
Workgroup: BESS Workgroup
Internet-Draft: draft-ietf-bess-evpn-lsp-ping-10
Published: 8 May 2023
Intended Status: Standards Track
Expires: 9 November 2023
Authors: P. Jain A. Sajassi S. Salam S. Boutros G. Mirsky
Cisco Cisco Cisco Ciena Ericsson

LSP-Ping Mechanisms for EVPN and PBB-EVPN

Abstract

LSP Ping is a widely deployed Operation, Administration, and Maintenance mechanism in MPLS networks. This document describes mechanisms for detecting data plane failures using LSP Ping in MPLS based Ethernet VPN (EVPN) and Provider Backbone Bridging with EVPN (PBB-EVPN) networks.

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 9 November 2023.

Copyright Notice

Copyright (c) 2023 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction
 2. Specification of Requirements
 3. Terminology
 4. Proposed Target FEC Stack Sub-TLVs
 - 4.1. EVPN MAC/IP Sub-TLV
 - 4.2. EVPN Inclusive Multicast Sub-TLV
 - 4.3. EVPN Ethernet Auto-Discovery Sub-TLV
 - 4.3.1. Ethernet TAG Value
 - 4.3.2. Per-ES EVPN Auto-Discovery Route with different RDs
 - 4.3.3. EVPN VPWS
 - 4.4. EVPN IP Prefix Sub-TLV
 5. Encapsulation of OAM Ping Packets
 6. Operations
 - 6.1. Unicast Data-plane connectivity checks
 - 6.2. Inclusive Multicast Data-plane Connectivity Checks
 - 6.2.1. Ingress Replication
 - 6.2.2. Using P2MP P-tree
 - 6.2.3. Controlling Echo Responses when using P2MP P-tree
 - 6.3. EVPN Aliasing Data-plane connectivity check
 - 6.4. EVPN IP Prefix (RT-5) Data-plane connectivity check
 7. Security Considerations
 8. IANA Considerations
 - 8.1. Sub-TLV Type
 - 8.2. Proposed new Return Codes
 9. Acknowledgments
 10. Normative References
- Authors' Addresses

1. Introduction

[RFC7432] describes MPLS based EVPN technology. An EVPN comprises CE(s) connected to PE(s). The PEs provide layer-2 EVPN among the CE(s) over the MPLS core infrastructure. In EVPN networks, the PEs advertise the MAC addresses learned from the locally connected CE(s), along with 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).

[RFC7623] describes the use of Provider Backbone Bridging [802.1ah] with EVPN. PBB-EVPN maintains the Customer MAC (C-MAC) learning in data plane and only advertises Provider Backbone MAC (B-MAC) addresses in control plane using BGP.

Procedures for simple and efficient mechanisms to detect data plane failures using LSP Ping in MPLS network are well defined in [RFC8029] and [RFC6425]. The basic idea for the LSP Ping mechanism is to send a MPLS Echo Request packet along the same data path as data packets belonging to the same Forwarding Equivalent Class (FEC). The Echo Request packet carries the FEC being verified in the Target FEC Stack TLV [RFC8029]. Once the Echo Request packet reaches the end of the MPLS path, it is sent to the control plane of the egress PE. The Echo Request packet contains sufficient information to verify the correctness of data plane operations and validate data plane against the control plane. The Egress PE sends the results of the validation in an Echo Reply packet to the originating PE of the Echo Request packet.

This document defines procedures to detect data plane failures using LSP Ping in MPLS networks deploying EVPN and PBB-EVPN. This document defines four new Sub-TLVs for Target FEC Stack TLV with the purpose of identifying the FEC on the Egress PE.

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

AD: Auto Discovery

B-MAC: Backbone MAC Address

BUM: Broadcast, Unknown Unicast or Multicast

CE: Customer Edge Device

C-MAC: Customer MAC Address

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

FEC: Forwarding Equivalent Class

G-ACh: Generic Associated Channel

GAL: G-ACh Label

OAM: Operations, Administration, and Maintenance

P2MP: Point-to-Multipoint

PBB-EVPN: Provider Backbone Bridge with EVPN

PE: Provider Edge Device

VPWS: Virtual Private Wire Service

4. Proposed Target FEC Stack Sub-TLVs

This document introduces four new Target FEC Stack Sub-TLVs that are included in the MPLS Echo Request packet. The Echo Request packets are used for connectivity check in data plane in EVPN and PBB-EVPN networks. The target FEC stack Sub-TLVs MAY be used to validate that an identifier for a given EVPN is programmed at the target node.

4.1. EVPN MAC/IP Sub-TLV

The EVPN MAC/IP Sub-TLV identifies the target MAC, MAC/IP binding for ARP/ND, or IP address for an EVPN Instance Identifier (EVI) under test at an egress PE. This Sub-TLV is included in the Echo Request sent by a EVPN/PBB-EVPN PE to a Peer PE.

The EVPN MAC/IP Sub-TLV fields are derived from the MAC/IP Advertisement route defined in Section 7.2 in [\[RFC7432\]](#) and have the format as shown in Figure 1. The fields of EVPN MAC/IP Sub-TLV should be set according to the following that is consistent with [\[RFC7432\]](#) and [\[RFC7623\]](#):

- The Route Distinguisher (RD) field is a 10-octet field and is set to the RD of the MAC-VRF on the Peer PE.
- The Ethernet TAG ID field can be 0 or a valid VLAN ID for EVPN VLAN-aware bundle service [\[RFC7432\]](#). For PBB-EVPN, the value of this field is always 0 as per Section 5.2 of [\[RFC7623\]](#).

- The Ethernet Segment Identifier field is 10-octet field. For EVPN, it is set to 0 for singlehomed ES or to a valid ESI ID for a multihomed ES. For PBB-EVPN, the Ethernet Segment Identifier field must be set to either 0 (for single-homed segments or multihomed segments with per-I-SID load-balancing) or to MAX-ESI (for multihomed segments with per-flow load-balancing) as describe in Section 5.2 in [RFC7623].
- The MAC Addr Len field specifies the MAC length in bits. Only 48 bit MAC Addresses are supported as this document follows MAC address length supported by [RFC7432].
- 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 either set to 32 for IPv4 or 128 for IPv6 address.
- The Must Be Zero fields are set to 0. The receiving PE should ignore the Must Be Zero fields.

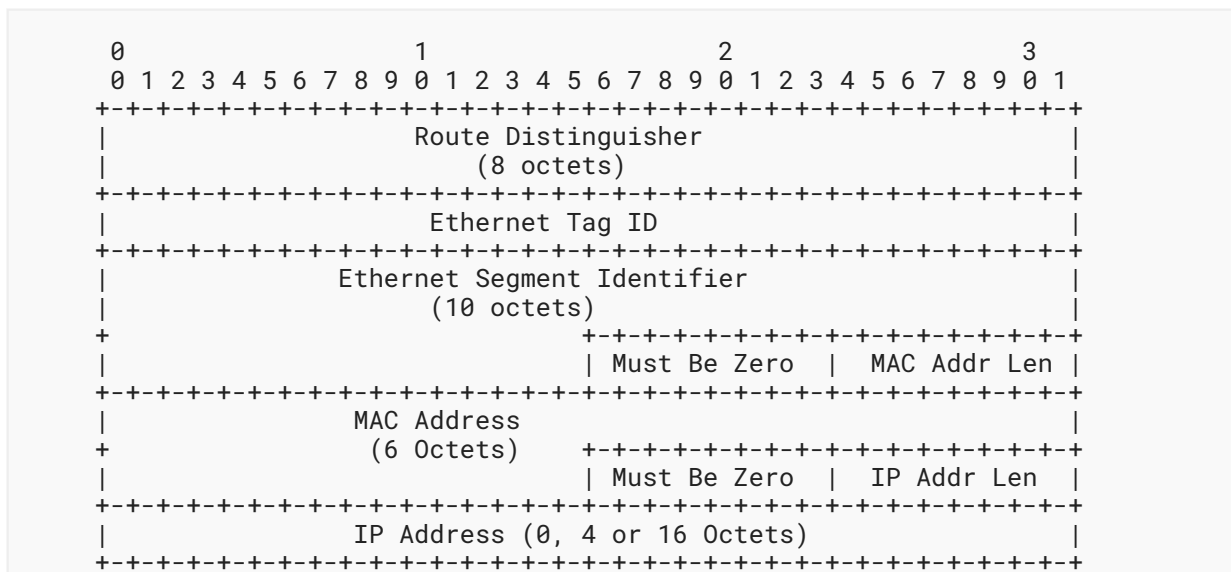


Figure 1: EVPN MAC Sub-TLV format

The MPLS Echo Request is sent by the ingress PE using the EVPN MPLS label(s) associated with the MAC/IP advertisement route announced by the egress PE and the MPLS transport label(s) to reach the egress PE.

In EVPN, MAC/IP Advertisement has multiple personality and it is used for the following cases:

- This route with only 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 forwarding, and IP forwarding in case of symmetric IRB.

When MPLS Echo Request is sent by an ingress PE, the contents of Echo Request and the egress PE mode of operation (i.e., IRB mode or L2 mode) along with EVPN MPLS label of the packet, determine which of the above three cases is this Echo Request for. When the egress PE receives MAC/IP Sub-TLV containing only MAC address, the egress PE validates the MAC state and forwarding. When the egress PE receives MAC/IP Sub-TLV containing both MAC and IP addresses and if the EVPN label 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 EVPN label points to an IP-VRF, then the egress PE validates IP state and forwarding. Any other combinations, such as the egress PE receiving MAC/IP Sub-TLV containing only MAC address but with EVPN label pointing to an IP-VRF, should be considered invalid and Echo Reply should be sent with the appropriate return code by the egress PE to the ingress PE.

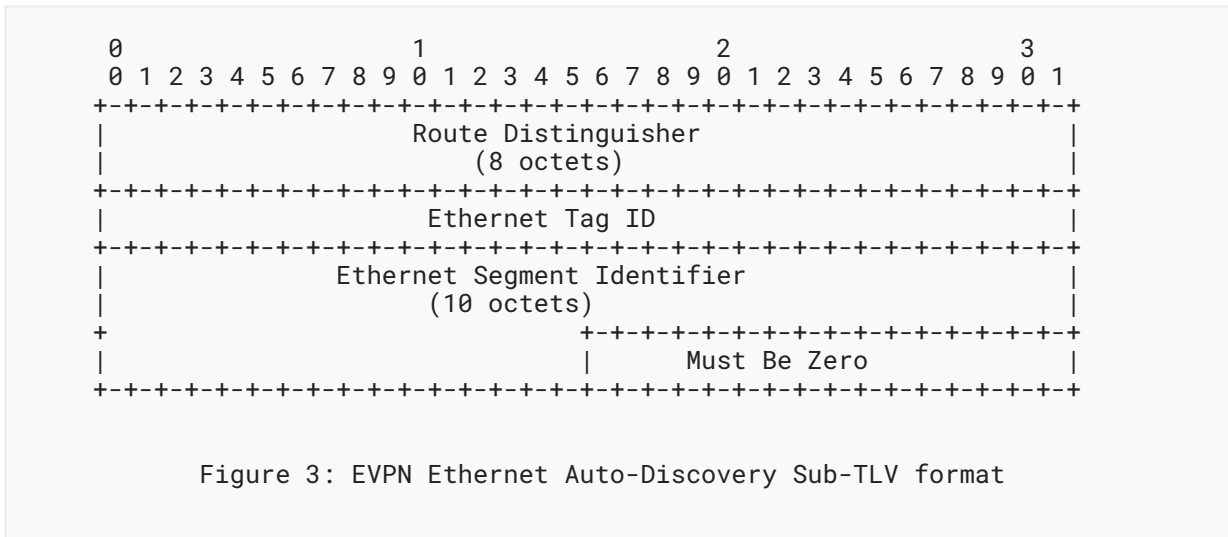
4.2. EVPN Inclusive Multicast Sub-TLV

The EVPN Inclusive Multicast Sub-TLV fields are based on the EVPN Inclusive Multicast Tag route defined in [RFC7432] Section 7.3. This TLV is included in the Echo Request sent to the EVPN peer PE by the originator of request to verify the multicast connectivity state on the peer PE(s) in EVPN and PBB-EVPN networks.

The EVPN Inclusive Multicast Sub-TLV has the format as shown in Figure 2. The fields of this Sub-TLV should be set according to the following that is consistent with [RFC7432] and [RFC7623]:

- The Route Distinguisher (RD) field is a 10-octet field and is set to the RD of the MAC-VRF on the Peer PE.
- For EVPN, the Ethernet TAG ID field can be set to 0 or a valid VLAN ID for EVPN VLAN-aware bundle service [RFC7432]. For PBB-EVPN, the value of this field is set to Service Instance Identifier (I-SID) value as per Section 5.3 of [RFC7623].
- The IP Addr Len field specifies length of the Originating Router's IP Addr field in bits and it is either set to 32 for IPv4 or 128 for IPv6 address.
- The Originating Router's IP Addr field is set to IPv4 or IPv6 address of the peer PE.

- The Ethernet TAG ID field can be 0, MAX-ET or a valid VLAN ID as described in Section 4.3.1 below.
- The Ethernet Segment Identifier field is 10-octet field and is set to 0 for singlehomed 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.



4.3.1. Ethernet TAG Value

EVPN Ethernet AD Sub-TLV can be sent in the context of per-Ethernet Segment (ES) or per-EVI. When an operator performs a connectivity check for the BUM L2 service, Echo Request packet sent, MAY contain EVPN Ethernet AD Sub-TLV to emulate traffic coming from a multihomed site. In this case, the EVPN Ethernet AD Sub-TLV is added in per-ES context. When Echo Request packet is sent for the connectivity check for EVPN Aliasing state, the context for the EVPN Ethernet AD Sub-TLV is per-EVI.

Ethernet Tag field value in EVPN Ethernet AD Sub-TLV, MUST be set according to the context:

- For per-ES context, the Ethernet Tag field in the sub-TLV MUST be set to the reserved MAX-ET value [RFC7432]
- For per-EVI context, the Ethernet Tag field in the sub-TLV MUST be set to the non-reserved value

4.3.2. Per-ES EVPN Auto-Discovery Route with different RDs

Section 8.2 of [RFC7432] specifies that a per-ES EVPN AD 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 AD Sub-TLV, MUST be the RD value received for the EVI in the per-ES EVPN AD Route.

4.3.3. EVPN VPWS

LSP Ping can also be used to detect data plane failures for EVPN Virtual Private Wire Service (VPWS) described in [RFC8214]. The Echo Request packet carries EVPN Ethernet AD Sub-TLV with fields populated from the EVPN Ethernet AD per EVI route announced by the egress PE for the EVPN VPWS under test. The Echo Request is sent by the ingress PE using the EVPN MPLS label associated with the EVPN Ethernet AD route announced by the egress PE and the MPLS transport label(s) to reach the egress PE.

The egress PE process the Echo Request packet and perform checks for the EVPN Ethernet AD Sub-TLV present in the Target FEC Stack TLV as described in Section 4.4 in [RFC8029] and respond according to [RFC8029] processing rule. Egress PE can identify that the Echo Request is for EVPN VPWS instance as EVI (identified by the RD) for EVPN VPWS is different from that for EVPN. The egress PE will use the information from the EVPN Ethernet AD Sub-TLV in the Target FEC Stack TLV and validate the VLAN state for the EVPN VPWS under test. For the success case, the egress PE will reply with Return Code 3 - "Replying router is an egress for the FEC at stack-depth".

4.4. EVPN IP Prefix Sub-TLV

The EVPN IP Prefix Sub-TLV identifies the IP Prefix for an EVI under test at a peer PE.

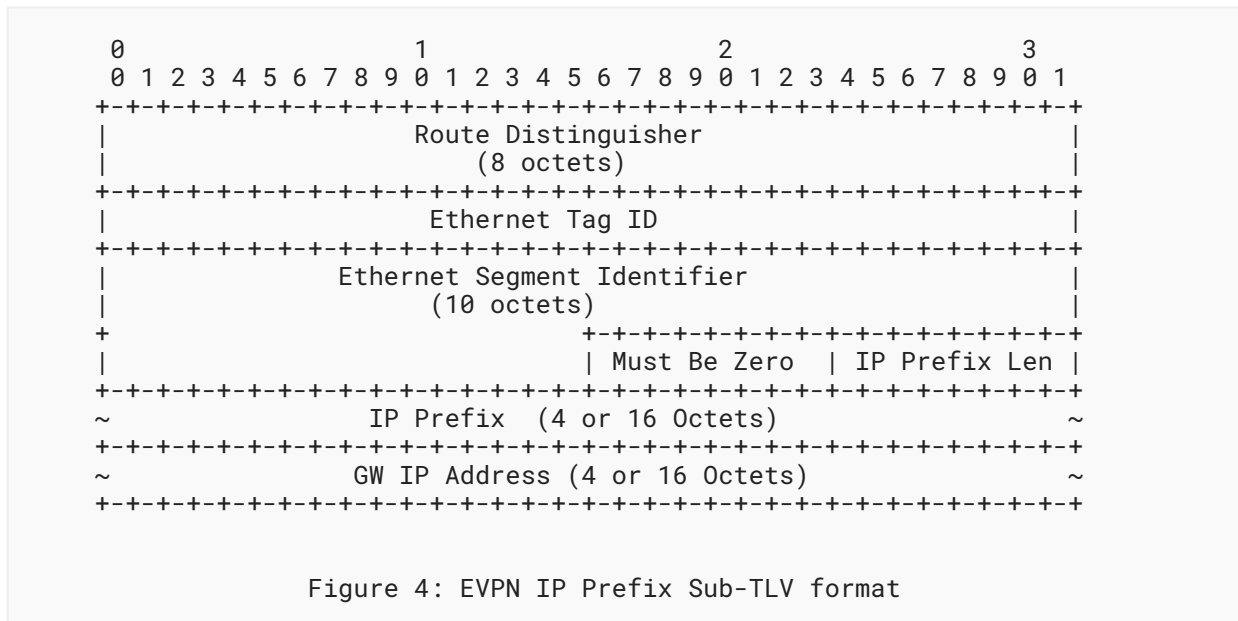
The EVPN IP Prefix Sub-TLV fields are derived from the IP Prefix Route (RT-5) advertisement defined in [RFC9136]. This sub-TLV applies to only EVPN.

The EVPN IP Prefix Sub-TLV has the format as shown in Figure 4. The total length of this sub-TLV can 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 EVPN IP Prefix Sub-TLV should be set according to the following that is consistent with [RFC9136]:

- 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 10-octet field and is set to a valid ESI ID if ESI is used as Overlay Index as per Section 3.1 of [RFC9136]. Otherwise the Ethernet Segment Identifier field is set to all 0s.
- The IP Prefix Len field specifies the number of bits in the IP Prefix field. It is set to 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 16-octet IPv6 address (with trailing 0 bits to make 128 bits in all).

- The Gateway (GW) IP Address field is set to a 4-octet IPv4 address or 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 Must Be Zero field is set to 0. The receiving PE should ignore the Must Be Zero field.



The MPLS Echo Request is sent by the ingress PE using the EVPN MPLS label(s) associated with the IP Prefix route announced by the egress PE and the MPLS transport label(s) to reach the egress PE.

5. Encapsulation of OAM Ping Packets

The MPLS Echo Request IP/UDP packets MUST be encapsulated with the Transport and EVPN Label(s) followed by the Generic Associated Channel Label (GAL) [RFC5586] which is the bottom most label. The GAL is followed by a Generic Associated Channel Header carrying IPv4(0x0021) or IPv6(0x0057) Channel Type. The code points for IPv4 and IPv6 channels are defined in Generic Associated Channel (G-ACh) Parameters by IANA.

6. Operations

6.1. Unicast Data-plane connectivity checks

Figure 5 is an example of a PBB-EVPN network. CE1 is dual-homed to PE1 and PE2. Assume, PE1 announced a MAC route with RD 192.0.2.1:00 and B-MAC 00-AA-00-BB-00-CC and with MPLS label 16001 for EVI 10. Similarly, PE2 announced a MAC route with RD 203.0.113.2:00 and B-MAC 00-AA-00-BB-00-CC and with MPLS label 16002.

On PE3, when an operator performs a connectivity check for the B-MAC address 00-AA-00-BB-00-CC on PE1, the operator initiates an LSP Ping request with the target FEC stack TLV containing EVPN MAC Sub-TLV in the Echo Request packet. The Echo Request packet is sent with the {Transport Label(s) to reach PE1, EVPN Label = 16001, GAL} MPLS label stack and IP ACH Channel header. Once the Echo Request packet reaches PE1, PE1 will use the GAL and the IP ACH Channel header to determine that the packet is IPv4 or IPv6 OAM Packet. The PE1 will process the packet and perform checks for the EVPN MAC Sub-TLV present in the Target FEC Stack TLV as described in Section 4.4 in [RFC8029] and respond according to [RFC8029] processing rules.

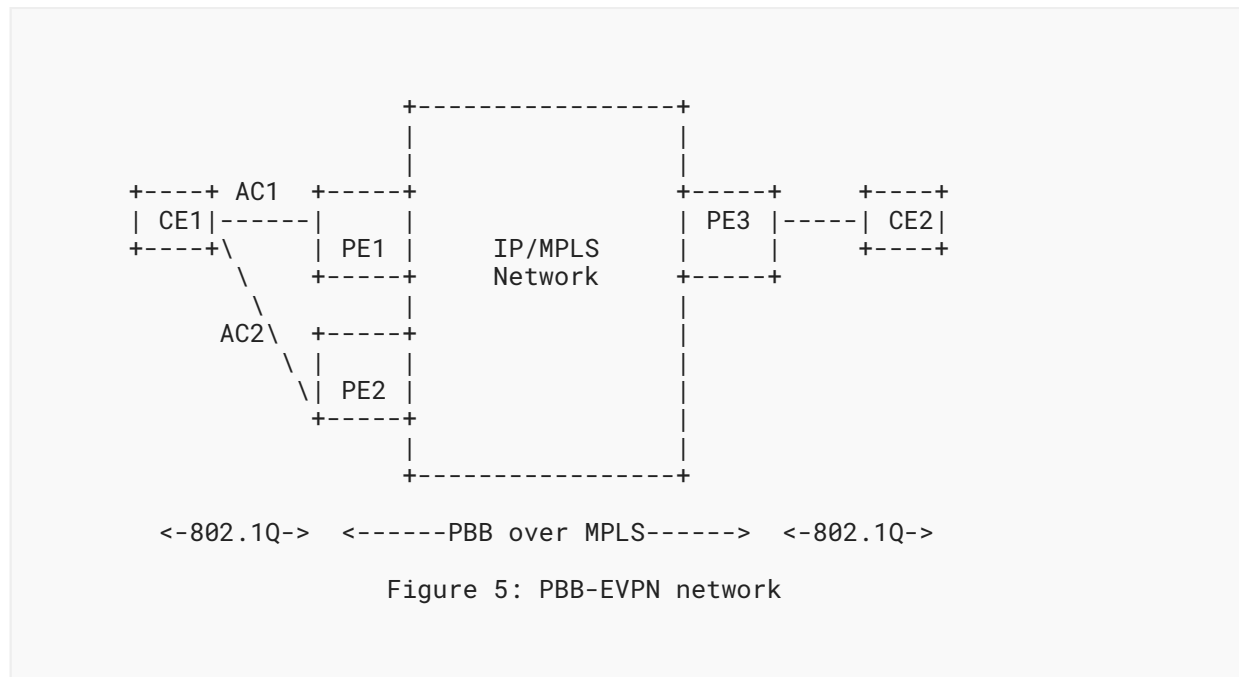


Figure 5: PBB-EVPN network

Similarly, on PE3, when an operator performs a connectivity check for the B-MAC address 00-AA-00-BB-00-CC on PE2, the operator initiates an LSP Ping request with the target FEC stack TLV containing EVPN MAC Sub-TLV in the Echo Request packet. The Echo Request packet is sent with the {MPLS transport Label(s) to reach PE2, EVPN Label = 16002, GAL} MPLS label stack and IP ACH Channel header.

LSP Ping operation for unicast data plane connectivity checks in EVPN, are similar to those described above for PBB-EVPN except that the checks are for C-MAC addresses instead of B-MAC addresses.

In EVPN network, an operator can also perform MAC state test using aliasing label for the MAC to verify the MAC state on the egress multihoming PE that did not learn the MAC from the multihomed CE on a local ESI but has announced Ethernet AD per-EVI and per-ESI routes for the ESI. This is due to the fact that MAC state on multihoming PEs that did not learn the MAC locally, get created from EVPN MAC/IP route advertisement from the multihoming PE that has learned the CE's MAC address locally.

6.2. Inclusive Multicast Data-plane Connectivity Checks

6.2.1. Ingress Replication

Assume PE1 announced an Inclusive Multicast route for EVI 10, with RD 192.0.2.1:00, Ethernet Tag (ISID 10), PMSI tunnel attribute Tunnel type set to ingress replication and downstream assigned inclusive multicast MPLS label 17001. Similarly, PE2 announced an Inclusive Multicast route for EVI 10, with RD 203.0.113.2:00, Ethernet Tag (ISID 10), PMSI tunnel attribute Tunnel type set to ingress replication and downstream assigned inclusive multicast MPLS label 17002.

Given CE1 is dual-homed to PE1 and PE2, assume that PE1 is the DF for ISID 10 for the port corresponding to the ESI 11aa.22bb.33cc.44dd.5500.

When an operator at PE3 initiates a connectivity check for the inclusive multicast on PE1, the operator initiates an LSP Ping request with the target FEC stack TLV containing EVPN Inclusive Multicast Sub-TLV in the Echo Request packet. The Echo Request packet is sent with the {Transport Label(s) to reach PE1, EVPN Incl. Multicast Label = 17001, GAL} MPLS label stack and IP ACH Channel header. Once the Echo Request packet reaches PE1, PE1 will use the GAL and the IP ACH Channel header to determine that the packet is IPv4 or IPv6 OAM Packet. The packet will have EVPN Inclusive multicast label. PE1 will process the packet and perform checks for the EVPN Inclusive Multicast Sub-TLV present in the Target FEC Stack TLV as described in Section 4.4 in [RFC8029] and respond according to [RFC8029] processing rules. For the success case, PE1 will reply with Return Code 3 - "Replying router is an egress for the FEC at stack-depth".

Similarly, an operator at PE3, may initiate an LSP Ping to PE2 with the target FEC stack TLV containing EVPN Inclusive Multicast sub-TLV in the Echo Request packet. The Echo Request packet is sent with the {transport Label(s) to reach PE2, EVPN Incl. Multicast Label = 17002, GAL} MPLS label stack and IP ACH Channel header. Once the Echo Request packet reaches PE2, PE2 will use the GAL and the IP ACH Channel header to determine that the packet is IPv4 or IPv6 OAM Packet. The processing on PE2 will be similar to the that on PE1 as described above and for the success case, PE2 will reply with Return Code 3 - "Replying router is an egress for the FEC at stack-depth" as per [RFC8029].

In case of EVPN, in the Echo Request packet, an EVPN Ethernet AD Sub-TLV and the associated MPLS Split Horizon Label above the GAL in the MPLS label stack, may be added to emulate traffic coming from a multihomed site. The Split Horizon label is used by leaf PE(s) attached to the same multihomed site to not forward packets back to the multihomed site. If the behavior on a leaf PE is to not forward the packet to the multihomed site on the ESI identified by EVPN Ethernet AD Sub-TLV because of Split Horizon filtering, the PE will reply with a Return Code indicating that "Replying router is egress for the FEC at the stack depth. In addition, the BUM packets are dropped on the ES corresponding to the ESI received in EVPN Ethernet AD Sub-TLV because of the Split Horizon Group filtering" as described in Section 8.

6.2.2. Using P2MP P-tree

Both inclusive P-Tree and aggregate inclusive P-tree can be used in EVPN or PBB-EVPN networks.

When using an inclusive P-tree arrangement, p2mp p-tree transport label itself is used to identify the L2 service associated with the Inclusive Multicast Route, this L2 service could be a customer Bridge, or a Provider Backbone Bridge.

For an Inclusive P-tree arrangement, when an operator performs a connectivity check for the multicast L2 service, the operator initiates an LSP Ping request with the target FEC stack TLV containing EVPN Inclusive Multicast Sub-TLV in the Echo Request packet. The Echo Request packet is sent over P2MP LSP with the {P2MP P-tree label, GAL} MPLS label stack and IP ACH Channel header.

When using Aggregate Inclusive P-tree, a PE announces an upstream assigned MPLS label along with the P-tree ID, so both the p2mp p-tree MPLS transport label and the upstream MPLS label can be used to identify the L2 service.

For an Aggregate Inclusive P-tree arrangement, when an operator performs a connectivity check for the multicast L2 service, the operator initiates an LSP Ping request with the target FEC stack TLV containing EVPN Inclusive Multicast Sub-TLV in the Echo Request packet. The Echo Request packet is sent over P2MP LSP using the IP-ACH Control channel with the {P2MP P-tree label, EVPN Upstream assigned Multicast Label, GAL} MPLS label stack and IP ACH Channel header.

The Leaf PE(s) of the p2mp tree will process the packet and perform checks for the EVPN Inclusive Multicast Sub-TLV present in the Target FEC Stack TLV as described in Section 4.4 in [RFC8029] and respond according to [RFC8029] processing rules. For the success case, the Leaf PE will reply with Return Code 3 - "Replying router is an egress for the FEC at stack-depth".

In case of EVPN, in the Echo Request packet, an EVPN Ethernet AD Sub-TLV and the associated MPLS Split Horizon Label above the GAL in MPLS label stack, may be added to emulate traffic coming from a multihomed site. In case of p2mp p-tree, the Split Horizon Label is upstream assigned and is received by all the leaf PEs of the p2mp-ptree. The Split Horizon label is used by leaf PE(s) attached to the same multihomed site not to forward packets back to the multihomed site. If the behavior on a leaf PE is to not forward the packet to the multihomed site on the ESI in EVPN Ethernet AD Sub-TLV because of Split Horizon filtering, the PE will reply with a Return Code indicating that "Replying router is egress for the FEC at the stack depth. In addition, the BUM packets are dropped on the ES corresponding to the ESI received in EVPN Ethernet AD Sub-TLV because of the Split Horizon Group filtering" as described in Section 8. If the leaf PE does not have the ESI identified in the EVPN Ethernet AD Sub-TLV, the PE can reply with a Return Code indicating that "Replying router is egress for the FEC at the stack depth. In addition, the BUM packets are forwarded because there is no ES corresponding to the ESI received in EVPN Ethernet AD Sub-TLV".

6.2.3. Controlling Echo Responses when using P2MP P-tree

The procedures described in [RFC6425] for preventing congestion of Echo Responses (Echo Jitter TLV) and limiting the Echo Reply to a single egress node (Node Address P2MP Responder Identifier TLV) can be applied to LSP Ping in EVPN and PBB-EVPN when using P2MP P-trees for broadcast, multicast, and unknown unicast traffic.

6.3. EVPN Aliasing Data-plane connectivity check

Assume PE1 announced an Ethernet AD per-EVI Route with the ESI set to CE1 system ID and MPLS label 19001, and PE2 an Ethernet AD per-EVI Route with the ESI set to CE1 system ID and MPLS label 19002.

At PE3, when an operator performs a connectivity check for the aliasing aspect of the EVPN Ethernet AD route on PE1, the operator initiates an LSP Ping request with the target FEC stack TLV containing EVPN Ethernet AD Sub-TLV in the Echo Request packet. The Echo Request packet is sent with the {Transport label(s) to reach PE1, EVPN Ethernet AD Label 19001, GAL} MPLS label stack and IP ACH Channel header.

When PE1 receives the packet it will process the packet and perform checks for the EVPN Ethernet AD Sub-TLV present in the Target FEC Stack TLV as described in Section 4.4 in [RFC8029] and respond according to [RFC8029] processing rules.

6.4. EVPN IP Prefix (RT-5) Data-plane connectivity check

Assume PE1 in Figure 5, announced an IP Prefix Route (RT-5) with an IP prefix reachable behind CE1 and MPLS label 20001. When an operator on PE3 performs a connectivity check for the IP prefix on PE1, the operator initiates an LSP Ping request with the target FEC stack TLV containing EVPN IP Prefix Sub-TLV in the Echo Request packet. The Echo Request packet is sent with the {Transport label(s) to reach PE1, EVPN IP Prefix Label 20001 } MPLS label stack.

When PE1 receives the packet it will process the packet and perform checks for the EVPN IP Prefix Sub-TLV present in the Target FEC Stack TLV as described in Section 4.4 in [RFC8029] and respond according to [RFC8029] processing rules.

7. Security Considerations

The proposal introduced in this document does not introduce any new security considerations beyond that already apply to [RFC7432], [RFC7623] and [RFC6425]. Furthermore, the security considerations discussed in [RFC8029] apply to this document and need to be considered. As described in [RFC8029], these security considerations are:

- Denial-of-Service (DoS) attack, by sending MPLS echo requests/replies to LSRs and thereby increasing their workload.
- Obfuscating the state of the MPLS data-plane liveness by spoofing, hijacking, replaying, or otherwise tampering with MPLS echo Requests and replies.
- Obtaining information about the network by an unauthorized source using an LSP ping.

There are mitigations described in [RFC8029]. The same mitigations can be applied to the LSP Ping procedures described in this draft and thus this document doesn't require additional security considerations beyond the one described in [RFC8029].

The proposal does not introduce any new privacy concerns because these TLVs contain the same information that are present in data packets and EVPN routes.

8. IANA Considerations

8.1. Sub-TLV Type

This document defines four new Sub-TLV type to be included in Target FEC Stack TLV (TLV Type 1, 16 and 21) [RFC9041] in Echo Request and Echo Reply messages in EVPN and PBB-EVPN network.

IANA is requested to assign lowest 4 free values for the four Sub-TLVs listed below from the "Standards Action" (0-16383) range, in the "Sub-TLVs for TLV Types 1, 16, and 21" sub-registry, in the "TLVs" registry in the "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" name space:

- EVPN MAC/IP Sub-TLV
- EVPN Inclusive Multicast Sub-TLV
- EVPN Auto-Discovery Sub-TLV
- EVPN IP Prefix Sub-TLV

8.2. Proposed new Return Codes

[RFC8029] defines values for the Return Code field of Echo Reply. This document proposes two new Return Codes, which SHOULD be included in the Echo Reply message by a PE in response to Echo Request message in EVPN and PBB-EVPN network.

IANA is requested to assign 2 lowest free values for the 2 Return Codes listed below from the "Return Codes" registry in the "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" name space:

- Replying router is egress for the FEC at the stack depth. In addition, the BUM packets are dropped on the ES corresponding to the ESI received in EVPN Ethernet AD Sub-TLV because of the Split Horizon Group filtering.
- Replying router is egress for the FEC at the stack depth. In addition, the BUM packets are forwarded because there is no ES corresponding to the ESI received in EVPN Ethernet AD Sub-TLV.

9. Acknowledgments

The authors would like to thank Loa Andersson, Alexander Vainshtein, Ron Sdayoor, Jim Guichard, Lars Eggert, John Scudder, Eric Vyncke, Warren Kumari, Patrice Brissette and Weiguo Hao for their valuable comments.

10. 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>>.
- [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>>.
- [RFC5586] Bocci, M., Ed., Vigoureux, M., Ed., and S. Bryant, Ed., "MPLS Generic Associated Channel", RFC 5586, DOI 10.17487/RFC5586, June 2009, <<https://www.rfc-editor.org/info/rfc5586>>.
- [RFC6425] Saxena, S., Ed., Swallow, G., Ali, Z., Farrel, A., Yasukawa, S., and T. Nadeau, "Detecting Data-Plane Failures in Point-to-Multipoint MPLS - Extensions to LSP Ping", RFC 6425, DOI 10.17487/RFC6425, November 2011, <<https://www.rfc-editor.org/info/rfc6425>>.
- [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>>.
- [RFC8029] Kompella, K., Swallow, G., Pignataro, C., Ed., Kumar, N., Aldrin, S., and M. Chen, "Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures", RFC 8029, DOI 10.17487/RFC8029, March 2017, <<https://www.rfc-editor.org/info/rfc8029>>.
- [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>>.
- [RFC8214] Boutros, S., Sajassi, A., Salam, S., Drake, J., and J. Rabadan, "Virtual Private Wire Service Support in Ethernet VPN", RFC 8214, DOI 10.17487/RFC8214, August 2017, <<https://www.rfc-editor.org/info/rfc8214>>.
- [RFC9041] Andersson, L., Chen, M., Pignataro, C., and T. Saad, "Updating the MPLS Label Switched Paths (LSPs) Ping Parameters IANA Registry", RFC 9041, DOI 10.17487/RFC9041, July 2021, <<https://www.rfc-editor.org/info/rfc9041>>.
- [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>>.

Authors' Addresses

Parag Jain

Cisco

Canada

Email: paragj@cisco.com**Ali Sajassi**

Cisco

United States of America

Email: sajassi@cisco.com**Samer Salam**

Cisco

Canada

Email: ssalam@cisco.com**Sami Boutros**

Ciena

United States of America

Email: sboutros@ciena.com**Greg Mirsky**

Ericsson

United States of America

Email: gregimirsky@gmail.com