BESS Workgroup
Internet-Draft
Intended status: Standards Track
Expires: 3 September 2026

J. Rabadan, Ed.
S. Sathappan
Nokia
A. Sajassi
Cisco
W. Lin
Juniper
L.A. Burdet
Cisco

2 March 2026

EVPN Inter-Domain Option-B Solution
draft-rabadan-bess-evpn-inter-domain-opt-b-08

Abstract

An EVPN Inter-Domain interconnect solution is required when two or more sites of the same Ethernet Virtual Private Network (EVPN) are connected to different IGP domains or Autonomous Systems (AS) and need to communicate. The Inter-Domain Option-B connectivity model is one of the most popular solutions for this type of EVPN connectivity. While several documents address specific aspects of this interconnect approach, none provide a comprehensive overview of how Inter-Domain Option-B connectivity affects EVPN procedures. This document examines the behavior of EVPN procedures in an Inter-Domain Option-B network and proposes solutions to the identified issues.

Table of Contents

1. Introduction

An EVPN Inter-Domain interconnect solution is required if two or more sites of the same Ethernet Virtual Private Network (EVPN) [I-D.ietf-bess-rfc7432bis] are connected to different IGP domains or Autonomous Systems (AS) and need to communicate. In general, there are different types of EVPN Inter-Domain models, classified based on the procedures implemented on the Border Routers that interconnect the domains. The industry typically groups these models into three categories:

* EVPN Service Interworking Solution: also referred to as the Service Gateway solution, this approach uses Border Routers that instantiate Virtual Routing and Forwarding tables (MAC-VRFs and/or IP-VRFs) and perform lookups (after decapsulating the transport headers) to forward packets between domains. [RFC9014], [I-D.ietf-bess-evpn-vpws-gateway] and [I-D.ietf-bess-evpn-ipvpn-interworking] specify the Service Gateway solution for EVPN ELAN, VPWS and Layer-3 services, respectively.

* Inter-Domain Option-B Solution: described in [RFC8365] section 10, this solution interconnects EVPN services by using Border Routers that rewrite the EVPN BGP next hops and program swap operations for the VNIs or MPLS labels (depending on whether the encapsulation is NVO-based or MPLS-based). The "Option-B" term refers to the resemblance of this model to the Multi-AS "type B" interconnect for IP-VPN defined in [RFC4364]. However, in this case, the procedures apply to the EVPN address family. This solution does not require the instantiation of Virtual Routing and Forwarding tables (VRFs) on the Border Routers.

* Inter-Domain Transport Solution: refers to any Inter-Domain solution that provides connectivity purely at the transport layer, without instantiating VRFs, rewriting EVPN BGP next hops, or programming swap operations for EVPN service identifiers (such as VNIs or MPLS service labels) on the Border Routers. An example is the Inter-AS Option-C model described in [RFC4364] section 10 subsection "c" - with the distinction that here, the procedures would be applied to EVPN routes rather than VPN-IPv4 or VPN-IPv6 routes.

The Inter-Domain Option-B connectivity model is one of the most widely adopted solutions for inter-domain EVPN connectivity because it provides domain isolation (avoiding the need to leak PE loopbacks between domains) while also eliminating the requirement to instantiate VRFs on Border Routers. While several documents reference this type of interconnect solution and define certain aspects of it, none provide a comprehensive summary of the impact of Inter-Domain Option-B connectivity on EVPN procedures. This document examines the behavior of EVPN procedures in an Inter-Domain Option-B network for:

* Multi-Homing

* EVPN E-Tree

* BUM and IP Multicast forwarding using Ingress Replication or Point-to-Multi-Point tunnels

* Other EVPN services and including Network Virtualization Overlay (NVO) encapsulations or MPLS-based encapsulations

and provides guidelines for the identified issues.

1.1. Terminology and Conventions

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.

* 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 BD, then the Ethernet segment is defined to be operating in All-Active redundancy mode.

* BD: Broadcast Domain. An EVI may be comprised of one BD (VLAN-based or VLAN Bundle services) or multiple BDs (VLAN-aware Bundle services). This document makes use of the term "BD" as described in [RFC9625] section 1.1.4.

* BR: Border Router, router that provides connectivity between domains, typically an Area Border Router (ABR) or Autonomous System Border Router (ASBR).

* BUM traffic: Broadcast, Unknown unicast and Multicast traffic.

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

* DF and non-DF: Designated Forwarder and non Designated Forwarder. In an Ethernet Segment, the Designated Forwarder PE or Service Gateway forwards unicast and BUM traffic. The non-Designated Forwarder PE or Service Gateway blocks BUM traffic (if working in All-Active redundancy mode) or unicast and BUM (if working in Single-Active redundancy mode).

* E-PE: Egress PE.

* Ethernet Segment (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’.

* Ethernet Segment Identifier (ESI): A unique non-zero identifier that identifies an Ethernet segment is called an ’Ethernet Segment Identifier’.

* 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 a PE. In VLAN-based or VLAN Bundle modes [I-D.ietf-bess-rfc7432bis] a BD is equivalent to a MAC-VRF.

* MPLS and non-MPLS NVO tunnels: refer to Multi-Protocol Label Switching (or the absence of it) Network Virtualization Overlay tunnels. Network Virtualization Overlay tunnels use an IP encapsulation for overlay frames, where the source IP address identifies the ingress PE (or ingress Border Router) and the destination IP address the egress PE (or egress Border Router).

* I-PE: Ingress PE.

* IP-VRF: A VPN Routing and Forwarding table for IP routes on an PE. In this document, an IP-VRF is an instantiation of a layer 3 EVPN service in a PE as per [RFC9135][RFC9136].

* IRB: Integrated Routing and Bridging

* IRB Interface: Integrated Bridging and Routing Interface. A virtual interface that connects the Bridge Table and the IP-VRF on an NVE.

* PE: Provider Edge device. In this document a PE can be a Leaf node in a Data Center or a traditional Provider Edge router in an MPLS network.

* 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 BD, then the Ethernet segment is defined to be operating in Single-Active redundancy mode.

* PMSI: Provider Multicast Service Interface.

* SBD: Supplementary Broadcast Domain, a special BD that has an IRB interface to an IP-VRF and it is used in the Optimized Inter-Subnet Multicast model, as described in [RFC9625].

* SR-MPLS SID: Segment Routing MPLS Segment IDentifier.

* SRv6 SID: Segment Routing for IPv6 Segment IDentifier.

* VRF: A generic Virtual Routing and Forwarding table, used in this document to indicate the instantiation of an EVPN service onto a PE. This service can be any supported EVPN service such as layer-2 multipoint services [I-D.ietf-bess-rfc7432bis], EVPN VPWS [RFC8214], EVPN E-Tree [RFC8317], PBB-EVPN [RFC7623], or Layer-3 services as defined in [RFC9135] or [RFC9136].

* VPWS: EVPN Virtual Private Wire Service, as in [RFC8214].

2. EVPN Inter-Domain Option-B General Procedures

The EVPN Inter-Domain Option-B procedures are applied on Border Routers that interconnect the domains. The Ingress and Egress PEs should be configured and operated in the same way they are when communicating with other PEs within their domain. The typical deployments are illustrated in Figure 1 and Figure 2. Figure 1 illustrates an Inter-Domain example where each domain corresponds to a separate IGP instance. In this example, Border Routers BR-1 and

BR-2 establish direct BGP EVPN sessions both between each other and with the Ingress PE (I-PE) and Egress PE (E-PE), respectively. However, Route Reflectors may also be present within each domain. The procedures described in this document remain unchanged regardless of whether Route Reflectors are used in the domains. Note that in this document, the term VRF is used generically and may refer to either a MAC-VRF or an IP-VRF, unless explicitly specified otherwise.

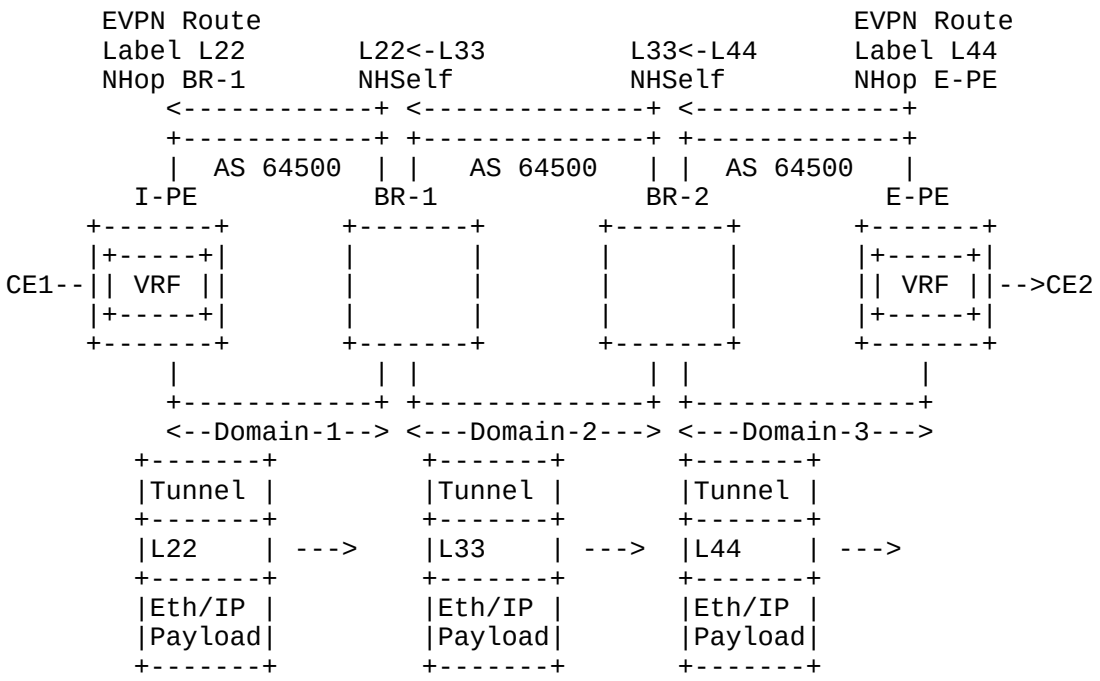

Figure 1: EVPN Inter-Domain Option-B scenario for IGP domains

This document describes also the Inter-Domain Option-B aspects in scenarios such as the one portrayed in Figure 2, where the Border Routers connect different Autonomous Systems. As in the example shown in Figure 1 the procedures remain unchanged if the domains use Route Reflectors.

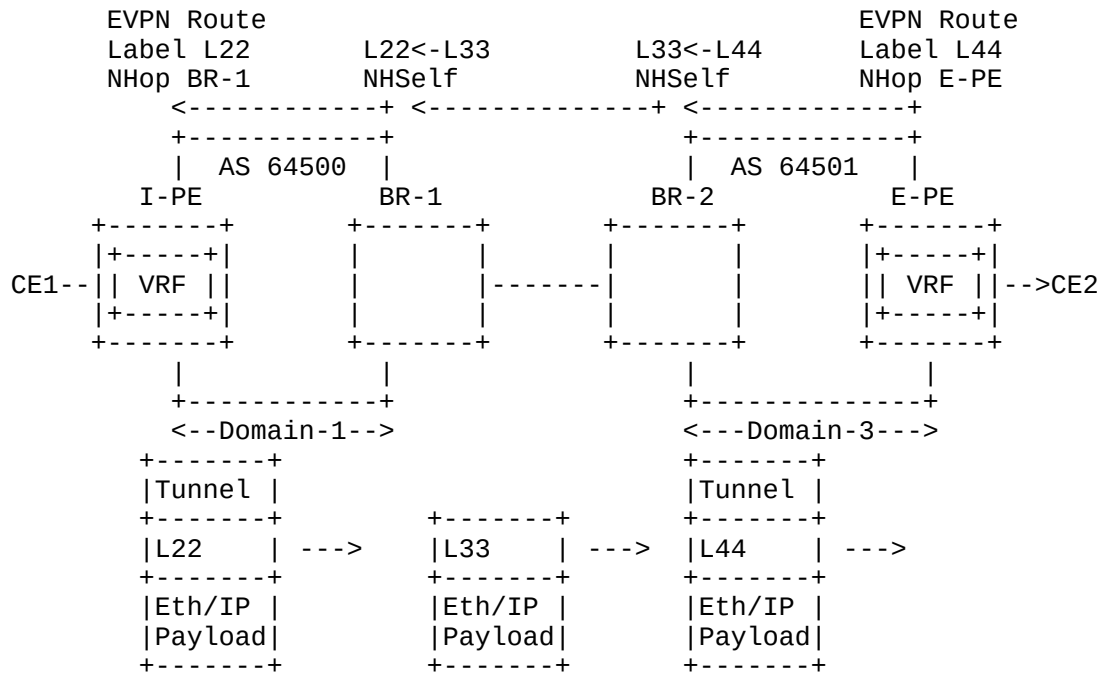

Figure 2: EVPN Inter-Domain Option-B scenario for Multi-AS Backbones

In either Figure 1 or Figure 2, this Inter-Domain Option-B solution involves the redistribution of EVPN routes from domain to domain by the Border Routers. A Border Router learns all the EVPN routes of its own domain - typically via IBGP from the Egress PE or as a client from the domain's Route Reflector - and readvertises those routes to the neighboring Border Router(s), via EBGP or IBGP.

When redistributing EVPN routes to the adjacent Border Routers or Route Reflectors within the adjacent domain, the Border Router updates the Next Hop IP address to its own address and assigns a newly generated EVPN label in the BGP MP_REACH_NLRI message. In effect, this means that the Border Router programs a label swap operation in the data path for the EVPN label. For example, packets received by BR-1 with EVPN label L22 are looked up and switched to the interface to the next domain or Border Router, now with EVPN label L33. The EVPN label in this document can be a 20-bit label (i.e., an MPLS label or Segment Routing MPLS Segment Identifier) or a 24-bit label (i.e., a VNI label for non-MPLS NVO tunnels).

For EVPN routes with 20-bit labels, if the Border Router receives the EVPN route via IBGP, the route is resolved to a transport MPLS or SR-MPLS tunnel that provides reachability to the Egress PE or the

adjacent Border Router. The imported EVPN route is considered valid and redistributed only if the Next Hop is resolved to such a transport tunnel. When the Border Router receives the EVPN route via single-hop EBGP, the next hop is resolved to a local interface associated to the next hop, and packets matching the Forwarding Information Base entry for that route are forwarded with a single label in the label stack, i.e. the swapped EVPN label.

In Inter-Domain Option-B scenarios where transport tunnels in the domains are NVO-based tunnels, the EVPN routes advertised from the egress PEs (and redistributed by the Border Routers) use either 20-bit labels (for MPLS NVO tunnels, e.g., MPLSoGRE) or 24-bit labels (for non-MPLS NVO tunnels, e.g., VXLAN). In these cases, the Border Routers not only swap the label (e.g., VNI) for the NVO packets that they route, but also change the source and destination IP address of the router IP header. Specifically, when the Border Router forwards packets into an adjacent domain, the outer source IP address is set to a local IP address of the Border Router, and the outer destination IP address is given by the next hop of the EVPN route that created the Forwarding Information Base entry.

Key attributes of this solution include:

* Border Routers maintain isolation between domains; for example, BR-2 does not leak the E-PE's loopback address into other domains.

* Border Routers do not require VRFs to be explicitly configured.

* Because VRFs are not used, Border Routers must import all EVPN routes from their domain(s), regardless of Route Targets, and then re-advertise them to adjacent domains, potentially applying RIB-IN or RIB-OUT policies to select which routes to redistribute.

This solution does not impose any changes or requirements on Ingress or Egress PEs, or Route Reflectors. All procedures are implemented solely on the Border Routers and remain transparent to the Ingress and Egress PEs.

[RFC8365] section 10.2 is the existing specification for Inter-Domain Option-B in case EVPN uses encapsulations with 20-bit or 24-bit labels, and in particular for the scenario in Figure 2. This document clarifies that the same procedures and considerations also apply to the scenario shown in Figure 1. Although the general operation of the Border Routers on the received EVPN routes is described above, Section 2.1 further clarifies the expected behavior for each EVPN route type.

2.1. Border Router procedures on EVPN routes

The Border Router behavior described in Section 2 can be summarized in the following tasks performed on each received EVPN BGP UPDATE:

* The Border Router accepts any EVPN route from the Border Routers and PEs to which it is connected (possibly filtering some of the routes via RIB-IN import policies).

* It extracts the EVPN label of each EVPN route, either from the NLRI (Network Layer Reachability Information) or from an attribute included in the BGP UPDATE.

* It programs an EVPN label swap operation in the data path, which switches the extracted EVPN label to a locally generated new EVPN label for the same EVPN route.

* It readvertises the EVPN route (assuming the operation is allowed by policy) with:

  a. Next Hop Self, i.e., a new IP address owned by the Border Router itself

  b. The locally generated EVPN label for the route

However, there are important subtleties in handling certain EVPN route types that must be clarified to ensure interoperability across implementations. We distinguish between EVPN Labeled Routes and EVPN Unlabeled Routes.

2.1.1. EVPN Labeled Routes

EVPN Labeled Routes are those that carry EVPN Labels or demultiplexors in the NLRI or an attribute of the BGP UPDATE. If those EVPN Labels are used in the Forwarding Information Base of the Border Router to forward packets between domains, the Label is extracted and added to the Forwarding Information Base associated to a swap operation. If these EVPN Labels are not used to forward packets between domains, but they indicate certain properties of the route, e.g.,: ESI Labels or E-Tree Labels, then the Labels are not extracted, programmed or changed when the route is readvertised. The previous statements MUST be applied to existing and future EVPN route types in Inter-Domain Option-B networks. As an example:

a. Ethernet Auto-Discovery per Ethernet Segment Route (or route type 1 per ES)

Defined in [I-D.ietf-bess-rfc7432bis], this route signals the multi-homing mode information, as well as the value of the ESI label, encoded in the ESI Label extended community. It is used for fast convergence in case of multi-homed PE failures, via the "Mass Withdraw per Ethernet Segment" procedure. When used with an ESI of zero, the route is used to advertised a Leaf Label in the E-Tree extended community [RFC8317]. The Leaf Label is used by the Ingress PE when forwarding BUM traffic generated from a Leaf Attachment Circuit. Both labels, ESI label and Leaf label, are not used for packet forwarding at the Border Router and therefore the Border Router does not extract them. The Border Router MUST preserve the content of the ESI label or the E-Tree extended community when readvertising the route to the adjacent domain. Although the next hop self operation is performed on the route by the Border Router, none of the NLRI fields are changed when readvertising the route to the adjacent domain.

b. Ethernet Auto-Discovery per EVPN Instance Route (or route type 1 per EVI)

Defined in [I-D.ietf-bess-rfc7432bis], this route signals the forwarding information associated to the local EVPN-VPWS Attachment Circuit [RFC8214], and when used with a non-zero ESI, it also performs the Aliasing and Backup procedures for multi-homing in EVPN services. The EVPN label encoded in the NLRI of this route is used when forwarding packets, hence the label must be extracted by the Border Router and programmed in the Forwarding Information Base for a swap operation. Besides the next hop self operation and the new valid label to be encoded in the route, the Border Router does not change any other field of the route. This includes the content of the EVPN Layer-2 Attributes extended community advertised with the route. [RFC8214] section 4 discusses the Inter-domain Option-B solution for EVPN-VPWS.

c. MAC/IP Advertisement Route (or route type 2)

Defined in [I-D.ietf-bess-rfc7432bis], this route advertises forwarding information for MAC and IP addresses that are used by the Ingress PE to populate the layer-2 Forwarding Information Base, the Address Resolution Protocol or Neighbor Discovery tables [RFC9161] or even the layer-3 Forwarding Information Base [RFC9135]. The route’s NLRI contains a mandatory EVPN label, Label1, and an optional Label2. In addition to the next hop self operation, a Border Router that receives a route type 2, with only Label1, needs to extract Label1 from the NLRI, program its value in the Forwarding Information Base, and generate a new valid label that is encoded in Label1 when redistributing the

route to the adjacent domain. If the received route type 2 contains a value for both, Label1 and Label2, the Border Router needs to program two separate entries in the Forwarding Information Base (for the value in Label1 and the value in Label2) and generate two valid Label1 and Label2 values. The rest of the information in the route, including EVPN extended communities and Default Gateway extended community, is preserved by the Border Router when readvertising. This method at the Border Router is applied irrespective of the Egress PE using an EVPN label per VRF, EVPN label per Ethernet Segment or EVPN label per MAC address. However, using a label per VRF on the Egress PEs has the least impact on the Border Routers Forwarding Information Base scale, compared to label per MAC or label per Ethernet Segment.

d. Inclusive Multicast Ethernet Tag Route (or route type 3)

Also defined in [I-D.ietf-bess-rfc7432bis], this route is used for the auto-discovery of the remote PEs attached to the same Broadcast domain, as well as the creation of the flooding tree used to forward BUM traffic by the PEs attached to the same Broadcast Domain. The route type 3 does not contain any EVPN label in its NLRI. The Provider Tunnel (P-Tunnel) identification is carried in the PMSI Tunnel Attribute. When used for Ingress Replication or Assisted Replication tunnel types, the PMSI Tunnel Attribute contains an EVPN Label (downstream allocated) that is extracted by the Border Router and programmed in the Forwarding Information Base in the same way as for the EVPN labels in the routes above. The Border Router generates a valid new label that is encoded in the PMSI Tunnel Attribute of the route readvertised to the adjacent domain. In addition to the next hop self and label swap operation, the Border Router preserves all the fields in the NLRI (including the Originating Router’s IP Address) and the attributes of the routes, including the Tunnel Identifier of the PMSI Tunnel Attribute and the Layer 2 Attributes extended community. When the route type 3 uses a P-Tunnel different than Ingress Replication, the Border Router should carry out the segmentation procedures specified in [RFC9572].

e. IP Prefix Route (or route type 5)

Specified in [RFC9136], this route allows the Egress PEs to advertise the IPv4 or IPv6 prefixes that they have learned locally in their IP-VRF. The route’s NLRI contains an EVPN label that the Option-B Border Router needs to extract and program in the Forwarding Information Base, along with a label swap operation. Beyond performing next hop self and generating a new valid EVPN label for the IP Prefix route readvertised to the

adjacent domain, the Border Router does not modify any of the fields in the NLRI and preserves all the attributes along with the route, including EVPN extended communities.

As shown in [RFC9136] Table 1, valid EVPN IP Prefix routes may be advertised either with a zero label or with a non-zero valid label. In all cases, the Border Routers update the next hop to their own address. However, generating a new valid EVPN label and re-advertising it in the IP Prefix route sent to the adjacent domain is required only when the received IP Prefix route carries a non-zero valid label. When the received EVPN label has a value of zero, the re-advertised IP Prefix route MUST also retain a label value of zero.

f. Per-Region I-PMSI A-D Route (or route type 9)

Used for P-Tunnel Segmentation on Border Routers, its definition and procedures are described in [RFC9572].

g. S-PMSI A-D Route (or route type 10)

Also defined in [RFC9572], the Border Router should follow the same procedures as for the Inclusive Multicast Ethernet Tag Route above.

2.1.2. EVPN Unlabeled Routes

Examples or EVPN Unlabeled Routes are:

* Ethernet Segment Route (or route type 4)

* Selective Multicast Ethernet Tag Route (or route type 6)

* Multicast Membership Report Synch Route (or route type 7)

* Multicast Leave Synch Route (or route type 8)

* Leaf Auto-Discovery Route (or route type 11)

The Border Router receiving these routes simply redistributes the routes to the adjacent domain with a next hop of itself, and preserving all the attributes that the routes contain.

3. EVPN Inter-Domain Option-B and Multi-Homing

This section summarizes the issues of the Inter-Domain Option-B associated to EVPN Multi-Homing. Figure 3 illustrates the use of multi-homing in an Inter-Domain Option-B example.

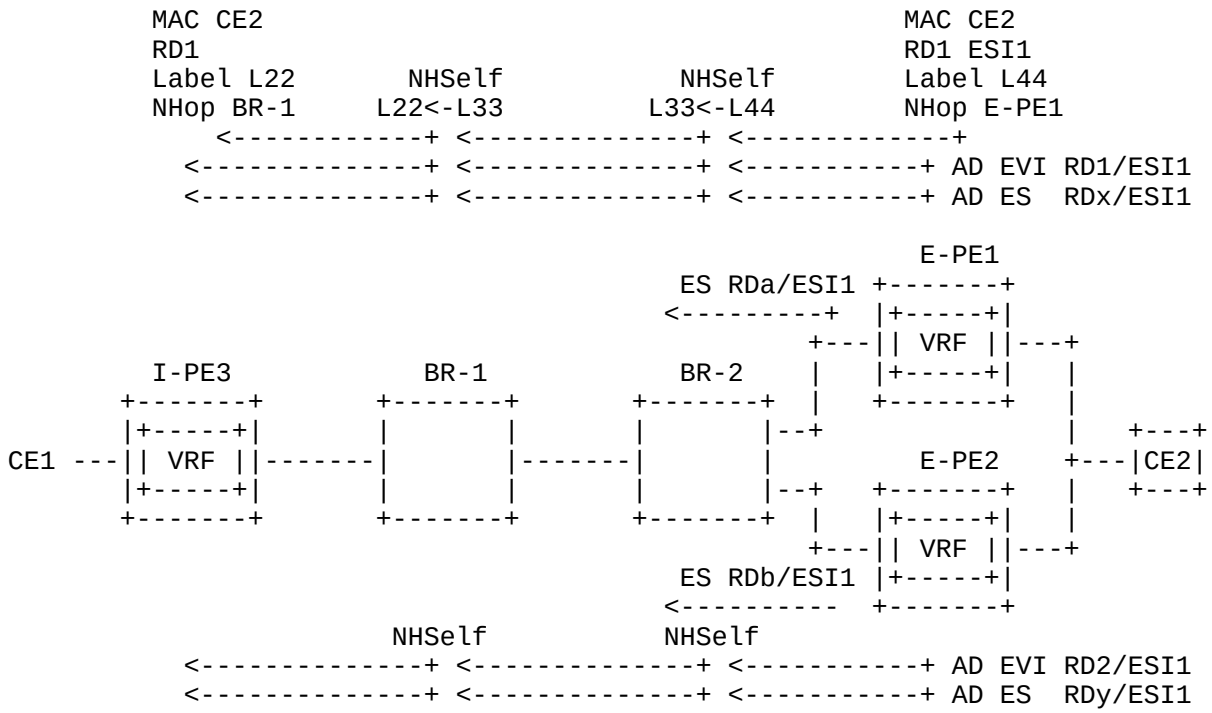

Figure 3: EVPN Inter-Domain Option-B and multi-homing

The Border Router rewriting the EVPN multi-homing routes next hop has an impact on the EVPN multi-homing procedures that follow:

* Mass Withdrawal

* Aliasing and Backup Path procedures

* Designated Forwarder Election and AC-Influenced Capability

* Split Horizon Filtering

This section defines a new sub-TLV — the Egress PE Identifier sub-TLV — for the Tunnel Encapsulation TLV of the BGP Tunnel Encapsulation Attribute [RFC9012], and describes its use to address the issues associated with the EVPN procedures described above in inter-domain Option B networks.

3.1. The Egress PE Identifier in the BGP Tunnel Encapsulation Attribute

This document defines a new Sub-TLV, the Egress PE Identifier Sub-TLV, to be carried in the Tunnel Encapsulation TLV of the BGP Tunnel Encapsulation Attribute [RFC9012]. IANA is requested to assign a new Type Code (TBDxx) for this Sub-TLV. The Value field of this Sub-TLV is composed of three subfields, as illustrated in Figure 4:

1. a Reserved subfield

2. a two-octet Address Family subfield

3. an Address subfield, whose length depends upon the Address Family value

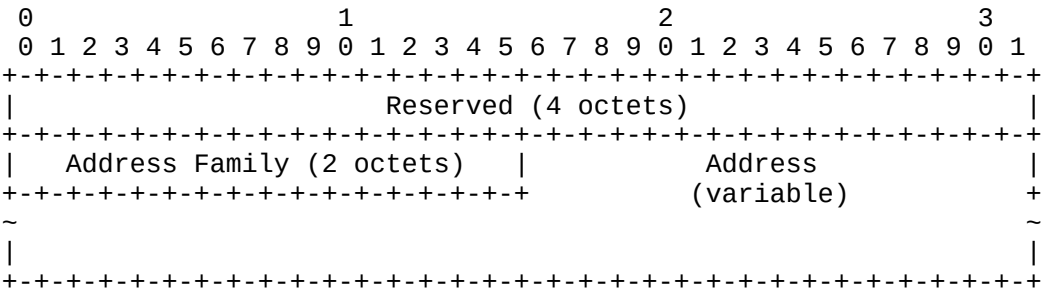

Figure 4: Egress PE Identifier Sub-TLV

The subfields are encoded as follows:

* The Reserved subfield SHOULD be originated as zero. It MUST be disregarded on receipt, and it MUST be propagated unchanged.

* The Address Family subfield contains a value from IANA's "Address Family Numbers" registry [IANA-ADDRESS-FAM]. This document assumes that the Address Family is either IPv4 or IPv6; use of other address families is outside the scope of this document.

* If the Address Family subfield contains the value for IPv4, the Address subfield MUST contain an IPv4 address (a /32 IPv4 prefix).

* If the Address Family subfield contains the value for IPv6, the Address subfield MUST contain an IPv6 address (a /128 IPv6 prefix).

* In a given BGP UPDATE, the address family (IPv4 or IPv6) of an Egress PE Identifier sub-TLV is independent of the address family of the UPDATE itself, or the address family of any of the fields

in the MP_REACH_NLRI attribute included in the UPDATE. For example, an UPDATE whose NLRI is an EVPN IP Prefix route including an IPv4 prefix may have a Tunnel Encapsulation attribute containing Egress PE Identifier sub-TLVs that contain IPv6 addresses. Also, different tunnels represented in the Tunnel Encapsulation attribute may have Egress PE Identifiers of different address families.

* The lengths of the sub-TLV's Value field permitted in this document are:

  - 10, if the Address Family subfield contains the value for IPv4.

  - 22, if the Address Family subfield contains the value for IPv6.

* The IPv4 or IPv6 address carried in the Egress PE Identifier sub-TLV uniquely identifies the PE within the network and SHOULD match the value of the PE's Originating Router's IP Address.

The following error handling rules apply to the Egress PE Identifier Sub-TLV:

* The Egress PE Identifier Sub-TLV MAY be included in any BGP UPDATE message whose AFI/SAFI is among those defined in [RFC9012]. However, this document specifies its use only with AFI/SAFI 25/70 (EVPN). The use of this Sub-TLV with other AFI/SAFI combinations is outside the scope of this document.

* When the Tunnel Encapsulation attribute is included in an EVPN UPDATE message, each TLV MAY contain an Egress PE Identifier sub-TLV. If present, only a single Egress PE Identifier sub-TLV MUST appear within that TLV. If a TLV has more than one Egress PE Identifier sub-TLV, the TLV is treated as if it had a malformed Egress PE Identifier sub-TLV.

* The Address value included in the Egress PE Identifier sub-TLV MUST be a valid address or the sub-TLV is considered malformed. If the IPv4 or IPv6 Address contained in the Address subfield is valid but not reachable, the sub-TLV is not considered to be malformed.

* Malformed Egress PE Identifier sub-TLVs are treated as unrecognized sub-TLVs. If the route carrying the Tunnel Encapsulation Attribute is propagated with the attribute, the unrecognized Egress PE Identifier sub-TLV MUST remain in the attribute.

* Other than the above, the Egress PE Identifier Sub-TLV follows the validation rules in [RFC9012] section 13.

Tunnel TLVs with which the Egress PE Identifier sub-TLV MAY be advertised include:

* MPLS (10), MPLS in GRE Encapsulation (11)

* VXLAN (8), NVGRE (9), VXLAN GPE (12), MPLS in UDP (13), Geneve (19)

Other tunnel TLVs are outside the scope of this document.

The Egress PE Identifier sub-TLV MAY coexist with other Sub-TLVs in the same Tunnel TLV, such as the Tunnel Egress Endpoint Sub-TLV.

3.2. Mass Withdraw in Inter-Domain Option-B

This section reviews the mass withdraw issue in Inter-Domain Option-B networks, and updates the EVPN procedures to support mass withdraw.

3.2.1. Problem Statement

The limitations of the mass withdraw procedures when the multi-homed egress PEs and the ingress PEs are in different domains are explained in [RFC8365] section 10.2.2.

As a refresher, consider the example in Figure 3, where CE2 is multi-homed to egress PE1 and PE2 (on Ethernet Segment ES1 with identifier ESI1), and the ingress PE3 resides in a different domain. In this scenario, only E-PE1 advertises the MAC/IP route for MAC CE2, whereas both E-PE1 and E-PE2 advertise the A-D per ES and A-D per EVI routes for ESI1.

Because the Border Routers rewrite the next hops of all routes, I-PE3 cannot correlate the MAC/IP Advertisement route with the A-D per ES route advertised from the same E-PE, since the only mechanism in [I-D.ietf-bess-rfc7432bis] to correlate A-D per ES and MAC/IP Advertisement routes advertised from the same E-PE is the route next hop. For example, if the link from CE2 to E-PE1 fails, E-PE1 sends a MP_UNREACH_NLRI message for the A-D per ES route and A-D per EVI route for ESI1. The messages get to I-PE3 and are processed, however, I-PE3 is unable to correlate the withdrawn A-D per ES route with the MAC/IP Advertisement route for CE2. Consequently, it does not perform a mass withdraw of the MACs associated with ESI1, as long as at least one A-D per ES route for ESI1 remains. Note that the Route Distinguishers of the MAC/IP Advertisement and A-D per ES routes advertised by E-PE1 are different, so they cannot be associated by RD alone.

As also explained in [RFC8365] section 10.2.2, a "mass withdraw per EVI" is possible though, because the A-D per EVI routes and MAC/IP Advertisement routes advertised from the same PE and ES can be correlated based on the Route Distinguisher. In Figure 3, if the link between CE2 and E-PE1 fails, I-PE3 receives the A-D per EVI route withdrawal from E-PE1 and can withdraw all the MACs related to the MAC/IP Advertisement routes that match the Route Distinguisher of the A-D per EVI route, i.e., RD1 in the example, hence MAC CE2 is flushed on I-PE3. Although this example focuses on MAC address withdrawal, the same issue also affects IP Prefixes as described in [I-D.ietf-bess-evpn-ip-aliasing].

This document assumes that “mass withdraw per EVI” is the default behavior supported by all PEs and Border Routers that do not implement this specification. When "mass withdraw per EVI" is used, unique RDs MUST be used on all the PEs attached to the same EVI.

3.2.2. Mass Withdraw in Inter-Domain Option-B procedures

The (per Ethernet Segment) mass withdraw limitation imposed by Border Routers (for MAC and IP Prefix routes) is addressed by including a unique IP address of the egress PE in the Address subfield of the Egress PE Identifier Sub-TLV within the BGP Tunnel Encapsulation Attribute. To support mass withdraw, this Sub-TLV MUST be advertised along with the Ethernet A-D per ES, MAC/IP Advertisement, and IP Prefix routes. The IP address encoded in the Egress PE Identifier Sub-TLV is an IPv4 or IPv6 address that uniquely identifies the egress PE.

The Egress PE Identifier may be included in the advertised EVPN routes either by the Egress PE itself or by a Border Router:

* When the egress PE advertises the Egress PE Identifier, the value of the Address subfield SHOULD match the BGP Next Hop value used in the corresponding EVPN routes.

* When a Border Router attaches an Egress PE Identifier to an imported route before readvertising it, the Address subfield value MUST correspond to the BGP next-hop address of the imported route, and not to the Border Router's own BGP next-hop address.

A Border Router MUST NOT modify the Egress PE Identifier sub-TLV when readvertising an EVPN route that already carries it.

When the egress PE or the Border Router include the Egress PE Identifier in the BGP Tunnel Encapsulation Attribute of the Ethernet A-D per ES, MAC/IP Advertisement, or IP Prefix routes, the ingress PE can correlate all routes originating from the same egress PE by comparing the common Egress PE Identifier IP address present in those routes. This mechanism enables mass withdraw per ES in Inter-Domain Option-B networks.

If the ingress PE does not receive the Egress PE Identifier sub-TLV in the received Ethernet A-D per ES routes for a given Ethernet Segment, the ingress PE attempts to correlate the A-D per ES routes with the other EVPN routes for the Ethernet Segment based on the RD Administrator Subfield of the routes, as follows:

* EVPN routes type 2 and 5 can be correlated with the A-D per ES routes from the same PE based on the Administrator subfield of the Route Distinguishers (RDs).

* That is, in Figure 3:

  - Suppose E-PE1 advertises the A-D per ES route with Route Distinguisher RDx = <RD1:0> and the MAC/IP Advertisement route with <RD1:1>, where "RD1" is the Administrator subfield of the Route Distinguisher.

  - E-PE2 allocates "RD2" as Administrator subfield for its A-D per ES and MAC/IP Advertisement routes.

  - Now, when the A-D per ES route from E-PE1 is withdrawn, I-PE3 can perform a mass withdraw operation by assuming that all the MACs from the MAC/IP Advertisement routes with RD1 as Administrator subfield were advertised by the same E-PE1 that failed and withdrew the A-D per ES route.

* This same approach can also be applied to support mass withdraw of IP Prefix routes.

If the ingress PE does not receive the Egress PE Identifier sub-TLV on the A-D per ES routes for the Ethernet Segment, and the RD Administrator Subfields of A-D per ES and routes type 2 or 5 do not match, the ingress PE can only perform "mass withdraw per EVI".

3.3. Aliasing and Backup Path Procedures

The Aliasing and Backup Path procedures work in an Inter-Domain Option-B solution as per [RFC8365], section 10.2. That is, since EVPN MAC/IP Advertisement routes and A-D per EVI routes are both advertised on a per Broadcast Domain basis and they use the same Route Distinguisher and route target, the receiving ingress PE can associate them together to determine the BGP paths available for the MAC (multiple aliasing paths in case of all-active mode, or one active and one backup in case of single-active mode). Different paths can still be distinguished unambiguously even if all traverse the same Border Router.

Although the Aliasing and Backup Path procedures themselves are not affected, it is important to note that the ingress PE installs the MAC from an EVPN MAC/IP Advertisement route (with non-reserved ESI), only if the associated set of Ethernet A-D per ES routes are received from the same egress PE ([I-D.ietf-bess-rfc7432bis], section 9.2.2). Due to the same issues described in Section 3.2.1, the ingress PE cannot determine if the received MAC/IP Advertisement route and the received set of Ethernet A-D per ES routes were originated by the same egress PE.

This document specifies the following approach in the ingress PE to address this resolution issue:

1. If present in the routes, the ingress PE uses the Egress PE Identifier approach in Section 3.2.2 to correlate MAC/IP Advertisement routes and A-D per ES routes, and then resolve the MAC/IP Advertisement route as in ([I-D.ietf-bess-rfc7432bis].

2. If the Egress PE Identifier sub-TLV is not present in the A-D per ES and MAC/IP Advertisement routes for the Ethernet Segment, the ingress PE uses a "loose" resolution for the MAC/IP Advertisement route - that is:

    * The ingress PE considers the MAC/IP Advertisement route (with a non-reserved ESI) resolved if, and only if, at least one Ethernet A-D per ES route has been received with the same ESI and the same next hop as the MAC/IP Advertisement route (and it is assumed that its route target set contains the route target of the MAC/IP Advertisement route).

3.4. Designated Forwarder Election and AC-Influenced Capability

On an all-active Ethernet Segment, the Designated Forwarder is the PE router responsible for sending Broadcast, Unknown Unicast, and Multicast (BUM) traffic to a multi-homed Customer Edge (CE) device, in the <ES, Ethernet Tag> for which the PE is elected. If the Ethernet Segment works in single-active mode or port-active mode, the Designated Forwarder is the PE router that sends all traffic to a multi-homed CE [RFC8584]. When a CE is multi-homed to two or more PEs sitting in different domains, the Designated Forwarder candidate list is still created normally. The Designated Forwarder Election is unaffected by the Border Routers next hop self operation on the ES routes. This is because the candidate list is created out of the Originating Router's IP Address of the ES routes (which is not modified by the Border Routers) rather than the ES route next hops [RFC8584]. However, the Attachment Circuit Influenced Designated Forwarder (AC-Influenced DF Election) capability [RFC8584] is affected by the next hop self operation on the Border Routers.

When the AC-Influenced DF Election capability is enabled on all the PEs attached to the Ethernet Segment, the Designated Forwarder candidate list needs to be pruned based on the presence of the A-D per ES and A-D per EVI routes for a given candidate. That is, even if E-PE1's ES route is received Figure 3, E-PE2 cannot add E-PE1 to the Designated Forwarder candidate list for <ES1, BD1> until the valid A-D per ES and A-D per EVI routes (for ES1 and BD1) are received and identified as originating from E-PE1. However, because BR-2 rewrites the next hop of the A-D routes, E-PE2 cannot rely on the next hop to identify which routes came from E-PE1. This issue is similar to the challenge described in Section 3.2.1 for mass withdraw, except here the PE must correlate the ES route and A-D per ES/EVI routes to the same PE of origin.

This document assumes that, in case the PEs attached to the same Ethernet Segment are deployed in different domains, operators may choose one of the following alternatives:

* Disable the AC-Influenced Designated Forwarder capability on all PEs attached to the Ethernet Segment, or

* Enable the AC-Influenced Designated Forwarder capability on all PEs attached to the Ethernet Segment, and correlate the received A-D per ES/EVI routes with their corresponding Originating Router's IP Address using the method described in Section 3.2.2.

3.5. Split Horizon Filtering

The Split Horizon Filtering is a fundamental part of the EVPN multi-homing procedures to avoid BUM looped frames to go back to the multi-homed CE. As described in [RFC9746] there are two Split Horizon Filtering Types: ESI label based and Local Bias. Which mechanism applies depends on the transport tunnel used for the EVPN BUM packets, and some tunnel types may support both approaches.

If two or more PEs belonging to the same Ethernet Segment are deployed in different domains, the procedures performed by the Border Routers may affect the Split Horizon Filtering mechanisms. Specifically:

1. If the multi-homed PEs use an ESI label based Split Horizon Filtering Type:

    a. Whether the PEs use upstream or downstream allocated ESI labels (for P2MP/MP2MP or Ingress Replication, respectively), the PEs in the Ethernet Segment need to correlate the identity of the PE advertising the ESI label with the Inclusive Multicast Ethernet Tag routes advertised by the same PE. This presents the same challenge described in Section 3.2.1, around identifying the origin of the A-D per ES route — only now the receiving PE must correlate A-D per ES routes with Type 3 routes rather than Type 2 or 5. In this case, the solution in Section 3.2.2 SHOULD be used: the ingress PE correlates the received A-D per ES routes Egress PE Identifier with the received Originating IP of the Inclusive Multicast Ethernet Tag routes.

    b. The use of ESI labels allocated from a Domain-wide Common Block (DCB) with the same label assigned to all PEs attached to the Ethernet Segment, can simplify the procedures. If that is the case, the ingress PE can program the received ESI label without the need to correlate the received A-D per ES routes with the Inclusive Multicast Ethernet Tag routes.

    c. Additionally, the Border Routers must preserve the ESI label when routing packets between domains.

2. If the multi-homed PEs use Local Bias as the Split Horizon Filtering Type:

    a. The Border Router cannot change the outer source IP address of the IP tunnel, so that the egress PE can still identify the source PE. Note this may not be possible in many implementations.

The above considerations can significantly influence Inter-Domain Option-B designs. Therefore, operators should carefully analyze the capabilities of their Border Routers and PEs before deploying CEs multi-homed to PEs located in different domains.

4. Inter-Domain Option-B and Load Balancing Procedures

This section discusses the impact of Inter-Domain Option-B Border Router procedures on load balancing related mechanisms, such as the use of the Flow Label or Control Word for MPLS tunnels (see [I-D.ietf-bess-rfc7432bis] section 18), or the source UDP port for NVO tunnels which provides entropy for load balancing traffic on the core routers. VXLAN [RFC7348] is an example of NVO tunnel type that uses the source UDP port to provide entropy.

4.1. Flow Label

The use of the Flow Label and its signaling is described in [I-D.ietf-bess-rfc7432bis] section 18.1. The ingress PE pushes the Flow Label only on EVPN-encapsulated known unicast packets that are forwarded to egress PEs which previously advertised their Flow Label support by setting the F-bit in their Inclusive Multicast Ethernet Tag routes. When programming the data path for a given MAC, the ingress PE must therefore enable the use of the Flow Label if the MAC/IP Advertisement route originated from the same PE that advertised the Inclusive Multicast Ethernet Tag route with the F-bit set. To achieve this, the ingress PE correlates the MAC/IP Advertisement route and the Inclusive Multicast Ethernet Tag route based on their matching Route Distinguishers.

The Flow Label MUST be preserved by the Border Routers when they receive EVPN-encapsulated packets containing a Flow Label, to ensure that EVPN packets for the same flow are forwarded following the same path within each domain.

4.2. Control Word

The signaling of the Control Word in the Inclusive Multicast Ethernet Tag routes (C-bit) is described in [I-D.ietf-bess-rfc7432bis] section 7.11. As in the case described in Section 4.1, when a Border Router rewrites the next hops of the MAC/IP Advertisement and Inclusive Multicast Ethernet Tag routes, the ingress PE needs to identify the egress PE based on the matching Route Distinguisher of the two routes. Also, if included in the received EVPN-encapsulated packets, the Control Word MUST be preserved by the Border Routers so that no packet reordering happens for flows forwarded into an adjacent domain.

4.3. Source UDP port

If ingress and egress PEs use NVO tunnels [RFC8365], i.e., IP tunnels, the ingress PE typically encodes a per-flow hash value into the the outer tunnel source UDP port of the EVPN-encapsulated packets. Examples of tunnel types that use the outer source UDP port as an entropy field include VXLAN, GENEVE, or MPLSoUDP. The Border Routers between the ingress and egress PEs MUST preserve the value of the source UDP port so that EVPN-encapsulated packets for the same flow are forwarded following the same path within each domain.

5. Inter-Domain Option-B and Layer-2 MTU

In the same way the support for Flow Label or Control Word is signaled, the egress PE’s supported layer-2 MTU (Maximum Transfer Unit) is indicated in the Layer-2 MTU field of the EVPN Layer-2 Attributes extended community advertised along with the Inclusive Multicast Ethernet Tag route ([I-D.ietf-bess-rfc7432bis], section 7.11.1). The Border Router(s) between ingress an egress PEs do not modify any of the advertised attributes, and therefore the layer-2 MTU value is propagated end to end up to the ingress PE. In general, the layer-2 MTU configured in all PEs attached to the same EVPN service SHOULD match, irrespective of the domain where they reside. In case MTUs are different in the different domains, [I-D.ietf-bess-rfc7432bis] allows the signaling a layer-2 MTU of zero from the egress PE, which is not checked at the ingress PE and ensures the EVPN destination is properly programmed at this ingress PE.

6. E-Tree Considerations

[RFC8317], or Ethernet-Tree in EVPN networks, describes two areas that are impacted by the presence of an Inter-Domain Option-B Border Router between ingress and egress PEs: the use of composite tunnels for BUM traffic and the egress PE filtering of BUM traffic originated from a Leaf Attachment Circuit.

6.1. E-Tree Composite Tunnels

A composite tunnel is tunnel type used by the Root PE to simultaneously indicate a P2MP tunnel in the transmit direction and an Ingress Replication tunnel in the receive direction for BUM traffic. For this reason, an Inclusive Multicast Ethernet Tag route for a composite tunnel comprises both, a downstream allocated EVPN label for Ingress replication, and a P2MP tunnel identifier. The EVPN label is extracted by the Border Router and programmed in the Forwarding Information Base, as described in Section 2.1.1 bullet "d". Since the Ingress Replication procedures apply, the Border

Router generates a valid new label that is encoded in the (composite type) PMSI Tunnel Attribute of the route readvertised to the adjacent domain. Also, as described in Section 2.1.1, the segmentation procedures in [RFC9572] are followed for the encoded P2MP tunnel in the same PMSI Tunnel Attribute.

6.2. Egress Filtering of BUM Traffic Originated from a Leaf Attachment Circuit

E-Tree operation in EVPN networks requires filtering traffic that originates from Leaf Attachment Circuits. While the ingress PE can decide whether to forward known unicast leaf traffic based on whether the destination MAC address belongs to a Leaf Attachment Circuit, the filtering of BUM traffic must be performed by the egress PE. To enable this, the egress PE advertises a Leaf Label along with an Ethernet A-D per ES route (with an ESI value of zero). The egress PE then relies on the ingress PE to include this Leaf Label when sending Leaf-originated BUM traffic [RFC8317].

If ingress and egress PEs are located in different domains of an Inter-Domain Option-B network, the ingress PE cannot correlate the received Inclusive Multicast Ethernet Tag route and A-D per ES route (which carries the Leaf Label) as being advertised by the same egress PE. Due to this limitation in identifying the egress PE’s Leaf Label, the ingress PE is unable to push the Leaf Label below the EVPN multicast label for a given egress PE. The issue is illustrated in Figure 5.

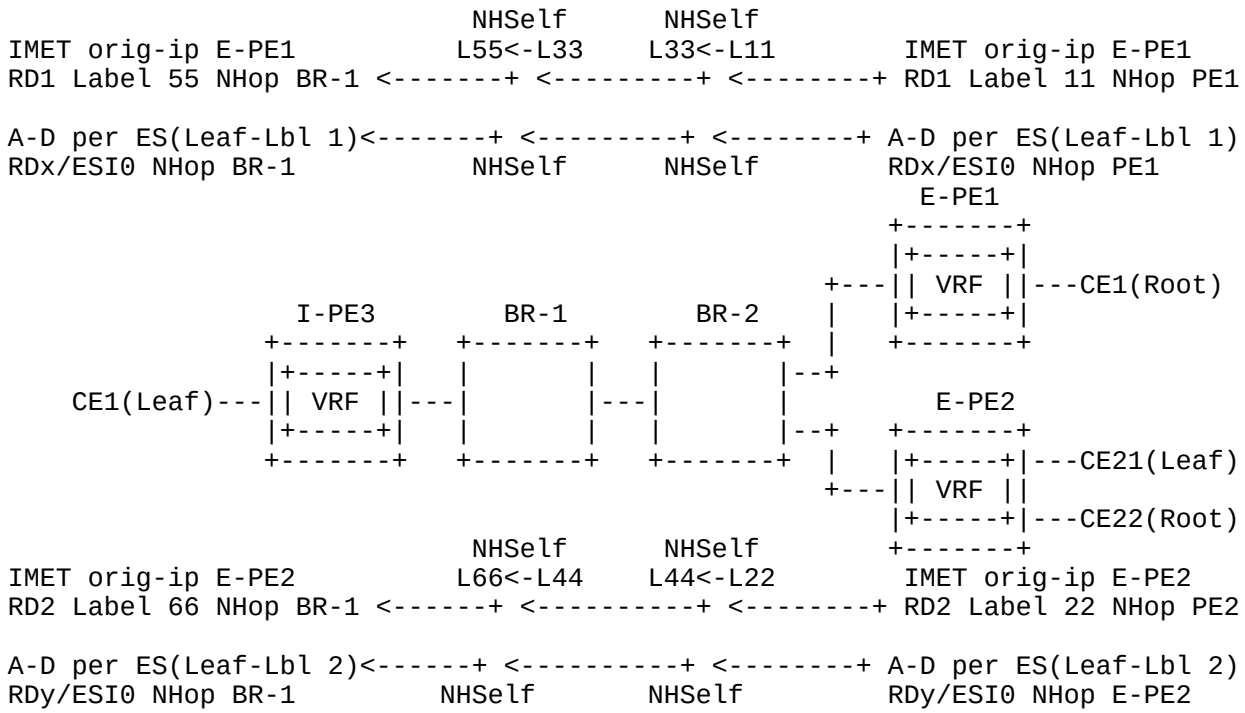

Figure 5: EVPN Inter-Domain Option-B and Leaf BUM filtering

Suppose the egress PEs and ingress PE are in a different domains, as shown in Figure 5, and that I-PE3 needs to forward EVPN-encapsulated BUM traffic from Leaf CE1, using Ingress Replication. I-PE3 receives Inclusive Multicast Ethernet Tag routes and A-D per ES routes from the two egress PEs, however, I-PE3 is unable to determine which Leaf Label to push when sending EVPN-encapsulated BUM traffic to E-PE1 or E-PE2. This issue arises because the A-D per ES routes can no longer be associated with their corresponding Inclusive Multicast routes based on the next hop, since all four routes in the example are received from the same next hop. This section proposes several solutions to address the issue. An implementation following this specification SHOULD use the Egress PE Identifier solution Section 6.2.1 and MAY use the Domain-wide Common Block Leaf Label solution Section 6.2.2 or the Source MAC-Based Egress filtering solution Section 6.2.3.

6.2.1. Identification of the PE of Origin based on the Egress PE Identifier

A way to solve the issue with E-Tree and the egress filtering of Leaf BUM traffic is to identify and correlate the Inclusive Multicast Ethernet Tag routes and A-D per ES routes (with ESI of zero) that were originated by the same egress PE. To achieve this, the method in Section 3.2.2 can be used.

In this case, however, the identification process is applied to correlate Inclusive Multicast Ethernet Tag routes with A-D per ES routes, rather than correlating MAC/IP Advertisement routes with A-D per ES routes. To identify an Inclusive Multicast Ethernet Tag route and an A-D per ES route advertised by the same egress PE, the Inclusive Multicast Ethernet Tag route Originating IP and the Egress PE Identifier in the A-D per ES route MUST match.

6.2.2. Domain-wide Common Block Leaf Labels

The use of Leaf Labels allocated from a Domain-wide Common Block (DCB) and the same Leaf label value used by all the PEs attached to the E-Tree EVPN service simplify the procedures. In this approach, all egress PEs advertise the same Leaf label in their A-D per ES routes for ESI of zero, and this Label value matches the local Leaf label on the ingress PE.

The ingress PE can then program the allocated Leaf label for all the destination egress PEs, without needing to correlate the received Inclusive Multicast and A-D per ES routes. This approach assumes all PEs in the Broadcast Domain allocate the same Leaf label. If the ingress PE detects any inconsistency in the signaled Leaf label - meaning that at least one PE of the Broadcast Domain advertises a different label than the local Leaf label - the ingress PE SHOULD NOT program the Leaf label when sending traffic to the egress PEs.

6.2.3. Source MAC-based Egress Filtering

Another potential solution is to use source MAC-based egress filtering, instead of Leaf label-based egress filtering for EVPN-encapsulated BUM traffic. If the ingress PE receives two or more A-D per ES routes (with ESI of zero) with the same next hop, then it does not program any of the received Leaf labels and forwards EVPN-encapsulated BUM packets with the EVPN label and without any Leaf label. Assuming that the ingress PE has previously advertised the local Leaf MAC addresses, when the BUM packets get to the egress PE, a source MAC lookup in the MAC-VRF will determine if the BUM packet is coming from a Leaf or a Root Attachment Circuit.

Taking the example shown in Figure 5, I-PE3 advertises CE1's MAC as a Leaf MAC in a route type 2, and hence CE1's MAC is programmed in E-PE1 and E-PE2 as Leaf. Since I-PE3 receives two A-D per ES routes (with ESI of zero) from the same next hop, I-PE3 determines that it cannot program the received Leaf labels, and therefore I-PE3 forwards BUM packets from CE1 to E-PE1 and E-PE2 with their corresponding Inclusive Multicast labels and without any Leaf label.

When the packets arrive at the egress PEs, E-PE1 and E-PE2 perform a source MAC lookup in their MAC-VRFs. Since CE1's MAC is marked as a Leaf MAC, E-PE1 and E-PE2 can filter the packets correctly. For example, E-PE2 forwards the traffic only to CE22 (root) and not to CE21 (leaf).

7. Inter-Domain Option-B and PBB-EVPN

Provider Backbone Bridging EVPN [RFC7623] is also supported in Inter-Domain Option-B. The following considerations apply:

* PBB-EVPN does not have any of the issues described in Section 3. This is due to the fact that PBB-EVPN multi-homing procedures do not rely on Ethernet A-D per ES or per EVI routes at all.

* PBB-EVPN does not have any of the issues described in Section 6 either, for the same reason. For E-Tree egress filtering of the EVPN-encapsulated BUM packets (so that they are only forwarded to local Root Attachment Circuits and not Leaf Attachment Circuits), PBB-EVPN relies on the source B-MAC identification at the egress PE. The procedures are not impacted by the presence of a Border Router between ingress and egress PEs.

* Also, this document assumes that the [I-D.ietf-bess-rfc7432bis] procedures to signal Flow Label, Control Word or Layer-2 MTU, do not apply to PBB-EVPN networks, hence there are no issues derived from those components.

8. Security Considerations

The solutions described here are mostly based on existing specifications, and as such, this document inherits the security considerations detailed in each of the normative reference documents.

In addition, this specification defines a new identifier for egress PEs, conveyed by means of the Egress PE Identifier sub-TLV. This identifier is used to uniquely identify the egress PE and to support the procedures described in this document. Because the Egress PE Identifier sub-TLV directly influences route correlation and forwarding behavior, its integrity is critical. An attacker capable

of intercepting EVPN routes and modifying the Egress PE Identifier sub-TLV could disrupt the procedures specified in this document. For example, altering the identifier may result in incorrect route association, misdirected traffic, failure of mass withdraw procedures, or unintended traffic drops. Therefore, the mechanisms used to distribute EVPN routes carrying the Egress PE Identifier sub-TLV need to provide adequate protection against route tampering, including the use of existing BGP security mechanisms and transport protection where applicable.

9. IANA Considerations

IANA is requested to allocate a new Sub-TLV code point from the "BGP Tunnel Encapsulation Attribute Sub-TLVs" registry under the "Border Gateway Protocol (BGP) Tunnel Encapsulation" grouping. The value is requested to be allocated from the 1-63 range, which requires "Standards Action".

Value	Description	Reference
TBDxx	Egress PE Identifier	[This document]

10. Contributors

11. Acknowledgments

The authors would like to thank Jeffrey Zhang and Alexander Vainshtein for his review and comments.

12. Annex - The EVPN Instance RD solution

For completeness, this section documents an alternative solution for correlating A-D per-ES routes with other EVPN routes that was discussed in earlier versions of this document. Although this is not the solution specified in this document, it is included in this annex because some implementations make use of it.

This solution is based on the E-PEs using the same Route Distinguisher on A-D per ES routes and routes type 2 or 5. The A-D per ES routes are normally advertised per <ES, EVI-set>, where an EVI-set is a group of EVPN Instances, each represented by a different route target included in the route. Because of this, the A-D per ES route cannot use the Route Distinguisher of an existing VRF on the PE, but instead requires a unique Route Distinguisher that is not assigned to any EVPN Instance (instantiated in a VRF). However, suppose each EVI-set is composed of only a single EVI. In this case, the A-D per ES routes are advertised per <ES, EVI>, resulting in a

separate A-D per ES route for each EVPN Instance (or VRF). Under this model, the A-D per ES routes can now use the Route Distinguisher assigned to the corresponding EVPN Instance (or VRF), which is the same Route Distinguisher used by the Type 2 or Type 5 routes for that EVI.

If that is the case, now the A-D per ES routes can use the Route Distinguisher assigned to the EVPN Instance (or VRF), which is the same one used by the routes type 2 or 5 for the EVI. Since A-D per ES routes are - with this solution - advertised per <ES, EVI>, this is really a "mass withdraw per EVI" solution, similar to the one described in Section 3.2 in terms of efficiency. However, the advantage of this solution is that the A-D per ES routes are REQUIRED, while A-D per EVI routes are OPTIONAL [I-D.ietf-bess-rfc7432bis] and may not be used in a given EVI.

13. References

13.1. 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>.

[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>.

[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>.

[RFC9136] Rabadan, J., Ed., Henderickx, W., Drake, J., Lin, W., and A. Sajassi, "IP Prefix Advertisement in Ethernet VPN (EVPN)", RFC 9136, DOI 10.17487/RFC9136, October 2021, <https://www.rfc-editor.org/info/rfc9136>.

[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>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <https://www.rfc-editor.org/info/rfc4364>.

[I-D.ietf-bess-rfc7432bis]
Sajassi, A., Burdet, L. A., Drake, J., and J. Rabadan, "BGP MPLS-Based Ethernet VPN", Work in Progress, Internet-Draft, draft-ietf-bess-rfc7432bis-13, 24 June 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-rfc7432bis-13>.

[RFC9014] Rabadan, J., Ed., Sathappan, S., Henderickx, W., Sajassi, A., and J. Drake, "Interconnect Solution for Ethernet VPN (EVPN) Overlay Networks", RFC 9014, DOI 10.17487/RFC9014, May 2021, <https://www.rfc-editor.org/info/rfc9014>.

[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>.

[RFC8317] Sajassi, A., Ed., Salam, S., Drake, J., Uttaro, J., Boutros, S., and J. Rabadan, "Ethernet-Tree (E-Tree) Support in Ethernet VPN (EVPN) and Provider Backbone Bridging EVPN (PBB-EVPN)", RFC 8317, DOI 10.17487/RFC8317, January 2018, <https://www.rfc-editor.org/info/rfc8317>.

[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>.

[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>.

[RFC9625] Lin, W., Zhang, Z., Drake, J., Rosen, E., Ed., Rabadan, J., and A. Sajassi, "EVPN Optimized Inter-Subnet Multicast (OISM) Forwarding", RFC 9625, DOI 10.17487/RFC9625, August 2024, <https://www.rfc-editor.org/info/rfc9625>.

[I-D.ietf-bess-evpn-ipvpn-interworking]
Rabadan, J., Sajassi, A., Rosen, E. C., Drake, J., Lin, W., Uttaro, J., and A. Simpson, "Interconnecting EVPN and IPVPN Domains", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-ipvpn-interworking-16, 28 February 2026, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-ipvpn-interworking-16>.

[RFC9012] Patel, K., Van de Velde, G., Sangli, S., and J. Scudder, "The BGP Tunnel Encapsulation Attribute", RFC 9012, DOI 10.17487/RFC9012, April 2021, <https://www.rfc-editor.org/info/rfc9012>.

13.2. Informative References

[RFC9161] Rabadan, J., Ed., Sathappan, S., Nagaraj, K., Hankins, G., and T. King, "Operational Aspects of Proxy ARP/ND in Ethernet Virtual Private Networks", RFC 9161, DOI 10.17487/RFC9161, January 2022, <https://www.rfc-editor.org/info/rfc9161>.

[RFC7348] Mahalingam, M., Dutt, D., Duda, K., Agarwal, P., Kreeger, L., Sridhar, T., Bursell, M., and C. Wright, "Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", RFC 7348, DOI 10.17487/RFC7348, August 2014, <https://www.rfc-editor.org/info/rfc7348>.

[RFC9572] Zhang, Z., Lin, W., Rabadan, J., Patel, K., and A. Sajassi, "Updates to EVPN Broadcast, Unknown Unicast, or Multicast (BUM) Procedures", RFC 9572, DOI 10.17487/RFC9572, May 2024, <https://www.rfc-editor.org/info/rfc9572>.

[I-D.ietf-bess-evpn-vpws-gateway] Rabadan, J., Sathappan, S., Prabhu, V., Lin, W., and P. Brissette, "Ethernet VPN Virtual Private Wire Services Gateway Solution", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-vpws-gateway-01, 2 March 2026, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-vpws-gateway-01>.

[I-D.ietf-bess-evpn-ip-aliasing] Sajassi, A., Rabadan, J., Pasupula, S., Krattiger, L., and J. Drake, "EVPN Support for L3 Fast Convergence and Aliasing/Backup Path", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-ip-aliasing-03, 7 May 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-ip-aliasing-03>.

[RFC9746] Rabadan, J., Nagaraj, K., Lin, W., and A. Sajassi, "BGP EVPN Multihoming Extensions for Split-Horizon Filtering", RFC 9746, DOI 10.17487/RFC9746, March 2025, <https://www.rfc-editor.org/info/rfc9746>.

[IANA-ADDRESS-FAM]
"IANA Address Family Numbers",
<https://www.iana.org/assignments/address-family-numbers/>.

Authors' Addresses

Jorge Rabadan (editor)
Nokia
520 Almanor Avenue
Sunnyvale, CA 94085
United States of America
Email: jorge.rabadan@nokia.com

Senthil Sathappan
Nokia
520 Almanor Avenue
Sunnyvale, CA 94085
United States of America
Email: senthil.sathappan@nokia.com

Ali Sajassi
Cisco
Email: sajassi@cisco.com

Wen Lin
Juniper
Email: wlin@juniper.net

Luc Andr Burdet
Cisco
Email: lburdet@cisco.com