

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 4 September 2025

J. Rabadan, Ed.
O. Dornon
V. Prabhu
Nokia
A. Nichol
Arista
Z. Zhang
W. Lin
Juniper Networks
3 March 2025

EVPN Multicast Forwarding for EVPN to EVPN Gateways
draft-rabnic-bess-evpn-mcast-eeg-05

Abstract

This document proposes an EVPN (Ethernet Virtual Private Network) extension to allow IP multicast forwarding on Service Gateways that interconnect two or more EVPN domains.

Status of This Memo

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

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

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

This Internet-Draft will expire on 4 September 2025.

Copyright Notice

Copyright (c) 2025 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components

extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	2
1.1. Terminology and Conventions	3
1.2. Multicast Layer-2 Interconnect on EVPN to EVPN Gateways (EEG)	5
1.3. Multicast Layer-3 Interconnect on EVPN to EVPN Gateways (EEG)	7
2. Layer-2 EVPN to EVPN Gateway Procedures	11
3. Layer-3 EVPN to EVPN Gateway Procedures	14
4. Security Considerations	17
5. IANA Considerations	17
6. Contributors	18
7. Acknowledgments	18
8. References	18
8.1. Normative References	18
8.2. Informative References	19
Authors' Addresses	20

1. Introduction

This document proposes an EVPN extension to allow IP multicast forwarding on Service Gateways that interconnect two or more EVPN domains. The extensions are applied in two separate scenarios:

- * EVPN Layer-2 Interconnect

- * EVPN Layer-3 Interconnect

In Layer-2 Interconnect scenarios, this document extends the procedures in [RFC9251] so that IP Multicast can be forwarded efficiently in Service Gateways that Interconnect two or more EVPN domains. Note that [RFC9014], in sections 4.4 and 4.6, specifies the Service Gateway solution for EVPN layer-2 multi-point services, including the procedures for layer-2 unicast and Broadcast, Unknown unicast and Multicast (BUM) traffic, however, it does not specify procedures to optimize the forwarding of IP Multicast on the Service Gateways that interconnect domains that use EVPN IGMP or MLD proxy procedures.

In Layer-3 Interconnect scenarios, this document extends the procedures in [RFC9625], so that Service Gateways that interconnect two or more EVPN layer-3 domains as in [I-D.ietf-bess-evpn-ipvpn-interworking] for IP unicast traffic can

forward Inter-Subnet Multicast traffic efficiently across EVPN layer-3 domains. [RFC9625] defines procedures to support efficient Inter-Subnet Multicast forwarding on Service Gateways that interconnect EVPN domains to MVPN [RFC6513] [RFC6514] or PIM domains [RFC7761]. This document completes the procedures to support efficient Inter-Subnet Multicast forwarding on Service Gateways that interconnect EVPN domains to other EVPN domains.

In both scenarios, we refer to the Service Gateway that implements this specification as an EVPN to EVPN Gateway (EEG).

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.
- * 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).
- * EEG: EVPN to EVPN Gateway, or a Service Gateway that interconnects two or more EVPN domains for the purpose of forwarding IP Multicast traffic.
- * ES: Ethernet Segment. When a customer site (device or network) is connected to one or more PEs via a set of Ethernet links, then that set of links is referred to as an Ethernet segment.

- * ESI: Ethernet Segment Identifier. 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.
- * EVPN domain: two PEs are in the same EVPN domain if they are attached to the same service and the packets between them do not require a data path lookup of the inner frame (e.g., in the BD of a MAC-VRF) in any intermediate router. A Service Gateway in this document interconnects two or more EVPN domains and is configured with a domain identifier per EVPN domain (referred to as EVPN domain-ID). The EVPN domain-ID is encoded in the D-PATH attribute as specified in [I-D.ietf-bess-evpn-dpath] for Layer-2 interconnect scenarios and in [I-D.ietf-bess-evpn-ipvpn-interworking] for Layer-3 interconnect scenarios.
- * I-ES: Interconnect Ethernet Segment or a special Ethernet Segment used to provide multi-homing on Service Gateways that follow the procedures in [RFC9014].
- * 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.
- * IIF: Incoming Interface. Refers to the Layer-3 interface in the IP-VRF that is identified as the one receiving a particular IP Multicast flow.
- * IP-VRF: A VPN Routing and Forwarding table for IP routes on an NVE/PE. The IP routes could be populated by any routing protocol, E.g., EVPN, IP-VPN and BGP PE-CE IP address families. An IP-VRF is also an instantiation of a layer 3 VPN in a PE.
- * 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.
- * OIF list and OIF entry: Outgoing Interface list or entry. Refers to the list of interfaces or interface entry in the IP-VRF (Layer-3 OIF) or BD (Layer-2 OIF) that are identified as output interfaces for a given multicast group.

- * 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.
- * SBD: Supplementary Broadcast Domain, a especial 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].
- * SMET route: Selective Multicast Ethernet Tag route, as defined in [RFC9251].
- * 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.

1.2. Multicast Layer-2 Interconnect on EVPN to EVPN Gateways (EEG)

This section describes an example of the first use-case for which this document specifies extensions.

Consider a pair of multi-homing Service Gateways (EEG1 and EEG2) that interconnect EVPN domain 1:1 and domain 2:2, as illustrated in Figure 1 (with 1:1 and 2:2 being the respective domain-IDs [I-D.ietf-bess-evpn-dpath]). In addition to EEG1 and EEG2, PE1 and PE2 are also attached to EVPN domain 1:1 (with 1:1 being the EVPN domain-ID), and PE3, PE4 and PE5 are also attached to domain 2:2. Source S1, Receiver-1, Receiver-2 and Receiver-3 are all connected to the same IP subnet and EVPN Broadcast Domain BD1. For unicast traffic, EEG1 and EEG2 follow the procedures in [RFC9014] sections 4.4 and 4.6. In particular, the Interconnect Ethernet Segment I-ES1 is instantiated in EEG1 and EEG2 and determines the redundancy and forwarding of the traffic between the two domains. In this example, EEG1 is the Designated Forwarder and EEG2 the non-Designated Forwarder router in I-ES1. 'Encap1' and 'encap2' in Figure 1 refer to any possible encapsulation that is supported by EVPN and BD1 uses to transmit or receive packets; for instance: MPLS, Segment Routing MPLS (SR-MPLS), VXLAN or SRv6. The procedures in this document apply irrespective of the combination of encapsulations being used in the EVPN domains that the EEG routers are interconnecting. In this scenario, IP Multicast sources and receivers can be attached to either domain and the EEG routers must be able to forward IP Multicast traffic efficiently across domains. PE1, PE2, PE3, PE4 and PE5 follow the procedures of [RFC9251], and they are not aware of the fact that they may be attached to different EVPN domains.

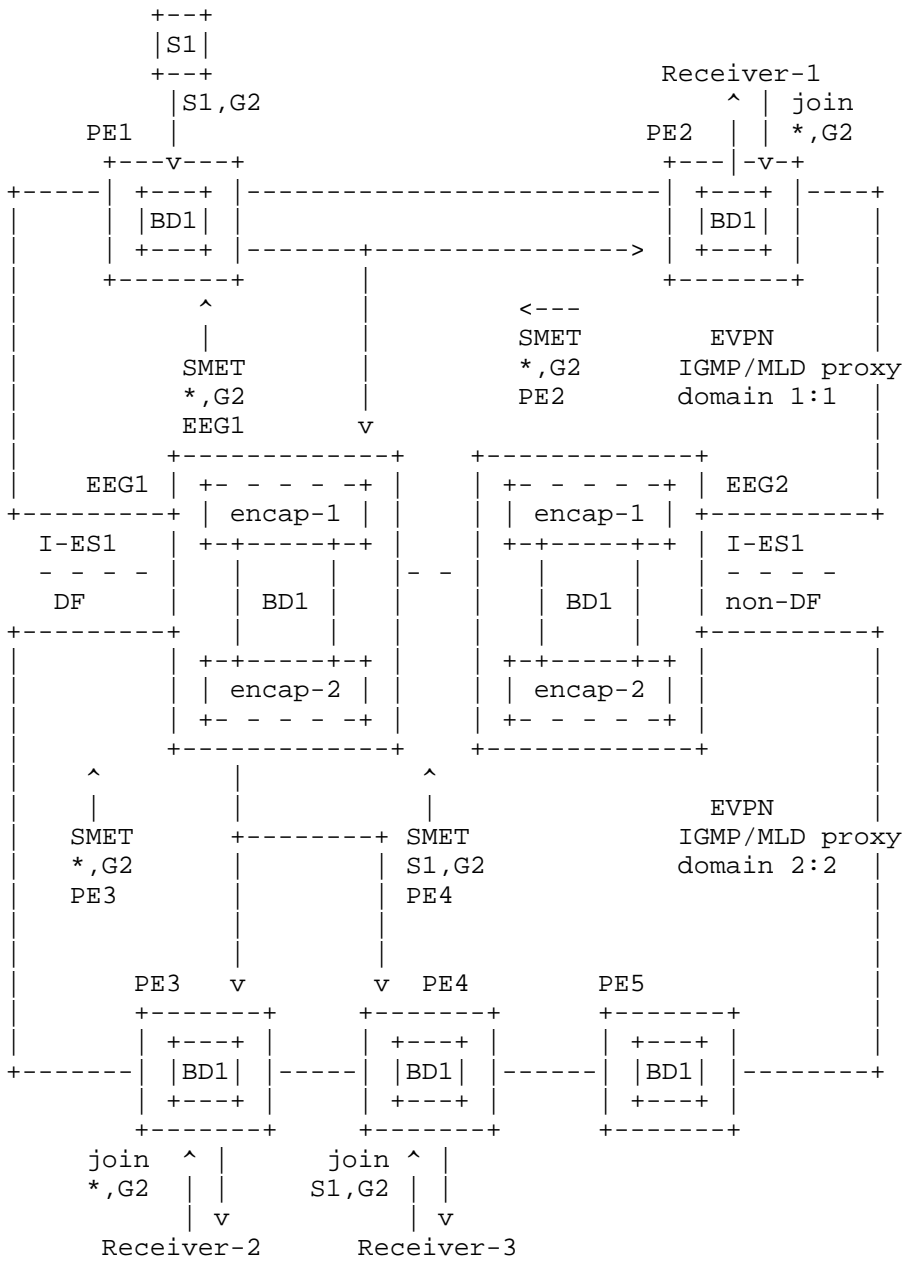


Figure 1: Layer-2 EEGs

Suppose S1 (with source IP address S1) sends IP multicast traffic to group G2, and Receiver-1 and Receiver-2 issue an IGMP (or MLD) join (*,G2). Receiver-3 sends an IGMP (or MLD) join (S1,G2). With the extensions in this document:

- * The EEG routers import the EVPN Selective Multicast Ethernet Tag (SMET) routes issued by the PE routers in the domains that they interconnect.
- * They perform a proxy function for the multicast groups learned from the imported SMET routes of a domain and advertise the resulting proxy membership to the other domains through SMET routes, using their own IP address as the Originator IP of the SMET route. As an example in Figure 1, EEG1 and EEG2 import the SMET routes for (*,G2) and (S1,G2) that they receive from PE3 and PE4, respectively. EEG1 and EEG2 create state for the two groups and the I-ES1 Designated Forwarder, i.e., EEG1, advertises an SMET route for (*,G2) using its own network parameters for the destination domain (EEG1's Originator IP, Route Distinguisher and Route Target of domain 1:1). Although not represented in Figure 1, the EEG routers also import the SMET route for (*,G2) issued by PE2 in domain 1:1. Upon the OIF creation for PE2, EEG1 triggers the advertisement of an SMET route for (*,G2) into domain 2:2 with its own network parameters in domain 2:2.
- * The EEG routers advertise the minimal set of SMET routes needed to attract traffic for a given multicast group. For example, even if the EEG routers receive SMET routes for both (*,G2) and (S1,G2), EEG1 only advertises an SMET route for (*,G2), as this is the minimal set necessary to receive traffic for any flow destined to G2.
- * The EEG routers do not require the use of Multicast Membership Report Synch or Multicast Leave Synch routes [RFC9251] to synchronize the multicast states received via SMET routes from each domain. This is because all EEG routers within a domain import the same set of SMET routes.
- * The EEG routers forward IP multicast traffic between domains in the same way BUM traffic is forwarded in an Interconnect Ethernet Segment in [RFC9014], that is, only the Designated Forwarder EEG forwards IP multicast traffic from sources in one domain to the other domains.

1.3. Multicast Layer-3 Interconnect on EVPN to EVPN Gateways (EEG)

This section describes an example of the second use-case for which this document specifies extensions.

Similar to Figure 1 consider a pair of multi-homing Service Gateways (EEG1 and EEG2) that interconnect EVPN domain 1:1 and domain 2:2 that are now EVPN OISM domains [RFC9625] for the same tenant, as illustrated in Figure 2. Note that Figure 2 is an example; the procedures described in this document apply regardless of whether the PEs are connected to the same or different Broadcast Domains. Sources and receivers may be attached to any PE or Broadcast Domain in the network and can reside in either the same or different subnets. The IP-VRF of the EEG routers imports EVPN IP Prefix routes [RFC9136] from one domain, install the routes in the IP-VRF and export the routes into EVPN IP Prefix routes into the other domains. In order to do that, the EEG nodes follow the gateway procedures in [I-D.ietf-bess-evpn-ipvpn-interworking]. The unicast routes in the IP-VRF of the EEG routers are used to create IIF entries in the layer-3 multicast states. In case the same IP prefix is received in two different EVPN IP Prefix routes, one from each EVPN domain, regular best path selection determines what EVPN IP Prefix route is selected and therefore what route is installed and exported into the other domain.

The encapsulations used in the EVPN domains can be any possible encapsulation that is supported by EVPN, for instance, MPLS, Segment Routing MPLS (SR-MPLS), VXLAN or SRv6. The procedures in this document apply irrespective of the combinations of encapsulations being used in the EVPN domains that the EEG routers are interconnecting. In this scenario, IP Multicast sources and receivers can be attached to either domain and the EEG routers must be able to forward IP Multicast traffic efficiently across domains. PE1, PE2, PE3, PE4 and PE5 follow the procedures of [RFC9625]. We assume PE1 and PE2 are attached to the Supplementary Broadcast Domain SBD1, whereas PE3, PE4 and PE5 are attached to the Supplementary Broadcast Domain SBD2. In this model, existing EVPN OISM PEs are unaware that certain sources or receivers are part of a different EVPN OISM Domain. The existing EVPN OISM nodes run only their standard [RFC9625] procedures and are entirely unaware of the remote EVPN OISM domains. Interworking is achieved by having some of the EVPN OISM PEs function as EVPN to EVPN Gateways (EEGs) running OISM procedures in all the domains they interconnect, as detailed in this document.

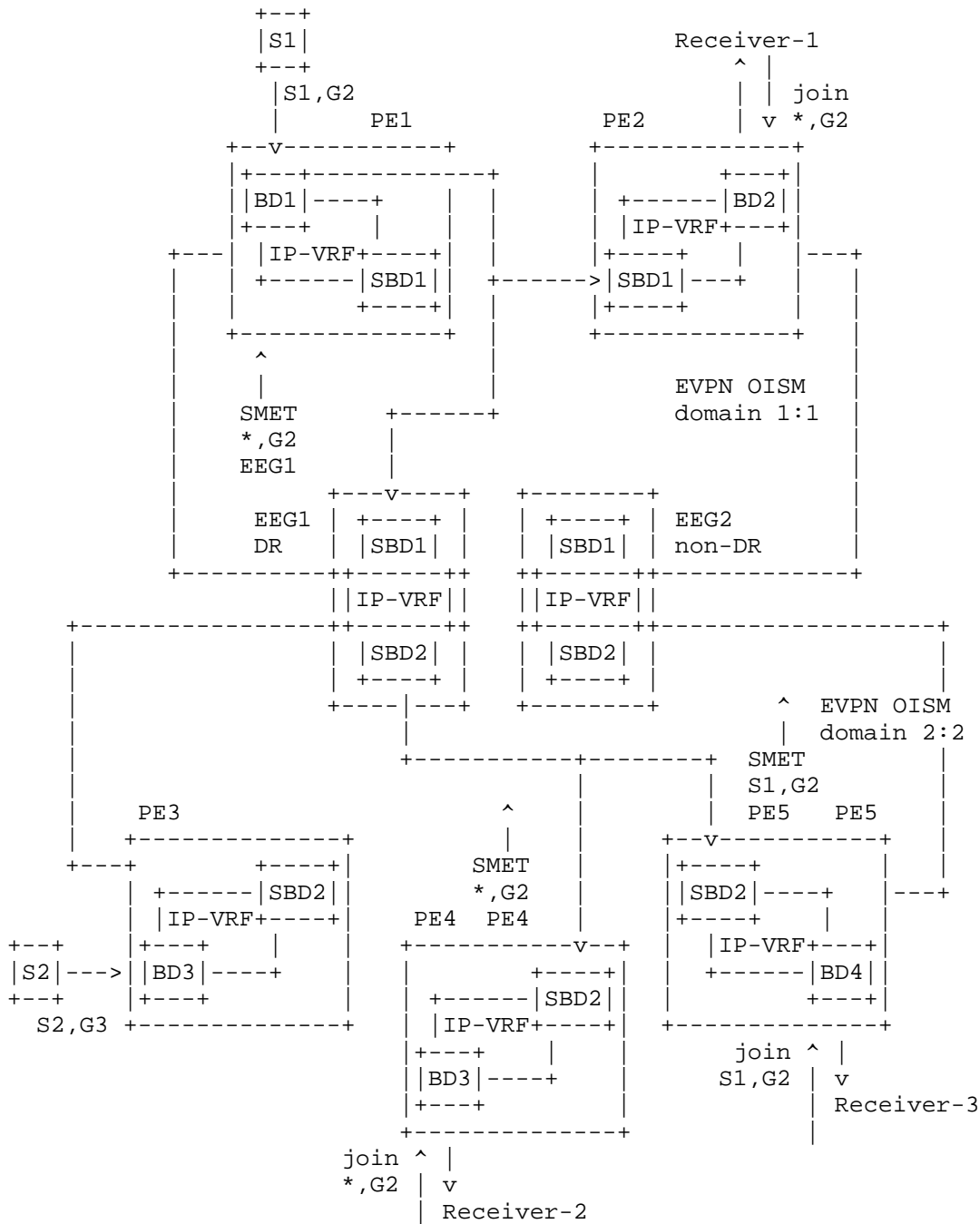


Figure 2: Layer-3 EEGs

In the example of Figure 2, suppose S1 (with source IP address S1) sends IP multicast traffic for group G2, and Receiver-1 and Receiver-2 issue an IGMP (or MLD) join (*,G2). Receiver-3 sends an IGMP (or MLD) join (S1,G2). With the extensions in this document:

- * The EEG routers import the SMET routes issued by the PE routers in the domains that they interconnect. Since the PEs on both domains follow [RFC9625] and are attached to the Supplementary Broadcast Domain of their respective OISM domain (SBD1 and SBD2), the EEG routers must be attached to the Supplementary Broadcast Domains of both domains that they interconnect, SBD1 and SBD2.
- * Out of the received SMET routes in one domain, the EEG routers create layer-2 OIF entries in the Supplementary Broadcast Domain of that domain, in addition to layer-3 IIF and OIF entries in the IP-VRF. The procedures to create layer-2 and layer-3 state in the Supplementary Broadcast Domain and IP-VRF of the EEG routers out of SMET routes follow the same procedures as in [RFC9625] for MVPN to EVPN Gateways (MEGs), only that the EEGs do not generate MVPN C-multicast routes, but SMET routes to attract the traffic for the group from the other EVPN OISM domain.
- * In case of EEG redundancy, that is, more than one EEG are attached to the same two EVPN OISM domains as in Figure 2, the EEG routers need to select the Supplementary Broadcast Domain Designated Router (SBD-DR) in each of the SBDs. The procedures to select an SBD-DR are described in Section 3. The selected SBD-DR has the First Hop Router function in the Supplementary Broadcast Domain where it is selected. In the example of Figure 2, if EEG1 is the SBD-DR for SBD2, EEG advertises an SMET route for (*,G2) in order to attract the multicast flow to G2 and forward it to domain 2:2. EEG2 is a non-Designated Router in SBD2, therefore EEG2 does not issue an SMET route for (*,G2) and it does not forward multicast traffic for G2 into domain 2:2 (EEG2 does not add the SBD2 IRB interface to the layer-3 OIF list).
- * The EEG routers distribute the minimum set of SMET routes between domains to attract the traffic for a given multicast group. As an example, in spite of the EEG routers receiving SMETs routes for (*,G2) and (S1,G2), EEG1 (the SBD2 Designated Router) only advertises an SMET route for (*,G2) since that is the minimum set required to attract the traffic for any flow to G2.
- * In [RFC9625] the Supplementary Broadcast Domain IRB interface is used in the OIF list only for traffic from external sources. This document extends the procedures so that an EEG router can be attached to multiple SBDs of the same IP-VRF and the Supplementary Broadcast Domain IRB can be added in the OIF list for a group,

when the IIF for the group is the IRB of another SBD attached to the same IP-VRF. In the example of Figure 2, EEG1 adds SBD2 IRB in the layer-3 OIF list for (S1,G2) and SBD1 IRB is the IIF for the same group.

2. Layer-2 EVPN to EVPN Gateway Procedures

This section provides the specification for EVPN to EVPN Gateways (EEGs) when configured for layer-2 interconnect, as in the use case of Section 1.2.

1. An EEG configured for layer-2 interconnect of two or more domains MUST support the procedures in [RFC9014] in sections 4.4 or 4.6 for unicast and BUM traffic forwarding. In addition, this specification uses the concept of the EVPN domain in [I-D.ietf-bess-evpn-dpath]:
 - a. An EEG interconnecting two EVPN domains of the same BD "redistributes" EVPN MAC/IP routes and carries out a proxy function for EVPN SMET routes between the domains.
 - b. In this document, the EEG proxy function for SMET routes refers to importing SMET routes in one domain of the BD, creating OIF entries within that domain, and exporting the SMET routes to the other domain(s) of the BD - provided that no other SMET route for the same multicast group has already been exported.
2. An EEG MUST import SMET routes received for the BD to which the EEG is attached.
 - a. An "SMET route received for" a BD in this context means that the SMET route has the route target that matches the BD import route target in one of the EVPN domains and it is a valid route based on the SMET definition in [RFC9251].
 - b. The imported SMET routes create layer-2 OIF entries for a given multicast group in the EVPN domain, and received multicast traffic for that group in a different EVPN domain of the BD will be forwarded using the multicast tree created by the imported EVPN Inclusive Multicast Ethernet Tag routes as in [RFC9251], or the multicast tree created by the EVPN Selective Provider Multicast Service Interface Auto Discovery routes (S-PMSI A-D routes as in [RFC9572]).

3. Upon receiving and importing an SMET route in a domain, an EEG that is not part of an Interconnect Ethernet Segment MUST perform the proxy function for that SMET route into the other domain(s) of the Broadcast Domains, as follows:
 - a. When doing proxy of SMET routes, the EEG MUST set its own IP address in the Originator IP field of the NLRI and MUST use its own route distinguisher for the domain.
 - b. An EEG with two domains in the same BD SHOULD use different route distinguishers when exporting routes into different domains and MAY use different route targets for different domains.
 - c. The proxy SMET route MUST preserve the Ethernet Tag ID, Multicast Source and Group information as well as the Flags of the SMET routes for which it provides the proxy function.
 - d. The proxy function on the EEG also includes "aggregation" of (S,G) and (*,G) states of the same IGMP/MLD version. That is, when at least one (*,G) for a group G has been imported via SMET route in one domain, only the SMET route for (*,G) is exported to the other domain. The (S,G) SMET routes for the same group G, which specify individual sources, SHOULD NOT be exported to the other domain. In other words, the minimum set of SMET routes for a group and version is distributed between domains.
4. Two or more EEG routers attached to the same two EVPN domains of a BD SHOULD use an Interconnect Ethernet Segment (I-ES) [RFC9014] to handle the redundancy and avoid multicast duplication, as follows:
 - a. Upon receiving and importing SMET routes in a domain, the I-ES Designated Forwarder MUST proxy the SMET routes to the other domain, and forward the multicast traffic between domains, assuming that it has OIF entries for the group in the domain of destination.
 - b. The non-Designated Forwarder MAY do proxy of the SMET routes but MUST NOT forward multicast traffic between domains as per [RFC9014], irrespective of the existence of OIF entries created by the received SMET routes. The operator can decide, by configuration, whether the non-Designated Forwarder exports SMET routes, depending on the trade-off between additional traffic and faster convergence in case of failure of the Designated Forwarder EEG.

5. In case of two or more EEG routers are attached to the same two EVPN domains of a BD, a control plane loop may be produced if the non-Designated Forwarder does proxy of the received SMET routes from the peer EEG into the other domain. In order to avoid that, the D-PATH [I-D.ietf-bess-evpn-dpath] attribute SHOULD be used as follows:
 - a. An EEG doing proxy of SMET routes between domains SHOULD add or modify the D-PATH BGP attribute [I-D.ietf-bess-evpn-dpath] in the exported SMET route, by prepending the domain-ID of the source domain (domain in which the route is imported).
 - b. If the EEG is doing proxy of multiple received SMET routes which (some or all) already contain the D-PATH attribute, the resulting proxy SMET route MUST contain the best D-PATH of all the contributing SMET routes. The "best" D-PATH is the shortest D-PATH in terms of number of domain-IDs, where no D-PATH means a length of zero. In case two routes with the same number of domain-IDs are left in the selection, a route with the numerically lowest left-most Domain-ID is preferred. In addition, the EEG prepends the domain-ID indicated in point 'a'. As an example, if EEG1 in Figure 1 receives three SMET routes, route 1 with no D-PATH, route 2 with D-PATH (3:3) and route 3 with D-PATH (4:4,3:3), when doing proxy, EEG1 selects the best D-PATH, i.e., zero length D-PATH, and when exporting into domain 1:1, EEG1 adds the D-PATH with domain 2:2 (as per point 'a').
 - c. Upon importing an SMET route, an EEG SHOULD NOT proxy an SMET route into another domain if the route contains a D-PATH with at least one domain-ID that is locally configured in any of the domains of the BD.
6. EVPN Multicast Membership Report Synch or Multicast Leave Synch routes [RFC9251] for the Interconnect Ethernet Segment MUST NOT be generated or imported.
7. An EEG router MAY support local sources and receivers attached to the BD. Local sources/receivers are considered to be part of a "local" domain in the BD, as described in [I-D.ietf-bess-evpn-dpath] for local Attachment Circuits on the Service Gateways.
 - a. Local receivers sending IGMP/MLD membership reports create OIF entries on the connected EEG, and the EEG MUST do proxy of the state into all the EVPN domains for which the EEG is Designated Forwarder. The EEG MAY do also proxy of the SMET routes into the EVPN domains for which the EEG is non-

Designated Forwarder. That is, consider two EEG routers attached to the same two EVPN domains of the same BD as in Figure 1, and EEG1 being the Designated Forwarder router of the Interconnect Ethernet Segment in the domain 2:2, and a local receiver is attached to EEG2. Assuming EEG2 did not export an SMET for a group G earlier, upon receiving a membership report from the local receiver, EEG2 MUST export an SMET route for G into domain 1:1. EEG2 MAY optionally export an SMET route into domain 2:2. SMET routes exported for local receivers SHOULD include the D-PATH attribute with the domain-ID associated with the local domain.

- b. Consider a local source for group G connected to an Interconnect Ethernet Segment non-Designated Forwarder EEG, and a receiver on one of the domains the EEG is interconnecting, e.g., domain 1:1. In this case, the EEG receives an SMET route from domain 1:1 and also from the Designated Forwarder EEG in a different domain. Even if the EEG has OIF entries for domain 1:1, the EEG MUST NOT send multicast traffic to domain 1:1 due to its non-Designated Forwarder state. This prevents the EEG from sending duplicate traffic to the receiver on domain 1:1.
- c. Local sources and receivers MAY be attached to Ethernet Segments. In this case, the EEG follows the procedures in [RFC9251] for synchronizing multicast state and other procedures.

- 8. This specification is compatible with [I-D.ietf-bess-evpn-redundant-mcast-source] section 4 (Warm Standby solution) irrespective of the sources being attached to the same or different EVPN domains.

3. Layer-3 EVPN to EVPN Gateway Procedures

This section provides the specification for EVPN to EVPN Gateways (EEGs) when configured for layer-3 interconnect, as in the use case of Section 1.2. It is important to note that this specification uses a Supplementary Broadcast Domain (SBD) per domain that the EEG is interconnecting as a way to explain the procedures as simplified as possible, however, any implementation that uses a single SBD per tenant, and produces the same control plane and data plane behavior from an external standpoint, is considered compliant with this document.

- 1. An EEG configured for layer-3 interconnect of two or more domains MUST support the gateway procedures in [I-D.ietf-bess-evpn-ipvpn-interworking] section 8, for IP unicast

forwarding between two EVPN domains. To differentiate an EVPN domain in Section 2 from an EVPN domain in a layer-3 interconnect context, this section refers to EVPN domains "of an IP-VRF" on the EEG, as opposed to EVPN domains "of a BD" in Section 2.

- a. An EEG interconnecting two EVPN domains of the same IP-VRF "redistributes" EVPN IP Prefix (and/or MAC/IP) routes and EVPN SMET routes between the domains. "Redistribution" of SMET routes between domains of the same IP-VRF, in this document, refers to the procedures related to importing the SMET route, programing the associated multicast state in the SBD and exporting the SMET route into a different domain.
 - b. When performing this redistribution of SMET routes, the EEG exports the minimum set of SMET routes to attract the traffic for a given multicast group. That is, when at least one (*,G) for a group G has been imported via SMET route in one domain, only an SMET route for (*,G) is exported to the other domain and the (S,G) SMET routes for the same group G (but with specific sources) SHOULD NOT be exported to the other domain.
2. An EEG creates one SBD instance per domain the IP-VRF is interconnecting. The SBD concept is specified in [RFC9625], only that this specification allows more than one SBD per IP-VRF on the EEG routers.
 - a. Each SBD attached to the same IP-VRF SHOULD use different route distinguisher and MAY use a different set of route targets when exporting SMET routes.
 - b. An SBD imports and exports SMET routes as per the procedures in [RFC9625].
 - c. Also this document extends [RFC9625] so that the SBD IRB can be added to the IP-VRF layer-3 OIF list for a group, when the IIF for the group is the IRB of another SBD attached to the same IP-VRF.

3. An EEG originates an EVPN Inclusive Multicast Ethernet Tag route for each SBD to which the IP-VRF is attached. We refer to these routes as SBD-IMET routes and they include a Multicast Flags Extended Community with the EEG Flag set. In addition, the SBD-IMET routes SHOULD also carry a Designated Election Extended Community, as described in [RFC9625] for the SBD-IMET routes on MVPN to EVPN Gateways (MEGs). After collecting all the SBD-IMET routes with the EEG flag set (including the local one), the EEG MUST perform a Designated Router election. The Designated Router election for an SBD:
 - a. MUST follow the procedures of [RFC9625] section 6.1.2.4
 - b. MUST be done only among EEGs that are attached to the same set of SBDs. That is, suppose the local EEG is attached to SBD1 and SBD2. Upon receiving an SBD-IMET route for SBD1 from a remote EEG1, EEG1 is considered a candidate Designated Router for SBD1 only if an SBD-IMET route for SBD2 is also received from EEG1.
 - c. produces a single winner for the SBD, referred to as the SBD-DR (Supplementary Broadcast Domain Designated Router). Upon programming a multicast group G, the SBD-DR in one SBD is responsible for redistributing the SMET route for G received on that SBD into the common set of SBDs of the same IP-VRF (common set to all candidate EEGs which run the election for that SBD).
4. A non-Designated Router for the SBD (non-SBD-DR) MAY redistribute SMET routes between domains but it MUST NOT add the SBD IRB for which it is non-SBD-DR as layer-3 OIF entry. The operator can decide via configuration whether the non-SBD-DR router redistributes SMET routes to other domains. This is a trade-off between attracting unnecessary traffic and speeding up convergence in case of a failure on the SBD-DR.
5. An SBD-DR MUST be selected for each SBD to which the EEG is attached, however, the SBD-DR election MAY be run into only one of the SBDs of the IP-VRF, and the same SBD-DR EEG derived for all SBDs of the IP-VRF.
 - a. For example, if SBD1 and SBD2 are SBDs of the same IP-VRF in EEG1 and EEG2, an implementation MAY run the SBD-DR election only in the context of SBD1 and extrapolate the result to SBD2. That is, if EEG1 is the SBD-DR for SBD1, EEG1 will be the SBD-DR for SBD2 as well.

- b. An implementation that runs the SBD-DR election in only one of the SBDs of the IP-VRF MUST set the EEG Flag and carry the Designated Election Extended Community only in the IMET routes for the SBD(s) that run SBD-DR election. On reception, if an EEG is attached to two (or more) SBDs of the same IP-VRF and receives an IMET per SBD from the redundant EEG, but only one IMET route has the EEG Flag set, the EEG MUST apply the SBD-DR election result to all the SBDs of the IP-VRF.
 6. An EEG router MAY support local sources and receivers, that are attached to Broadcast Domains (BDs) that have IRB interfaces into the IP-VRF of the EEG. Procedures for local sources and receivers follow [RFC9625].
 7. The MEG (MVPN to EVPN Gateway), PEG (PIM to EVPN Gateway) [RFC9625] and EEG (EVPN to EVPN Gateway) procedures MAY be used for the same tenant on the same Service Gateways.
 8. This specification is compatible with [I-D.ietf-bess-evpn-redundant-mcast-source] section 4 (Warm Standby solution) irrespective of the sources being attached to the same or different EVPN domains.
- #### 4. Security Considerations

This document extends the procedures of [RFC9251] and [RFC9625], in the scenarios described by [RFC9014] and [I-D.ietf-bess-evpn-ipvpn-interworking]. Therefore it inherits all the Security Considerations described in all those specifications. In addition, this document now allows the distribution of SMET routes across EVPN domains, and therefore provides a new tool for an attacker to be able to leak SMET routes into a remote EVPN domain that could not receive SMET routes from a remote domain prior to this specification. An attacker that manages to leak SMET routes into remote domains, may attract multicast traffic that may not be leaked otherwise into the local domain.

5. IANA Considerations

This document requests a new Flag in the subregistry called "Multicast Flags Extended Community", under the "Border Gateway Protocol (BGP) Extended Communities" registry, to indicate EEG support along with the IMET routes.

Bit	Name	Reference

TBD	EVPN to EVPN Gateway (EEG)	This document

Figure 3

6. Contributors

7. Acknowledgments

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>>.
- [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>>.
- [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>>.
- [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-10, 24 July 2024, <<https://datatracker.ietf.org/doc/html/draft-ietf-bess-rfc7432bis-10>>.
- [RFC9251] Sajassi, A., Thoria, S., Mishra, M., Patel, K., Drake, J., and W. Lin, "Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Proxies for Ethernet VPN (EVPN)", RFC 9251, DOI 10.17487/RFC9251, June 2022, <<https://www.rfc-editor.org/info/rfc9251>>.
- [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>>.
- [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, "EVPN Interworking with IPVPN", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-ipvpn-interworking-13, 4 February 2025, <<https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-ipvpn-interworking-13>>.

[I-D.ietf-bess-evpn-dpath]

Rabadan, J., Sathappan, S., Gautam, M., Brissette, P., and W. Lin, "Domain Path (D-PATH) for Ethernet VPN (EVPN) Interconnect Networks", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-dpath-02, 3 March 2025, <<https://datatracker.ietf.org/api/v1/doc/document/draft-ietf-bess-evpn-dpath/>>.

8.2. Informative References

- [RFC6513] Rosen, E., Ed. and R. Aggarwal, Ed., "Multicast in MPLS/BGP IP VPNs", RFC 6513, DOI 10.17487/RFC6513, February 2012, <<https://www.rfc-editor.org/info/rfc6513>>.
- [RFC6514] Aggarwal, R., Rosen, E., Morin, T., and Y. Rekhter, "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", RFC 6514, DOI 10.17487/RFC6514, February 2012, <<https://www.rfc-editor.org/info/rfc6514>>.
- [RFC7761] Fenner, B., Handley, M., Holbrook, H., Kouvelas, I., Parekh, R., Zhang, Z., and L. Zheng, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", STD 83, RFC 7761, DOI 10.17487/RFC7761, March 2016, <<https://www.rfc-editor.org/info/rfc7761>>.
- [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-redundant-mcast-source]
- Rabadan, J., Kotalwar, J., Sathappan, S., Zhang, Z. J., and W. Lin, "Multicast Source Redundancy in EVPN Networks", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-redundant-mcast-source-15, 14 February 2025, <<https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-redundant-mcast-source-15>>.

Authors' Addresses

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

Olivier Dornon
Nokia
Email: olivier.dornon@nokia.com

Vinod Prabhu
Nokia
Email: vinod.prabhu@nokia.com

Alex Nichol
Arista
Email: anichol@arista.com

Zhaohui Zhang
Juniper Networks
Email: zzhang@juniper.net

Wen Lin
Juniper Networks
Email: wlin@juniper.net