Authors:
| P. Jain | A. Sajassi | S. Salam | S. Boutros | G. Mirsky |
| *Cisco* | *Cisco* | *Cisco* | *Ciena* | *Ericsson* |

RFC 9489
Label Switched Path (LSP) Ping Mechanisms for EVPN and Provider Backbone Bridging EVPN (PBB-EVPN)

Abstract

Label Switched Path (LSP) Ping is a widely deployed Operation, Administration, and Maintenance (OAM) 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 EVPN (PBB-EVPN) networks.

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/rfc9489.

Copyright Notice

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).

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

Procedures for simple and efficient mechanisms to detect data plane failures using LSP Ping in MPLS networks are well defined in [RFC8029] and [RFC6425]. The basic idea for the LSP Ping mechanism is to send an 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 the 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 the 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

A-D: Auto-Discovery

B-MAC: Backbone MAC

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

FEC: Forwarding Equivalent Class

G-ACh: Generic Associated Channel

GAL: G-ACh Label

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

ND: Neighbor Discovery

OAM: Operations, Administration, and Maintenance

P2MP: Point-to-Multipoint

PBB-EVPN: Provider Backbone Bridging EVPN

PE: Provider Edge device

VPWS: Virtual Private Wire Service

4. 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 checks in the 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 EVI under test at an egress PE. This sub-TLV is included in the Echo Request sent by an EVPN/PBB-EVPN PE to a peer PE.

The fields of the EVPN MAC/IP sub-TLV are derived from the MAC/IP Advertisement route defined in Section 7.2 of [RFC7432] and have the format shown in Figure 1. The fields of the EVPN MAC/IP sub-TLV 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 [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 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. 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 described in Section 5.2 of [RFC7623].
- 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 [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 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.

```
 0                   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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Route Distinguisher                       |
|                        (8 octets)                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Ethernet Tag ID                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Ethernet Segment Identifier                   |
|                        (10 octets)                            |
+                         +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         | Must Be Zero  | MAC Addr Len |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 MAC Address                                   |
+                (6 octets)      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         | Must Be Zero  | IP Addr Len  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 IP Address (0, 4 or 16 octets)               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
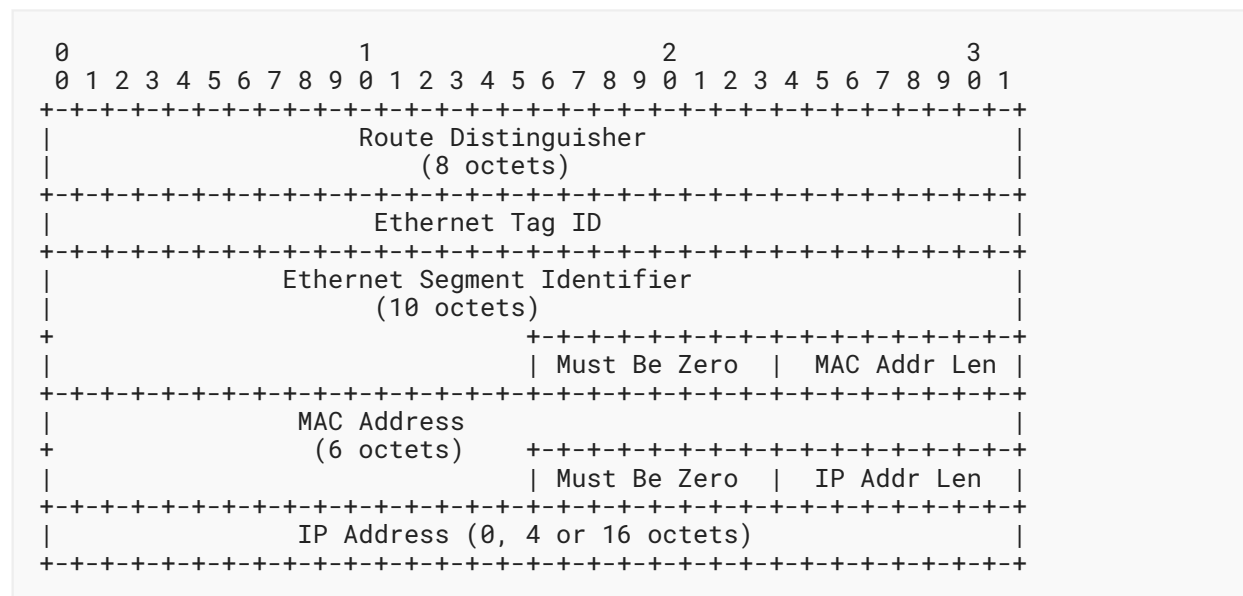
Figure 1: EVPN MAC/IP Sub-TLV Format

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, 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).

When an MPLS Echo Request is sent by an ingress PE, the contents of the 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 three cases above this Echo Request is for. When the egress PE receives the EVPN MAC/IP sub-TLV containing only the MAC address, the egress PE validates the MAC state and forwarding. When the egress PE receives the EVPN 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 (e.g., the egress PE receiving the EVPN MAC/IP sub-TLV containing only the MAC address but with the EVPN label pointing to an IP-VRF) should be considered invalid, and the egress PE should send an Echo Reply with the appropriate Return Code to the ingress PE.

4.2. EVPN Inclusive Multicast Sub-TLV

The fields of the EVPN Inclusive Multicast sub-TLV are based on the EVPN Inclusive Multicast Tag route defined in Section 7.3 of [RFC7432]. This TLV is included in the Echo Request sent to the EVPN peer PE by the originator of the 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 shown in Figure 2. The fields of this sub-TLV should be set according to the following, which 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 the Service Instance Identifier (I-SID) value as per Section 5.3 of [RFC7623].
- The IP Addr Len field specifies the length of the Originating Router's IP Addr field in bits and is set to either 32 for IPv4 addresses or 128 for IPv6 addresses.
- The Originating Router's IP Addr field is set to the IPv4 or IPv6 address of the peer PE.
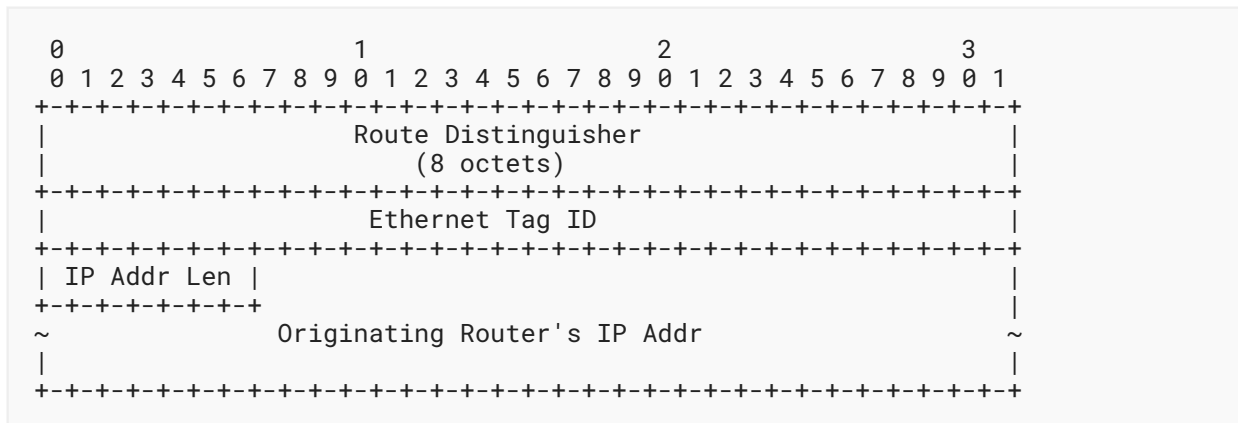
```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Route Distinguisher                       |
|                        (8 octets)                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Ethernet Tag ID                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| IP Addr Len  |                                                |
+-+-+-+-+-+-+-+-+                                                |
~                  Originating Router's IP Addr                 ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: EVPN Inclusive Multicast Sub-TLV Format

BUM traffic can be sent using ingress replication or P2MP P-tree in EVPN and PBB-EVPN networks. When using ingress replication, the Echo Request is sent using a label stack of [Transport label, Inclusive Multicast label] to each egress PE participating in EVPN or PBB-EVPN. The Inclusive Multicast label is the downstream-assigned label announced by the egress PE to which the Echo Request is being sent. The Inclusive Multicast label is the inner label in the MPLS label stack.

When using P2MP P-tree in EVPN or PBB-EVPN, the Echo Request is sent using a P2MP P-tree transport label for the Inclusive P-tree arrangement or using a label stack of [P2MP P-tree Transport label, upstream-assigned EVPN Inclusive Multicast label] for the Aggregate Inclusive P2MP P-tree arrangement as described in Section 6.

In an EVPN network, to emulate traffic coming from a multihomed site, an additional EVPN Ethernet A-D sub-TLV in the Target FEC Stack TLV and an ESI Split Horizon Group MPLS label as the bottom label are also included in the Echo Request packet. When using P2MP P-tree, the ESI Split Horizon Group MPLS label is upstream assigned. Please see Section 6.2.2 for operations using P2MP P-trees.

4.3. EVPN Ethernet Auto-Discovery (A-D) Sub-TLV

The fields in the EVPN Ethernet A-D sub-TLV are based on the EVPN Ethernet A-D route advertisement defined in Section 7.1 of [RFC7432]. The EVPN Ethernet A-D sub-TLV only applies to EVPN.

The EVPN Ethernet A-D sub-TLV has the format shown in Figure 3. The fields of this sub-TLV should be set according to the following, which is consistent with [RFC7432]:

• 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 4.3.2 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 4.3.1.

- 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.
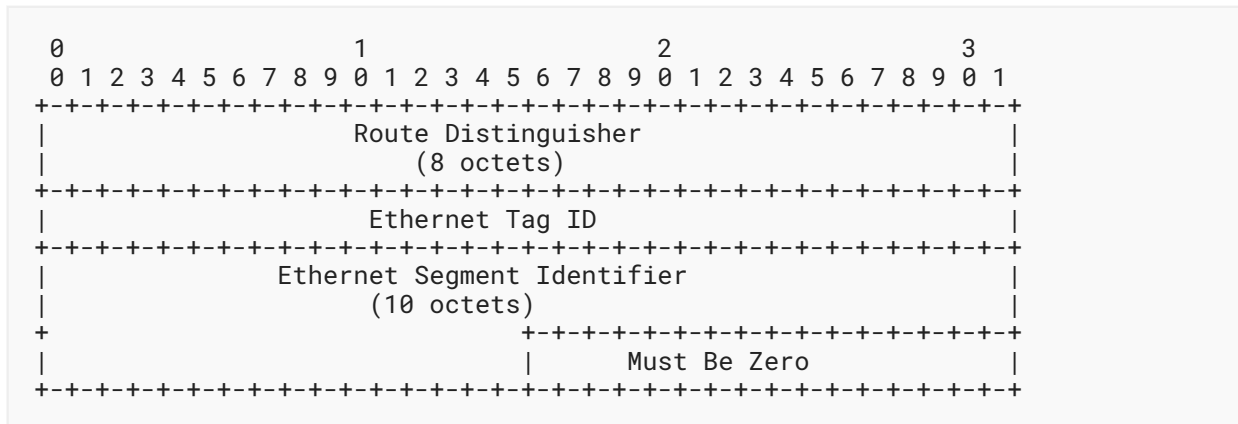
```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Route Distinguisher                       |
|                        (8 octets)                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Ethernet Tag ID                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Ethernet Segment Identifier                     |
|                       (10 octets)                             |
+                         +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         |           Must Be Zero              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: EVPN Ethernet A-D Sub-TLV Format

4.3.1. Ethernet Tag Value

The EVPN Ethernet A-D sub-TLV 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 Echo Request packet is sent and **MAY** contain the EVPN Ethernet A-D sub-TLV to emulate traffic coming from a multihomed site. In this case, the EVPN Ethernet A-D sub-TLV is added in the per-ES context. When an Echo Request packet is sent for the connectivity check for EVPN Aliasing state, the context for the EVPN Ethernet A-D sub-TLV is per-EVI.

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

- For the per-ES context, the Ethernet Tag field in the sub-TLV **MUST** be set to the reserved MAX-ET value [RFC7432].
- For the 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 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 sub-TLV **MUST** be the RD value received for the EVI in the per-ES EVPN A-D route.

4.3.3. EVPN VPWS

LSP Ping can also be used to detect data plane failures for the EVPN VPWS described in [RFC8214]. The Echo Request packet carries the EVPN Ethernet A-D sub-TLV with fields populated from the EVPN Ethernet A-D 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 A-D route announced by the egress PE and the MPLS transport label(s) to reach the egress PE.

The egress PE processes the Echo Request packet and performs checks for the EVPN Ethernet A-D sub-TLV present in the Target FEC Stack TLV as described in Section 4.4 of [RFC8029] and responds according to processing rules in [RFC8029]. The egress PE can identify that the Echo 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 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 <RSC>").

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 only applies to EVPN.

The EVPN IP Prefix sub-TLV has the format shown in Figure 4. The total length (not shown) of this sub-TLV **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 sub-TLV should be set according to the following, which 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 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.
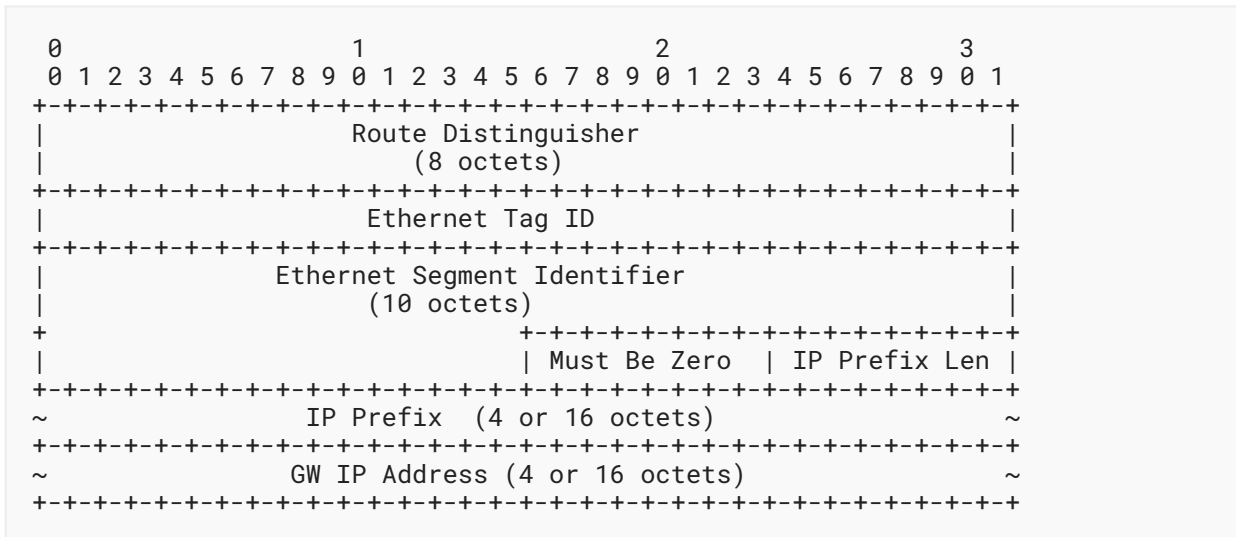- The Must Be Zero field is set to 0. The receiving PE should ignore the Must Be Zero field.

```
 0                   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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Route Distinguisher                       |
|                        (8 octets)                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Ethernet Tag ID                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Ethernet Segment Identifier                   |
|                       (10 octets)                             |
+                          +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          | Must Be Zero  | IP Prefix Len |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                IP Prefix  (4 or 16 octets)                    ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                GW IP Address (4 or 16 octets)                 ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: EVPN IP Prefix Sub-TLV Format

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 GAL [RFC5586], which is the bottommost label. The GAL is followed by a G-ACh header carrying the IPv4(0x0021) or IPv6(0x0057) Channel Type. The code points for IPv4 and IPv6 channels are defined in the "Generic Associated Channel (G-ACh) Parameters" IANA registry.

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 that 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 the EVPN MAC/IP 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 if the packet is an IPv4 or IPv6 OAM packet. The PE1 will process the packet and perform checks for the EVPN MAC/IP sub-TLV present in the Target FEC Stack TLV as described in Section 4.4 of [RFC8029] and respond according to the processing rules in [RFC8029].

```
                      +----------------+
                      |                |
                      |                |
      +----+ AC1  +-----+            +-----+    +----+
      | CE1|------|     |            | PE3 |-----| CE2|
      +----+\     | PE1 |   IP/MPLS  |     |    +----+
           \      +-----+   Network  +-----+
            \        |                  |
         AC2\  +-----+                  |
             \ |     |                  |
              \| PE2 |                  |
               +-----+                  |
                  |                     |
                  +----------------+
                      |                |
                      +----------------+

     <-802.1Q->  <-------PBB over MPLS------>  <-802.1Q->
```
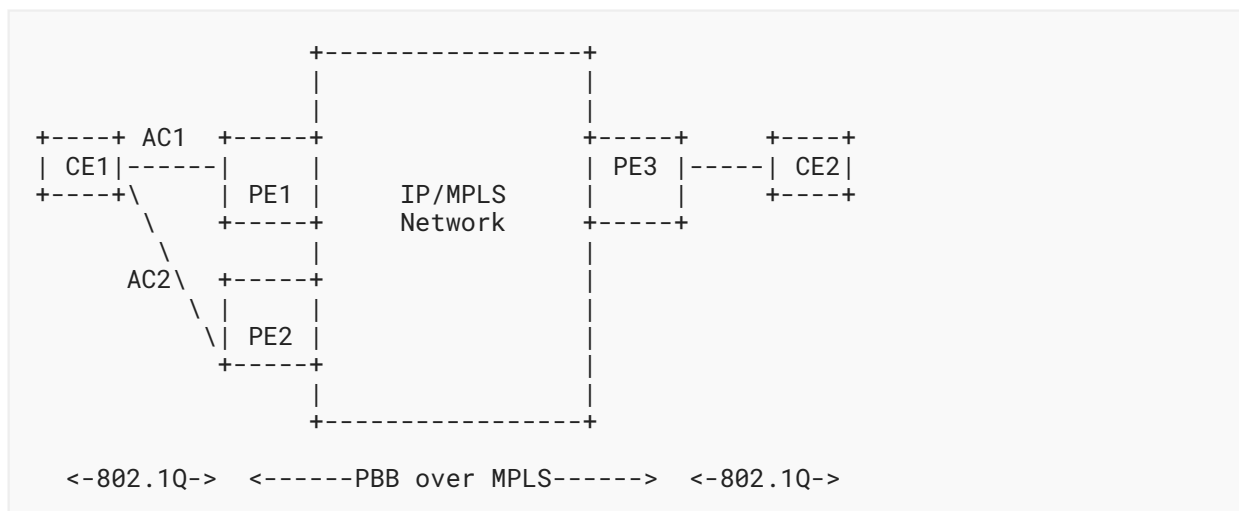
Figure 5: PBB-EVPN Network

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 the EVPN MAC/IP 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 operations 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 networks, an operator can also perform a MAC state test using an 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 A-D 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 the 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 Inclusive 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 if the packet is an IPv4 or IPv6 OAM packet. The packet will have the 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 of [RFC8029] and respond according to the processing rules in [RFC8029]. For the success case, PE1 will reply with Return Code 3 ("Replying router is an egress for the FEC at stack-depth <RSC>").

Similarly, an operator at PE3 may initiate an LSP Ping to PE2 with the Target FEC Stack TLV containing the 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 Inclusive 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 if the packet is an IPv4 or IPv6 OAM packet. The processing on PE2 will be similar to that on PE1 as described above. For the success case, PE2 will reply with Return Code 3 ("Replying router is an egress for the FEC at stack-depth <RSC>") as per [RFC8029].

In an Echo Request packet for EVPN, a combination of an EVPN Ethernet A-D sub-TLV and the associated MPLS Split Horizon label, immediately preceding the GAL in the MPLS label stack, may be used 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 prevent forwarding of 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 the EVPN Ethernet A-D sub-TLV because of Split Horizon filtering, the PE will reply with Return Code 37 (see Section 8) and drop the BUM packets on the ES corresponding to the ESI received in the EVPN Ethernet A-D sub-TLV because of the Split Horizon Group filtering.

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, the 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 the EVPN Inclusive Multicast sub-TLV in the Echo Request packet. The Echo Request packet is sent over P2MP LSP with the {P2MP P-tree Transport label, GAL} MPLS label stack and IP ACH Channel header.

When using an Aggregate Inclusive P-tree arrangement, 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 the 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 Transport label, EVPN upstream-assigned Multicast label, GAL} MPLS label stack and IP ACH Channel header.

The leaf PE(s) of the P2MP P-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 of [RFC8029] and respond according to the processing rules in [RFC8029]. For the success case, the leaf PE will reply with Return Code 3 ("Replying router is an egress for the FEC at stack-depth <RSC>").

In an Echo Request packet for EVPN, a combination of an EVPN Ethernet A-D sub-TLV and the associated MPLS Split Horizon label, immediately preceding the GAL in the MPLS label stack, may be used to emulate traffic coming from a multihomed site. When using P2MP P-tree, the Split Horizon label is upstream assigned and is received by all the leaf PEs of the P2MP P-tree. The Split Horizon label is used by leaf PE(s) attached to the same multihomed site so that packets will not be forwarded 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 the EVPN Ethernet A-D sub-TLV because of Split Horizon filtering, the PE will reply with Return Code 37 (see Section 8) and drop the BUM packets on the ES corresponding to the ESI received in the EVPN Ethernet A-D sub-TLV because of the Split Horizon Group filtering. If the leaf PE does not have the ESI identified in the EVPN Ethernet A-D sub-TLV, the PE **MAY** reply with Return Code 38 (see Section 8), and the BUM packets are forwarded because there is no ES corresponding to the ESI received in the EVPN Ethernet A-D 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 (P2MP Responder Identifier TLV with either the IPv4 Node Address P2MP Responder sub-TLV or the IPv6 Node Address P2MP Responder sub-TLV) can be applied to LSP Ping in EVPN and PBB-EVPN when using P2MP P-trees for BUM traffic.

6.3. EVPN Aliasing Data Plane Connectivity Check

Assume PE1 announced an Ethernet A-D per-EVI route with the ESI set to CE1 system ID and MPLS label 19001. Additionally, assume PE2 announced an Ethernet A-D 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 A-D route on PE1, the operator initiates an LSP Ping request with the Target FEC Stack TLV containing the EVPN Ethernet A-D sub-TLV in the Echo Request packet. The Echo Request packet is sent with the {Transport label(s) to reach PE1, EVPN Ethernet A-D 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 A-D sub-TLV present in the Target FEC Stack TLV as described in Section 4.4 of [RFC8029] and respond according to the processing rules in [RFC8029].

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 the 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 of [RFC8029] and respond according to the processing rules in [RFC8029].

7. Security Considerations

This document does not introduce any new security considerations beyond those that apply in [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:

- A Denial-of-Service (DoS) attack by sending MPLS Echo Requests/Replies to Label Switching Routers (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 through 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 document; thus, this document doesn't require additional security considerations beyond the ones described in [RFC8029].

This document 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 types to be included in the Target FEC Stack TLV (TLV types 1, 16, and 21) [RFC9041] in Echo Request and Echo Reply messages in EVPN and PBB-EVPN networks.

IANA has assigned the following values from the "Standards Action" (0-16383) range in the "Sub-TLVs for TLV Types 1, 16, and 21" subregistry within the "TLVs" registry of the "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" name space.

Sub-Type	Sub-TLV Name	Reference
42	EVPN MAC/IP	RFC 9489
43	EVPN Inclusive Multicast	RFC 9489
44	EVPN Ethernet Auto-Discovery	RFC 9489
45	EVPN IP Prefix	RFC 9489

Table 1

8.2. New Return Codes

[RFC8029] defines values for the Return Code field of Echo Reply messages. This document defines two new Return Codes that **SHOULD** be included in the Echo Reply message by a PE in response to an Echo Request message in EVPN and PBB-EVPN networks.

IANA has assigned the following values in the "Return Codes" registry of the "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" name space.

Value	Meaning	Reference
37	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 the EVPN Ethernet Auto-Discovery sub-TLV because of the Split Horizon Group filtering.	RFC 9489
38	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 the EVPN Ethernet Auto-Discovery sub-TLV.	RFC 9489

Table 2

9. 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>.

Acknowledgments

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