BESS
Internet-Draft
Intended status: Standards Track
Expires: 6 May 2026

Y. Liu
ZTE

2 November 2025

Data Plane Failure Detection Mechanisms for EVPN over SRv6
draft-liu-bess-srv6-evpn-validation-02

Abstract

This document proposes extension for ICMPv6 to detect data plane failures for EVPN over SRv6.

Status of This Memo

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

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

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

This Internet-Draft will expire on 6 May 2026.

Table of Contents

1. Introduction

[RFC7432] describes MPLS-based EVPN technology. An EVPN comprises one or more Customer Edge devices (CEs) connected to one or more Provider Edge devices (PEs). The PEs provide Layer 2 (L2) EVPN among the CE(s) over the MPLS core infrastructure. In EVPN networks, the PEs advertise the Media Access Control (MAC) addresses learned from the locally connected CE(s), along with the MPLS label, to remote PE(s) in the control plane using multiprotocol BGP [RFC4760]. EVPN enables multihoming of CE(s) connected to multiple PEs and load balancing of traffic to and from multihomed CE(s).

[RFC9252] defines procedures and messages for SRv6-based BGP services, including Layer 3 VPN and EVPN over SRv6. To support SRv6-based EVPN overlays, one or more SRv6 Service SIDs are advertised with Route Types 1, 2, 3, and 5. The SRv6 Service SID(s) per Route Type is advertised in SRv6 L2/L3 Service TLVs within the BGP Prefix-SID attribute which is attached to MP-BGP NLRIs. The existing NLRIs of MPLS-based EVPN are reused instead of defining new ones, the NLRI encodings over SRv6 core are similar with those for MPLS, the only difference is that the MPLS labels in the NLRIs carry part of the SRv6 Service SIDs or they are set to Implicit NULL as described in [RFC9252] Section 4.

For MPLS EVPN, [RFC9489] defines procedures to detect data plane failures using LSP Ping in MPLS networks deploying EVPN. For EVPN over SRv6, the requirements to detect data plane failures are similar. This document proposes extension for ICMPv6 to fulfill such requirements for EVPN over SRv6.

2. Specification of Requirements

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

A-D:Auto-Discovery

BUM: Broadcast, Unknown Unicast, and Multicast

CE: Customer Edge device

C-MAC: Customer MAC

DF: Designated Forwarder

ES: Ethernet Segment

ESI: Ethernet Segment Identifier

EVI: EVPN Instance Identifier that globally identifies the EVPN Instance

EVPN: Ethernet Virtual Private Network

MAC-VRF: A Virtual Routing and Forwarding table for MAC addresses on a PE

PE: Provider Edge device

VPWS: Virtual Private Wire Service

4. ICMPv6 Messages

[draft-liu-6man-icmp-verification] introduces the mechanism to verify the data plane message in IPv6/SRv6 networks by extending ICMPv6 messages. Two new types of ICMPv6 validation messages, ICMPv6 Validation Request and ICMPv6 Validation Reply are defined. Like any other ICMPv6 message, the messages are encapsulated in an IPv6 header.

For ease of reading, the format of ICMPv6 Validation Request is shown in Figure 1. As per [RFC4884], the Extension Structure contains one Extension Header followed by one or more ICMP Extension Objects.

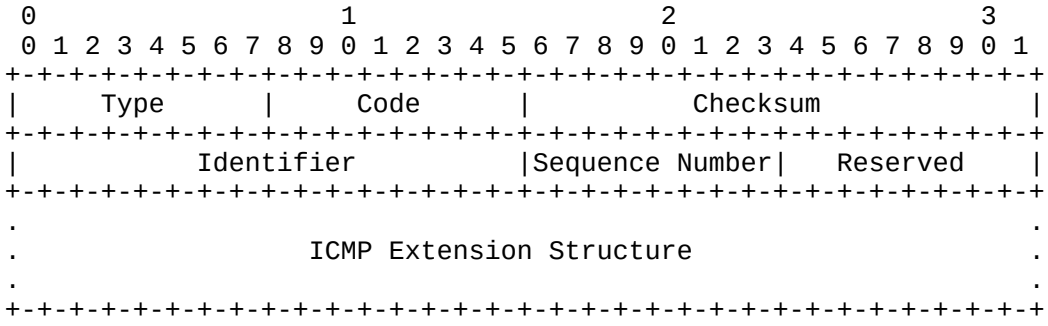

Figure 1: Validation Request

When applied in the ICMPv6 Validation Request message, a new type of ICMP Extension Object, Validation Information Object, is defined in [draft-liu-6man-icmp-verification] to carry the information related with the SRv6 SID to be verified. The format of ICMP Extension Object is shown in figure 2.

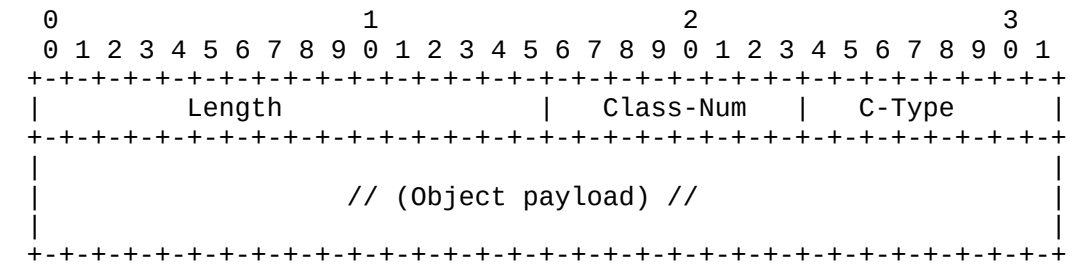

Figure 2: Validation Information Object

In this object, the C-Type is used to indicate the type of the information that needs to be verified which is carried in the object payload. The new values of of C-Type and the corresponding object payload defined in this document for EVPN are given below:

C-Type	Object Payload
10	EVPN MAC/IP
11	EVPN Inclusive Multicast
12	EVPN Ethernet Auto-Discovery
13	EVPN IP Prefix

The detailed formats and usages of these objects are described in section 5.

The format of ICMPv6 Validation Reply defined in [draft-liu-6man-icmp-verification] is shown in Figure 3, the value of the Code field indicates the validation result.

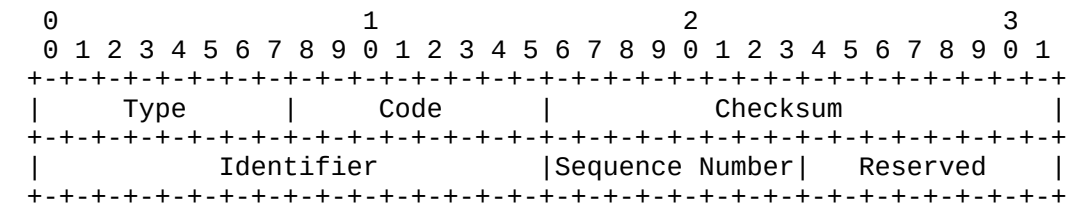

Figure 3: ICMPv6 Validation Reply

The ICMPv6 Validation Request packets are used for connectivity checks in the data plane in EVPN networks. The Validation Information Objects can be used to validate that an identifier for a given EVPN is programmed at the target node.

The ICMPv6 Validation Request for EVPN can be sent with the SRv6 Service SID as the IPv6 destination address without SRH encapsulated, or when the ICMPv6 Validation Request is sent with SRH the SRv6 Service SID is set the last segment of the SRH.

Once the ICMPv6 Validation Request reaches the target egress PE, the egress PE, as the final destination of the IPv6 packet, will proceed to process the next header in the packet, i.e, the ICMPv6 Validation Request. Then the PE will perform checks for the information present in the Validation Information Object, that is, the PE will check whether the information carried in the Validation Information Object is programmed locally, and whether it is valid. If the above two conditions are both met, the egress PE will generate an ICMPv6 Validation Reply with Code 0 ("Validation passed"). Otherwise, the return code is 3 ("Information mismatch"), which indicates the EVPN information carried in the ICMPv6 Validation Request is not reachable from the egress PE.

5. Validation Information Objects

This document introduces several new Validation Information Objects that can be carried in the ICMPv6 Validation Request.

5.1. EVPN MAC/IP Object

The EVPN MAC/IP Object identifies the target MAC, MAC/IP binding for ARP/ND, or IP address for an EVI under test at an egress PE. This Object is included in the ICMPv6 Validation Request sent by an EVPN PE to a peer PE. the format of the EVPN MAC/IP Object is shown in figure 4.

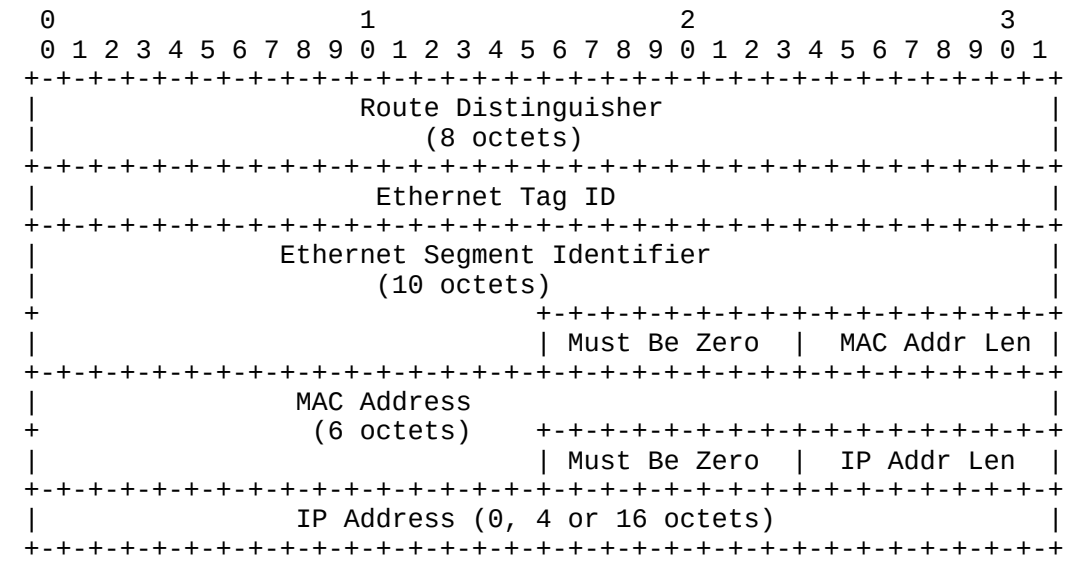

Figure 4: EVPN MAC/IP Object

The fields of the EVPN MAC/IP Object are derived from the MAC/IP Advertisement route defined in Section 7.2 of [RFC7432]. The fields of the EVPN MAC/IP Object should be set according to the following, which is consistent with [RFC7432] and [RFC7623]:

* The Ethernet Tag ID field can be 0 or a valid VLAN ID for EVPN VLAN-aware bundle service [RFC7623].

* The Ethernet Segment Identifier field is a 10-octet field. For EVPN, it is set to 0 for a single-homed ES or to a valid ESI ID for a multihomed ES.

* The MAC Addr Len field specifies the MAC length in bits. Only 48-bit MAC addresses are supported as this document follows the MAC address length supported by [RFC7623].

* The MAC Address field is set to the 6-octet MAC address.

* The IP Address field is optional. When the IP Address field is not present, the IP Addr Len field is set to 0. When the IP Address field is present, the IP Addr Len field is in bits and is set to either 32 for IPv4 addresses or 128 for IPv6 addresses.

* The Must Be Zero fields are set to 0. The receiving PE should ignore the Must Be Zero fields.

As described in [RFC9489] section 4.1. In EVPN, the MAC/IP Advertisement route has multiple uses and is used for the following cases:

* This route with only a MAC address and MPLS Label1 is used for populating MAC-VRF and performing MAC forwarding.

* This route with MAC and IP addresses and only MPLS Label1 is used for populating both MAC-VRF and ARP/ND tables (for ARP suppression) as well as for performing MAC forwarding.

* This route with MAC and IP addresses and both MPLS Label1 and Label2 is used for populating MAC-VRF and IP-VRF tables as well as for both MAC and IP forwarding in the case of symmetric Integrated Routing and Bridging (IRB).

The above descriptions are still applicable for SRv6 EVPN, the only difference is that MPLS Label1 and Label2 are replaced by SRv6 L2 Service SID enclosed in an SRv6 L2 Service TLV and SRv6 L3 Service SID enclosed in an SRv6 L3 Service TLV separately.

When an ICMPv6 Echo Request is sent by an ingress PE, the contents of the ICMPv6 Validation Request and the egress PE mode of operation (i.e., IRB mode or L2 mode) along with SRv6 Service SID of the packet determine which of the three cases above this Echo Request is for. When the egress PE receives the EVPN MAC/IP Object containing only the MAC address, the egress PE validates the MAC state and forwarding. When the egress PE receives the EVPN MAC/IP Object containing both MAC and IP addresses and if the SRv6 Service SID points to a MAC-VRF, then the egress PE validates the MAC state and forwarding. If the egress PE is not configured in symmetric IRB mode, it also validates ARP/ND state. However, if the SRv6 Service SID points to an IP-VRF, then the egress PE validates IP state and forwarding. Any other combinations (e.g., the egress PE receiving the EVPN MAC/IP Object containing only the MAC address but with the SRv6 Service SID pointing to an IP-VRF) should be considered invalid, and the egress PE should send an ICMPv6 Validation Reply with the appropriate Code to the ingress PE.

5.2. EVPN Inclusive Multicast Object

Inclusive Multicast Ethernet Tag Route over SRv6 Core are described in [RFC9252] section 6.3. [draft-ietf-bess-mvpn-evpn-sr-p2mp] further describes how to realized P-Tunnels by SRv6 P2MP trees.

The multicast connectivity state validation for EVPN over SRv6 will described in further version of the draft.

5.3. EVPN Ethernet Auto-Discovery (A-D) Object

The fields in the EVPN Ethernet A-D Object are based on the EVPN Ethernet A-D route advertisement defined in Section 7.1 of [RFC7432]. RFC9252 section 6.1 describes EVPN Ethernet A-D route over SRv6 Core. Ethernet A-D routes are Route Type 1, as defined in [RFC7432], and may be used to achieve split-horizon filtering, fast convergence, and aliasing. EVPN Route Type 1 is also used in EVPN-VPWS as well as in EVPN-flexible cross-connect, mainly to advertise point-to-point service IDs.

The EVPN Ethernet A-D Object only applies to EVPN.

The EVPN Ethernet A-D Object has the format shown in Figure 5. The fields of this Object should be set according to the following, which is consistent with [RFC7432] and [RFC9252]:

* The Route Distinguisher (RD) field is a 10-octet field and is set to the RD of the MAC-VRF on the peer PE. Please see Section 5.3.2 for the case when a per-ES A-D route is announced with different RDs.

* The Ethernet Tag ID field can be 0, MAX-ET, or a valid VLAN ID as described in Section 5.3.1.

* The Ethernet Segment Identifier field is a 10-octet field and is set to 0 for a single-homed ES or to a valid ESI ID for a multihomed ES.

* The Must Be Zero field is set to 0. The receiving PE should ignore the Must Be Zero field.

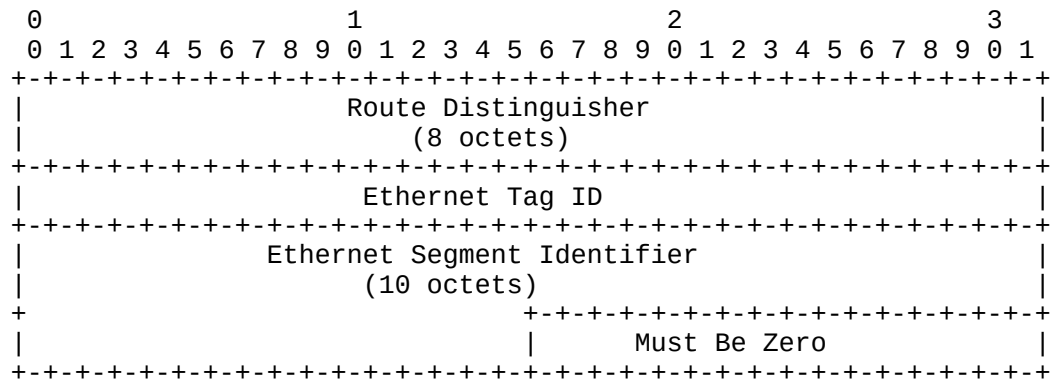

Figure 5: EVPN Ethernet A-D Object

5.3.1. Ethernet Tag Value

The EVPN Ethernet A-D Object can be sent in the context of per-ES or per-EVI. When an operator performs a connectivity check for the BUM L2 service, an ICMPv6 Validation Request is sent with the EVPN Ethernet A-D Object to emulate traffic coming from a multihomed site. In this case, the EVPN Ethernet A-D Object is added in the per-ES context. When an ICMPv6 Validation Request is sent for the connectivity check for EVPN Aliasing state, the context for the EVPN Ethernet A-D Object is per-EVI.

The Ethernet Tag field value in the EVPN Ethernet A-D Object MUST be set according to the context:

* For the per-ES context, the Ethernet Tag field in the Object MUST be set to the reserved MAX-ET value [RFC7432].

* For the per-EVI context, the Ethernet Tag field in the Object MUST be set to the non-reserved value.

5.3.2. Per-ES EVPN Auto-Discovery Route with Different RDs

Section 8.2 of [RFC7432] specifies that a per-ES EVPN A-D route for a given multihomed ES may be advertised more than once with different RD values because many EVIs may be associated with the same ES and Route Targets for all these EVIs may not fit in a single BGP Update message. In this case, the RD value used in the EVPN Ethernet A-D Object MUST be the RD value received for the EVI in the per-ES EVPN A-D route.

5.3.3. EVPN VPWS

This mechanism can also be used to detect data plane failures for the EVPN VPWS ([RFC8214]) over SRv6 described in [RFC9252] section 6.1.2. The ICMPv6 Validation Request carries the EVPN Ethernet A-D Object with fields populated from the EVPN Ethernet A-D per-EVI route announced by the egress PE for the EVPN VPWS under test. The ICMPv6 Validation Request is sent by the ingress PE using the SRv6 service SID associated with the EVPN Ethernet A-D route announced by the egress PE and the transport encapsulations (e.g., SRv6, IP) to reach the egress PE.

The egress PE processes the ICMPv6 Validation Request packet and performs checks for the EVPN Ethernet A-D Object. The egress PE can identify that the ICMPv6 Validation Request is for the EVPN VPWS instance as EVI (identified by the RD) for EVPN VPWS is different from EVI assigned for EVPN. The egress PE will use the information from the EVPN Ethernet A-D Object and validate the VLAN state for the EVPN VPWS under test. For the success case, the egress PE will reply with Code 0 ("Validation passed").

5.4. EVPN IP Prefix Object

The EVPN IP Prefix Object identifies the IP prefix for an EVI under test at a peer PE.

EVPN Route Type 5 is used to advertise IP address reachability through MP-BGP to all other PEs in a given EVPN instance as defined in [RFC9136]. RFC9252 section 6.5 describes IP Prefix Route over SRv6 Core.

The EVPN IP Prefix Object fields are derived from the IP Prefix route (RT-5) advertisement defined in [RFC9136] and [RFC9252]. This Object only applies to EVPN.

The EVPN IP Prefix Object has the format shown in Figure 6. The total length (not shown) of this Object MUST be either 32 bytes (if IPv4 addresses are carried) or 56 bytes (if IPv6 addresses are carried). The IP prefix and gateway IP address MUST be from the same IP address family, as described in Section 3.1 of [RFC9136].

The fields of the EVPN IP Prefix Object should be set according to the following, which is consistent with [RFC9136] and [RFC9252]:

* The Route Distinguisher (RD) field is a 10-octet field and is set to the RD of the IP-VRF on the peer PE.

* The Ethernet Tag ID field can be 0 or a valid VLAN ID for EVPN VLAN-aware bundle service [RFC7432].

* The Ethernet Segment Identifier field is a 10-octet field and is set to a valid ESI ID if the ESI is used as an Overlay Index as per Section 3.1 of [RFC9136]. Otherwise, the Ethernet Segment Identifier field is set to 0.

* The IP Prefix Len field specifies the number of bits in the IP Prefix field. It is set to a value between 0 and 32 for IPv4 or between 0 to 128 for IPv6.

* The IP Prefix field is set to a 4-octet IPv4 address (with trailing 0 bits to make 32 bits in all) or a 16-octet IPv6 address (with trailing 0 bits to make 128 bits in all). The address family of this field is inferred from the sub-TLV length field, as discussed above.

* The Gateway (GW) IP Address field is set to a 4-octet IPv4 address or a 16-octet IPv6 address if it’s used as an Overlay Index for the IP prefixes. If the GW IP Address is not being used, it must be set to 0 as described in Section 3.1 of [RFC9136]. The address family of this field is inferred from the sub-TLV length field, as discussed above.

* The Must Be Zero field is set to 0. The receiving PE should ignore the Must Be Zero field.

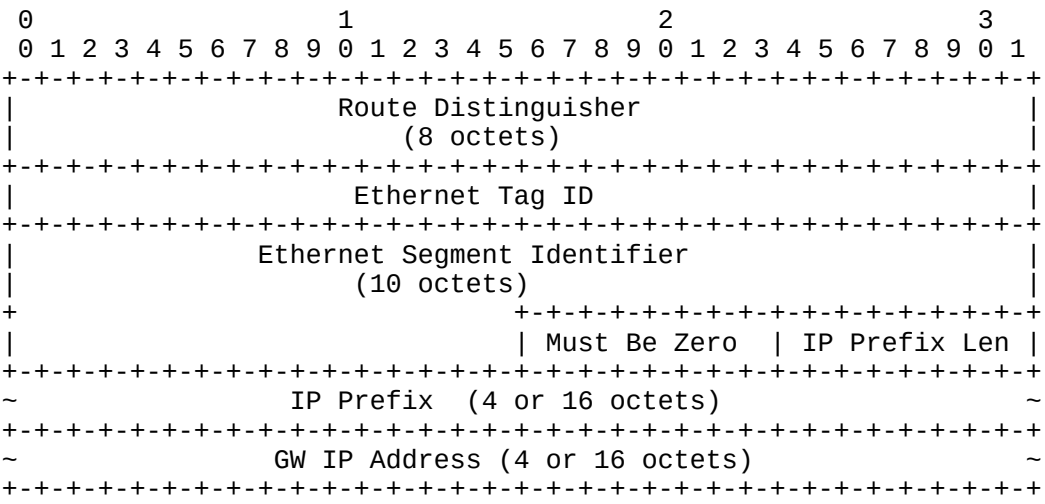

Figure 6: EVPN IP Prefix Object

The ICMPv6 Validation Request is sent by the ingress PE using the SRv6 service SID associated with the IP Prefix route announced by the egress PE and the possible transport encapsulations(e.g, SRv6 segment list) to reach the egress PE.

6. Operations

6.1. Unicast Data Plane Connectivity Checks

Figure 7 is an example of a EVPN network. CE1 is dual-homed to PE1 and PE2. Assume that PE1 announced a MAC route with RD 192.0.2.1:00 and C-MAC 00-AA-00-BB-00-CC and with SRv6 Sevice SID A:2:101:: for EVI 10. Similarly, PE2 announced a MAC route with RD 203.0.113.2:00 and C-MAC 00-AA-00-BB-00-CC and with SRv6 Sevice SID B:2:101::.

On PE3, when an operator performs a connectivity check for the C-MAC address 00-AA-00-BB-00-CC on PE1, the operator initiates an ICMPv6 Validation Request containing the EVPN MAC/IP Object. The ICMPv6 Validation Request packet is sent with the {SRv6 segment list to reach PE1, SRv6 Service SID = A:2:101::} . Once the ICMPv6 Validation Request packet encapsulated with SRH reaches PE1, PE1 as the final destination of the IPv6 packet, will proceed to process the next header in the packet, i.e, the ICMPv6 Validation Request. Then PE1 will process the packet and perform checks for the EVPN MAC/IP Object and return the ICMPv6 Validation Reply with the code indicating the validation result.

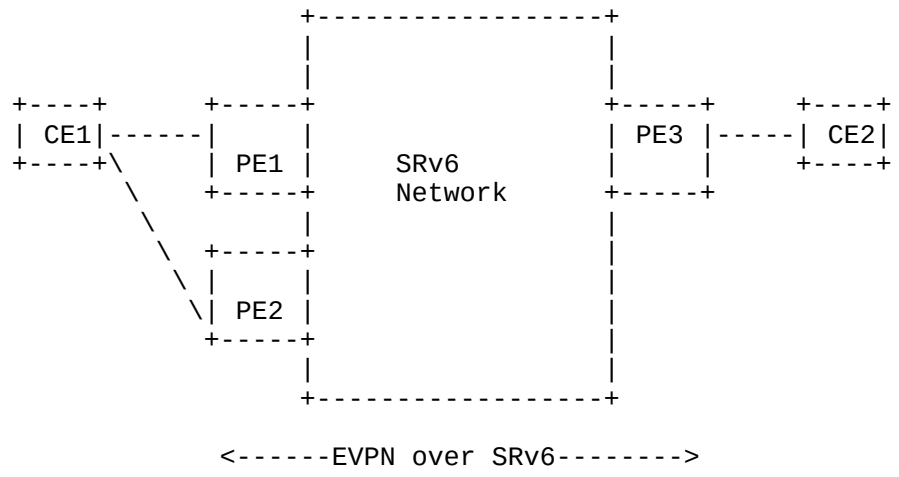

Figure 7: EVPN Network

6.2. Inclusive Multicast Data Plane Connectivity Checks

To be completed.

6.3. EVPN Aliasing Data Plane Connectivity Check

Still taking the network in Figure 7 as an example, assume PE1 announced an Ethernet A-D per-EVI route with the ESI set to CE1 system ID and SRv6 Service SID A:2:101::. Additionally, assume PE2 announced an Ethernet A-D per-EVI route with the ESI set to CE1 system ID and SRv6 Service SID B:2:101::.

At PE3, when an operator performs a connectivity check for the aliasing aspect of the EVPN Ethernet A-D route on PE1, the operator initiates an ICMPv6 Validation Request with the EVPN Ethernet A-D Object. The ICMPv6 Validation Request packet is sent with the SRH{SRv6 Segment List to reach PE1, SRv6 Service SID A:2:101::} and IPv6 header.

When PE1 receives the packet, it will process the packet and perform checks for the EVPN Ethernet A-D Object present in the packet and return the ICMPv6 Validation Reply with the code indicating the validation result.

6.4. EVPN IP Prefix (RT-5) Data Plane Connectivity Check

Assume PE1 in Figure 7 announced an IP Prefix route (RT-5) with an IP prefix reachable behind CE1 and SRv6 Sevice SID A:2:101::. When an operator on PE3 performs a connectivity check for the IP prefix on PE1, the operator initiates an ICMPv6 Validation Request with the EVPN IP Prefix Object included. The ICMPv6 Validation Request packet is sent with the SRH{SRv6 Segment List to reach PE1, SRv6 Service SID A:2:101::} and IPv6 header.

When PE1 receives the packet, it will process the packet and perform checks for the EVPN IP Prefix Object present in the packet and return the ICMPv6 Validation Reply with the code indicating the validation result.

7. IANA Considerations

TBA

8. Security Considerations

Security considerations discussed in [RFC4443], [RFC4884] and [RFC9252] apply to this document.

To protect against unauthorized sources using validation request messages to obtain network information, it is RECOMMENDED that implementations provide a means of checking the source addresses of validation request messages against an access list before accepting the message.

The validation mechanism SHOULD be only used in the limited domain. The validation request contains the control plane information, policies should be implemented on the edge devices of the domain to prevent the information from being leaked into other domains.

In order to protect local resources, implementations SHOULD rate-limit incoming ICMP Request messages.

This document does not introduce any new privacy concerns because these Objects contain the same information that are present in data packets and EVPN routes.

9. References

9.1. Normative References

[I-D.liu-6man-icmp-verification]
Liu, Y. and Y. Liu, "Extending ICMPv6 for SRv6-related Information Validation", Work in Progress, Internet-Draft, draft-liu-6man-icmp-verification-08, 17 October 2025, <https://datatracker.ietf.org/doc/html/draft-liu-6man-icmp-verification-08>.

[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>.

[RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", STD 89, RFC 4443, DOI 10.17487/RFC4443, March 2006, <https://www.rfc-editor.org/info/rfc4443>.

[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>.

[RFC4884] Bonica, R., Gan, D., Tappan, D., and C. Pignataro, "Extended ICMP to Support Multi-Part Messages", RFC 4884, DOI 10.17487/RFC4884, April 2007, <https://www.rfc-editor.org/info/rfc4884>.

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

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

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[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>.

[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>.

9.2. Informative References

[RFC9489] Jain, P., Sajassi, A., Salam, S., Boutros, S., and G. Mirsky, "Label Switched Path (LSP) Ping Mechanisms for EVPN and Provider Backbone Bridging EVPN (PBB-EVPN)", RFC 9489, DOI 10.17487/RFC9489, November 2023, <https://www.rfc-editor.org/info/rfc9489>.

Author’s Address

Yao Liu
ZTE
Nanjing
China
Email: liu.yao71@zte.com.cn