Stream: Internet Engineering Task Force (IETF)
RFC: 9784
Category: Standards Track
Published: June 2025
ISSN: 2070-1721
Authors: A. Sajassi, *Cisco Systems*; P. Brissette, *Cisco Systems*; R. Schell, *Independent*; J. Drake, *Independent*; J. Rabadan, *Nokia*

RFC 9784
Virtual Ethernet Segments for EVPN and Provider Backbone Bridge EVPN

Abstract

Ethernet VPN (EVPN) and Provider Backbone Bridge EVPN (PBB-EVPN) introduce a comprehensive suite of solutions for delivering Ethernet services over MPLS/IP networks. These solutions offer advanced multihoming capabilities. Specifically, they support Single-Active and All-Active redundancy modes for an Ethernet Segment (ES), which is defined as a collection of physical links connecting a multihomed device or network to a set of Provider Edge (PE) devices. This document extends the concept of an ES by allowing an ES to be associated with a set of Ethernet Virtual Circuits (EVCs), such as VLANs, or other entities, including MPLS Label Switched Paths (LSPs) or pseudowires (PWs). This extended concept is referred to as virtual Ethernet Segments (vESes). This document lists the requirements and specifies the necessary extensions to support vES in both EVPN and PBB-EVPN.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc9784.

Copyright Notice

Table of Contents

1. Introduction

Ethernet VPN (EVPN) [RFC7432] and Provider Backbone Bridge EVPN (PBB-EVPN) [RFC7623] introduce a comprehensive suite of solutions for delivering Ethernet services over MPLS/IP networks. These solutions offer advanced multihoming capabilities. Specifically, they support Single-Active and All-Active redundancy modes for an Ethernet Segment (ES). As defined in [RFC7432], an ES represents a collection of Ethernet links that connect a customer site to one or more Provider Edge (PE) devices.

This document extends the concept of an ES by allowing an ES to be associated with a set of Ethernet Virtual Circuits (EVCs) (such as VLANs) or other entities, including MPLS Label Switched Paths (LSPs) or pseudowires (PWs). This extended concept is referred to as virtual Ethernet Segments (vESes). This document lists the requirements and specifies the necessary extensions to support vES in both EVPN and PBB-EVPN. The scope of this document includes PBB-EVPN [RFC7623], EVPN over MPLS [RFC7432], and EVPN over IP [RFC8365]; however, it excludes EVPN over SRv6 [RFC9252].

1.1. 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.

1.2. vESes in Access Ethernet Networks

Some service providers (SPs) seek to extend the concept of physical Ethernet links in an ES to encompass EVCs, wherein multiple EVCs (such as VLANs) can be aggregated onto a single physical External Network-Network Interface (ENNI). An ES composed of a set of EVCs rather than physical links is referred to as a vES. Figure 1 illustrates two PE devices (PE1 and PE2), each with an ENNI aggregating several EVCs. Some of these EVCs on a given ENNI can be associated with vESes. For instance, the multihomed vES depicted in Figure 1 consists of EVC4 on ENNI1 and EVC5 on ENNI2.

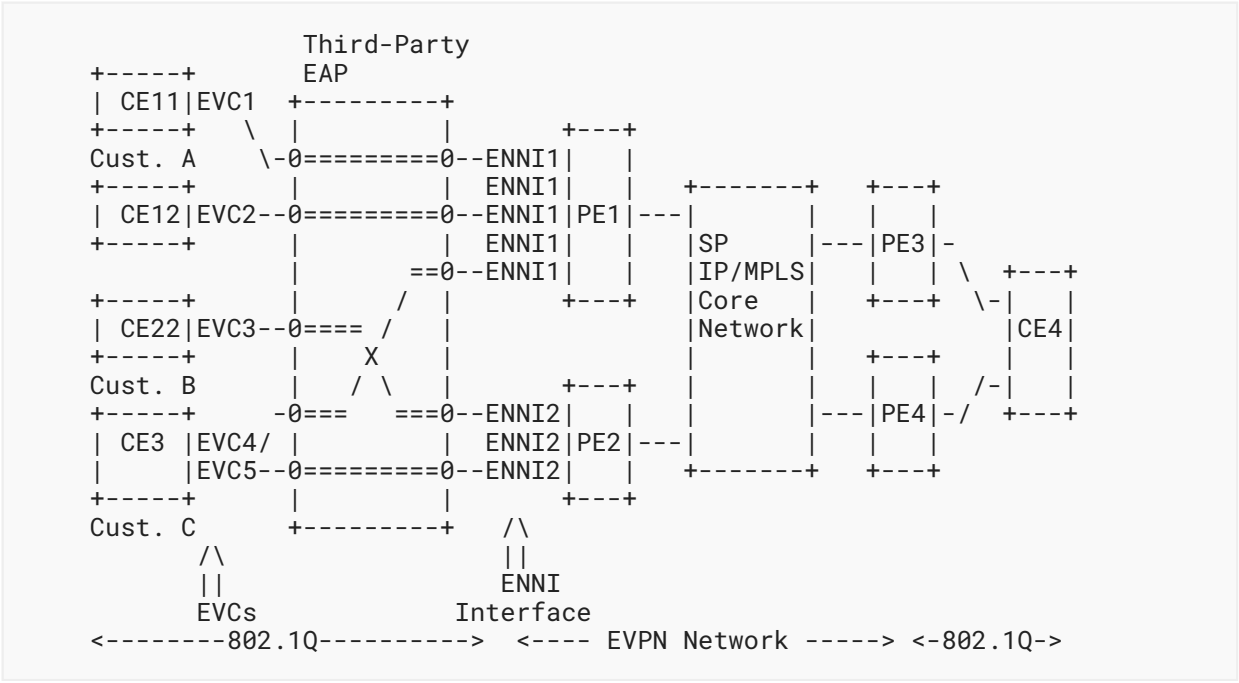

Figure 1: Single-Homed Devices and a Dual-Homed Device/Network on the Same ENNI

ENNIs are commonly used to reach remote customer sites via independent Ethernet access networks or third-party Ethernet Access Providers (EAPs). ENNIs can aggregate traffic from many vESes (e.g., hundreds to thousands), where each vES is represented by its associated EVC on that ENNI. As a result, ENNIs and their associated EVCs are key elements of SP external boundaries that are carefully designed and closely monitored. As a reminder, the ENNI is the demarcation between the SP (IP/MPLS core network) and the third-party Ethernet Access Provider.

To meet customers' Service Level Agreements (SLAs), SPs build redundancy via multiple EVPN PEs and across multiple ENNIs (as shown in Figure 1), where a given vES can be multihomed to two or more EVPN PE devices (on two or more ENNIs) via their associated EVCs. Just like physical ESs in the solutions described in [RFC7432] and [RFC7623], these vESes can be single-homed or multihomed ESs, and when multihomed, they can operate in either Single-Active or All-Active redundancy modes. In a typical SP external-boundary scenario (e.g., with an EAP), an ENNI can be associated with several thousands of single-homed vESes, several hundreds of Single-Active vESes, and tens or hundreds of All-Active vESes. The specific figures used throughout this document reflect the relative quantities (hundreds, thousands, etc.) of various elements as understood at the time of writing.

1.3. vESes in Access MPLS Networks

Other SPs want to extend the concept of physical links in an ES to individual PWs or to MPLS LSPs in Access MPLS networks, i.e., a vES consisting of a set of PWs or a set of LSPs. Figure 2 illustrates this concept.

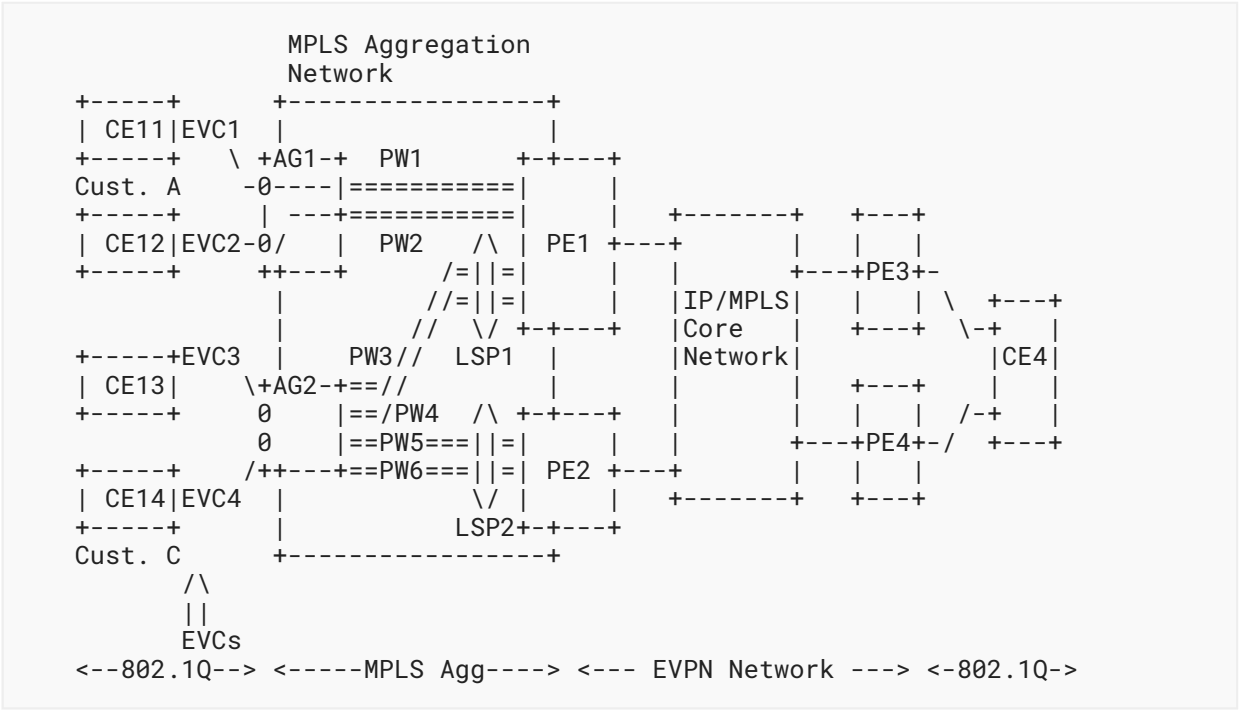

Figure 2: A Dual-Homed and Single-Homed Network on MPLS Aggregation Networks

In certain scenarios, SPs utilize MPLS Aggregation Networks that are managed by separate administrative entities or third-party organizations to gain access to their own IP/MPLS core network infrastructure. This situation is depicted in Figure 2.

In such scenarios, a vES is defined as a set of individual PWs when aggregation is not feasible. If aggregation is possible, the vES can be associated with a group of PWs that share the same unidirectional LSP pair, where the LSP pair consists of the ingress and egress LSPs between the same endpoints.

In Figure 2, EVC3 is connected to a VPWS instance in AG2 that is connected to PE1 and PE2 via PW3 and PW5, respectively. EVC4 is connected to another VPWS instance on AG2 that is connected to PE1 and PE2 via PW4 and PW6, respectively. Since the PWs for the two VPWS instances can be aggregated into the same LSP pair going to and coming from the MPLS network, a common vES can be defined for the four mentioned PWs. In Figure 2, LSP1 and LSP2 represent the two LSP pairs between PE1 and AG2 and between PE2 and AG2, respectively. The vES consists of these two LSP pairs (LSP1 and LSP2), and each LSP pair has two PWs. This vES will be shared by two separate EVPN instances (e.g., EVI-1 and EVI-2) in the EVPN network. PW3 and PW4 are associated with EVI-1 and EVI-2, respectively, on PE1, and PW5 and PW6 are associated with EVI-1 and EVI-2, respectively, on PE2.

In some cases, the aggregation of PWs that share the same LSP pair may not be possible. For instance, if PW3 were terminated into a third PE, e.g., PE3, instead of PE1, the vES would need to be defined on each PW on each PE.

For MPLS/IP access networks where a vES represents a set of LSP pairs or a set of PWs, this document extends the Single-Active multihoming procedures defined in [RFC7432] and [RFC7623] to accommodate vES. The extension of vES to support All-Active multihoming in MPLS/IP access networks is beyond the scope of this document.

This document defines the concept of a vES and specifies the additional extensions necessary to support a vES in accordance with [RFC7432] and [RFC7623]. Section 3 enumerates the set of requirements for a vES. Section 4 details the extensions for a vES applicable to EVPN solutions, including those specified in [RFC7432] and [RFC7209]. These extensions are designed to meet the requirements listed in Section 3. Section 4 also provides an overview of the solution, while Section 5 addresses failure handling, recovery, scalability, and fast convergence of [RFC7432] and [RFC7623] for vESes.

2. Terminology

AC: Attachment Circuit

B-MAC: Backbone MAC Address

CE: Customer Edge

C-MAC: Customer/Client MAC Address

DF: Designated Forwarder

ENNI: External Network-Network Interface

ES: Ethernet Segment

ESI: Ethernet Segment Identifier

Ethernet A-D: Ethernet Auto-Discovery

EVC: Ethernet Virtual Circuit [MEF63]

EVI: EVPN Instance

EVPN: Ethernet VPN

I-SID: Service Instance Identifier (24 bits and global within a PBB network; see [RFC7080]).

MAC: Media Access Control

PBB: Provider Backbone Bridge

PBB-EVPN: Provider Backbone Bridge EVPN

PE: Provider Edge

VPWS: Virtual Private Wire Service

Single-Active (SA) Redundancy Mode: When only a single PE, among a group of PEs attached to an ES, is allowed to forward traffic to/from that ES, the ES is defined as operating in Single-Active redundancy mode.

All-Active (AA) Redundancy Mode: When all PEs attached to an ES are allowed to forward traffic to/from that ES, the ES is defined as operating in All-Active redundancy mode.

3. Requirements

This section describes the requirements specific to vES for EVPN and PBB-EVPN solutions. These requirements are in addition to the ones described in [RFC8214], [RFC7432], and [RFC7623].

3.1. Single-Homed and Multihomed vES

A PE device **MUST** support the following types of vESes:

(R1a) The PE **MUST** handle single-homed vESes on a single physical port, such as a single ENNI.

(R1b) The PE **MUST** support a combination of single-homed vESes and Single-Active multihomed vESes simultaneously on a single physical port, such as a single ENNI. Throughout this document, Single-Active multihomed vESes will be referred to as "Single-Active vESes".

(R1c) The PE **MAY** support All-Active multihomed vESes on a single physical port. Throughout this document, All-Active multihomed vESes will be referred to as "All-Active vESes".

(R1d) The PE **MAY** support a combination of All-Active vESes along with other types of vESes on a single physical port.

(R1e) A multihomed vES, whether Single-Active or All-Active, can span across two or more ENNIs on any two or more PEs.

3.2. Local Switching

Many vESes of different types can be aggregated on a single physical port on a PE device and some of these vESes can belong to the same service instance (e.g., EVI). This translates into the need for supporting local switching among the vESes for the same service instance on the same physical port (e.g., ENNI) of the PE.

(R2a) A PE device that supports the vES function **MUST** support local switching among different vESes associated with the same service instance on a single physical port. For instance, in Figure 1, PE1 must support local switching between CE11 and CE12, which are mapped to two single-homed vESes on ENNI1. In the case of Single-Active vESes, the local switching is performed among active EVCs associated with the same service instance on the same ENNI.

3.3. EVC Service Types

A physical port, such as an ENNI of a PE device, can aggregate numerous EVCs, each associated with a vES. An EVC may carry one or more VLANs. Typically, an EVC carries a single VLAN and is therefore associated with a single broadcast domain. However, there are no restrictions preventing an EVC from carrying multiple VLANs.

(R3a) An EVC can be associated with a single broadcast domain, such as in a VLAN-based service or a VLAN bundle service.

(R3b) An EVC **MAY** be associated with several broadcast domains, such as in a VLAN-aware bundle service.

Similarly, a PE can aggregate multiple LSPs and PWs. In the case of individual PWs per vES, a PW is typically associated with a single broadcast domain, although there are no restrictions preventing a PW from carrying multiple VLANs if the PW is configured in Raw mode.

(R3c) A PW can be associated with a single broadcast domain, such as in a VLAN-based service or a VLAN bundle service.

(R3d) A PW **MAY** be associated with several broadcast domains, such as in a VLAN-aware bundle service.

3.4. Designated Forwarder (DF) Election

Section 8.5 of [RFC7432] specifies the default procedure for DF election in EVPN, which is also applied in [RFC7623] and [RFC8214]. [RFC8584] elaborates on additional procedures for DF election in EVPN. These DF election procedures are performed at the granularity of (ESI, Ethernet Tag). In the context of a vES, the same EVPN default procedure for DF election is applicable but at the granularity of (vESI, Ethernet Tag). In this context, the Ethernet Tag is represented by an I-SID in PBB-EVPN and by a VLAN ID (VID) in EVPN. As described in [RFC7432], this default procedure for DF election at the granularity of (vESI, Ethernet Tag) is also known as "service carving." The goal of service carving is to evenly distribute the DFs for different vESes among various PEs, thereby ensuring an even distribution of traffic across the PEs. The following requirements are applicable to the DF election of vESes for EVPN and PBB-EVPN.

(R4a) A PE that supports vES function **MUST** support a vES with m EVCs among n ENNIs belonging to p PEs in any arbitrary order, where n >= p >= m >=2. For example, if there is a vES with 2 EVCs and there are 5 ENNIs on 5 PEs (PE1 through PE5), then vES can be dual-homed to PE2 and PE4, and the DF election must be performed between PE2 and PE4.

(R4b) Each vES **MUST** be identified by its own virtual ESI (vESI).

3.5. EVC Monitoring

To detect the failure of an individual EVC and subsequently perform DF election for its associated vES as a result of this failure, each EVC should be monitored independently.

(R5a) Each EVC **SHOULD** be independently monitored for its operational health.

(R5b) A failure in a single EVC, among many aggregated on a single physical port or ENNI, **MUST** trigger a DF election for its associated vES.

3.6. Failure and Recovery

(R6a) Failure and failure recovery of an EVC for a single-homed vES **SHALL NOT** impact any other EVCs within its service instance or any other service instances. In other words, for PBB-EVPN, it **SHALL NOT** trigger any MAC flushing within both its own I-SID and other I-SIDs.

(R6b) In case of All-Active vES, failure and failure recovery of an EVC for that vES **SHALL NOT** impact any other EVCs within its service instance or any other service instances. In other words, for PBB-EVPN, it **SHALL NOT** trigger any MAC flushing within both its own I-SID and other I-SIDs.

(R6c) Failure and failure recovery of an EVC for a Single-Active vES **SHALL** impact only its own service instance. In other words, for PBB-EVPN, MAC flushing **SHALL** be limited to the associated I-SID only and **SHALL NOT** impact any other I-SIDs.

(R6d) Failure and failure recovery of an EVC for a Single-Active vES **MUST** only impact C-MACs associated with a multihomed device/network for that service instance. In other words, MAC flushing **MUST** be limited to a single service instance (I-SID in the case of PBB-EVPN) and only C-MACs for a Single-Active multihomed device/network.

3.7. Fast Convergence

Since many EVCs (and their associated vESes) are aggregated via a single physical port (e.g., ENNI), when there is a failure of that physical port, it impacts many vESes and equally triggers many ES route withdrawals. Formulating, sending, receiving, and processing such large numbers of BGP messages can introduce delay in DF election and convergence time. As such, it is highly desirable to have a mass-withdraw mechanism similar to the one in [RFC7432] for withdrawing many Ethernet A-D per ES routes.

(R7a) There **SHOULD** be a mechanism equivalent to EVPN mass withdraw such that upon an ENNI failure, only a single BGP message to the PEs is needed to trigger DF election for all impacted vESes associated with that ENNI.

4. Solution Overview

The solutions described in [RFC7432] and [RFC7623] are leveraged as is, with the modification that the ESI assignment is performed for an EVC or a group of EVCs or LSPs/PWs instead of a link or a group of physical links. In other words, the ESI is associated with a vES (hereby referred to as the "vESI").

In the EVPN solution, the overall procedures remain consistent, with the primary difference being the handling of physical port failures that can affect multiple vESes. Sections 5.1 and 5.3 describe the procedures for managing physical port or link failures in the context of EVPN. In a typical multihomed setup, MAC addresses learned behind a vES are advertised using the ESI associated with the vES (i.e., the vESI). EVPN aliasing and mass-withdraw operations are conducted with respect to the vES identifier. Specifically, the Ethernet A-D routes for these operations are advertised using the vESI instead of the ESI.

For the PBB-EVPN solution, the main change is with respect to the B-MAC address assignment, which is performed in a similar way to what is described in Section 6.2.1.1 of [RFC7623], with the following refinements:

- One shared B-MAC address **SHOULD** be used per PE for the single-homed vESes. In other words, a single B-MAC is shared for all single-homed vESes on that PE.
- One shared B-MAC address **SHOULD** be used per PE, per physical port (e.g., ENNI) for the Single-Active vESes. In other words, a single B-MAC is shared for all Single-Active vESes that share the same ENNI.
- One shared B-MAC address **MAY** be used for all Single-Active vESes on that PE.
- One B-MAC address **SHOULD** be used per set of EVCs representing an All-Active vES. In other words, a single B-MAC address is used per vES for All-Active scenarios.
- A single B-MAC address **MAY** also be used per vES, per PE for Single-Active scenarios.

4.1. EVPN DF Election for vES

The service carving procedures for vESes are almost the same as the procedures outlined in Section 8.5 of [RFC7432] and in [RFC8584], except that ES is replaced with vES.

For the sake of clarity and completeness, the default DF election procedure of [RFC7432] is repeated below with the necessary changes:

1. When a PE discovers the vESI or is configured with the vESI associated with its attached vES, it advertises an ES route with the associated ES-Import Route Target extended community attribute.
2. The PE then starts a timer (default value = 3 seconds) to allow the reception of ES routes from other PE nodes connected to the same vES. This timer value **MUST** be the same across all PEs connected to the same vES.

3. When the timer expires, each PE builds an ordered list of the IP addresses of all the PE nodes connected to the vES (including itself), in increasing numeric value. Each IP address in this list is extracted from the "Originator Router's IP address" field of the advertised ES route. Every PE is then given an ordinal indicating its position in the ordered list, starting with 0 as the ordinal for the PE with the numerically lowest IP address. The ordinals are used to determine which PE node will be the DF for a given EVPN instance on the vES using the following rule: Assuming a redundancy group of N PE nodes, the PE with ordinal i is the DF for an EVPN instance with an associated Ethernet Tag value of V when (V mod N) = i. It should be noted that using the "Originator Router's IP address" field in the ES route to get the PE IP address needed for the ordered list allows a CE to be multihomed across different ASes, if such need ever arises.
4. The PE that is elected as a DF for a given EVPN instance will unblock traffic for that EVPN instance. Note that the DF PE unblocks all traffic in both ingress and egress directions for Single-Active vESes and unblocks multi-destination in the egress direction for All-Active multihomed vESes. All non-DF PEs block all traffic in both ingress and egress directions for Single-Active vESes and block multi-destination traffic in the egress direction for All-Active vESes.

In case of an EVC failure, the affected PE withdraws its corresponding ES route if there are no more EVCs associated to the vES in the PE. This will re-trigger the DF election procedure on all the PEs in the redundancy group. For PE node failure, or upon PE commissioning or decommissioning, the PEs re-trigger the DF election procedure across all affected vESes. In case of a Single-Active scenario, when a service moves from one PE in the redundancy group to another PE because of DF re-election, the PE (which ends up being the elected DF for the service) **MUST** trigger a MAC address flush notification towards the associated vES if the multihoming device is a bridge or the multihoming network is an Ethernet bridged network.

For LSP-based and PW-based vES, the non-DF PE **SHOULD** signal PW-status 'standby' to the Aggregation PE (e.g., AG1 and AG2 in Figure 2), and a new DF PE **MAY** send a Label Distribution Protocol (LDP) MAC withdraw message as a MAC address flush notification. It should be noted that the PW-status is signaled for the scenarios where there is a one-to-one mapping between EVI (EVPN instance) and the PW.

4.2. Grouping and Route Coloring for vES

Physical ports (e.g., ENNI) that aggregate many EVCs are 'colored' to enable the grouping schemes described below.

By default, the MAC address of the corresponding port (e.g., ENNI) is used to represent the 'color' of the port, and the EVPN Router's MAC Extended Community defined in [RFC9135] is used to signal this color.

The difference between coloring mechanisms for EVPN and PBB-EVPN is that the extended community is advertised with the Ethernet A-D per ES route for EVPN, whereas the extended community is advertised with the B-MAC route for PBB-EVPN.

The subsequent sections detailing Grouping of Ethernet A-D per ES routes and Grouping of B-MAC addresses will be essential for addressing port failure handling, as discussed in Sections 5.3, 5.4, and 5.5.

4.2.1. EVPN Route Coloring for vES

When a PE discovers the vESI or is configured with the vESI associated with its attached vES, an ES route and Ethernet A-D per ES route are generated using the vESI identifier.

These ES and Ethernet A-D per ES routes specific to each vES are colored with an attribute representing their association to a physical port (e.g., ENNI).

The corresponding port 'color' is encoded in the EVPN Router's MAC Extended Community defined in [RFC9135] and advertised along with the ES and Ethernet A-D per ES routes for this vES. The color (which is the MAC address of the port) **MUST** be unique.

The PE also constructs a special Grouping Ethernet A-D per ES route that represents all the vESes associated with the port (e.g., ENNI). The corresponding port 'color' is encoded in the ESI field. For this encoding, Type 3 ESI (Section 5 of [RFC7432]) is used with the MAC field set to the color (MAC address) of the port and the 3-octet local discriminator field set to 0xFFFFFF.

The ESI label extended community (Section 7.5 of [RFC7432]) is not relevant to Grouping Ethernet A-D per ES route. The label value is not used for encapsulating Broadcast, Unknown Unicast, and Multicast (BUM) packets for any split-horizon function. The ESI label extended community **MUST NOT** be added to Grouping Ethernet A-D per ES route and **MUST** be ignored on receiving the PE.

The Grouping Ethernet A-D per ES route is advertised with a list of Route Targets corresponding to the affected service instances. If the number of associated Route Targets exceeds the capacity of a single route, multiple Grouping Ethernet A-D per ES routes are advertised accordingly as specified in Section 8.2 of [RFC7432].

4.2.2. PBB-EVPN Route Coloring for vES

In PBB-EVPN, particularly when there are large numbers of service instances (i.e., I-SIDs) associated with each EVC, the PE device **MAY** assign a color attribute to each vES B-MAC route, indicating their association with a physical port (e.g., an ENNI).

The corresponding port 'color' is encoded in the EVPN Router's MAC Extended Community defined in [RFC9135] and advertised along with the B-MAC for this vES in PBB-EVPN.

The PE **MAY** also construct a special Grouping B-MAC route that represents all the vESes associated with the port (e.g., ENNI). The corresponding port 'color' is encoded directly into this special Grouping B-MAC route.

5. Failure Handling and Recovery

There are several failure scenarios to consider such as:

A: CE uplink port failure

B: Ethernet Access Network failure

C: PE access-facing port or link failure

D: PE node failure

E: PE isolation from IP/MPLS network

The solutions specified in [RFC7432], [RFC7623], and [RFC8214] provide protection against failures as described in these respective references. In the context of these solutions, the presence of vESes introduces an additional failure scenario beyond those already considered, specifically the failure of individual EVCs. Addressing vES failure scenarios necessitates the independent monitoring of EVCs or PWs. Upon detection of failure or service restoration, appropriate DF election and failure recovery mechanisms must be executed.

[RFC7023] is used for monitoring EVCs, and upon failure detection of a given EVC, the DF election procedure per Section 4.1 is executed. For PBB-EVPN, some extensions are needed to handle the failure and recovery procedures of [RFC7623] to meet the above requirements. These extensions are described in the next section.

[RFC4377] and [RFC6310] are used for monitoring the status of LSPs and/or PWs associated to vES.

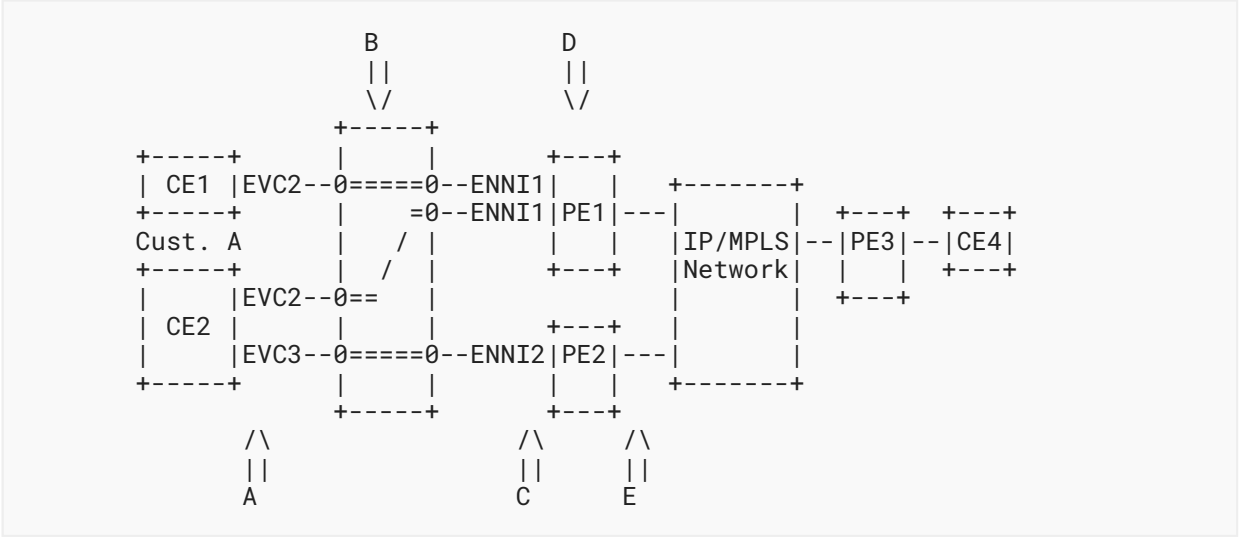

Figure 3: Failure Scenarios A, B, C, D, and E

5.1. EVC Failure Handling for Single-Active vES in EVPN

In [RFC7432], when a DF PE connected to a Single-Active multihomed ES loses connectivity to the segment, due to link or port failure, it signals the remote PEs to invalidate all MAC addresses associated with that ES. This is done by means of a mass-withdraw message, by withdrawing the

Ethernet A-D per ES route. It should be noted that for dual-homing use cases where there is only a single backup path, MAC invalidating can be avoided by the remote PEs as they can update their next hop associated with the affected MAC entries to the backup path per the procedure described in Section 8.2 of [RFC7432].

In case of an EVC failure that impacts a single vES, this same EVPN procedure is used. In this case, the mass withdraw is conveyed by withdrawing the Ethernet A-D per vES route carrying the vESI representing the failed EVC. Upon receiving this message, the remote PEs perform the same procedures outlined in Section 8.2 of [RFC7432].

5.2. EVC Failure Handling for a Single-Active vES in PBB-EVPN

In [RFC7432], when a PE connected to a Single-Active ES loses connectivity to the segment, due to link or port failure, it signals the remote PE to flush all C-MAC addresses associated with that ES. This is done by updating the advertised B-MAC route's MAC Mobility extended community.

In case of an EVC failure that impacts a single vES, if the above PBB-EVPN procedure is used, it results in excessive C-MAC flushing because a single physical port can support a large number of EVCs (and their associated vESes); therefore, updating the advertised B-MAC corresponding to the physical port, with MAC Mobility extended community, will result in flushing C-MAC addresses not just for the impacted EVC but for all other EVCs on that port.

To reduce the scope of C-MAC flushing to only the impacted service instances (the service instance(s) impacted by the EVC failure), the PBB-EVPN C-MAC flushing needs to be adapted on a per-service-instance basis (i.e., per I-SID). [RFC9541] introduces a B-MAC/I-SID route where the existing PBB-EVPN B-MAC route is modified to carry an I-SID, instead of a NULL value, in the "Ethernet Tag ID" field. To the receiving PE, this field indicates flushing all C-MAC addresses associated with that I-SID for that B-MAC. This C-MAC flushing mechanism per I-SID **SHOULD** be used in case of an EVC failure impacting a vES. Since an EVC typically maps to a single broadcast domain and thus a single service instance, the affected PE only needs to advertise a single B-MAC/I-SID route. However, if the failed EVC carries multiple VLANs each with its own broadcast domain, then the affected PE needs to advertise multiple B-MAC/I-SID routes, i.e., one route for each VLAN (broadcast domain), meaning one route for each I-SID. Each B-MAC/I-SID route basically instructs the remote PEs to perform flushing for C-MACs corresponding to the advertised B-MAC only for the advertised I-SID.

The C-MAC flushing based on a B-MAC/I-SID route works fine when there are only a few VLANs (e.g., I-SIDs) per EVC. However, if the number of I-SIDs associated with a failed EVC is large, then it is **RECOMMENDED** to assign a B-MAC per vES, and upon EVC failure, the affected PE simply withdraws this B-MAC message to other PEs.

5.3. Port Failure Handling for Single-Active vESes in EVPN

When many EVCs are aggregated via a single physical port on a PE, where each EVC corresponds to a vES, then the port failure impacts all the associated EVCs and their corresponding vESes. If the number of EVCs corresponding to the Single-Active vESes for that physical port is in the

thousands, then thousands of service instances are impacted. Therefore, the propagation of failure in BGP needs to address all these impacted service instances. In order to achieve this, the following extensions are added to the baseline EVPN mechanism:

1. The PE **MAY** color each Ethernet A-D per ES route for a given vES, as described in Section 4.2.1. The PE **SHOULD** use the MAC physical port by default. The receiving PEs take note of this color and create a list of vESes for this color.
2. The PE **MAY** advertise a special Grouping Ethernet A-D per ES route for that color, which represents all the vESes associated with the port.
3. Upon a port failure (e.g., an ENNI failure), the PE **MAY** send a mass-withdraw message by withdrawing the Grouping Ethernet A-D per ES route.
4. When this message is received, the remote PE **MAY** detect the special vES mass-withdraw message by identifying the Grouping Ethernet A-D per ES route. The remote PEs **MAY** then access the list of vESes created per item 1 for the specified color and locally initiate MAC address invalidating procedures for each of the vESes in the list.

In scenarios where a logical ENNI is used, the above procedure equally applies. The logical ENNI is represented by a Grouping Ethernet A-D per ES route where the Type 3 ESI and the 6 bytes used in the ENNI's ESI MAC address field are used as a color for the vESes as described above and in Section 4.2.1.

5.4. Port Failure Handling for Single-Active vESes in PBB-EVPN

When many EVCs are aggregated via a single physical port on a PE, where each EVC corresponds to a vES, then the port failure impacts all the associated EVCs and their corresponding vESes. If the number of EVCs corresponding to the Single-Active vESes for that physical port is in the thousands, then thousands of service instances (I-SIDs) are impacted. In such failure scenarios, the following two MAC flushing mechanisms per [RFC7623] can be performed.

1. If the MAC address of the physical port is used for PBB encapsulation as B-MAC SA, then upon the port failure, the PE **MUST** use the EVPN MAC route withdrawal message to signal the flush.
2. If the PE's shared MAC address is used for PBB encapsulation as B-MAC SA, then upon the port failure, the PE **MUST** re-advertise this MAC route with the MAC Mobility extended community to signal the flush.

The first method is recommended because it reduces the scope of flushing the most.

As noted above, the advertisement of the extended community along with the B-MAC route for coloring purposes is optional and only recommended when there are many vESes per physical port and each vES is associated with a very large number of service instances (i.e., a large number of I-SIDs).

If there are large numbers of service instances (i.e., I-SIDs) associated with each EVC, and if there is a B-MAC assigned per vES as recommended in the above section, then in order to handle port failure efficiently, the following extensions are added to the baseline PBB-EVPN mechanism:

1. Each vES **MAY** be colored with a MAC address representing the physical port like the coloring mechanism for EVPN. In other words, each B-MAC representing a vES is advertised with the 'color' of the physical port per Section 4.2.2. The receiving PEs take note of this color being advertised along with the B-MAC route, and for each such color, they create a list of vESes associated with this color.
2. The PE **MAY** advertise a special Grouping B-MAC route for that color (consisting of a port MAC address by default), which represents all the vESes associated with the port.
3. Upon a port failure (e.g., ENNI failure), the PE **MAY** send a mass-withdraw message by withdrawing the Grouping B-MAC route.
4. When this message is received, the remote PE **MAY** detect the special vES mass-withdraw message by identifying the Grouping B-MAC route. The remote PEs **MAY** then access the list created in (1) above for the specified color and flush all C-MACs associated with the failed physical port.

5.5. Fast Convergence in EVPN and PBB-EVPN

As described above, when many EVCs are aggregated via a physical port on a PE, and where each EVC corresponds to a vES, the port failure impacts all the associated EVCs and their corresponding vESes. Two actions must be taken as the result of such a port failure:

- For EVPN, initiate the mass-withdraw procedure for all vESes associated with the failed port to invalidate MACs and for PBB-EVPN to flush all C-MACs associated with the failed port across all vESes and the impacted I-SIDs
- Use DF election for all impacted vESes associated with the failed port

Section 5.3 already describes how to perform a mass withdraw for all affected vESes and invalidate MACs using a single BGP withdrawal of the Grouping Ethernet A-D per ES route. Section 5.4 describes how to only flush C-MAC addresses associated with the failed physical port (e.g., optimum C-MAC flushing) as well as, optionally, the withdrawal of a Grouping B-MAC route.

This section describes how to perform DF election in the most optimal way, e.g., by triggering DF election for all impacted vESes (which can be very large) among the participating PEs via a single BGP message as opposed to sending a large number of BGP messages (one per vES). This section assumes that the MAC flushing mechanism described in Section 5.4 and route coloring are used.

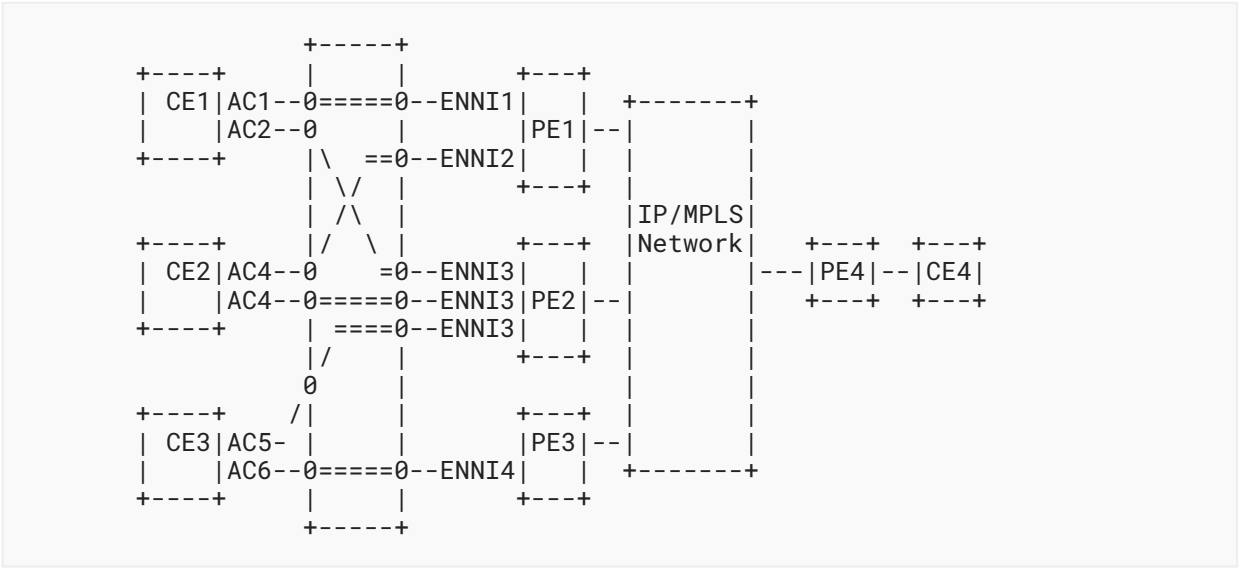

Figure 4: Fast Convergence Upon ENNI Failure

As discussed in Section 4.2, it is highly desirable to have a mass-withdraw mechanism similar to the one in [RFC7432]. Although such an optimization is desirable, it is **OPTIONAL**. If the optimization is implemented, the following procedures are used:

1. When a vES is configured, the PE advertises the ES route for this vES with a color that corresponds to the associated physical port.
2. All receiving PEs within the redundancy group record this color and compile a list of vESes associated with it.
3. Additionally, the PE advertises a Grouping Ethernet A-D per ES route for EVPN, and a Grouping B-MAC route for PBB-EVPN, which corresponds to the color and vES grouping.
4. In the event of a port failure, such as an ENNI failure, the PE withdraws the previously advertised Grouping Ethernet A-D per ES route or Grouping B-MAC route associated with the failed port. The PE should prioritize sending these Grouping route withdrawal messages over the withdrawal of individual vES routes affected by the failure. For instance, as depicted in Figure 4, when the physical port associated with ENNI3 fails on PE2, it withdraws the previously advertised Grouping Ethernet A-D per ES route. Upon receiving this withdrawal message, other multihoming PEs (such as PE1 and PE3) recognize that the vESes associated with CE1 and CE3 are impacted, based on the associated color, and thus initiate the DF election procedure for these vESes. Furthermore, upon receiving this withdrawal message, remote PEs (such as PE4) initiate the failover procedure for the vESes associated with CE1 and CE3 and switch to the other PE for each vES redundancy group.
5. On reception of Grouping Ethernet A-D per ES route or Grouping B-MAC route withdrawal, other PEs in the redundancy group initiate DF election procedures across all their affected vESes.

6. The PE with the physical port failure (ENNI failure) sends a vES route withdrawal for every impacted vES. Upon receiving these messages, the other PEs clear up their BGP tables. It should be noted that the vES route withdrawal messages are not used for executing DF election procedures by the receiving PEs when Grouping Ethernet A-D per ES route or Grouping B-MAC route withdrawal has been previously received.

6. Security Considerations

All the security considerations in [RFC7432] and [RFC7623] apply directly to this document because this document leverages the control and data plane procedures described in those documents.

This document does not introduce any new security considerations beyond that of [RFC7432] and [RFC7623] because advertisements and the processing of ES routes for vES in this document follow that of physical ES in those RFCs.

7. IANA Considerations

This document has no IANA actions.

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 1997, <https://www.rfc-editor.org/info/rfc2119>.

[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>.

[RFC7623] Sajassi, A., Ed., Salam, S., Bitar, N., Isaac, A., and W. Henderickx, "Provider Backbone Bridging Combined with Ethernet VPN (PBB-EVPN)", RFC 7623, DOI 10.17487/RFC7623, September 2015, <https://www.rfc-editor.org/info/rfc7623>.

[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>.

[RFC8365] Sajassi, A., Ed., Drake, J., Ed., Bitar, N., Shekhar, R., Uttaro, J., and W. Henderickx, "A Network Virtualization Overlay Solution Using Ethernet VPN (EVPN)", RFC 8365, DOI 10.17487/RFC8365, March 2018, <https://www.rfc-editor.org/info/rfc8365>.

[RFC9135] Sajassi, A., Salam, S., Thoria, S., Drake, J., and J. Rabadan, "Integrated Routing and Bridging in Ethernet VPN (EVPN)", RFC 9135, DOI 10.17487/RFC9135, October 2021, <https://www.rfc-editor.org/info/rfc9135>.

[RFC9541] Rabadan, J., Ed., Sathappan, S., Nagaraj, K., Miyake, M., and T. Matsuda, "Flush Mechanism for Customer MAC Addresses Based on Service Instance Identifier (I-SID) in Provider Backbone Bridging EVPN (PBB-EVPN)", RFC 9541, DOI 10.17487/RFC9541, March 2024, <https://www.rfc-editor.org/info/rfc9541>.

8.2. Informative References

[MEF63] Metro Ethernet Forum, "Subscriber Ethernet Services Definitions", MEF Standard, MEF 6.3, November 2019, <https://www.mef.net/resources/mef-6-3-subscriber-ethernet-service-definitions>.

[RFC4377] Nadeau, T., Morrow, M., Swallow, G., Allan, D., and S. Matsushima, "Operations and Management (OAM) Requirements for Multi-Protocol Label Switched (MPLS) Networks", RFC 4377, DOI 10.17487/RFC4377, February 2006, <https://www.rfc-editor.org/info/rfc4377>.

[RFC6310] Aissaoui, M., Busschbach, P., Martini, L., Morrow, M., Nadeau, T., and Y. Stein, "Pseudowire (PW) Operations, Administration, and Maintenance (OAM) Message Mapping", RFC 6310, DOI 10.17487/RFC6310, July 2011, <https://www.rfc-editor.org/info/rfc6310>.

[RFC7023] Mohan, D., Ed., Bitar, N., Ed., Sajassi, A., Ed., DeLord, S., Niger, P., and R. Qiu, "MPLS and Ethernet Operations, Administration, and Maintenance (OAM) Interworking", RFC 7023, DOI 10.17487/RFC7023, October 2013, <https://www.rfc-editor.org/info/rfc7023>.

[RFC7080] Sajassi, A., Salam, S., Bitar, N., and F. Balus, "Virtual Private LAN Service (VPLS) Interoperability with Provider Backbone Bridges", RFC 7080, DOI 10.17487/RFC7080, December 2013, <https://www.rfc-editor.org/info/rfc7080>.

[RFC7209] Sajassi, A., Aggarwal, R., Uttaro, J., Bitar, N., Henderickx, W., and A. Isaac, "Requirements for Ethernet VPN (EVPN)", RFC 7209, DOI 10.17487/RFC7209, May 2014, <https://www.rfc-editor.org/info/rfc7209>.

[RFC8584] Rabadan, J., Ed., Mohanty, S., Ed., Sajassi, A., Drake, J., Nagaraj, K., and S. Sathappan, "Framework for Ethernet VPN Designated Forwarder Election Extensibility", RFC 8584, DOI 10.17487/RFC8584, April 2019, <https://www.rfc-editor.org/info/rfc8584>.

[RFC9252] Dawra, G., Ed., Talaulikar, K., Ed., Raszuk, R., Decraene, B., Zhuang, S., and J. Rabadan, "BGP Overlay Services Based on Segment Routing over IPv6 (SRv6)", RFC 9252, DOI 10.17487/RFC9252, July 2022, <https://www.rfc-editor.org/info/rfc9252>.

Acknowledgements

The authors would like to thank Mei Zhang, Jose Liste, and Luc André Burdet for their reviews of this document and their feedback.

Authors' Addresses

Ali Sajassi
Cisco Systems
Email: sajassi@cisco.com

Patrice Brissette
Cisco Systems
Email: pbrisset@cisco.com

Rick Schell
Independent
Email: rick_schell@outlook.com

John E. Drake
Independent
Email: je_drake@yahoo.com

Jorge Rabadan
Nokia
Email: jorge.rabadan@nokia.com