BESS Workgroup
Internet-Draft
Intended status: Standards Track
Expires: 20 December 2025

J. Rabadan, Ed.
Nokia
A. Sajassi, Ed.
Cisco
E. Rosen
Individual
J. Drake
Independent
W. Lin
Juniper
J. Uttaro
Independent
A. Simpson
Nokia

18 June 2025

EVPN Interworking with IPVPN
draft-ietf-bess-evpn-ipvpn-interworking-14

Abstract

Ethernet Virtual Private Network (EVPN) is used as a unified control plane for tenant network intra and inter-subnet forwarding. When a tenant network spans not only EVPN domains but also domains where BGP VPN-IP or IP families provide inter-subnet forwarding, there is a need to specify the interworking aspects between BGP domains of type EVPN, VPN-IP and IP, so that the end-to-end tenant connectivity can be accomplished. This document specifies how EVPN interworks with VPN-IPv4/VPN-IPv6 and IPv4/IPv6 BGP families for inter-subnet forwarding. The document also addresses the interconnect of EVPN domains for Inter-Subnet Forwarding routes. In addition, this specification defines a new BGP Path Attribute called D-PATH (Domain PATH) that protects gateways against control plane loops. D-PATH modifies the BGP best path selection for multiprotocol BGP routes of SAFI 128 and EVPN IP Prefix routes, and therefore this document updates the BGP best path selection specification, but only for IPVPN and EVPN families.

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/.

Copyright Notice

Table of Contents

1. Introduction and Problem Statement

EVPN is used as a unified control plane for tenant network intra and inter-subnet forwarding. When a tenant network spans not only EVPN domains but also domains where BGP VPN-IP or IP families provide inter-subnet forwarding, there is a need to specify the interworking aspects between the different families, so that the end-to-end tenant connectivity can be accomplished. This document specifies how EVPN should interwork with VPN-IPv4/VPN-IPv6 and IPv4/IPv6 BGP families for inter-subnet forwarding. The document also addresses the interconnect of an EVPN domain to another EVPN domain for Inter-Subnet Forwarding routes. In addition, this specification defines a new BGP Path Attribute called D-PATH (Domain PATH) that protects gateways against control plane loops. Loops are created when two (or more) redundant gateway PEs interconnect two domains and exchange inter-subnet forwarding routes. For instance, if PE1 and PE2 are redundant gateway PEs interconnecting an IPVPN and an EVPN domain, gateway PE1 receives a VPN-IP route to prefix P and propagates the route into an EVPN IP Prefix to P. If gateway PE2 receives the EVPN IP Prefix route, it cannot propagate the route back to the IPVPN domain, or it would create a loop for prefix P.

D-PATH modifies the BGP best path selection for multiprotocol BGP routes of SAFI 128 and EVPN IP Prefix routes, and therefore this document updates the BGP best path selection procedures in [RFC4271] for IPVPN and EVPN families.

EVPN supports the advertisement of IPv4 or IPv6 prefixes in two different route types:

* Route Type 2 - EVPN MAC/IP Advertisement route (only for /32 and /128 host routes), as described by [RFC9135].

* Route Type 5 - EVPN IP Prefix route, as described by [RFC9136].

When interworking with other BGP address families (AFIs/SAFIs) for inter-subnet forwarding, the IP prefixes in those two EVPN route types must be propagated to other domains using different SAFIs. Some aspects of that propagation must be clarified. Examples of these aspects or procedures across BGP families are: route selection, loop prevention or BGP Path attribute propagation. The Interworking PE concepts are defined in Section 3, and the rest of the document describes the interaction between Interworking PEs and other PEs for end-to-end inter-subnet forwarding.

The interworking procedures in this document always require creating an IP-VRF on the interworking PE. When connecting different domains, the interworking PE follows these steps: it receives routes from one

domain (along with that domain' s encapsulation parameters), installs them in the IP-VRF route table, and then reoriginates the routes with the encapsulation parameters of the adjacent domain before advertising them. This reorigination process ensures that the procedures remain independent of the specific transport tunnels used in each domain, as it functions as a service interworking solution.

2. Conventions used in this document

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.

3. Terminology and Interworking PE Components

This section summarizes the terminology related to the "Interworking PE" concept that will be used throughout the rest of the document.

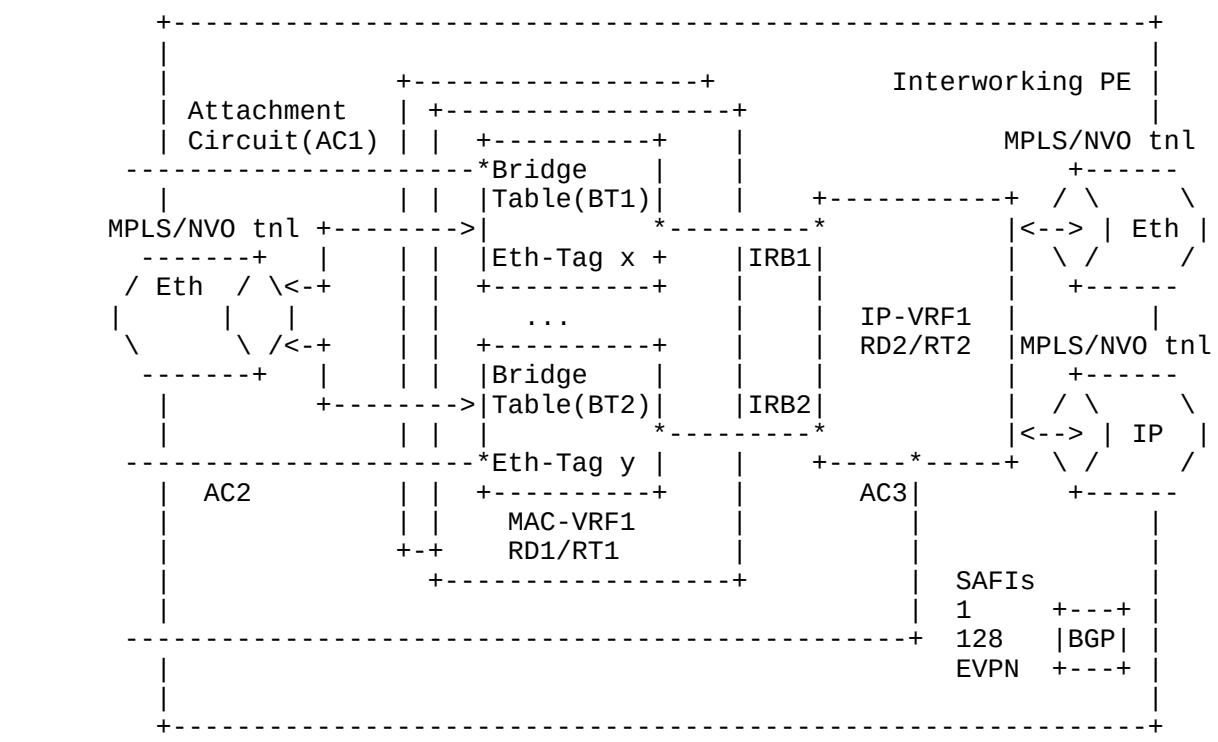

Figure 1: EVPN-IPVPN Interworking PE

* ISF SAFI: Inter-Subnet Forwarding (ISF) SAFI is a MP-BGP Sub-Address Family that advertises reachability for IP prefixes and can be used for inter-subnet forwarding within a given tenant network. The ISF SAFIs are 1 (including IPv4 and IPv6 AFIs), 128 (including IPv4 and IPv6 AFIs) and 70 (EVPN, including only AFI 25). This document uses the following terms interchangeably: ISF SAFI 1 or BGP IP, ISF SAFI 128 or IPVPN, ISF SAFI 70 or EVPN.

* ISF route: a route for a given prefix, whose ISF SAFI may change as it transits different domains. BGP IP routes as in [RFC4760] [RFC8950], IPVPN routes as in [RFC4364], [RFC4659], or EVPN IP Prefix routes as in [RFC9136], are considered ISF routes in this document.

* IP-VRF: an IP Virtual Routing and Forwarding table, as defined in [RFC4364]. Route Distinguisher and Route Target(s) are required properties of an IP-VRF. An IP-VRF is programmed with ISF routes.

* MAC-VRF: a MAC Virtual Routing and Forwarding table, as defined in [RFC7432]. It is also the instantiation of an EVI (EVPN Instance) in a PE. Route Distinguisher and Route Target(s) are required properties and they are normally different than the ones defined in the associated IP-VRF (if there is an associated IP-VRF linked to a Bridge Table of the MAC-VRF, via IRB interface).

* BT: a Bridge Table, as defined in [RFC7432]. A BT is the instantiation of a Broadcast Domain in a PE. When there is a single Broadcast Domain in a given EVI, the MAC-VRF in each PE will contain a single BT. When there are multiple BTs within the same MAC-VRF, each BT is associated to a different Ethernet Tag. The EVPN routes specific to a BT, will indicate which Ethernet Tag the route corresponds to.

  Example: In Figure 1, MAC-VRF1 has two BTs: BT1 and BT2. Ethernet Tag x is defined in BT1 and Ethernet Tag y in BT2.

* AC: Attachment Circuit or logical interface associated to a given BT or IP-VRF. To determine the AC on which a packet arrived, the PE will examine the combination of a physical port and VLAN tags (where the VLAN tags can be individual c-tags, s-tags or ranges of both).

  Example: In Figure 1, AC1 is associated to BT1, AC2 to BT2 and AC3 to IP-VRF1.

* IRB: Integrated Routing and Bridging interface. It refers to the logical interface that connects a BT to an IP-VRF and allows to forward packets with destination in a different subnet.

* MPLS/NVO tunnel: It refers to a tunnel that can be MPLS or NVO-based (Network Virtualization Overlays) and it is used by MAC-VRFs and IP-VRFs. Irrespective of the type, the tunnel may carry an Ethernet or an IP payload. MAC-VRFs can only use tunnels with Ethernet payloads (setup by EVPN), whereas IP-VRFs can use tunnels with Ethernet (setup by EVPN) or IP payloads (setup by EVPN or IPVPN). IPVPN-only PEs have IP-VRFs but they cannot send or receive traffic on tunnels with Ethernet payloads.

  Example: Figure 1 shows an MPLS/NVO tunnel that is used to transport Ethernet frames to/from MAC-VRF1. The PE determines the MAC-VRF and BT the packets belong to based on the EVPN label (MPLS or VNI). Figure 1 also shows two MPLS/NVO tunnels being used by IP-VRF1, one carrying Ethernet frames and the other one carrying IP packets.

* RT-2: Route Type 2 or MAC/IP route, as per [RFC7432].

* RT-5: Route Type 5 or IP Prefix route, as per [RFC9136].

* NVE: Network Virtualization Edge router.

* Domain: Two PEs are in the same domain if they are attached to the same tenant and the packets between them do not require a data path IP lookup (in the tenant space) in any intermediate router. A gateway PE is always configured with multiple DOMAIN-IDs. The domain boundaries are not limited to an Autonomous System or an IGP instance. The PEs in a domain can all be part of the same or different Autonomous System, and an Autonomous System can also contain multiple domains.

  Example 1: Figure 2 depicts an example where Tenant Systems TS1 and TS2 belong to the same tenant, and they are located in different Data Centers that are connected by gateway PEs (see the gateway PE definition later). These gateway PEs use IPVPN in the WAN. When TS1 sends traffic to TS2, the intermediate routers between PE1 and PE2 require a tenant IP lookup in their IP-VRFs so that the packets can be forwarded. In this example there are three different domains. The gateway PEs connect the EVPN domains to the IPVPN domain.

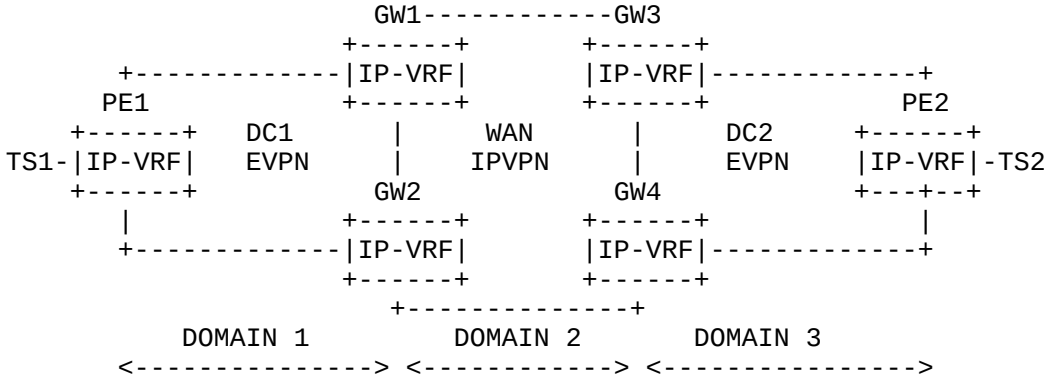

Figure 2: Multiple domain DCI example

Example 2: Figure 3 illustrates a similar example, but PE1 and PE2 are now connected by a BGP-LU (BGP Labeled Unicast) tunnel, and they have a BGP peer relationship for EVPN. Contrary to Example 1, there is no need for tenant IP lookups on the intermediate routers in order to forward packets between PE1 and PE2. Therefore, there is only one domain in the network and PE1/PE2 belong to it.

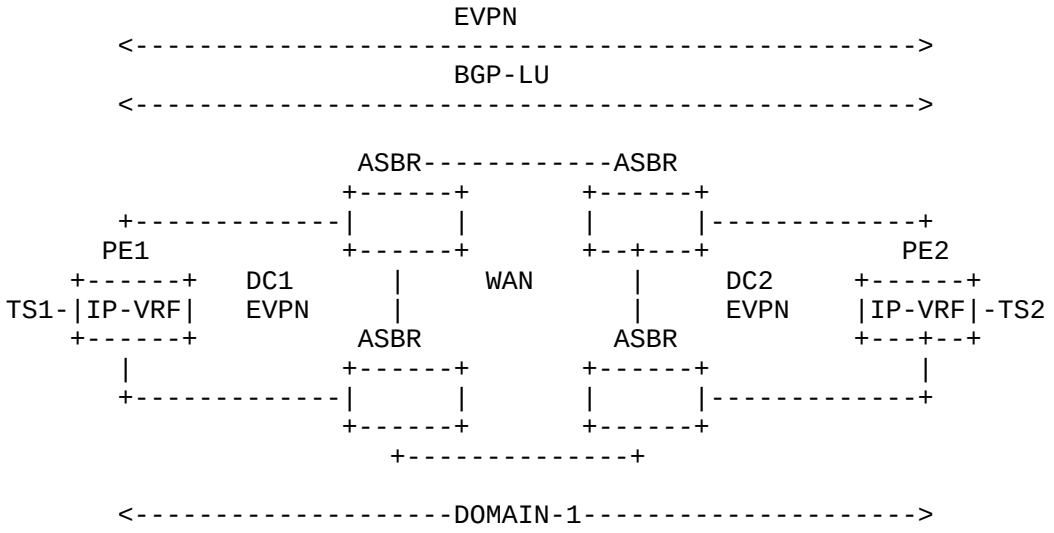

Figure 3: Single domain DCI example

* Regular Domain: a domain in which a single control plane ISF SAFI, i.e., BGP IP, IPVPN or EVPN, is used. A Regular Domain is composed of regular PEs, see below. In Figure 2 and Figure 3, above, all domains are regular domains.

* Composite Domain: a domain in which multiple control plane ISF SAFIs, i.e., BGP IP, IPVPN and/or EVPN, are used and which is composed of regular PEs and composite PEs, see below.

* Regular PE: a PE that is attached to a domain, either regular or composite, and which uses one of the control plane protocols (BGP IP, IPVPN or EVPN) operating in the domain.

* Interworking PE: a PE that may advertise a given prefix with an EVPN ISF route (EVPN MAC/IP Advertisement route or EVPN IP Prefix route) and/or an IPVPN ISF route and/or a BGP IP ISF route. An interworking PE has one IP-VRF per tenant, and zero, one or multiple MAC-VRFs per tenant. Each MAC-VRF may contain one or more BTs, where each BT may be attached to that IP-VRF via IRB. There are two types of Interworking PEs: composite PEs and gateway PEs. Both PE functions can be independently implemented per tenant and they may both be implemented for the same tenant.

  Example: Figure 1 shows an interworking PE of type gateway, where ISF SAFIs 1, 128 and 70 are enabled. IP-VRF1 and MAC-VRF1 are instantiated on the PE, and together provide inter-subnet forwarding for the tenant.

* Composite PE: an interworking PE that is attached to a composite domain and advertises a given prefix to an IPVPN peer with an IPVPN ISF route, to an EVPN peer with an EVPN ISF route, and to a route reflector with both an IPVPN and EVPN ISF route. A composite PE performs the procedures of Section 7.

  Example: Figure 4 shows an example where PE1 is a composite PE since PE1 has EVPN and another ISF SAFI enabled to the same route-reflector, and PE1 advertises a given IP prefix IPn/x twice, one using EVPN and another one using ISF SAFI 128. PE2 and PE3 are not composite PEs.

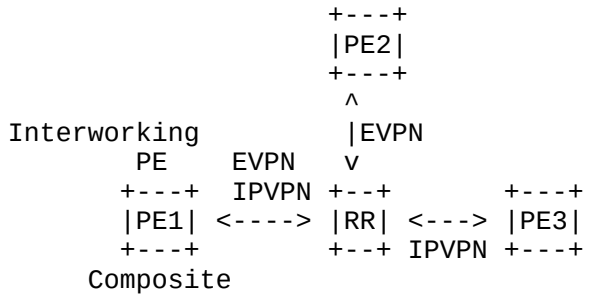

Figure 4: Interworking composite PE example

* Gateway PE: an interworking PE that is attached to two (or more) domains, each either regular or composite. The Gateway PE can have IBGP and/or EBGP peers on the domains that is connecting. Based on configuration, the Gateway PE does one of the following:

  - Propagates ISF routes of the same ISF SAFI, i.e., BGP IP, IPVPN or EVPN, between the two domains.

  - Propagates an ISF route received with an ISF SAFI to a domain that uses a different ISF SAFI. E.g., it propagates a received EVPN ISF route as an IPVPN ISF route in the other domain and vice versa. A gateway PE performs the procedures of Section 8.

  A gateway PE is always configured with multiple DOMAIN-IDs. The DOMAIN-ID is encoded in the Domain Path Attribute (D-PATH), and advertised along with ISF SAFI routes. Section 4 describes the D-PATH attribute.

  Example: Figure 5 illustrates an example where PE1 is a gateway PE since the EVPN and IPVPN SAFIs are enabled on different BGP peers, and a given local IP prefix IPn/x is sent to both BGP peers for the same tenant. PE2 and PE1 are in one domain and PE3 and PE1 are in another domain.

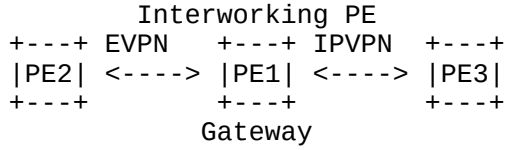

Figure 5: Interworking gateway PE example

* Composite/Gateway PE: an interworking PE that is both a composite PE and a gateway PE that is attached to two domains, one regular and one composite, and which does the following:

- Propagates an ISF route from the regular domain into the composite domain. Within the composite domain it acts as a composite PE.

- Propagates an ISF route from the composite domain into the regular domain. Within the regular domain it is propagated as an ISF route using the ISF SAFI for that domain.

This is particularly useful when a tenant network uses multiple ISF SAFIs (BGP IP, IPVPN and EVPN domains) and any-to-any connectivity is required. In this case end-to-end control plane consistency, when possible, is desired.

4. Domain Path Attribute (D-PATH)

The BGP Domain Path (D-PATH) attribute is an optional and transitive BGP path attribute.

Similar to AS_PATH, D-PATH is composed of a sequence of Domain segments. Each Domain segment is comprised of <domain segment length, domain segment value>, where the domain segment value is a sequence of one or more Domains, as illustrated in Figure 6. Each domain is represented by <DOMAIN-ID:ISF_SAFI_TYPE>.

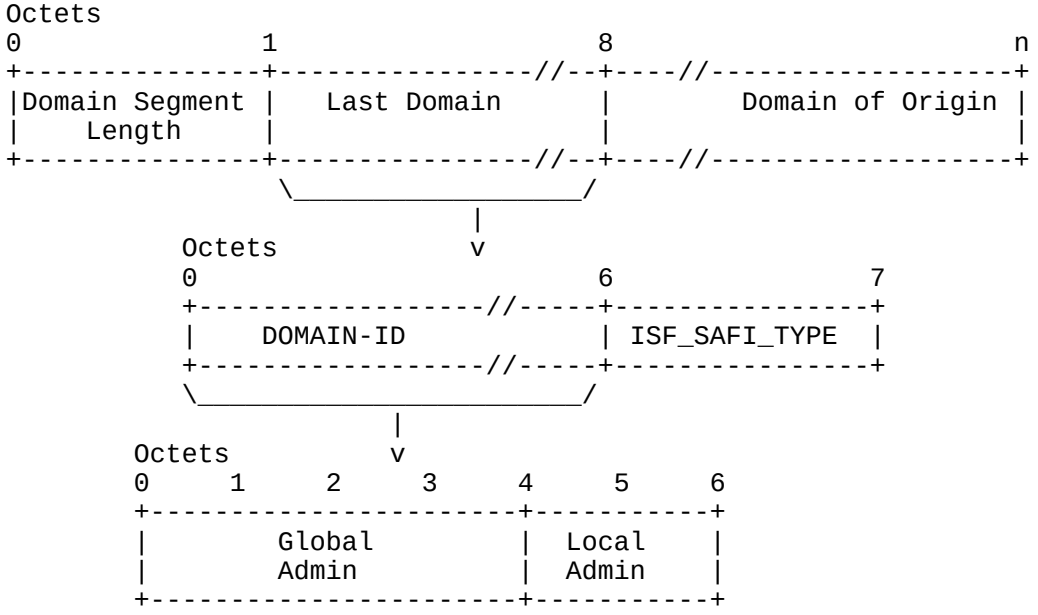

Figure 6: D-PATH Domain Segment

* The domain segment length field is a 1-octet field, containing the number of domains in the segment.

* DOMAIN-ID is a 6-octet field that represents a domain. It is composed of a 4-octet Global Administrator sub-field and a 2-octet Local Administrator sub-field. The Global Administrator sub-field MAY be filled with an Autonomous System Number (ASN, Public or Private), an IPv4 address, or any value that guarantees the uniqueness of the DOMAIN-ID (when the tenant network is connected to multiple Operators) and helps troubleshooting and debugging of D-PATH in ISF routes. The Local Administrator sub-field is any local 2-octet value, and its allocation or configuration is a local implementation matter. The representation of the Global Administrator and Local Administrator values as opaque unsigned integers is RECOMMENDED.

* ISF_SAFI_TYPE is a 1-octet field that indicates the Inter-Subnet Forwarding SAFI type in which a route was received, before the route is re-exported into a different domain. The ISF_SAFI_TYPE field is informational and does not have any impact on the loop detection or BGP Path selection procedures. The following types are assigned by this document:

Value	Type
0	Gateway PE local ISF route
70	EVPN
128	SAFI 128

Table 1

The BGP D-PATH attribute is supported on ISF routes of type IPVPN and EVPN and MUST NOT be advertised along with routes different from IPVPN and EVPN routes. By default, the BGP D-PATH attribute is not advertised and MUST be explicitly enabled by configuration on the Gateway PEs. In addition, D-PATH:

a. Identifies the sequence of domains, each identified by a <DOMAIN-ID:ISF_SAFI_TYPE> through which a given ISF route of type IPVPN or EVPN has passed.

    * This attribute list MAY contain one or more segments. Each segment’s Domain Segment Length MUST be equal or greater than one.

* The first entry in the list (leftmost) is the <DOMAIN-ID:ISF_SAFI_TYPE> from which a gateway PE is propagating an ISF IPVPN or EVPN route. The last entry in the list (rightmost) is the <DOMAIN-ID:ISF_SAFI_TYPE> from which a gateway PE received an ISF IPVPN or EVPN route without a D-PATH attribute (the Domain of Origin). Intermediate entries in the list are domains that the ISF IPVPN or EVPN route has transited.

* As an example, an ISF IPVPN or EVPN route received with a D-PATH attribute containing a domain segment of {length=2, <6500:2:IPVPN>,<6500:1:EVPN>} indicates that the route was originated in EVPN domain 6500:1, and propagated into IPVPN domain 6500:2.

* In order to minimize the number of segments in the D-PATH attribute, the local gateway PE prepends its own domain as the last element of the domain segment. If the act of prepending a new domain causes an overflow in the domain segment (i.e., more than 255 domains), the local gateway PE MUST prepend a new segment and prepend its own domain to this new segment.

b. It is added/modified by a gateway PE when propagating an update to a different domain (which runs the same or different ISF SAFI):

* A gateway PE’s IP-VRF, that connects two domains, belongs to two DOMAIN-IDs, e.g. 6500:1 for EVPN and 6500:2 for IPVPN.

* Whenever a prefix arrives at a gateway PE in a particular ISF SAFI route, if the gateway PE needs to export that prefix to a BGP peer, the gateway PE MUST prepend a <DOMAIN-ID:ISF_SAFI_TYPE> to the list of domains in the D-PATH of the received route, as long as the gateway PE works in Uniform-Propagation-Mode, as explained in Section 5.2.

* For instance, in an IP-VRF configured with DOMAIN-IDs 6500:1 for EVPN and 6500:2 for IPVPN, if an EVPN route for prefix P is received and P installed in the IP-VRF, the IPVPN route for P that is exported to an IPVPN peer will prepend the domain <6500:1:EVPN> to the previously received D-PATH attribute. Likewise, IP-VRF prefixes that are received from IP-VPN, will be exported to EVPN peers with the domain <6500:2:IPVPN> added to the segment.

* In the above example, if the EVPN route is received without D-PATH, the gateway PE will add the D-PATH attribute with one segment {length=1, <6500:1:EVPN>} when re-advertising to domain 6500:2.

* Within the Domain of Origin, the update does not contain a D-PATH attribute because the update has not passed through a gateway PE yet.

c. For a local ISF route, i.e., a configured route or a route learned from a local attachment circuit, a gateway PE has three choices:

1. It MAY advertise that ISF route without a D-PATH attribute into one or more of its configured domains, in which case the D-PATH attribute will be added by the other gateway PEs in each of those domains.

2. It MAY advertise that ISF route with a D-PATH attribute into one or more of its configured domains, in which case the D-PATH attribute in each copy of the ISF route is initialized with an ISF_SAFI_TYPE of 0 and the DOMAIN-ID of the domain with which the ISF route is associated.

3. It MAY advertise that ISF route with a D-PATH attribute that contains a configured domain specific to its local ISF routes into one or more of its configured domains, in which case the D-PATH attribute in each copy of the ISF route is initialized with a ISF_SAFI_TYPE of 0 and the DOMAIN-ID for the local ISF routes. This DOMAIN-ID MUST be globally unique and MAY be shared by two or more gateway PEs. Although the three options help detect control plane loops, this option 3 is RECOMMENDED, since it is the option that provides more information about the differentiated origin of the route (it uses a DOMAIN-ID and ISF_SAFI_TYPE that identifies the route as a local gateway route).

d. An ISF IPVPN or EVPN route received by a gateway PE with a D-PATH attribute that contains one or more of its locally associated DOMAIN-IDs (for the IP-VRF) is considered to be a looped ISF route for the purpose of re-exporting the route to the adjacent domain in a Gateway PE. The ISF route in this case MUST be flagged as "looped", MUST NOT be exported, and MAY be installed in the IP-VRF only in case there is no better route after the best path selection (Section 6). The ISF_SAFI_TYPE is irrelevant for the purpose of loop detection of an ISF route. In other words, an ISF route is considered as a looped route if it contains a D-PATH attribute with at least one DOMAIN-ID matching a local DOMAIN-ID, irrespective of the ISF_SAFI_TYPE of the DOMAIN-ID.

   For instance, in the example of Figure 2, gateway GW1 receives TS1 prefix in two different ISF routes:

   * In an EVPN IP Prefix route with next-hop PE1 and no D-PATH attribute.

   * In a SAFI 128 route with next-hop GW2 and D-PATH = {length=1, <6500:1:EVPN>}, assuming that DOMAIN-ID for domain 1 is 6500:1.

   Gateway GW1 flags the SAFI 128 route as "looped" (since 6500:1 is a local DOMAIN-ID in GW1) and it will not install it in the tenant IP-VRF, since the route selection process selects the EVPN IP Prefix route due to a shorter D-PATH attribute (Section 6). Gateway GW1 identifies the route as "looped" even if the ISF_SAFI_TYPE value is unknown to GW1, i.e., any value different from the ones specified in this document).

e. A DOMAIN-ID value on a gateway PE MAY be assigned for a peering domain or MAY be scoped for an individual tenant IP-VRF.

   * If allocated for a peering domain, the DOMAIN-ID applies to all tenant IP-VRFs for that domain.

* If allocated for a specific tenant IP-VRF, the processing of the received D-PATH and its propagation is in the context of the IP-VRF DOMAIN-ID. Route leaking is a use-case where a per-IP-VRF DOMAIN-ID assignment is necessary. Suppose gateways PE1 and PE2 are attached to two different tenant IP-VRFs, IP-VRF-1 and IP-VRF-2. ISF SAFI routes advertised by gateway PE1 for IP-VRF-1 are received on gateway PE2 with DOMAIN-ID 6500:1. If the routes are leaked from IP-VRF-1 into IP-VRF-2 on PE2, and re-advertised back to PE1 in the context of IP-VRF-2, PE1 will not identify the route as a looped route. This is because PE1 processes the route in the context of IP-VRF-2, where DOMAIN-ID 6500:1 is not a local DOMAIN-ID.

f. The number of domains in the D-PATH attribute indicates the number of gateway PEs that the ISF route update has transited. If one of the transit gateway PEs leaks a given ISF route between two local IP-VRFs, it MAY prepend a domain with a ISF_SAFI_TYPE of 0 for the leaked route when the route is exported into an ISF SAFI. In that case, the number of domains in the D-PATH attribute indicates the number of tenant IP-VRFs that the ISF route update has transited.

g. The following error-handling rules apply to the D-PATH attribute:

1. A received D-PATH attribute MUST be considered malformed if it contains a malformed Domain Segment.

2. A Domain Segment MUST be considered malformed in any of the following cases:

   * The length of the Domain Segment is longer than the D-PATH attribute that contains it.

   * After the last successfully parsed Domain Segment there are less than eight octets remaining.

   * The D-PATH attribute length is less than 8 octets.

   * For each contained Domain Segment, the Domain Segment length is one octet, containing the number of Domains in this segment, each of which are 7 octets in length. If the total length of the Domain Segment in octets (1 + 7 * number of Domains) exceeds the remaining length of the D-PATH attribute, the Domain Segment is malformed.

3. A PE receiving an UPDATE message with a malformed D-PATH attribute SHALL apply "treat-as-withdraw" [RFC7606].

4. Domains in the D-PATH attribute with unknown ISF_SAFI_TYPE values are accepted and not considered an error.

5. The D-PATH Path Attribute MUST NOT occur more than once in the BGP UPDATE's Path Attributes. If the D-PATH Path Attribute appears more than once in an UPDATE message, then all the occurrences of the attribute other than the first one are discarded and the UPDATE message will continue to be processed, as per [RFC7606].

6. D-PATH can be advertised with SAFI 128 and EVPN routes and MUST NOT be sent with any other AFI/SAFIs. If D-PATH is received along with routes of AFI/SAFI different from the IPVPN and EVPN families, the behavior treat-as-withdraw is applied [RFC7606].

h. The use of D-PATH is restricted to "walled garden" Virtual Private Networks, and the operator MUST NOT turn on the generation of D-PATH along with IPVPN and/or EVPN routes if there are CEs attached to a PE (of any domain in the Virtual Private Network) that are connected to the Internet. In addition, a Gateway PE MUST support the removal of the D-PATH attribute on import and on export, based on configuration.

5. BGP Path Attribute Propagation across Domains

Based on its configuration, a gateway PE is required to propagate an ISF route between two domains that use the same or different ISF SAFI. This requires a definition of what a gateway PE has to do with BGP Path Attributes attached to the ISF route that the gateway PE is propagating. This section specifies the BGP Path Attribute propagation modes that a gateway PE may follow when receives an ISF route with ISF SAFI-x, installs the route in the IP-VRF and exports the ISF route into ISF SAFI-y. ISF SAFI-x and SAFI-y values MAY be the same values.

5.1. No-Propagation-Mode

This is the default mode of operation for gateway PEs that re-export ISF routes from a domain into another domain. In this mode, the gateway PE will simply re-initialize the BGP Path Attributes when propagating an ISF route, as it would for direct or local IP prefixes. This model may be enough in those use-cases where, e.g., the EVPN domain is considered an "abstracted" CE and remote IPVPN/IP PEs don't need to consider the original EVPN Attributes for path computations.

Since this mode of operation does not propagate the D-PATH attribute either, redundant gateway PEs are exposed to routing loops. Those loops may be resolved by policies and the use of other attributes, such as the Route Origin extended community [RFC4360], however not all the loop situations may be identified.

5.2. Uniform-Propagation-Mode

In this mode, the gateway PE simply keeps accumulating or mapping certain key commonly used BGP Path Attributes when propagating an ISF route. This mode is typically used in networks where EVPN and IPVPN SAFIs are used seamlessly to distribute IP prefixes.

The following rules MUST be observed by the gateway PE when propagating BGP Path Attributes:

1. The gateway PE imports an ISF route in the IP-VRF and stores the original Path Attributes. The following set of Path Attributes SHOULD be propagated by the gateway PE when advertising the ISF route to a different domain (other BGP Path Attributes SHOULD NOT be propagated):

    * AS_PATH

    * D-PATH, only when advertising SAFI 128 and EVPN routes.

    * IBGP-only Path Attributes (when advertising to IBGP peers): LOCAL_PREF, ORIGINATOR_ID, CLUSTER_ID

    * MED

    * AIGP

    * Communities, Extended Communities, Large Communities and Wide Communities [I-D.ietf-idr-wide-bgp-communities], except in the exception cases detailed in point 4 of this section.

2. When propagating an ISF route to a different IBGP peer, the gateway PE SHOULD keep the AS_PATH of the originating ISF route and add it to the destination ISF SAFI without any modification. When re-advertising to an EBGP peer, the gateway PE SHOULD keep the AS_PATH of the originating ISF route and prepend the IP-VRF’s AS before sending the route.

3. When propagating an ISF route to IBGP peers, the gateway PE SHOULD keep the IBGP-only Path Attributes from the originating route to the re-advertised route. As the ISF route is re-originated, the route reflector function [RFC4456] is not required on the gateway PE.

4. As discussed in point 1, Communities, Extended Communities, Large Communities and Wide Communities SHOULD be preserved from the originating ISF route by the gateway PE. Exceptions of Extended Communities that SHOULD NOT be propagated are:

   a. BGP Encapsulation extended communities [RFC9012].

   b. Route Target extended communities. Route Targets are always initialized when readvertising an ISF route into a different domain, i.e., they are not propagated. The initialized Route Target in the re-advertised ISF route may or may not have the same value as the Route Target of the originating ISF route.

   c. All the extended communities of type EVPN.

   The gateway PE SHOULD NOT copy the above extended communities from the originating ISF route to the re-advertised ISF route.

5. For a given ISF route, only the BGP Path Attributes of the best path can be propagated to another ISF route. If multiple paths are received for the same route in an ISF SAFI, the BGP best path selection will determine what the best path is, and therefore the set of Path Attributes to be propagated. Even if Equal Cost Multi-Path (ECMP) is enabled on the IP-VRF by policy, only the BGP Path Attributes of the selected best path are propagated.

5.3. Aggregation of Routes and Path Attribute Propagation

Instead of propagating a high number of (host) ISF routes between domains, a gateway PE that receives multiple ISF routes from a domain MAY choose to propagate a single ISF aggregate route into a different domain. In this document, aggregation is used to combine the characteristics of multiple ISF routes in such way that a single aggregate ISF route can be propagated to the destination domain. Aggregation of multiple ISF routes of one ISF SAFI into an aggregate ISF route is only done by a gateway PE.

Aggregation on gateway PEs may use either the No-Propagation-Mode or the Uniform-Propagation-Mode explained in Section 5.1 and Section 5.2, respectively.

When using Uniform-Propagation-Mode, Path Attributes of the same type code MAY be aggregated according to the following rules:

* AS_PATH is aggregated based on the rules in [RFC4271]. The gateway PEs are not expected to receive AS_PATH attributes with path segments of type AS_SET [RFC9774]. Routes received with AS_PATH attributes including AS_SET path segments MUST NOT be aggregated.

* An ISF aggregate route SHOULD NOT be advertised unless all the contributing ISF routes have the same D-PATH DOMAIN-ID members, regardless of their order. If there is at least one contributing ISF route that has a different D-PATH DOMAIN-ID, the gateway PE SHOULD advertise each contributing ISF route with its own D-PATH (prepended with the gateway's domain). An implementation MAY override this behavior, via policy, to advertise an ISF aggregate route without D-PATH in case the contributing routes did not have the same D-PATH DOMAIN-ID members.

* The Community, Extended Community and Large Community attributes of the aggregate ISF route SHOULD contain all the Communities/ Extended Communities/Large Communities from all of the aggregated ISF routes, with the exceptions of the extended communities listed in Section 5.2 that are not propagated.

* For other attributes, rules in [RFC4271] are followed.

Assuming the aggregation can be performed (the above rules are applied), the operator should consider aggregation to deal with scaled tenant networks where a significant number of host routes exists. For example, large Data Centers.

6. Route Selection Process for ISF Routes

A PE may receive an IP prefix in ISF routes with different ISF SAFIs, from the same or different BGP peer. It may also receive the same IP prefix (host route) in an EVPN MAC/IP Advertisement route and EVPN IP Prefix route. A route selection algorithm across all ISF SAFIs is needed so that:

* Different gateway and composite PEs have a consistent and deterministic view on how to reach a given prefix.

* Prefixes advertised in EVPN and other ISF SAFIs can be compared based on path attributes commonly used by operators across networks.

* Equal Cost Multi-Path (ECMP) is allowed across EVPN and other ISF SAFI routes.

For a given prefix advertised in one or more non-EVPN ISF routes, the BGP best path selection procedure will produce a set of "non-EVPN best paths". For a given prefix advertised in one or more EVPN ISF routes, the BGP best path selection procedure will produce a set of "EVPN best paths". To support EVPN/non-EVPN ISF interworking in the context of the same IP-VRF receiving non-EVPN and EVPN ISF routes for the same prefix, it is then necessary to run a tie-breaking selection algorithm on the union of these two sets. This tie-breaking algorithm begins by considering all EVPN and other ISF SAFI routes, equally preferable routes to the same destination, and then selects routes to be removed from consideration. The process terminates as soon as only one route remains in consideration.

The route selection algorithm must remove from consideration the routes following the rules and the order defined in [RFC4271], with the following exceptions and in the following order:

1. Immediately after removing from consideration all routes that are not tied for having the highest Local Preference, any routes that do not have the shortest D-PATH are also removed from consideration. Routes with no D-PATH are considered to have a zero-length D-PATH. A BGP speaker MUST skip this rule for ISF SAFI routes that are not imported in an IP-VRF.

2. Then regular [RFC4271] selection criteria is followed.

3. At the end of the selection algorithm, if at least one route still under consideration is an EVPN MAC/IP Advertisement route, remove from consideration any EVPN IP Prefix routes.

4. If Steps 1-3 leave Equal Cost Multi-Paths (ECMP) between non-EVPN and EVPN paths, the EVPN path MUST be considered (and the non-EVPN path removed from consideration). However, if ECMP across ISF SAFIs is enabled by policy, and one EVPN path and one non-EVPN path remain at the end of step 3, both path types MUST be used.

The above process modifies the [RFC4271] selection criteria for multiprotocol BGP routes with SAFI 128 and EVPN IP Prefix routes to include the shortest D-PATH so that operators minimize the number of Gateways and domains through which packets need to be routed. D-PATH does not modify the selection process for routes different from SAFI 128 or EVPN routes (received routes with other SAFIs get a treat-as-withdraw behavior as described in Section 4).

Example 1 - PE1 receives the following routes for IP1/32, that are candidate to be imported into IP-VRF-1:

{SAFI=EVPN, RT-2, Local-Pref=100, AS-Path=(100,200)}
{SAFI=EVPN, RT-5, Local-Pref=100, AS-Path=(100,200)}
{SAFI=128, Local-Pref=100, AS-Path=(100,200)}

Selected route: {SAFI=EVPN, RT-2, Local-Pref=100, AS-Path=100,200] (due to step 3, and no ECMP).

Example 2 - PE1 receives the following routes for IP2/24, that are candidate to be imported into IP-VRF-1:

{SAFI=EVPN, RT-5, D-PATH=(6500:3:IPVPN), AS-Path=(100,200), MED=10}
{SAFI=128, D-PATH=(6500:1:EVPN,6500:2:IPVPN), AS-Path=(200), MED=200}

Selected route: {SAFI=EVPN, RT-5, D-PATH=(6500:3:IPVPN), AS-Path=(100,200), MED=10} (due to step 1).

7. Composite PE Procedures

As described in Section 3, composite PEs are typically used in tenant networks where EVPN and IPVPN are both used to provide inter-subnet forwarding within the same composite domain.

Figure 7 depicts an example of a composite domain, where PE1/PE2/PE4 are composite PEs (they support EVPN and IPVPN ISF SAFIs on their peering to the Route Reflector), and PE3 is a regular IPVPN PE.

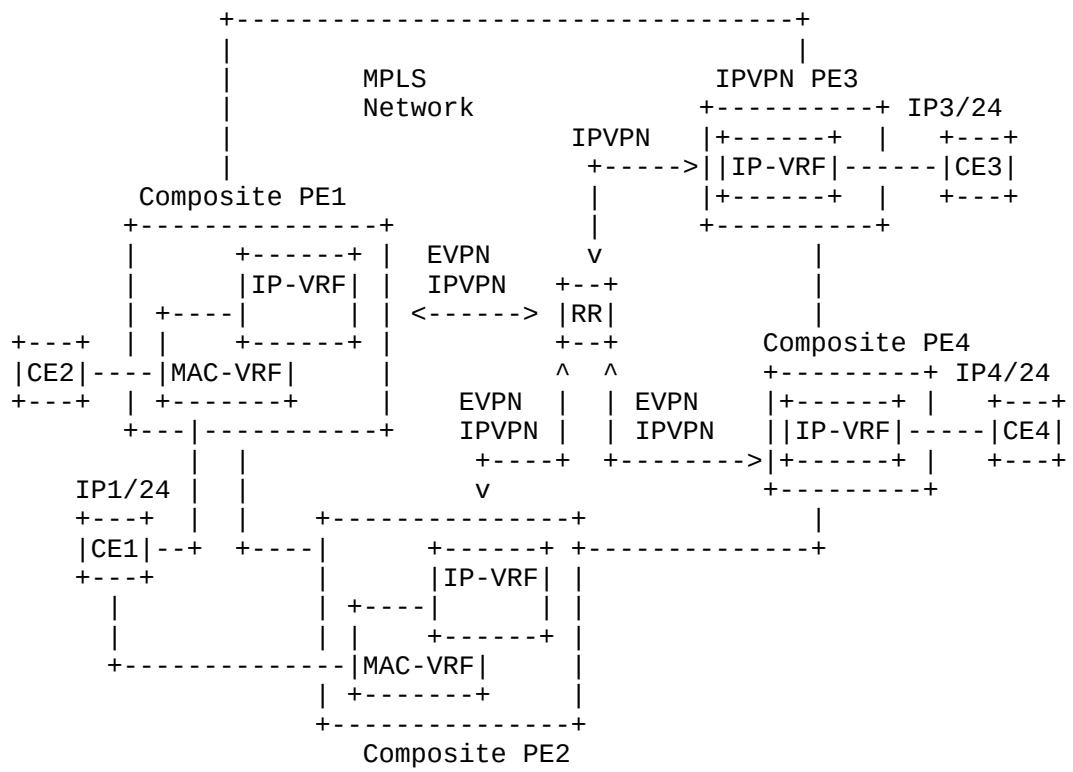

Figure 7: Composite PE example

In a composite domain with composite and regular PEs:

1. The composite PEs MUST advertise the same IP prefixes in each ISF SAFI to the Route Reflector (RR). For example, in Figure 7, the prefix IP1/24 is advertised by PE1 and PE2 to the Route Reflector in two separate NLRIs, one for AFI/SAFI 1/128 and another one for EVPN. If the two routes are advertised with the same set of attributes, the remote Composite PE selection process will pick up the EVPN route over the IPVPN route (as per Section 6). For this reason, prioritizing the announcement of the EVPN route prior to the IPVPN route is an OPTIONAL optimization, since, if the EVPN route arrives at the composite PE first, the selected route is not replaced by the IPVPN route later.

2. As an informative note, the Route Reflector does not forward EVPN routes to neighbors on which the EVPN SAFI is not enabled, and similarly, the Route Reflector does not forward IPVPN routes to neighbors on which the IPVPN SAFI is not enabled. For example, the Route Reflector does not forward EVPN routes to PE3 (since

the Route Reflector does not have the EVPN SAFI enabled on its BGP session to PE3), whereas the IPVPN routes are forwarded to all the PEs.

3. IPVPN PEs process and import IPVPN routes, as in [RFC4364]. As an example, PE3 receives only the IPVPN route for IP1/24 and resolves the BGP next-hop to an MPLS tunnel (with IP payload) to PE1 and/or PE2.

4. Composite PEs MUST process routes for the same prefix coming from different ISF SAFI routes, and perform route selection.

   * As an example, PE4 receives IP1/24 encoded in EVPN and another ISF SAFI route (EVPN IP Prefix route and IPVPN). The route selection follows the procedures in Section 6.

   * Assuming an EVPN route is selected, PE4 resolves the BGP next-hop to an MPLS tunnel (with Ethernet or IP payload) to PE1 and/or PE2. As described in Section 3, two EVPN PEs may use tunnels with Ethernet or IP payloads to connect their IP-VRFs, depending on the [RFC9136] model implemented.

   * The other composite PEs (PE1 and PE2) receive also the same IP prefix via EVPN and IPVPN SAFIs and they also follow the route selection in Section 6.

5. When a given route has been selected as the route for a particular packet, the transmission of the packet MUST be done according to the rules for that route’s AFI/SAFI.

6. As an informative note, in composite domains, such as the one in Figure 7, the EVPN advanced forwarding features will only be available to composite and EVPN PEs (assuming they select an EVPN IP Prefix route to forward packets for a given IP prefix), and not to IPVPN PEs. For example, assuming PE1 sends IP1/24 in an EVPN and an IPVPN route and the EVPN route is the best one in the selection, the recursive resolution of the EVPN IP Prefix routes can only be used in PE2 and PE4 (composite PEs), and not in PE3 (IPVPN PE). As a consequence of this, the indirection provided by the EVPN IP Prefix route recursive resolution and its benefits in a scaled network, will not be available in all the PEs in the network.

8. Gateway PE Procedures

Section 3 defines a gateway PE as an Interworking PE that is attached to two (or more) domains and propagates ISF routes between those domains. Examples of gateway PEs are Data Center gateways connecting domains that make use of EVPN and other ISF SAFIs for a given tenant. The gateway PE procedures in this document provide an interconnect solution for ISF routes and complement the gateway definition of [RFC9014], which focuses on the interconnect solution for Layer 2. This section applies to the interconnect of two domains that use different ISF SAFIs (e.g., EVPN to IPVPN), as well as the interconnect of two domains of the same ISF SAFI (e.g., EVPN to EVPN). Figure 8 illustrates a gateway PE use-case, in which PE1 and PE2 (and PE3/PE4) are gateway PEs interconnecting domains for the same tenant.

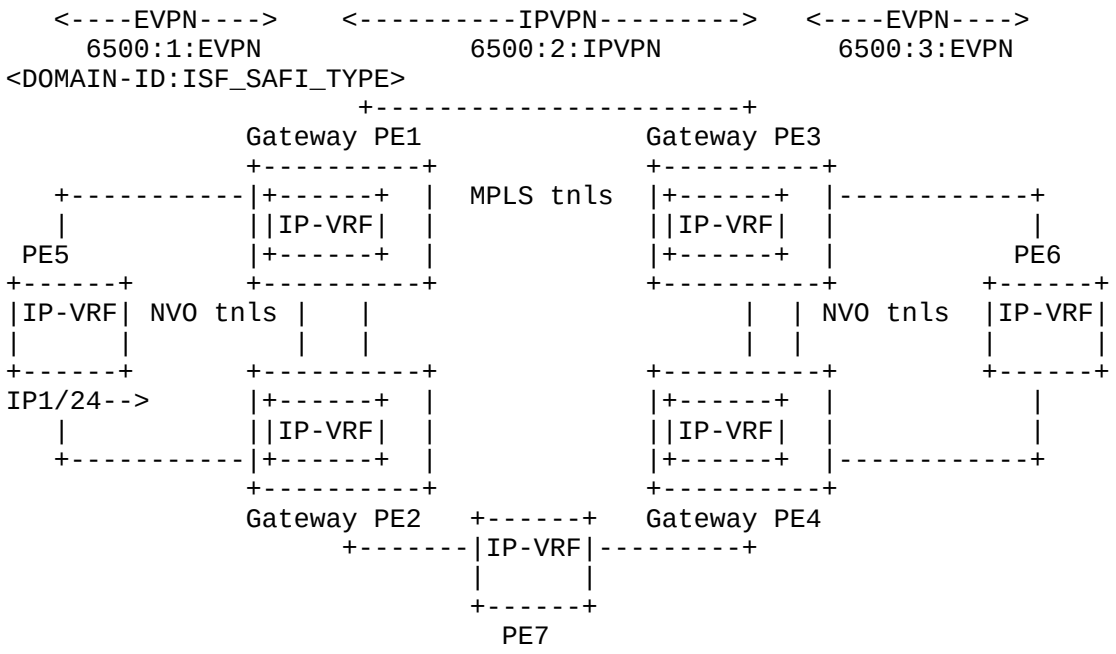

Figure 8: Gateway PE example

The procedures for a gateway PE enabled for ISF SAFI-x and ISF SAFI-y on the same IP-VRF follow:

1. A gateway PE that imports an ISF SAFI-x route to prefix P in an IP-VRF, MUST export P in ISF SAFI-y if:

a. P is installed in the IP-VRF - which means the SAFI-x route is well-formed, valid and the best one for P - and

b. PE has a BGP peer for SAFI-y (enabled for the same IP-VRF) and

c. The advertisement is allowed by policy and

d. ISF SAFI-x and ISF SAFI-y are any of the types defined in Section 3. Note that SAFI-x and SAFI-y MAY have the same value.

In the example of Figure 8, gateway PE1 and PE2 receive an EVPN IP Prefix route with IP1/24, install the prefix in the IP-VRF and re-advertise it using SAFI 128.

2. A gateway PE that receives an ISF SAFI-x route to prefix P in an IP-VRF MUST NOT export P in ISF SAFI-y if:

a. The ISF SAFI-x route is not well-formed or valid. Rules to determine if a route is well-formed or valid for a given ISF SAFI are defined by the specification of each ISF SAFI. As an example, an EVPN IP Prefix route received with non-zero ESI and GW IP values, at the same time, is not valid as per [RFC9136], section 3.2.

b. The ISF SAFI-x route contains a D-PATH attribute with one or more of the gateway PE’s locally associated domains for the IP-VRF. In this case the route is considered to be a looped ISF route, as described in Section 4 and hence MUST NOT be exported in ISF SAFI-y.

Once the gateway PE determines that P must be exported, P will be advertised using ISF SAFI-y as follows:

a. If Uniform-Propagation-Mode is enabled Section 5.2, the D-PATH attribute MUST be included if SAFI-y is equal to 128 or EVPN, so that loops can be detected in remote gateway PEs. When a gateway PE propagates an ISF route between domains, it MUST prepend a <DOMAIN-ID:ISF_SAFI_TYPE> to the received D-PATH attribute. The DOMAIN-ID and ISF_SAFI_TYPE fields refer to the domain over which the gateway PE received the IP prefix and the ISF SAFI of the route, respectively. If the received IP prefix route did not include any D-PATH attribute, the gateway IP MUST add the D-PATH when readvertising. The D-PATH in this case will have only one segment on the list, the <DOMAIN-ID:ISF_SAFI_TYPE> of the received route.

In the example of Figure 8, gateway PE1/PE2 receive the EVPN IP Prefix route with no D-PATH attribute since the route is originated at PE5. Therefore PE1 and PE2 will add the D-PATH attribute including <DOMAIN-ID:ISF_SAFI_TYPE> = <6500:1:EVPN>. Gateways PE3/PE4 will propagate the route again, now prepending their <DOMAIN-ID:ISF_SAFI_TYPE> = <6500:2:IPVPN>. PE6 receives the EVPN IP Prefix routes with D-PATH = {<6500:2:IPVPN>,<6500:1:EVPN>} and can use that information to make BGP path decisions.

b. The gateway PE MAY use the Route Distinguisher of the IP-VRF to readvertise P in the ISF SAFI-y.

c. The label allocation used by each gateway PE is a local implementation matter. The IP-VRF advertising IP prefixes for EVPN and another ISF SAFI may use a label per-VRF, per-prefix, etc.

d. The gateway PE MUST be able to use the same or different set of Route Targets per domain on the same IP-VRF. In particular, if different domains use different set of Route Targets for the same tenant, the gateway PE MUST be able to import and export routes with the different sets.

e. Even though Figure 8 only shows two domains per gateway PE, the gateway PEs may be connected to more than two domains.

f. There is no limitation of gateway PEs that a given IP prefix P can pass through until it reaches a given PE.

g. If the gateway PE uses Uniform-Propagation-Mode for BGP Path Attribute propagation, besides the processing of D-PATH described in point "a", the rules in Section 5.2 are followed.

h. As an informative note, if P was originated in an EVPN domain but traversed a different ISF SAFI domain (or domains), it will lose EVPN-specific attributes that are used in advanced EVPN procedures. For example, even if PE1 advertises IP1/24 along with a given non-zero ESI (for recursive resolution to that ESI), when PE6 receives the IP prefix in an EVPN route, the ESI value will be zero. This is because the route traverses an ISF SAFI domain that is different from EVPN.

9. Interworking Use-Cases

While Interworking PE networks may well be similar to the examples described in Section 7 and Section 8, in some cases a combination of both functions may be required. Figure 9 illustrates an example where the gateway PEs are also composite PEs, since not only they need to propagate ISF routes between domains (from EVPN SAFI to IPVPN and/or EVPN SAFIs), but they also need to interwork with IPVPN-only PEs in a domain with a mix of composite and IPVPN-only PEs.

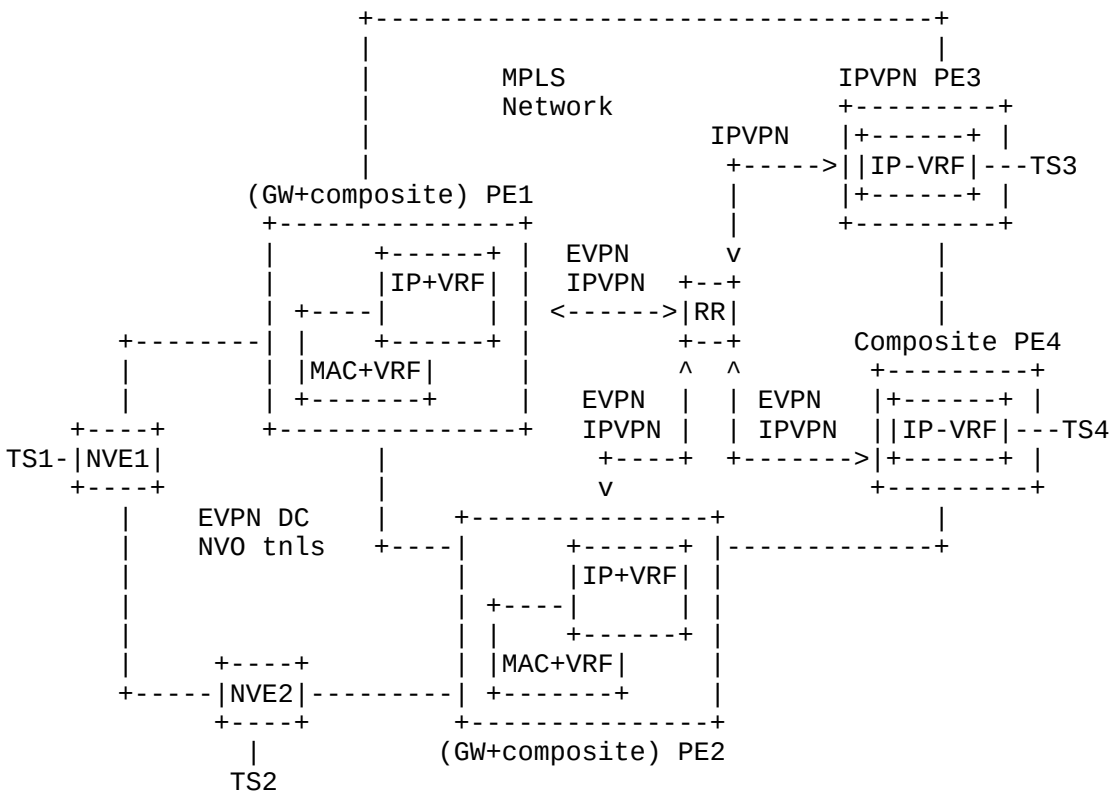

Figure 9: Gateway and composite combined functions - example

In the example above, PE1 and PE2 MUST follow the procedures described in Section 7 and Section 8. Compared to the example in Section 8, PE1 and PE2 now need to also propagate ISF routes from EVPN to EVPN, in addition to propagating prefixes from EVPN to IPVPN.

It is worth noting that PE1 and PE2 will receive TS4's IP prefix via IPVPN and EVPN IP Prefix routes. When readvertising to NVE1 and NVE2, PE1 and PE2 will consider the D-PATH rules and attributes of the selected route for TS4 (Section 6 describes the Route Selection Process).

10. BGP Error Handling on Interworking PEs

An Interworking PE (acting as gateway PE or composite PE) observes the following error-handling procedures for ISF routes:

* An UPDATE message for an ISF route containing a D-PATH attribute MUST follow the error-handling rules for D-PATH, as specified in Section 4.

* Any received UPDATE for an ISF route complies with the procedures in [RFC7606].

* The Interworking PEs do not introduce any new error-handling rules for UPDATES received with NLRIs and BGP Path Attributes defined in other specifications. Implementors should refer to the appropriate error handling documents for each of the supported route types. These include:

  BGP IP routes: [RFC4760], [RFC8950].

  IPVPN routes: [RFC4364], [RFC4659].

  EVPN MAC/IP routes (RT-2): [RFC7432], [RFC8365].

  EVPN IP Prefix routes (RT-5): [RFC9136].

  Received UPDATE messages to be programmed in IP-VRFs supporting Segment Routing for IPv6 data path (SRv6): [RFC9252].

If a gateway PE is set to propagate BGP Path Attributes for ISF routes across domains, the procedures in Section 5.2 guarantee that a BGP speaker does not receive UPDATES with well-formed but unexpected BGP Path Attributes. If a gateway PE fails to follow the propagation rules in Section 5.2 and propagates some BGP Path Attributes erroneously, the receiving PEs follow the specifications for the specific ISF route type and BGP Path Attribute. Some (but not all) examples follow:

* If the gateway PE erroneously propagates the Router's MAC Extended Community [RFC9135] from an EVPN domain to another EVPN domain, the receiving PE may find two EVPN Router's MAC extended communities in the same ISF route. In this case, the PE follows [RFC9135] and processes the first one (ignoring the second extended community).

* If the gateway PE erroneously propagates the BGP Encapsulation Extended Community (or equivalent Encapsulation TLV in the Tunnel Encapsulation Attribute) [RFC9012] from an EVPN domain to another EVPN domain, the receiving PE may find two BGP Encapsulation Extended Communities with different values in the same ISF route. The PE in this case follows [RFC8365], which allows multiple encapsulations being signaled in the route. As per [RFC9012], encapsulations advertised using the Tunnel Encapsulation attribute are considered equally with those advertised using the Encapsulation Extended Community.

* If the gateway PE erroneously propagates any EVPN extended community from an EVPN domain into an IPVPN domain, the receiving IPVPN PE ignores the EVPN extended communities, since their semantics do not apply to the IPVPN SAFI.

* If the gateway PE erroneously propagates a BGP Prefix-SID attribute with SRv6 Service TLVs [RFC9252] for an ISF route propagated between domains, the receiving PE follows [RFC9252] in case multiple SRv6 TLV instances are received.

11. Conclusion

This document describes the procedures required in PEs that process and advertise ISF routes for a given tenant. In particular, this document defines:

* A route selection algorithm so that a PE can determine what path to choose between EVPN paths and other ISF SAFI paths.

* A new BGP Path attribute called D-PATH that provides loop protection and visibility on the domains a particular route has traversed.

* The way BGP Path Attributes should be propagated between domains.

* The procedures that must be followed on Interworking PEs that behave as composite PEs, gateway PEs or a combination of both.

The above procedures provide an operator with the required tools to build large tenant networks that may span multiple domains, use different ISF SAFIs to handle IP prefixes, in a deterministic way and with routing loop protection.

12. Security Considerations

In general, the security considerations described in [RFC9136] and [RFC4364] apply to this document.

Section 4 introduces the use of the D-PATH attribute, which provides a loop prevention mechanism that is used by gateway PEs that propagate ISF IPVPN/EVPN routes between domains. A correct use of the D-PATH will prevent control plane and data plane loops in the network, however an incorrect configuration of the DOMAIN-IDs or an inconsistent support of D-PATH on the Gateway PEs may lead to the detection of false route loops, the blackholing of the traffic or may result in inconsistent and sub-optimal routing. An attacker may benefit of this transitive attribute to propagate the wrong domain information across multiple domains.

Section 4 restricts the use of D-PATH to IPVPN and EVPN routes in "walled garden" Virtual Private Networks. An upgraded PE removes D-PATH from the BGP Path Attributes before advertising an IP Prefix to a CE in a SAFI 1 route. However, if D-PATH is received by a non-upgraded IPVPN PE that has an attached CE connected to the Internet, the PE may incorrectly propagate the D-PATH attribute in a SAFI 1 route to the CE, and the D-PATH attribute may then escape out of the "walled garden" to the Internet. This may happen when the IPVPN PE re-exports a route directly, or via route leaking between IP-VRFs. Since D-PATH is a transitive attribute, if not upgraded to understand D-PATH, the CE may propagate the attribute to the Internet. However, since the attribute does not change the best path selection for SAFI 1 routes, D-PATH cannot create loops or inconsistent routing in the Internet. Upgraded Internet routers receiving the D-PATH attribute in a SAFI 1 route will apply the treat-as-withdraw behavior, as discussed in Section 4. As an additional security mechanism, a PE following this specification that receives an ISF EVPN or IPVPN route from a non-upgraded PE should discard the route via policy if the route contains the D-PATH attribute.

In addition, Section 5.2 introduces the propagation of BGP Path Attributes between domains on gateway PEs. Without this mode of propagation, BGP Path Attributes are re-initialized when re-exporting ISF routes into a different domain, and the operator does not have the end-to-end visibility of a given ISF route path. However, the Uniform Propagation mode introduces the capability of propagating BGP Path Attributes beyond the ISF SAFI scope. While this is a useful

tool to provide end-to-end visibility across multiple domains, it can also be used by an attacker to propagate wrong (although correctly formed) BGP Path Attributes that can influence the BGP path selection in remote domains. An implementation can also choose Section 5.1 (No-propagation mode) to minimize the risks derived from propagating incorrect attributes, but again, this mode of operation will prevent the receiver PE from seeing the attributes that the originator of the route intended to convey in the first place.

13. IANA Considerations

This document defines a new BGP path attribute known as the BGP Domain Path (D-PATH) attribute.

IANA has assigned a new attribute code type from the "BGP Path Attributes" subregistry under the "Border Gateway Protocol (BGP) Parameters" registry:

Path Attribute Value	Code	Reference
36	BGP Domain Path (D-PATH)	[This document]

14. Acknowledgments

The authors want to thank Russell Kelly, Dhananjaya Rao, Suresh Basavarajappa, Mallika Gautam, Senthil Sathappan, Arul Mohan Jovel, Naveen Tubugere, Mathanraj Petchimuthu, Eduard Vasilenko, Amit Kumar, Mohit Kumar, Lukas Krattiger, Gyan Mishra and Stephane Litkowski for their review and suggestions. Thanks to Sue Hares and Jeff Haas as well, for their detailed review to clarify the procedures of the D-PATH attribute.

15. Contributors

16. References

16.1. Normative References

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

[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>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

[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>.

[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>.

[RFC7606] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K. Patel, "Revised Error Handling for BGP UPDATE Messages", RFC 7606, DOI 10.17487/RFC7606, August 2015, <https://www.rfc-editor.org/info/rfc7606>.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <https://www.rfc-editor.org/info/rfc4760>.

[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>.

[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>.

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

16.2. Informative References

[RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", RFC 4360, DOI 10.17487/RFC4360, February 2006, <https://www.rfc-editor.org/info/rfc4360>.

[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>.

[RFC9774] Kumari, W., Sriram, K., Hannachi, L., and J. Haas, "Deprecation of AS_SET and AS_CONFED_SET in BGP", RFC 9774, DOI 10.17487/RFC9774, May 2025, <https://www.rfc-editor.org/info/rfc9774>.

[RFC4659] De Clercq, J., Ooms, D., Carugi, M., and F. Le Faucheur, "BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN", RFC 4659, DOI 10.17487/RFC4659, September 2006, <https://www.rfc-editor.org/info/rfc4659>.

[RFC8950] Litkowski, S., Agrawal, S., Ananthamurthy, K., and K. Patel, "Advertising IPv4 Network Layer Reachability Information (NLRI) with an IPv6 Next Hop", RFC 8950, DOI 10.17487/RFC8950, November 2020, <https://www.rfc-editor.org/info/rfc8950>.

[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>.

[I-D.ietf-idr-wide-bgp-communities] Raszuk, R., Haas, J., Lange, A., Decraene, B., Amante, S., and P. Jakma, "BGP Community Container Attribute", Work in Progress, Internet-Draft, draft-ietf-idr-wide-bgp-communities-12, 17 March 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-wide-bgp-communities-12>.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", RFC 4456, DOI 10.17487/RFC4456, April 2006, <https://www.rfc-editor.org/info/rfc4456>.

Authors' Addresses

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

A. Sajassi (editor)
Cisco
225 West Tasman Drive
San Jose, CA 95134
United States of America
Email: sajassi@cisco.com

E. Rosen
Individual
Email: erosen52@gmail.com

J. Drake
Independent
Email: je_drake@yahoo.com

W. Lin
Juniper
Email: wlin@juniper.net

J. Uttaro
Independent
Email: juttaro@ieee.org

A. Simpson
Nokia
Email: adam.1.simpson@nokia.com