Network Working Group
Internet Draft
Intended status: Standards Track
Expires: January 30, 2013

Parag Jain, Ed.
Sami Boutros
Samer Salam
Cisco Systems

July 9, 2012

LSP-Ping Mechanisms for E-VPN and PBB-EVPN
draft-jain-l2vpn-evpn-lsp-ping-00.txt

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 E-VPN 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), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

Table of Contents

1. Introduction

[EVPN] describes MPLS based Ethernet VPN (E-VPN) technology. An E-VPN comprises CE(s) connected to PE(s). The PEs provide layer 2 E-VPN among the CE(s) over the MPLS core infrastructure. In E-VPN 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. E-VPN enables multi-homing of CE(s) connected to multiple PEs and load balancing of traffic to and from multi-homed CE(s).

[PBBEVPN] describes the use of Provider Backbone Bridging [802.1ah] with E-VPN. 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 E-VPN 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. Conventions used in this document

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 RFC-2119 [RFC2119].

The term FEC-Type is used to refer to a tuple consisting of <FEC Element Type, Address Family>.

3. Terminology

B-MAC: Backbone MAC Address

CE: Customer Edge Device

C-MAC: Customer MAC Address

DF: Designated Forwarder

ESI: Ethernet Segment Identifier

EVI: E-VPN Instance

E-VPN: Ethernet Virtual Private Network

MPLS-OAM: MPLS Operations, Administration and Maintenance

P2MP: Point-to-Multipoint

PBB: Provider Backbone Bridge

PE: Provider Edge Device

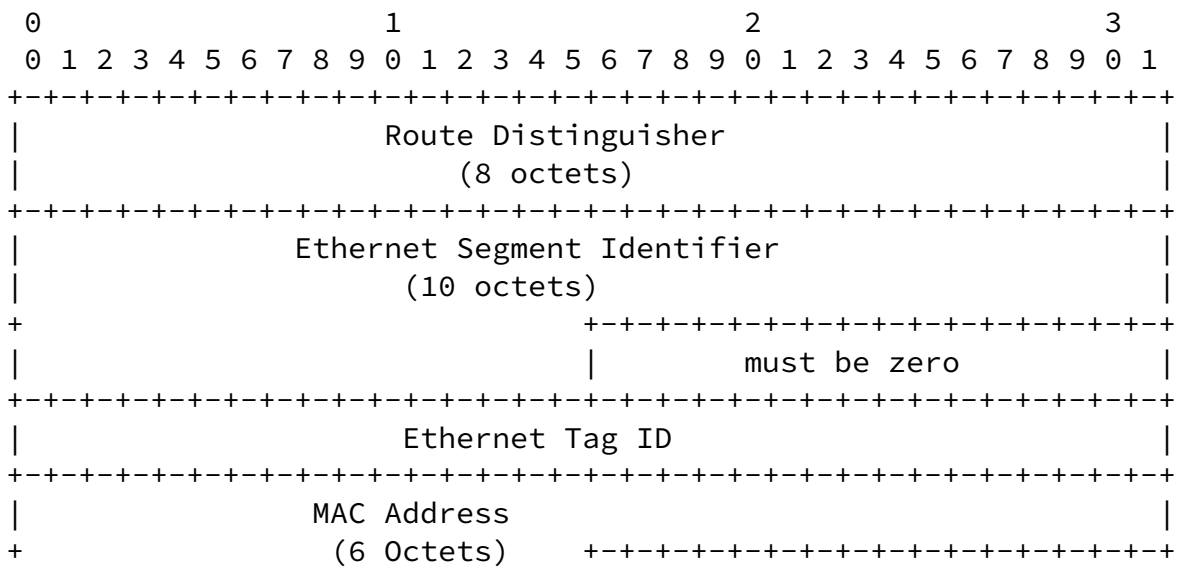
4. Proposed Target FEC Stack Sub-TLVs

This document introduces three 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 E-VPN and PBB-EVPN networks. These Target FEC Stack sub-TLVs are described next.

4.1. E-VPN MAC Sub-TLV

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

The E-VPN MAC sub-TLV fields are derived from the MAC advertisement route defined in [EVPN] 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.



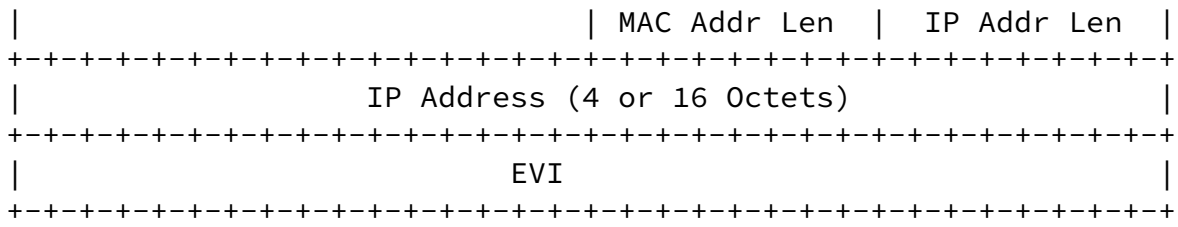


Figure 1: E-VPN MAC sub-TLV format

The LSP Ping echo request is sent using the E-VPN 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. E-VPN Inclusive Multicast Sub-TLV

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

The E-VPN Inclusive Multicast sub-TLV has the format as shown in Figure 2. This TLV is included in the echo request sent to the E-VPN peer PE by the originator of request to verify the multicast connectivity state on the peer PE(s) in E-VPN and PBB-EVPN.

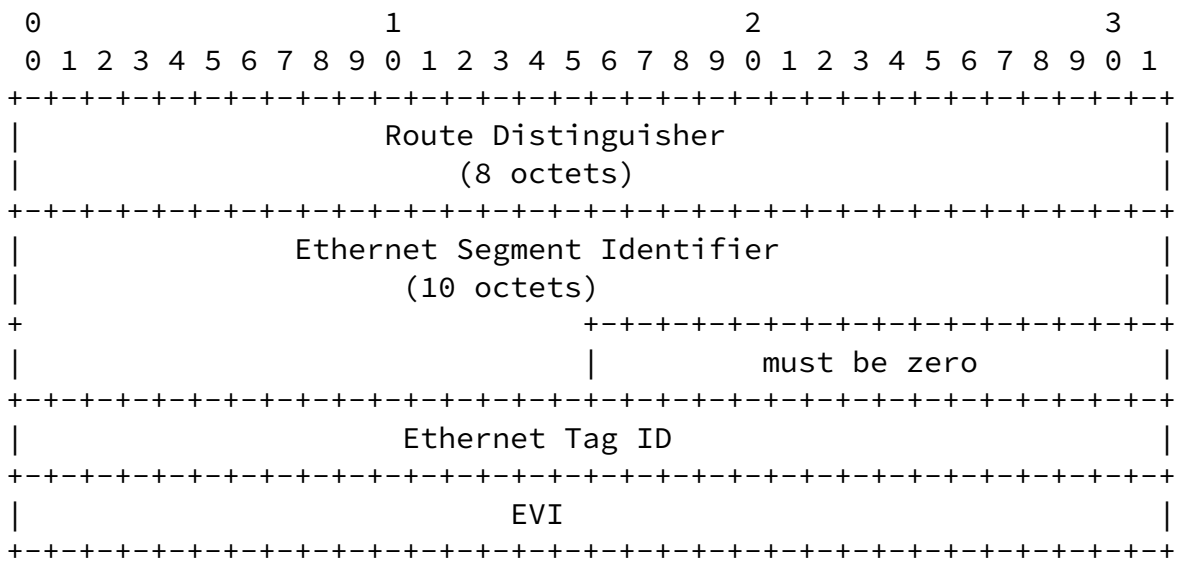


Figure 2: E-VPN Inclusive Multicast sub-TLV format

Broadcast, multicast and unknown unicast traffic can be sent using ingress replication or P2MP P-tree in E-VPN 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 E-VPN 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 E-VPN 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,

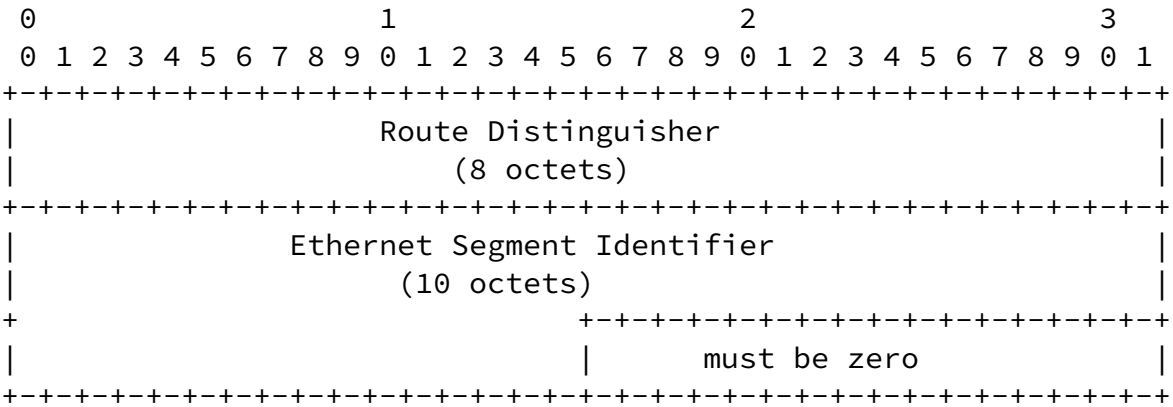
upstream assigned EVPN Inclusive Multicast label> for aggregate inclusive P2MP P-tree arrangement as described in Section 5.

In case of E-VPN, an additional, E-VPN 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 5.

4.3. E-VPN Auto-Discovery Sub-TLV

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

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



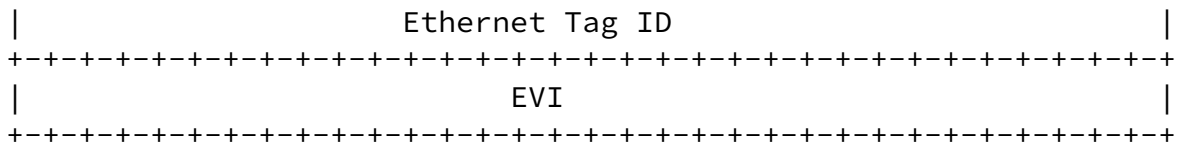


Figure 3: E-VPN Auto-Discovery sub-TLV format

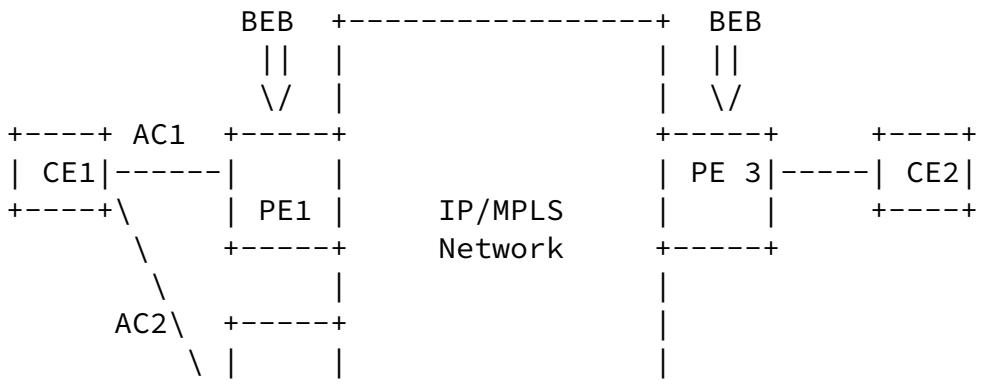
5. Operations

5.1. Unicast Data-plane connectivity checks

Figure 4 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 E-VPN MAC sub-TLV in the Echo Request packet. The Echo Request packet is sent with the {Transport Label(s) to reach PE1 + E-VPN Label = 16001} MPLS label stack. Once the echo request packet reaches PE1, it will process the packet and perform checks for the E-VPN 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.



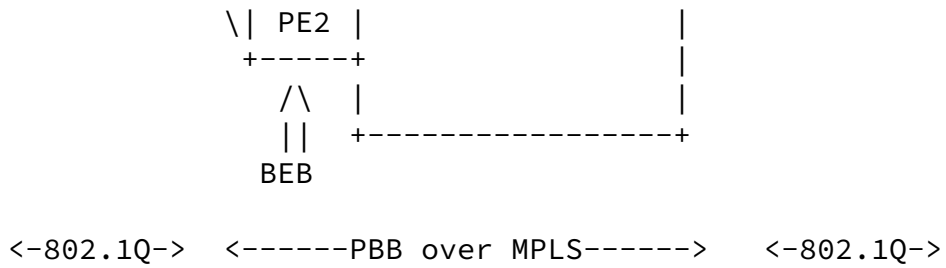


Figure 4: PBB EVPN network

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 E-VPN MAC sub-TLV in the echo request packet. The echo request packet is sent with the {MPLS transport Label(s) to reach PE2 + E-VPN Label = 16002} MPLS label stack.

LSP Ping operation for unicast data-plane connectivity checks in E-VPN, are similar to as described above for PBB-EVPN except that the checks are for C-MACs and not for B-MACs.

5.2. Inclusive Multicast Data-plane Connectivity Checks

5.2.1. 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 E-VPN Inclusive Multicast sub-TLV in the Echo Request packet. The Echo Request packet is sent with the {Transport Label(s) to reach PE1 + E-VPN Incl. Multicast Label = 17001} MPLS label stack. Once the packet reaches PE1, the packet will have E-VPN Inclusive multicast label. PE1 will process the packet and perform checks for the E-VPN 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 E-VPN Inclusive Multicast sub-TLV in the echo request packet. The echo request packet is sent with the {transport Label(s) to reach PE2 + E-VPN Incl. Multicast Label = 17002} MPLS label stack. 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 7.

In case of E-VPN, in the Echo Request packet, an Ethernet AD sub-TLV and the associated MPLS Split Horizon Label at the bottom of 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 7.

5.2.2. Using P2MP P-tree

Both inclusive P-Tree and aggregate inclusive P-tree can be used in E-VPN 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 E-VPN Inclusive Multicast sub-TLV in the echo request packet. The echo request packet is sent with the {P2MP P-tree label} MPLS label stack.

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 E-VPN Inclusive Multicast sub-TLV in the echo request packet. The echo request packet is sent with the {P2MP P-tree label + E-VPN Upstream assigned Multicast Label} MPLS label stack.

The Leaf PE(s) of the p2mp tree will process the packet and perform checks for the E-VPN 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 7.

In case of E-VPN, in the Echo Request packet, an Ethernet AD sub-TLV and the associated MPLS Split Horizon Label at the bottom of 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 7.

5.2.3. 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 E-VPN when using P2MP P-trees for broadcast, multicast and unknown unicast traffic.

5.3. E-VPN 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 E-VPN Ethernet AD sub-TLV in the echo request packet. The echo request packet is sent with the {Transport label(s) to reach PE1 + E-VPN Ethernet AD Label 19001} MPLS label stack.

When PE1 receives the packet it will process the packet and perform checks for the E-VPN 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. Security Considerations

The proposal introduced in this document does not introduce any new security considerations beyond that already apply to [EVPN], [PBBE VPN] and [RFC6425].

7. 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 E-VPN MAC route sub-TLV.

- o E-VPN Inclusive Multicast route sub-TLV

- o E-VPN Auto-Discovery 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.

8. References

8.1. Normative References

[EVPN] Aggarwal et al., "BGP MPLS Based Ethernet VPN", draft-ietf-l2vpn-evpn-00.txt, work in progress, February 2012.

[PBBEVPN] Sajassi et al., "PBB E-VPN", draft-ietf-l2vpn-pbb-evpn-03.txt, work in progress, March 2012.

[RFC4379] K. Kompella, G. Swallow, "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", RFC 4379, February 2006.

[RFC6425] Saxena, S et al, Detecting Data Plane Failures in Point-to-Multipoint Multiprotocol Label Switching (MPLS) - Extensions to LSP. RFC 6425, November 2011.

8.2. Informative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC2119, March 1997.

[RFC5085] T. Nadeau, et. al, "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires ", RFC 5085, December 2007.

[RFC6388] Minei, I., Kompella, K., Wijnands, I., and Thomas, B., "LDP Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths, RFC 6388, November 2011.

[RFC4875] Aggarwal, R., Papadimitriou, D., and Yasukawa, S., "Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs)", RFC 4875, May 2007.

9. Acknowledgments

The authors would like to thank Patrice Brissette for his valuable input and comments.

This document was prepared using 2-Word-v2.0.template.dot.

Authors' Addresses

Parag Jain
Cisco Systems, Inc.,
2000 Innovation Drive,
Kanata, ON K2K3E8, Canada.
E-mail: paragj@cisco.com

Sami Boutros
Cisco Systems, Inc.
3750 Cisco Way,
San Jose, CA 95134, USA.
E-mail: sboutros@cisco.com

Samer Salam
Cisco Systems, Inc.
595 Burrard Street, Suite 2123,
Vancouver, BC V7X 1J1, Canada.
E-mail: ssalam@cisco.com