

BESS Workgroup
Internet-Draft
Intended status: Standards Track
Expires: May 25, 2017

P. Jain
Cisco Systems, Inc.
S. Boutros
VmWare, Inc.
S. Salam
Cisco Systems, Inc.
November 21, 2016

LSP-Ping Mechanisms for EVPN and PBB-EVPN
draft-jain-bess-evpn-lsp-ping-04

Abstract

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 EVPN and 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 <http://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 May 25, 2017.

Copyright Notice

Copyright (c) 2016 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 (<http://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 Simplified BSD License text as described in [Section 4](#).e of

the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	2
2.	Specification of Requirements	3
3.	Terminology	3
4.	Proposed Target FEC Stack Sub-TLVs	3
4.1.	EVPN MAC Sub-TLV	4
4.2.	EVPN Inclusive Multicast Sub-TLV	4
4.3.	EVPN Auto-Discovery Sub-TLV	6
4.4.	EVPN IP Prefix Sub-TLV	6
5.	Encapsulation of OAM Ping Packets	7
6.	Operations	7
6.1.	Unicast Data-plane connectivity checks	7
6.2.	Inclusive Multicast Data-plane Connectivity Checks	9
6.2.1.	Ingress Replication	9
6.2.2.	Using P2MP P-tree	10
6.2.3.	Controlling Echo Responses when using P2MP P-tree	11
6.3.	EVPN Aliasing Data-plane connectivity check	11
6.4.	EVPN IP Prefix (RT-5) Data-plane connectivity check	11
7.	Security Considerations	12
8.	IANA Considerations	12
9.	Acknowledgments	12
10.	References	13
10.1.	Normative References	13
10.2.	Informative References	13
	Authors' Addresses	14

[1.](#) Introduction

[RFC7432] describes MPLS based Ethernet VPN (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, 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 multi-protocol BGP. EVPN enables multi-homing of CE(s) connected to multiple PEs and load balancing of traffic to and from multi-homed CE(s).

[RFC7623] describes the use of Provider Backbone Bridging [802.1ah] with EVPN. PBB-EVPN maintains the 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

[[RFC4379](#)][[RFC6425](#)]. This document defines procedures to detect data-plane failures using LSP Ping in MPLS networks deploying EVPN and PBB-EVPN. This draft defines 3 new Sub-TLVs for Target FEC Stack TLV with the purpose of identifying the FEC on the Peer PE.

2. Specification of Requirements

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [[RFC2119](#)].

3. Terminology

B-MAC: Backbone MAC Address

CE: Customer Edge Device

C-MAC: Customer MAC Address

DF: Designated Forwarder

ESI: Ethernet Segment Identifier

EVI: EVPN Instance Identifier that globally identifies the EVPN Instance

EVPN: Ethernet Virtual Private Network

MPLS-OAM: MPLS Operations, Administration and Maintenance

P2MP: Point-to-Multipoint

PBB: Provider Backbone Bridge

PE: Provider Edge Device

4. Proposed Target FEC Stack Sub-TLVs

This document introduces four new Target FEC Stack sub-TLVs that are included in the LSP-Ping Echo Request packet sent for detecting faults in data-plane connectivity in EVPN and PBB-EVPN networks. These Target FEC Stack sub-TLVs are described next.

All these TLVs contain 8 bytes EVI value which identifies the EVPN instance globally.

4.1. EVPN MAC Sub-TLV

The EVPN MAC sub-TLV is used to identify the MAC for an EVI under test at a peer PE.

The EVPN MAC sub-TLV fields are derived from the MAC advertisement route defined in [RFC4379] and has the format as shown in Figure 1. This TLV is included in the Echo Request sent to the Peer PE by the PE that is the originator of the request.

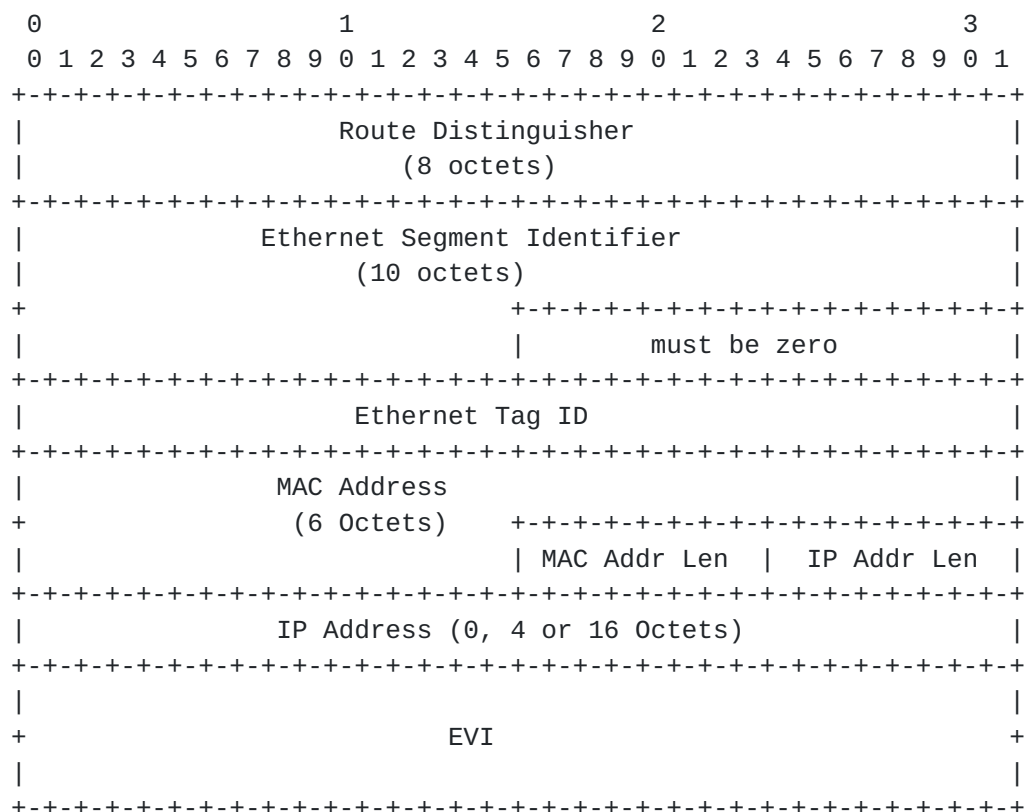


Figure 1: EVPN MAC sub-TLV format

The LSP Ping echo request is sent using the EVPN MPLS label(s) associated with the MAC route announced by a remote PE and the MPLS transport label(s) to reach the remote PE.

4.2. EVPN Inclusive Multicast Sub-TLV

The EVPN Inclusive Multicast sub-TLV fields are based on the EVPN Inclusive Multicast route defined in [RFC7432].

The EVPN Inclusive Multicast sub-TLV has the format as shown in Figure 2. 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.

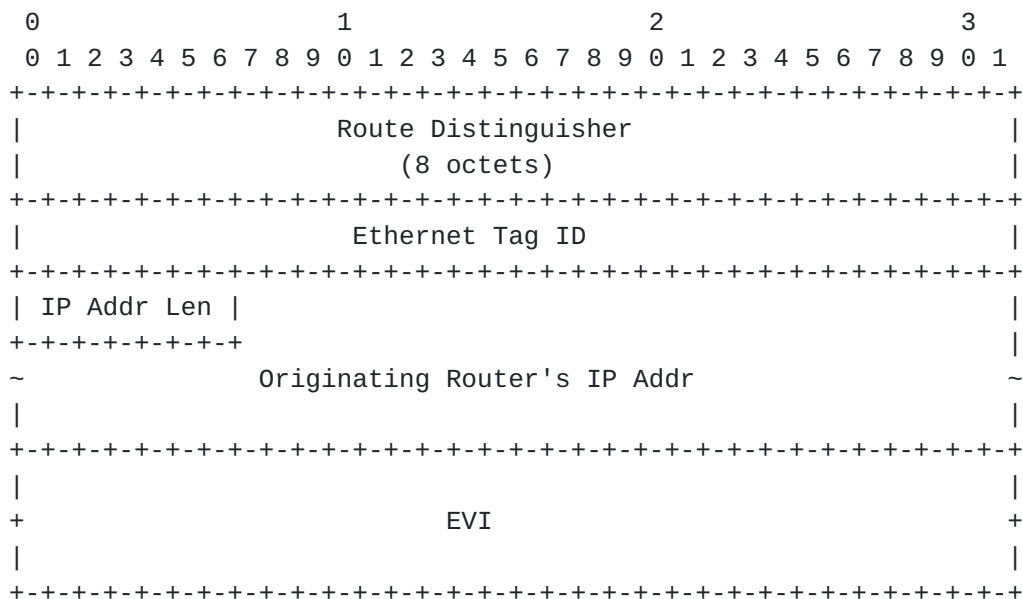


Figure 2: EVPN Inclusive Multicast sub-TLV format

Broadcast, multicast and unknown unicast traffic can be sent using ingress replication or P2MP P-tree in EVPN and PBB-EVPN network. In case of ingress replication, the Echo Request is sent using a label stack of [Transport label, Inclusive Multicast label] to each remote PE participating in EVPN or PBB-EVPN. The inclusive multicast label is the downstream assigned label announced by the remote 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 P2MP P-tree transport label for inclusive P-tree arrangement or using a label stack of [P2MP P-tree transport label, upstream assigned EVPN Inclusive Multicast label] for aggregate inclusive P2MP P-tree arrangement as described in [Section 6](#).

In case of EVPN, an additional, EVPN Auto-Discovery sub-TLV and ESI MPLS label as the bottom label, may also be included in the Echo Request as is described in [Section 6](#).

4.3. EVPN Auto-Discovery Sub-TLV

The EVPN Auto-Discovery (AD) sub-TLV fields are based on the Ethernet AD route advertisement defined in [RFC7432]. EVPN AD sub-TLV applies to only EVPN.

The EVPN AD sub-TLV has the format shown in Figure 3.

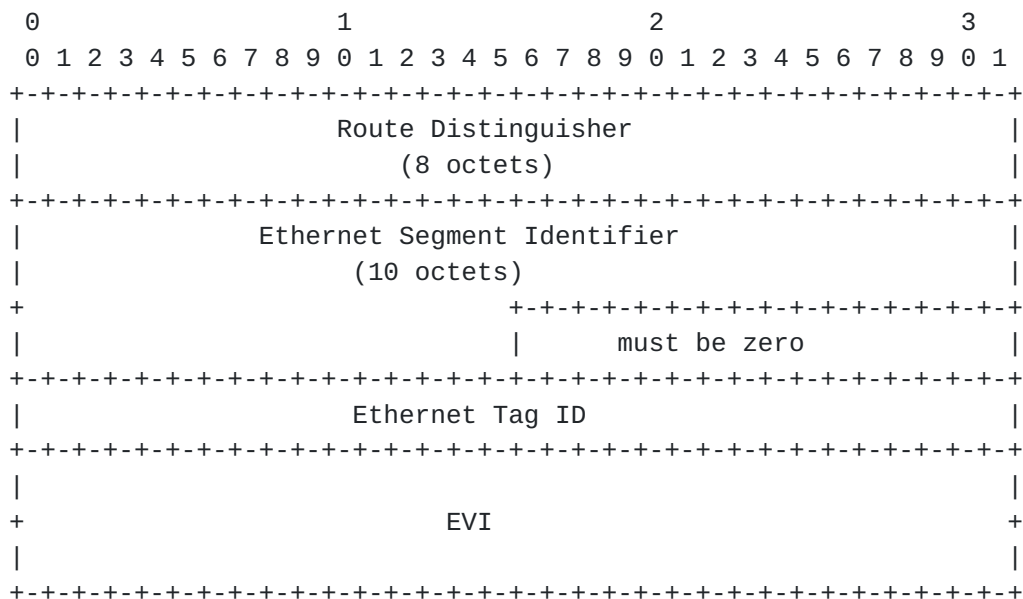


Figure 3: EVPN Auto-Discovery sub-TLV format

4.4. EVPN IP Prefix Sub-TLV

The EVPN IP Prefix sub-TLV is used to identify 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 [I-D.ietf-bess-evpn-prefix-advertisement] and has the format as shown in Figure 4. This TLV is included in the Echo Request sent to the Peer PE by the PE that is the originator of the request.

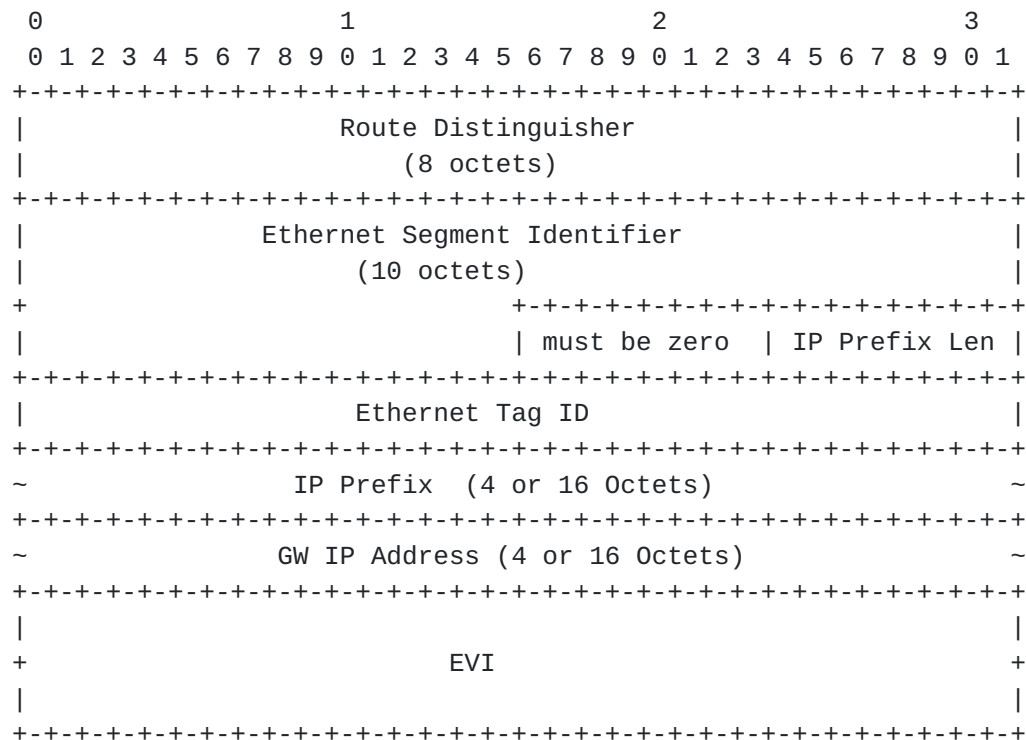


Figure 4: EVPN IP Prefix sub-TLV format

The LSP Ping echo request is sent using the EVPN MPLS label(s) associated with the IP Prefix route announced by a remote PE and the MPLS transport label(s) to reach the remote PE.

5. Encapsulation of OAM Ping Packets

The LSP Ping Echo request IPv4/UDP packets will be encapsulated with the Transport and EVPN Label(s) followed by the Generic Associated Channel Label (GAL) [[RFC6426](#)]. The GAL label will be followed by the Associated Channel Header (ACH) with the Pseudowire Associated Channel Type 16 bit value in the ACH set to IPv4 indicating that the carried packet is an IPv4 packet respectively.

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 1.1.1.1:00 and B-MAC 00aa.00bb.00cc and with MPLS label 16001 for EVI 10. Similarly PE2 announced a MAC route with RD 2.2.2.2:00 and B-MAC 00aa.00bb.00cc and with MPLS label 16002.

On PE3, when a operator performs a connectivity check for the B-MAC address 00aa.00bb.00cc 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 label and the IP ACH Channel header to determine that the packet is IPv4 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 \[RFC4379\]](#) and respond according to [\[RFC4379\]](#) processing rules.

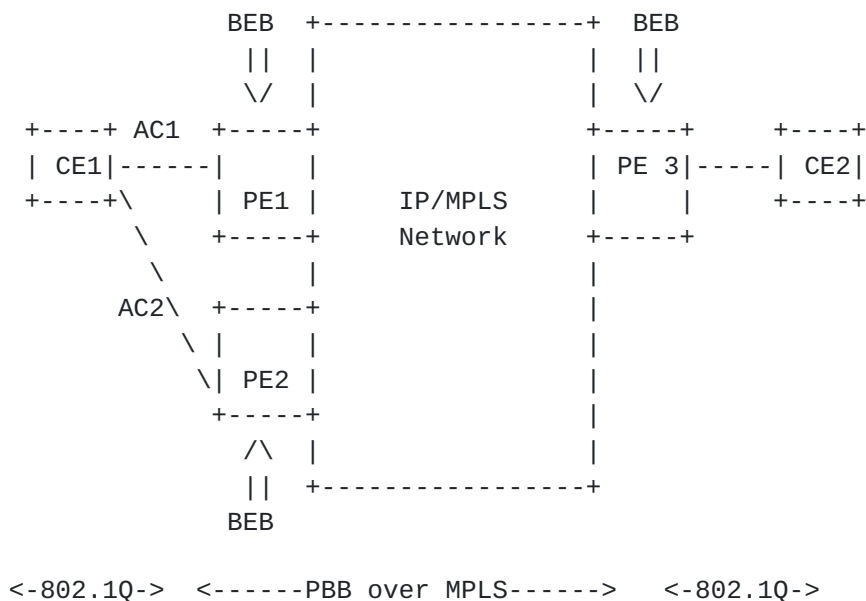


Figure 5: PBB EVPN network

Similarly, on PE3, when an operator performs a connectivity check for the B-MAC address 00aa.00bb.00cc 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 E-VPN, are similar to those described above for PBB-EVPN except that the checks are for C-MAC addresses instead of B-MAC addresses.

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 1.1.1.1:00, 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 2.2.2.2:00, 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 label and the IP ACH Channel header to determine that the packet is IPv4 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 \[RFC4379\]](#) and respond according to [\[RFC4379\]](#) processing rules.

Operator at PE3, may similarly also 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 label and the IP ACH Channel header to determine that the packet is IPv4 OAM Packet. Since PE2 is not the DF for ISID 10 for the port corresponding to the ESI value in the Inclusive Multicast sub-TLV in the Echo Request, PE2 will reply with special code indicating that FEC exists on the router and the behavior is to drop the packet because of not DF as described in [Section 8](#).

In case of EVPN, in the Echo Request packet, an Ethernet AD sub-TLV and the associated MPLS Split Horizon Label above the GAL label in the MPLS label stack, may be added to emulate traffic coming from a MH site, this label is used by leaf PE(s) attached to the same MH site not to forward packets back to the MH site. If the behavior on

a leaf PE is to drop the packet because of Split Horizon filtering, the PE2 will reply with special code indicating that FEC exists on the router and the behavior is to drop the packet because of Split Horizon 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, in that case 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 \[RFC4379\]](#) and respond according to [\[RFC4379\]](#) processing rules. A PE that is not the DF for the EVI on the ESI in the Inclusive Multicast sub-TLV, will reply with a special code indicating that FEC exists on the router and the behavior is to drop the packet because of not DF as described in [Section 8](#).

In case of EVPN, in the Echo Request packet, an Ethernet AD sub-TLV and the associated MPLS Split Horizon Label above the GAL Label in MPLS label stack, may be added to emulate traffic coming from a MH

site, this label is used by leaf PE(s) attached to the same MH site not to forward packets back to the MH site. If the behavior on a leaf PE is to drop the packet because of Split Horizon filtering, the PE2 will reply with special code indicating that FEC exists on the router and the behavior is to drop the packet because of Split Horizon Filtering as described in [Section 8](#).

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 PBB EVPN and 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 Auto discovery Route with the ESI set to CE1 system ID and MPLS label 19001, and PE2 an Ethernet Auto discovery Route with the ESI set to CE1 system ID and MPLS label 19002.

When an operator performs at PE3 a connectivity check for the aliasing aspect of the Ethernet AD route to 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 \[RFC4379\]](#) and respond according to [[RFC4379](#)] processing rules.

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

Assume PE1 announced an IP Prefix Route (RT-5) with an IP prefix reachable behind CE1 and MPLS label 19001. 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 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 IP Prefix sub-TLV present in the Target FEC Stack

TLV as described in [Section 4.4 in \[RFC4379\]](#) and respond according to [\[RFC4379\]](#) 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 [\[RFC6425\]](#).

8. IANA Considerations

This document defines 3 new sub-TLV type to be included in Target FEC Stack TLV (TLV Type 1) [\[RFC4379\]](#) in LSP Ping.

IANA is requested to assign a sub-TLV type value to the following sub-TLV from the "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Parameters - TLVs" registry, "TLVs and sub- TLVs" sub-registry:

- o EVPN MAC route sub-TLV
- o EVPN Inclusive Multicast route sub-TLV
- o EVPN Auto-Discovery Route sub-TLV
- o EVPN IP Prefix Route sub-TLV

Proposed new Return Codes

[\[RFC4379\]](#) 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 LSP Ping Echo Request message:

1. The FEC exists on the PE and the behavior is to drop the packet because of not DF.
2. The FEC exists on the PE and the behavior is to drop the packet because of Split Horizon Filtering.

9. Acknowledgments

The authors would like to thank Patrice Brissette and Weiguo Hao for their comments.

10. References

10.1. Normative References

- [I-D.ietf-bess-evpn-prefix-advertisement]
Rabadan, J., Henderickx, W., Palislaamovic, S., Isaac, A., Drake, J., Lin, W., and A. Sajassi, "IP Prefix Advertisement in EVPN", [draft-ietf-bess-evpn-prefix-advertisement-03](#) (work in progress), September 2016.
- [RFC4379] Kompella, K. and G. Swallow, "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", [RFC 4379](#), DOI 10.17487/RFC4379, February 2006, <<http://www.rfc-editor.org/info/rfc4379>>.
- [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, <<http://www.rfc-editor.org/info/rfc6425>>.
- [RFC6426] Gray, E., Bahadur, N., Boutros, S., and R. Aggarwal, "MPLS On-Demand Connectivity Verification and Route Tracing", [RFC 6426](#), DOI 10.17487/RFC6426, November 2011, <<http://www.rfc-editor.org/info/rfc6426>>.
- [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, <<http://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, <<http://www.rfc-editor.org/info/rfc7623>>.

10.2. Informative References

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

- [RFC4875] Aggarwal, R., Ed., Papadimitriou, D., Ed., and S. Yasukawa, Ed., "Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs)", [RFC 4875](#), DOI 10.17487/RFC4875, May 2007, <<http://www.rfc-editor.org/info/rfc4875>>.
- [RFC5085] Nadeau, T., Ed. and C. Pignataro, Ed., "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires", [RFC 5085](#), DOI 10.17487/RFC5085, December 2007, <<http://www.rfc-editor.org/info/rfc5085>>.
- [RFC6338] Giralt, V. and R. McDuff, "Definition of a Uniform Resource Name (URN) Namespace for the Schema for Academia (SCHAC)", [RFC 6338](#), DOI 10.17487/RFC6338, August 2011, <<http://www.rfc-editor.org/info/rfc6338>>.

Authors' Addresses

Parag Jain
Cisco Systems, Inc.
2000 Innovation Drive
Kanata, ON K2K-3E8
Canada

Email: paragj@cisco.com

Sami Boutros
VmWare, Inc.
USA

Email: sboutros@vmware.com

Samer Salam
Cisco Systems, Inc.
595 Burrard Street, Suite 2123
Vancouver, BC V7X 1J1
Canada

Email: ssalam@cisco.com

