Stream: Internet Engineering Task Force (IETF)
RFC: 9746
Updates: 7432, 8365
Category: Standards Track
Published: March 2025
ISSN: 2070-1721
Authors: J. Rabadan, Ed. (*Nokia*), K. Nagaraj (*Nokia*), W. Lin (*Juniper*), A. Sajassi (*Cisco*)

RFC 9746
BGP EVPN Multihoming Extensions for Split-Horizon Filtering

Abstract

An Ethernet Virtual Private Network (EVPN) is commonly used with Network Virtualization Overlay (NVO) tunnels as well as with MPLS and Segment Routing (SR) tunnels. The multihoming procedures in EVPN may vary based on the type of tunnel used within the EVPN Broadcast Domain. Specifically, there are two multihoming split-horizon procedures designed to prevent looped frames on multihomed Customer Edge (CE) devices: the Ethernet Segment Identifier (ESI) Label-based procedure and the local-bias procedure. The ESI Label-based split-horizon procedure is applied to MPLS-based tunnels such as MPLS over UDP (MPLSoUDP), while the local-bias procedure is used for other tunnels such as Virtual eXtensible Local Area Network (VXLAN) tunnels.

Current specifications do not allow operators to choose which split-horizon procedure to use for tunnel encapsulations that support both methods. Examples of tunnels that may support both procedures include MPLSoUDP, MPLS over GRE (MPLSoGRE), Generic Network Virtualization Encapsulation (Geneve), and Segment Routing over IPv6 (SRv6) tunnels. This document updates the EVPN multihoming procedures described in RFCs 7432 and 8365, enabling operators to select the split-horizon procedure that meets their specific requirements.

Status of This Memo

This is an Internet Standards Track document.

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

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc9746.

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

Ethernet Virtual Private Networks (EVPNs) are commonly used with the following tunnel encapsulations:

- Network Virtualization Overlay (NVO) tunnels, where the EVPN procedures are specified in [RFC8365]. MPLSoGRE [RFC4023], MPLSoUDP [RFC7510], Geneve [RFC8926], or VXLAN [RFC7348] tunnels are considered NVO tunnels.
- MPLS and Segment Routing over MPLS (SR-MPLS) tunnels, where the relevant EVPN procedures are specified in [RFC7432]. SR-MPLS tunneling is specified in [RFC8660].
- Segment Routing over IPv6 (SRv6) tunnels, where the relevant EVPN procedures are specified in [RFC9252]. SRv6 is specified in [RFC8402] and [RFC8754].

In this document, the term "split horizon" follows the definition in [RFC7432]. Split horizon refers to the EVPN multihoming procedure that prevents a Provider Edge (PE) from sending a frame back to a multihomed Customer Edge (CE) when that CE originated the frame in the first place.

EVPN multihoming procedures may vary depending on the type of tunnel utilized within the EVPN Broadcast Domain. Specifically, there are two multihoming split-horizon procedures employed to prevent looped frames on multihomed CE devices: the ESI Label-based procedure and the local-bias procedure.

The ESI Label-based split-horizon procedure is used for MPLS or MPLS over X (MPLSoX) tunnels, such as MPLSoUDP, and its procedures are detailed in [RFC7432]. Conversely, the local-bias procedure is used for IP-based tunnels, such as VXLAN tunnels, and it is described in [RFC8365].

1.1. Conventions and Terminology

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.

AC: Attachment Circuit

A-D per ES route: Auto-Discovery per Ethernet Segment route (as defined in [RFC7432]).

Arg.FE2: Refers to the ESI filtering argument used for split horizon as specified in [RFC9252].

BD: Broadcast Domain. Refers to an emulated Ethernet, such that two systems on the same BD will receive each other's BUM traffic. In this document, BD also refers to the instantiation of a BD on an EVPN PE. An EVPN PE can be attached to one or multiple BDs of the same tenant.

BUM: Broadcast, Unknown Unicast, and Multicast

CE: Customer Edge

DF: Designated Forwarder. As defined in [RFC7432], an ES may be multihomed (attached to more than one PE). An ES may also contain multiple BDs of one or more EVIs. For each such EVI, one of the PEs attached to the segment becomes that EVI's DF for that segment. Since a BD may belong to only one EVI, we can speak unambiguously of the BD's DF for a given segment.

ES: Ethernet Segment

ESI: Ethernet Segment Identifier

EVI: EVPN Instance

EVI-RT: EVI Route Target. Refers to a group of NVEs attached to the same EVI that will share the same EVI-RT.

Geneve: Generic Network Virtualization Encapsulation [RFC8926] (see tunnel type 19 in [TUNNEL-ENCAP]).

MPLS tunnels and non-MPLS NVO tunnels: Refers to Multiprotocol Label Switching (or the absence of it) Network Virtualization Overlay tunnels. NVO tunnels use an IP encapsulation for overlay frames, where the source IP address identifies the ingress NVE and the destination IP address identifies the egress NVE.

MPLSoUDP: Multiprotocol Label Switching over User Datagram Protocol [RFC7510] (see tunnel type 13 in [TUNNEL-ENCAP]).

MPLSoGRE: Multiprotocol Label Switching over Generic Network Encapsulation [RFC4023] (see tunnel type 11 in [TUNNEL-ENCAP]).

MPLSoX: Refers to MPLS over any IP encapsulation, for example, MPLSoUDP or MPLSoGRE.

NVE: Network Virtualization Edge

NVGRE: Network Virtualization Using Generic Routing Encapsulation [RFC7637] (see tunnel type 9 in [TUNNEL-ENCAP]).

PE: Provider Edge

RTs: Route Targets

VXLAN: Virtual eXtensible Local Area Network [RFC7348] (see tunnel type 8 in [TUNNEL-ENCAP]).

VXLAN-GPE: VXLAN Generic Protocol Extension [VXLAN-GPE] (see tunnel type 12 in [TUNNEL-ENCAP]).

SHT: Split-Horizon Type. Refers to the split-horizon method that a PE intends to use and advertises in an A-D per ES route.

SRv6: Segment Routing over IPv6 (see [RFC8402] and [RFC8754]).

TLV: Type-Length-Value

This document also assumes familiarity with the terminology of [RFC7432] and [RFC8365].

1.2. Split-Horizon Filtering and Tunnel Encapsulations

EVPN supports two split-horizon filtering mechanisms:

1. ESI Label-based split-horizon filtering [RFC7432]:

   When EVPN is employed for MPLS transport tunnels, an MPLS label facilitates split-horizon filtering to support All-Active multihoming. The ingress NVE device appends a label corresponding to the source ESI (the ESI label) during packet encapsulation. The egress NVE verifies the ESI label when attempting to forward a multi-destination frame through a local ES interface. If the ESI label matches the site identifier (the ESI) associated with that ES interface, then the packet is not forwarded. This mechanism effectively prevents forwarding loops for BUM traffic.

   ESI Label split-horizon filtering should also be utilized with Single-Active multihoming to prevent transient loops for in-flight packets when the egress NVE assumes the role of DF for an ES.

2. Local-bias filtering [RFC8365]:

   Since IP tunnels such as VXLAN or NVGRE do not support the ESI label or any MPLS label, an alternative split-horizon filtering procedure must be implemented for All-Active multihoming. This mechanism, known as local bias, relies on the source IP address of the tunnel to determine whether to forward BUM traffic to a local ES interface at the egress NVE.

   In summary and as specified in [RFC8365], each NVE tracks the IP address(es) of other NVEs with which it shares multihomed ESs. Upon receiving a BUM frame encapsulated in an IP tunnel, the egress NVE inspects the source IP address in the tunnel header, which identifies the ingress NVE. The egress NVE then filters out the frame on all local interfaces connected to ESs that are shared with the ingress NVE.

   Due to this behavior at the egress NVE, the ingress NVE is required to perform local replication to all directly attached ESs, regardless of the DF election state, for all BUM traffic ingressing from the access ACs. This local replication at the ingress NVE is the basis for the term "local bias".

   Local bias is not suitable for Single-Active multihoming, as the ingress NVE deactivates the ACs for which it is not the DF. Consequently, local replication to non-DF ACs cannot occur, leading to transient in-flight BUM packets being looped back to the originating site by newly elected DF egress NVEs.

[RFC8365] specifies that local bias is exclusively utilized for IP tunnels, while ESI Label-based split horizon is employed for IP-based MPLS tunnels. However, IP-based MPLS tunnels such as MPLSoGRE or MPLSoUDP are also categorized as IP tunnels and have the potential to support both procedures. These tunnels are capable of carrying ESI labels and also utilize a tunnel IP header in which the source IP address identifies the ingress NVE.

Similarly, certain IP tunnels (those that include an identifier for the source ES in the tunnel header) may also potentially support either procedure. Examples of such tunnels include Geneve and SRv6:

- In a Geneve tunnel, the source IP address identifies the ingress NVE; therefore, local bias is possible. Also, Section 4.1 of [EVPN-GENEVE] defines an Ethernet option Type-Length-Value (TLV) to encode an ESI label value.
- In an SRv6 tunnel, the source IP address identifies the ingress NVE. By default, and as outlined in [RFC9252], the ingress PE adds specific information to the SRv6 packet to enable the egress PE to identify the source ES of the BUM packet. This information is the ESI filtering argument (Arg.FE2) (see Section 6.1.1 of [RFC9252] and Section 4.12 of [RFC8986]) of the service Segment Identifier (SID) received on an A-D per ES route from the egress PE.

Table 1 presents various tunnel encapsulations along with their supported and default split-horizon methods. For Geneve, the default SHT is contingent upon the negotiation of the Ethernet Option with the Source ID TLV. In the case of SRv6, the default SHT is specified as ESI Label filtering in the table, as its behavior is analogous to that of ESI Label filtering. In this document, "ESI Label filtering" refers to the split-horizon filtering based on the presence of a source ES identifier in the tunnel header.

This document classifies the tunnel encapsulations used by EVPN into:

1. IP-based MPLS tunnels
2. MPLS and SR-MPLS tunnels
3. IP tunnels
4. SRv6 tunnels

Table 1 lists the encapsulations supported by this document. Any tunnel encapsulation not listed in Table 1 is out of scope. Tunnel encapsulations used by EVPN can be categorized into one of the four encapsulation groups mentioned above and support split-horizon filtering based on the following rules:

- IP-based MPLS tunnels and SRv6 tunnels are capable of supporting both split-horizon filtering methods.
- MPLS and SR-MPLS tunnels only support ESI Label-based split-horizon filtering.
- IP tunnels support local-bias split-horizon filtering and may also support ESI Label-based split-horizon filtering, provided they incorporate a mechanism to identify the source ESI in the header.

Tunnel Encapsulation	Default Split-Horizon Type (SHT)	Supports Local Bias	Supports ESI Label
MPLSoGRE (IP-based MPLS)	ESI Label filtering	Yes	Yes

Tunnel Encapsulation	Default Split-Horizon Type (SHT)	Supports Local Bias	Supports ESI Label
MPLSoUDP (IP-based MPLS)	ESI Label filtering	Yes	Yes
MPLS and SR-MPLS	ESI Label filtering	No	Yes
VXLAN (IP tunnels)	Local Bias	Yes	No
NVGRE (IP tunnels)	Local Bias	Yes	No
VXLAN-GPE (IP tunnels)	Local Bias	Yes	No
Geneve (IP tunnels)	Local Bias (if no ESI Lb), ESI Label (if ESI lb)	Yes	Yes
SRv6	ESI Label filtering	Yes	Yes

Table 1: Tunnel Encapsulations and Split-Horizon Types

The ESI Label method is applicable for both All-Active and Single-Active configurations, whereas the local-bias method is suitable only for All-Active configurations. Moreover, the ESI Label method is effective across different network domains, while local bias is constrained to networks where there is no change in the next hop between the NVEs attached to the same ES. Nonetheless, some operators favor the local-bias method due to its simplification of the encapsulation process, reduced resource consumption on NVEs, and the fact that the ingress NVE always forwards traffic locally to other interfaces, thereby decreasing the delay in reaching multihomed hosts.

This document extends the EVPN multihoming procedures to allow operators to select the preferred split-horizon method for a given NVO tunnel according to their specific requirements. The choice between local bias and ESI Label split horizon is now allowed (by configuration) for tunnel encapsulations that support both methods, and this selection is advertised along with the EVPN A-D per ES route. IP tunnels that do not support both methods, such as VXLAN or NVGRE, will continue to adhere to the procedures specified in [RFC8365]. Note that this document does not modify the local bias or the ESI Label split-horizon procedures themselves, just focuses on the signaling and selection of the split-horizon method to apply by the multihomed NVEs.

2. BGP EVPN Extensions

Extensions to EVPN are required to enable NVEs to advertise their preferred split-horizon method for a given ES. Figure 1 illustrates the ESI Label extended community (Section 7.5 of [RFC7432]), which is consistently advertised alongside the EVPN A-D per ES route. All NVEs connected to an ES advertise an A-D per ES route for that ES, including the extended community, which communicates information regarding the multihoming mode (either All-Active or Single-Active) and, if necessary, specifies the ESI Label to be utilized.

```
                     1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Type=0x06     | Sub-Type=0x01 | Flags(1 octet)|  Reserved=0   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Reserved=0   |          ESI Label                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: ESI Label Extended Community

[RFC7432] defines the low-order bit of the Flags octet (bit 0) as the "Single-Active" bit:

- A value of 0 means that the multihomed ES is operating in All-Active multihoming redundancy mode.
- A value of 1 means that the multihomed ES is operating in Single-Active multihoming redundancy mode.

Section 5 establishes a registry for the Flags octet, designating the "Single-Active" bit as the low-order bit of the newly defined Multihoming Redundancy Mode field.

2.1. The Split-Horizon Type

[RFC8365] does not include any explicit indication regarding the split-horizon method in the A-D per ES route. In this document, the split-horizon procedure defined in Section 8.3.1 of [RFC8365] is considered the default behavior, presuming that local bias is employed exclusively for IP tunnels, while ESI Label-based split horizon is used for IP-based MPLS tunnels. This document specifies that the two high-order bits in the Flags octet (bits 6 and 7) constitute the "Split-Horizon Type" or "SHT" field, where:

```
 7 6 5 4 3 2 1 0
+-+-+-+-+-+-+-+-+
|SHT|       |RED|
+-+-+-+-+-+-+-+-+
RED = "Multihoming Redundancy Mode" field (see Table 2)

SHT bit 7 6
-----------
        0 0  --> Default SHT
                 Backwards compatible with [RFC8365] and [RFC7432]
        0 1  --> Local Bias
        1 0  --> ESI Label-based filtering
        1 1  --> Unassigned
```

- SHT = 00 is backwards compatible with [RFC8365] and [RFC7432], and indicates that the advertising NVE intends to use the default or built-in SHT. The default SHT is shown in Table 1 for each encapsulation. An egress NVE that follows the [RFC8365] behavior and does not support this specification will ignore the SHT bits (which is equivalent to processing them as a value of 00).
- SHT = 01 indicates that the advertising NVE intends to use local-bias procedures in the ES for which the AD per-ES route is advertised.
- SHT = 10 indicates that the advertising NVE intends to use the ESI Label-based split-horizon method procedures in the ES for which the AD per-ES route is advertised.
- SHT = 11 is Unassigned.

2.2. Use of the Split-Horizon Type in A-D per ES Routes

The following behavior is observed:

- An SHT value of 01 or 10 **MUST NOT** be used with encapsulations that support only one SHT in Table 1, and **MAY** be used by encapsulations that support the two SHTs in Table 1.
- An SHT value different than 00 expresses the intent to use a specific split-horizon method, but does not reflect the actual operational SHT used by the advertising NVE, unless all the NVEs attached to the ES advertise the same SHT.
- In case of an inconsistency in the SHT value advertised by the NVEs attached to the same ES for a given EVI, all the NVEs **MUST** revert to the behavior in [RFC8365] and use the default SHT in Table 1, irrespective of the advertised SHT.
- An SHT different than 00 **MUST NOT** be set if the "Single-Active" bit is set. A received A-D per ES route where the "Single-Active" and SHT bits are different than zero **MUST** follow the treat-as-withdraw behavior in [RFC7606].
- The SHT **MUST** have the same value in each Ethernet A-D per ES route that an NVE advertises for a given ES and a given encapsulation (see Section 3 for NVEs supporting multiple encapsulations).

As an example, egress NVEs that support IP-based MPLS tunnels, such as MPLSoGRE or MPLSoUDP, will advertise A-D per ES routes for the ES along with the BGP Encapsulation Extended Community, as defined in [RFC9012]. This extended community indicates the encapsulation type (MPLSoGRE or MPLSoUDP) and may use the SHT value of 01 or 10 to signify the intent to use local bias or the ESI Label, respectively.

An egress NVE **MUST NOT** use an SHT value other than 00 when advertising an A-D per ES route with the following tunnel encapsulation types from [RFC9012]: VXLAN (type 8), NVGRE (type 9), MPLS (type 10), or no BGP Tunnel Encapsulation Extended Community at all. In all these cases, it is presumed that there is no choice for the split-horizon method; therefore, the SHT value **MUST** be set to 00. If a route with any of the mentioned encapsulation options is received and has an SHT value different than 00, it **SHOULD** apply the treat-as-withdraw behavior, per [RFC7606].

An egress NVE advertising A-D per ES route(s) for an ES with Geneve encapsulation (tunnel encapsulation type 19 in the BGP Tunnel Encapsulation attribute [RFC9012]) **MAY** use an SHT value of 01 or 10. A value of 01 indicates the intent to use local bias, regardless of the presence of an Ethernet option TLV with a non-zero Source-ID, as described in [EVPN-GENEVE]. A value of 10 indicates the intent to use ESI Label-based split horizon, and it is only valid if an Ethernet option TLV with a non-zero Source-ID is present. A value of 00 indicates the default behavior outlined in Table 1, which is to use local bias if:

a. no ESI Label is present in the Ethernet option TLV, or
b. there is no Ethernet option TLV.

Otherwise, the ESI Label split-horizon method is applied.

These procedures assume a single encapsulation supported in the egress NVE. Section 3 describes additional procedures for NVEs supporting multiple encapsulations.

2.3. The ESI Label Value in A-D per ES Routes

This document also updates [RFC8365] regarding the value that is advertised in the ESI Label field of the ESI Label extended community, as follows:

- The A-D per ES route(s) for an ES **MAY** have an ESI Label value of zero if the SHT value is 01. Section 2.2 specifies the scenarios where the SHT can be 01. An ESI Label value of zero eliminates the need to allocate labels in cases where they are not utilized, such as in the local-bias method.
- The A-D per ES route(s) for an ES **MAY** have an ESI Label value of zero for VXLAN or NVGRE encapsulations.

2.4. Backwards Compatibility with NVEs from RFC 8365

As discussed in Section 2.2, this specification is backwards compatible with the split-horizon filtering behavior in [RFC8365] and a non-upgraded NVE can be attached to the same ES as other NVEs supporting this specification.

An NVE maintains an administrative SHT value for an ES, which is advertised alongside the A-D per ES route, and an operational SHT value, which is the one actually used regardless of what the NVE has advertised. The administrative SHT matches the operational SHT if all the NVEs attached to the ES have the same administrative SHT.

This document assumes that an implementation of [RFC7432] or [RFC8365] that does not support the specifications in this document will ignore the values of all the Flags in the ESI Label extended community, except for the "Single-Active" bit. Based on this assumption, a non-upgraded NVE will disregard any SHT value other than 00. If an upgraded NVE receives at least one A-D per ES route for the ES with an SHT value of 00, it **MUST** revert its operational SHT to the default split-horizon method, as described in Table 1, irrespective of its administrative SHT.

For instance, consider an NVE attached to ES N that receives two A-D per ES routes for N from different NVEs, NVE1 and NVE2. If the route from NVE1 has an SHT value of 00 and the one from NVE2 has an SHT value of 01, the NVE **MUST** use the default split-horizon method specified in Table 1 as its operational SHT, regardless of its administrative SHT.

All NVEs attached to an ES with an operational SHT value of 10 **MUST** advertise a valid, non-zero ESI Label. If the operational SHT value is 01, the ESI Label **MAY** be zero. If the operational SHT value is 00, the ESI Label may be zero only if the default encapsulation supports local bias exclusively, and the NVEs do not require the presence of a valid, non-zero ESI Label.

If an NVE changes its operational SHT value from 01 (Local Bias) to 00 (Default SHT) due to the presence of a new non-upgraded NVE in the ES, and it previously advertised a zero ESI Label, it **MUST** send an update with a valid, non-zero ESI Label, unless all the non-upgraded NVEs in the ES support only local bias. For example, consider NVE1 and NVE2 using MPLSoUDP as encapsulation, attached to the same Ethernet Segment ES1, and advertising an SHT value of 01 (Local Bias) with a zero ESI Label value. Suppose NVE3, which does not support this specification, joins ES1 and advertises an SHT value of 00 (default). Upon receiving NVE3's A-D per ES route, NVE1 and NVE2 **MUST** update their A-D per ES routes for ES1 to include a valid, non-zero ESI Label value. The assumption here is that NVE3 only supports the default ESI Label-based split-horizon filtering.

3. Procedures for NVEs Supporting Multiple Encapsulations

As specified in [RFC8365], an NVE that supports multiple data plane encapsulations (e.g., VXLAN, NVGRE, MPLS, MPLSoUDP, Geneve) must indicate all supported encapsulations using BGP Encapsulation extended communities as defined in [RFC9012] for all EVPN routes. This section provides clarification on the multihoming split-horizon behavior for NVEs that advertise and receive multiple BGP Encapsulation extended communities along with the A-D per ES routes. This section uses the notation {x, y} (more than two encapsulations is possible too) to denote the encapsulations advertised in BGP Encapsulation extended communities (or the BGP Tunnel Encapsulation Attribute), where x and y represent different encapsulation values. When Geneve is one of the encapsulations, the tunnel type is indicated in either a BGP Encapsulation extended community or a BGP Tunnel Encapsulation Attribute.

It is important to note that an NVE **MAY** advertise multiple A-D per ES routes for the same ES, rather than a single route, with each route conveying a set of Route Targets (RTs). The total set of RTs associated with a given ES is referred to as the RT-set for that ES. Each of the EVIs represented in the RT-set will have its RT included in one, and only one, A-D per ES route for the ES. When multiple A-D per ES routes are advertised for the same ES, each route must have a distinct Route Distinguisher.

As per [RFC8365], an NVE that advertises multiple encapsulations in the A-D per ES route(s) for an ES **MUST** advertise encapsulations that use the same split-horizon filtering method in the same route. For example:

- An A-D per ES route for ES-x may be advertised with {VXLAN, NVGRE} encapsulations.
- An A-D per ES route for ES-y may be advertised with {MPLS, MPLSoUDP, MPLSoGRE} encapsulations (or a subset).
- However, an A-D per ES route for ES-z **MUST NOT** be advertised with {MPLS, VXLAN} encapsulations.

This document extends the described behavior as follows:

a. An A-D per ES route for ES-x may be advertised with multiple encapsulations, some of which support a single split-horizon method. In this case, the SHT value **MUST** be 00. For instance, encapsulations such as {VXLAN, NVGRE}, {VXLAN, Geneve}, or {MPLS, MPLSoGRE, MPLSoUDP} can be advertised in an A-D per ES route. In all these cases, the SHT value **MUST** be 00 and the treat-as-withdraw behavior [RFC7606] is applied in case of any other value.

b. An A-D per ES route for ES-y may be advertised with multiple encapsulations that all support both split-horizon methods. In this case, the SHT value **MAY** be 01 if the preferred method is local bias, or 10 if the ESI Label-based method is desired. For example, encapsulations such as {MPLSoGRE, MPLSoUDP, Geneve} (or a subset) **MAY** be advertised in an A-D per ES route with an SHT value of 01. The ESI Label value in this case **MAY** be zero.

c. If ES-z with an RT-set composed of (RT1, RT2, RT3.. RTn) supports multiple encapsulations requiring different split-horizon methods, a distinct A-D per ES route (or group of routes) per split-horizon method **MUST** be advertised. For example, consider an ES-z with n RTs, where:

    - the EVIs corresponding to (RT1..RTi) support VXLAN,
    - the ones for (RTi+1..RTm) (with i<m) support MPLSoUDP with local bias, and
    - the ones for (RTm+1..RTn) (with m<n) support Geneve with ESI Label-based split horizon.

    In this scenario, three groups of A-D per ES routes **MUST** be advertised for ES-z:

    - A-D per ES route group 1, including (RT1..RTi) with encapsulation {VXLAN} and an SHT value of 00. The ESI Label **MAY** be zero.
    - A-D per ES route group 2, including (RTi+1..RTm) with encapsulation {MPLSoUDP} and an SHT value of 01. The ESI Label **MAY** be zero.
    - A-D per ES route group 3, including (RTm+1..RTn) with encapsulation {Geneve} and an SHT value of 10. The ESI Label **MUST** have a valid, non-zero value, and the Ethernet option as defined in [RFC8926] **MUST** be advertised.

As per [RFC8365], it is the responsibility of the operator of a given EVI to ensure that all of the NVEs within that EVI support a common encapsulation. Failure to meet this condition may result in service disruption or failure.

4. Security Considerations

All the security considerations described in [RFC7432] are applicable to this document.

Additionally, this document modifies the procedures for split-horizon filtering as outlined in [RFC8365], offering operators a choice between local bias and ESI Label-based filtering for tunnels that support both methods. Misconfiguration of the desired SHT could lead to forwarding behaviors that differ from the intended configuration. Apart from this risk, this document describes procedures to ensure that all PE devices or NVEs connected to the same ES agree on a common SHT method, with a fallback to a default behavior in case of a mismatch in the SHT bits being advertised by any two PEs or NVEs in the ES. Consequently, unauthorized changes to the SHT configuration by an attacker on a single PE or NVE of the ES should not cause traffic disruption (as long as the SHT value is valid as per this document) but may result in alterations to forwarding behavior.

5. IANA Considerations

Per this document, IANA has created the "EVPN ESI Label Extended Community Flags" registry for the 1-octet Flags field in the ESI Label Extended Community [RFC7432], as follows:

Bit Position	Name	Reference
0-1	Multihoming Redundancy Mode	[RFC7432]
2-5	Unassigned	
6-7	Split-Horizon Type	RFC 9746

Table 2

IANA has also created the "Multihoming Redundancy Mode" registry for the related field of the "EVPN ESI Label Extended Community Flags". The registry has been populated with the following initial values:

Value	Multihoming Redundancy Mode	Reference
00	All-Active	[RFC7432]
01	Single-Active	[RFC7432]
10	Unassigned	

Value	Multihoming Redundancy Mode	Reference
11	Unassigned	

Table 3

Finally, IANA has created the "Split-Horizon Type" registry for the related field of the "EVPN ESI Label Extended Community Flags". The registry has been populated with the following initial values:

Value	Split-Horizon Type Value	Reference
00	Default SHT	RFC 9746
01	Local Bias	RFC 9746
10	ESI Label-based filtering	RFC 9746
11	Unassigned	

Table 4

New registrations in the "EVPN ESI Label Extended Community Flags", "Multihoming Redundancy Mode", and "Split-Horizon Type" registries will be made through the "IETF Review" procedure defined in [RFC8126]. These registries are located in the "Border Gateway Protocol (BGP) Extended Communities" registry group.

6. References

6.1. Normative References

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

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

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[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>.

[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>.

[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>.

6.2. Informative References

[EVPN-GENEVE] Boutros, S., Sajassi, A., Drake, J., Rabadan, J., and S. Aldrin, "EVPN control plane for Geneve", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-geneve-08, 5 July 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-geneve-08>.

[RFC4023] Worster, T., Rekhter, Y., and E. Rosen, Ed., "Encapsulating MPLS in IP or Generic Routing Encapsulation (GRE)", RFC 4023, DOI 10.17487/RFC4023, March 2005, <https://www.rfc-editor.org/info/rfc4023>.

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

[RFC7510] Xu, X., Sheth, N., Yong, L., Callon, R., and D. Black, "Encapsulating MPLS in UDP", RFC 7510, DOI 10.17487/RFC7510, April 2015, <https://www.rfc-editor.org/info/rfc7510>.

[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>.

[RFC7637] Garg, P., Ed. and Y. Wang, Ed., "NVGRE: Network Virtualization Using Generic Routing Encapsulation", RFC 7637, DOI 10.17487/RFC7637, September 2015, <https://www.rfc-editor.org/info/rfc7637>.

[RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/info/rfc8402>.

[RFC8660] Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with the MPLS Data Plane", RFC 8660, DOI 10.17487/RFC8660, December 2019, <https://www.rfc-editor.org/info/rfc8660>.

[RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <https://www.rfc-editor.org/info/rfc8754>.

[RFC8926] Gross, J., Ed., Ganga, I., Ed., and T. Sridhar, Ed., "Geneve: Generic Network Virtualization Encapsulation", RFC 8926, DOI 10.17487/RFC8926, November 2020, <https://www.rfc-editor.org/info/rfc8926>.

[RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 10.17487/RFC8986, February 2021, <https://www.rfc-editor.org/info/rfc8986>.

[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>.

[TUNNEL-ENCAP] IANA, "Border Gateway Protocol (BGP) Tunnel Encapsulation", <https://www.iana.org/assignments/bgp-tunnel-encapsulation>.

[VXLAN-GPE] Maino, F., Kreeger, L., and U. Elzur, "Generic Protocol Extension for VXLAN (VXLAN-GPE)", Work in Progress, Internet-Draft, draft-ietf-nvo3-vxlan-gpe-13, 4 November 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-nvo3-vxlan-gpe-13>.

Acknowledgments

The authors would like to thank Anoop Ghanwani, Gyan Mishra, and Jeffrey Zhang for their review and useful comments. Thanks to Gunter Van de Velde and Sue Hares as well, for their thorough review.

Authors' Addresses

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

Kiran Nagaraj
Nokia
520 Almanor Avenue
Sunnyvale, CA 94085
United States of America
Email: kiran.nagaraj@nokia.com

Wen Lin
Juniper Networks
Email: wlin@juniper.net

Ali Sajassi
Cisco Systems, Inc.
Email: sajassi@cisco.com