deletes the forwarding route for the L2VPN A-D route if it was installed beforehand.

b. For traffic coming from its local access interface to the remote PE direction, only one route can be installed for the

same local access interface. Forwarding should be based on the EVPN A-D route. The composite PE should update the forwarding route in a make-before-break fashion if the forwarding route for L2VPN A-D route has already been installed before the processing of the incoming EVPN A-D route.

- If a composite PE receives both L2VPN and EVPN A-D routes from a remote PE for the same p2p Ethernet service, and later on the remote PE is reverted back to a L2VPN only PE and withdraws its EVPN A-D route, the composite PE should also update the forwarding route accordingly in a make-before-break fashion:

  a. For the traffic coming from the remote PE to its local access interface direction, if the forwarding route for the L2VPN A-D route is not there, the composite PE should install the forwarding route for the L2VPN A-D route before it tears down the forwarding route for the EVPN A-D route.

  b. For the traffic coming from its local access interface to the remote PE direction, only one route can be installed for the same local access interface. The composite PE should update the forwarding route based on the L2VPN A-D route in a make-before-break fashion.

- Upon reception of an A-D per EVI route and an L2VPN route for the same P2P service, if both routes match the configured IDs, a composite PE MUST select the EVPN route and forward the traffic only to the EVPN PE, and not the L2VPN PE.

When a packet arrives at an egress PE, the VPN service label carried in the packet is used as the demultiplexer to identify the AC connecting to the destination CE.

6.4. Composite PE Procedures for All-Active Multi-Homing

Two or more Composite PEs MAY be attached to the same All-Active multi-homed CE and still be seamlessly integrated in an L2VPN

network. This is illustrated in Figure 2.

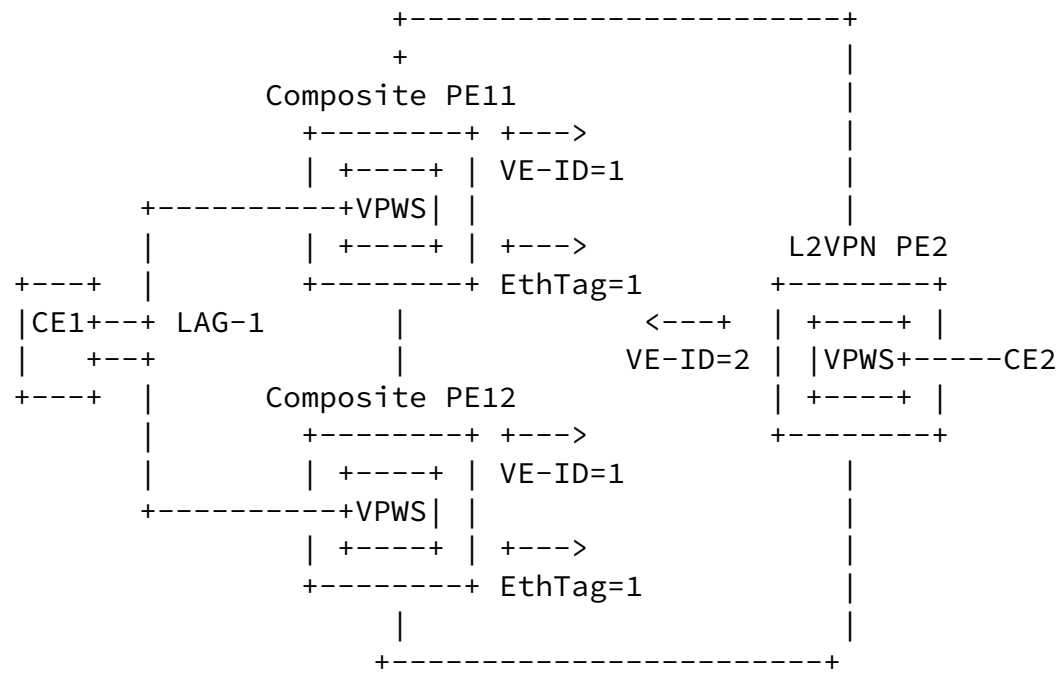

Figure 2: L2VPN and EVPN-VPWS integration with all-active multi-homed CEs

In the example of Figure 2, PE11 and PE12 are configured as composite PEs with the same local CE identifiers. That is, both PEs are configured with local VE-ID (1) and the same remote VE-ID (2). Also, both will be configured with the same EVPN local Ethernet Tag (1) and remote Ethernet Tag (2). As long as PE2 does not become a composite PE or an EVPN PE, it will not import the A-D per-EVI routes and will import the L2VPN routes only. PE2 will make a selection to use either PE11 or PE12 to setup an L2VPN VPWS service.

For example, assuming PE11 is selected, PE2 forwards the traffic coming from CE2 to PE11 (there is no per-flow load-balancing). In case of failure, upon receiving the L2VPN rout withdraw from PE11, PE2 will change its forwarding next-hop to PE12.

In the reverse direction, CE1 will perform per-flow load-balancing to PE11 and PE12. Both PEs will program their forwarding paths to send CE1 traffic to PE2.

The benefit of this solution is that when PE2 can be upgraded to an EVPN or composite PE, the P2P service can be migrated to EVPN VPWS with no changes on CE1 or PE11/PE12.

7. Extended Seamless Integration Solution for Multipoint Ethernet VPN

[RFC8560] specifies how EVPN and VPLS PEs can be seamlessly integrated into the same network, assuming the VPLS PEs use [RFC4761] or [RFC4762] procedures to setup the pseudowires to the remote VPLS PEs or composite PEs. [RFC8560] procedures consider that CEs can be multi-homed to two VPLS PEs, or to a group of composite PEs in a single-active or port-active Ethernet Segment. All-active multi-homing is out of scope.

This specification updates [RFC8560] in case All-Active multi-homing is used on two or more composite PEs of the same multipoint VPN service and the composite PEs and VPLS PEs use the BGP-VPLS [RFC4761] control and data plane procedures. Seamless integration and All-active multi-homing procedures for [RFC4762] VPLS is out of scope. This document also updates [RFC8560] to clarify the required MAC Flush procedures in case single-active/all-active/por-active multi-homing is used on the composite PEs.

7.1. All-Active Multi-Homing and Seamless Integration for BGP-VPLS services

All-active Ethernet Segments MAY be used in a VPLS service composed of composite and BGP-VPLS PEs. Ethernet Segments are an EVPN construct, hence only supported in composite PEs and not BGP-VPLS PEs. Figure 3 illustrates an example of the use of All-active Ethernet Segments and seamless integration between BGP-VPLS and EVPN.

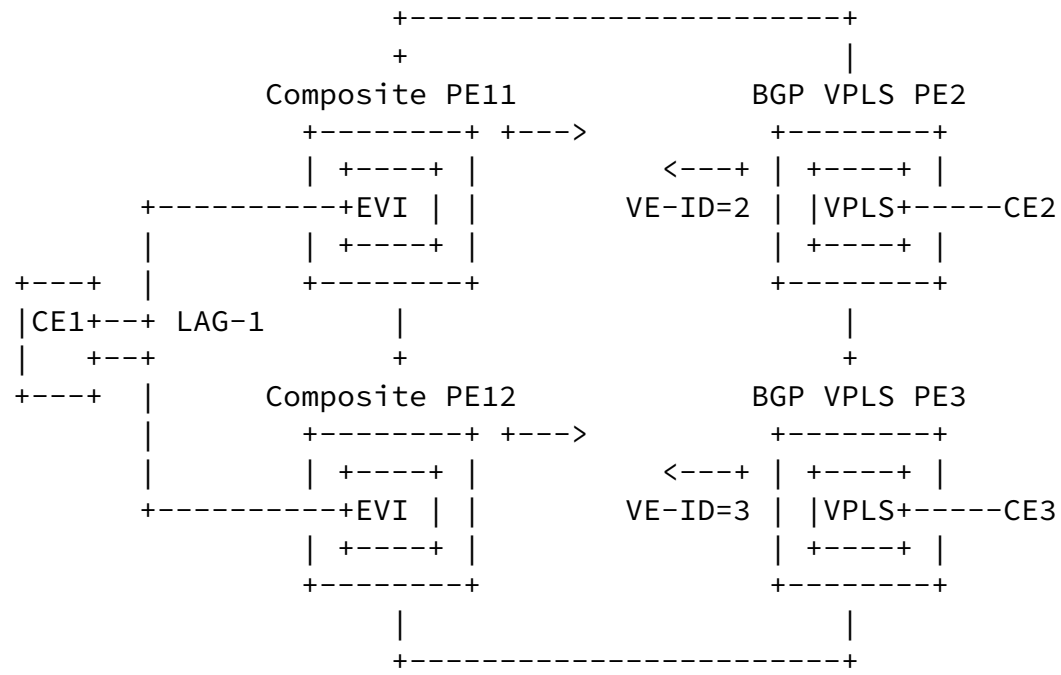

Figure 3: BGP-VPLS and EVPN PEs with all-active multi-homing

In Figure 3, the composite PEs will be provisioned for EVPN and All-active Multi-homing as specified in [RFC7432]. In addition, BGP-VPLS is enabled on the same services. PE11 and PE12 will therefore advertise the corresponding EVPN and BGP-VPLS routes. The EVPN routes are only imported by PE11 and PE12, whereas the BGP-VPLS routes are imported by all the PEs in the service.

In this case, the PEs MUST follow the procedures in [RFC8560] that are repeated below for the reader's benefit:

- The composite PEs MUST place their EVPN MP2P service tunnels and BGP-VPLS PWs in the same Split Horizon Group. I.e., traffic coming from a BGP-VPLS PW MUST NOT be forwarded to an EVPN tunnel.

- If two composite PEs successfully attempt to setup a BGP-VPLS PW and an EVPN tunnel, the BGP-VPLS pseudowire will be brought operationally down.

- The composite PEs will not advertise any MAC/IP routes for MAC

address learned on a BGP-VPLS PW that is part of the Split Horizon Group assigned to the EVPN tunnels.

In addition, this document updates [RFC8560] so that All-active multi-homing Ethernet Segments MAY exist in the composite PEs. If an all-active multi-homing ES is defined in a group of composite PEs, all the BDs associated to the LAG MUST support and follow the EVPN multi-homing procedures.

If a group of composite PEs work in all-active multihoming and another group of composite PEs work in single-active, normal EVPN procedure will be used between these two group of composite PEs.

If a group of composite PEs work in all-active multihoming and remote BGP-VPLS PEs work in single-active, BGP-VPLS procedure will be used between composite PEs and BGP-VPLS PEs.

When all-active multi-homing is used, a MAC flip-flopping effect will exist on the BGP-VPLS PEs. In Figure 3, this effect results in CE1's MAC moving between two different PWs in PE2 and PE3. E.g., at first CE1 may hash the traffic to PE11, and PE2 may learn CE1's MAC on its pseudowire PE2-PE11. Later, if CE1 hashes the traffic (for a different flow) to PE12, PE2 will move CE1's MAC to its PW PE2-PE12. This MAC move or "flip-flopping" can happen continuously and may have harmful consequences for the BGP-VPLS PEs. In some cases, the BGP-VPLS PEs will consider this to be a loop.

The solution to avoid the MAC "flip-flopping" is based on the support of "MAC Pinning" on the BGP-VPLS PEs, as follows:

- In Figure 3, the composite PEs and BGP-VPLS PEs will setup their PWs normally.

- The MAC flip-flopping effect would be avoided by enabling MAC Pinning on the PE2 and PE3 pseudowires.

- With MAC Pinning enabled, PE2 and PE3 will learn CE1's MAC on only one PW and will not be relearned in the same or different PW until the MAC ages out. E.g., consider CE1 hashes the first flow to PE11 and PE11 forwards to PE2. PE2 learns CE1's MAC on PW PE2-PE11. Since MAC Pinning is applied on that PW, subsequent frames arriving at PW PE2-PE12 with CE1's MAC will not trigger a

relearn process on PE2.

MAC Pinning is assumed to be supported by all the BGP-VPLS PEs in the network, therefore no upgrade is required on the BGP-VPLS PEs to support this specification.

7.2. Extensions for MAC Flush

Irrespective of the type of multi-homing used on the composite PEs, in case of a failure on the Ethernet Segment (node or link failure) the composite PEs MUST indicate the need to flush MAC addresses to the remote BGP-VPLS PEs.

E.g., in Figure 3, consider CE1's MAC is learned on PW PE2-PE11 (on PE2). If the link CE1-PE11 fails, PE2 will continue sending the

unicast traffic to CE1 using the PW to PE11, and therefore causing a blackhole until CE1's MAC ages out.

A MAC flush mechanism is required in order to speed up the convergence in case of ES failures. This requires some extensions to [RFC8560] and it will be added in future versions.

8. IANA Considerations

This document raises no new IANA request. There is no IANA actions.

9. Security Considerations

This document does not introduce any new security concern. This document inherits the same security as they are specified in [RFC6624] and [RFC8214].

10. Acknowledgements

The authors would like to thank Hitesh Mali and Vasu Venkatraman for their valuable comments and feedbacks. They would also like to thank John Drake for his review and support.

11. Contributors

In addition to the authors listed, the following individuals also

contributed to this document:

Vinod Prabhu, Nokia

12. References

12.1. Normative References

[EVPN-MH-PA] Brissette, P., Thoria, S., and A. Sajassi, "EVPN multi-homing port-active load-balancing", internet-draft draft-brissette-bess-evpn-mh-pa-04, March 2019.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC4761] Kompella, K., Ed. and Y. Rekhter, Ed., "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC 4761, DOI 10.17487/RFC4761, January 2007, <https://www.rfc-editor.org/info/rfc4761>.

[RFC6624] Kompella, K., Kothari, B., and R. Cherukuri, "Layer 2 Virtual Private Networks Using BGP for Auto-Discovery and Signaling", RFC 6624, DOI 10.17487/RFC6624, May 2012, <https://www.rfc-editor.org/info/rfc6624>.

[RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February 2015, <https://www.rfc-editor.org/info/rfc7432>.

[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>.

[RFC8214] Boutros, S., Sajassi, A., Salam, S., Drake, J., and J. Rabadan, "Virtual Private Wire Service Support in Ethernet VPN", RFC 8214, DOI 10.17487/RFC8214, August 2017, <https://www.rfc-editor.org/info/rfc8214>.

[RFC8560] Sajassi, A., Ed., Salam, S., Del Regno, N., and J. Rabadan, "Seamless Integration of Ethernet VPN (EVPN) with Virtual Private LAN Service (VPLS) and Their Provider Backbone Bridge (PBB) Equivalents", RFC 8560, DOI 10.17487/RFC8560, May 2019, <https://www.rfc-editor.org/info/rfc8560>.

12.2. Informative References

[RFC4448] Martini, L., Ed., Rosen, E., El-Aawar, N., and G. Heron, "Encapsulation Methods for Transport of Ethernet over MPLS Networks", RFC 4448, DOI 10.17487/RFC4448, April 2006, <https://www.rfc-editor.org/info/rfc4448>.

[RFC4762] Lasserre, M., Ed. and V. Kompella, Ed., "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", RFC 4762, DOI 10.17487/RFC4762, January 2007, <https://www.rfc-editor.org/info/rfc4762>.

Authors' Addresses

Wen Lin
Juniper Networks, Inc.

EMail: wlin@juniper.net

Bin Wen
Comcast

EMail: bin_wen@comcast.com

Voiteck Kozak
Comcast

EMail: voitek_kozak@comcast.com

Jorge Rabadan
Nokia

EMail: jorge.rabadan@nokia.com