Network Working Group
Internet Draft
Category: Standards Track
Expiration Date: May 2009

George Swallow
Cisco Systems, Inc.

Vanson Lim
Cisco Systems, Inc.

November 3, 2008

Proxy LSP Ping

draft-ietf-mpls-remote-lsp-ping-03.txt

Status of this Memo

Abstract

This document defines a means of remotely initiating Multiprocal Label Switched Protocol Pings on Label Switched Paths. A proxy ping request is sent to any Label Switching Routers along a Label Switched Path. The primary motivations for this facility are first to limit the number of messages and related processing when using LSP Ping in large Point-to-Multipoint LSPs, and second to enable leaf to root tracing.

Contents

1. Introduction

It is anticipated that very large Point-to-Multipoint (P2MP) Label Switched Paths (LSPs) will exist. Further it is anticipated that many of the applications for P2MP tunnels will require OAM that is both rigorous and scalable.

Suppose one wishes to trace a P2MP LSP to localize a fault which is affecting one egress or a set of egresses. Suppose one follows the normal procedure for tracing - namely repeatedly pinging from the root, incrementing the TTL by one after each three or so pings. Such a procedure has the potential for producing a large amount of processing at the P2MP-LSP midpoints and egresses. It also could produce an unwieldy number of replies back to the root.

One alternative would be to begin sending pings from points at or near the affected egress(es) and working backwards toward the root. The TTL could be held constant as say two, limiting the the number of responses to the number of next-next-hops of the point where a ping is initiated.

This document defines protocol extensions to MPLS ping [RFC4379] to allow a third party to remotely cause an MPLS echo request message to be sent down a Label Switched Path (LSP) or part of an LSP. The procedure described in the paragraphs above does require that the initiator know the previous-hop node to the one which was pinged on the prior iteration. This information is readily available in [P2MP-TE]. This also document provides a means for obtaining this information for [mLDP].

While the motivaton for this document came from multicast scaling concerns, its applicability may be wider. However other uses of this facility are beyond the scope of this document. In particular, the procedures defined in this document only allow testing of a FEC stack consisting of a single FEC. It also does not allow the initiator to specify the label assigned to that FEC, nor does it allow the initiator to cause any additional labels to be added to the label stack of the actual MPLS echo request message. Further the discussion is cauched in terms of multipoint LSPs.

1.1. Conventions

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 [KeyWords].

The term "Must Be Zero" (MBZ) is used in object descriptions for reserved fields. These fields MUST be set to zero when sent and ignored on receipt.

Based on context the terms leaf and egress are used interchangeably. Egress is used where consistency with [RFC4379] was deemed appropriate. Receiver is used in the context of receiving protocol messages.

[Note (to be removed after assignments occur): <tba> = to be assigned by IANA]

2. Proxy Ping Overview

This document defines a protocol interaction between a first node and a node which is part of an LSP to allow the first node to remotely initiate a an LSP ping for the LSP. Two new LSP Ping messages are defined for remote pinging, the MPLS proxy ping request and the MPLS proxy ping reply.

A remote ping operation on a P2MP LSP involves at least three LSRs; in some scenarios none of these are the ingress (root) or an egress (leaf) of the LSP.

We refer to these nodes with the following terms:

Initiator - the node which initiates the ping operation by sending an MPLS proxy ping request message

Proxy LSR - the node which is the destination of the MPLS proxy request message and potential initiator of the MPLS echo request

Receiver(s) - the receivers of the MPLS echo request messages

The initiator formats an MPLS proxy ping request message and sends it to the proxy LSR, a node it believes to be on the path of the LSP. This message specifies the MPLS echo request to be sent inband of the LSP. It may also request the proxy LSR to acknowledge the receipt of the proxy ping request message and/or respond with the address of the previous hop, i.e. the LSR upstream of it on this LSP.

The proxy LSR validates that it has a label mapping for the specified FEC and that it is authorized to send the specified MPLS echo request on behalf of the initiator. Depending on the Reply Mode carried in the header of the proxy ping request message and the above results an MPLS remote echo reply message might be sent back to the initiator. This message may also communicate the address of the previous hop.

If the proxy LSR has a label mapping for the FEC and and all authorization checks have passed, the proxy LSR formats an MPLS echo request. If the source address of the IP packet is not the initiator, it includes a Reply-to Address object containing the initiator's address. It then sends it inband of the LSP.

The receivers process the MPLS echo request as normal, sending their MPLS echo replies back to the initiator.

3. Proxy MPLS Echo Request / Reply Pprocedures

3.1. Procedures for the initiator

The initiator creates an MPLS proxy ping request message.

The message MUST contain a Target FEC Stack that describes the FEC being tested.

The message MUST contain a Proxy Echo Parameters object. In that object, the address type is set to either IPv4 or IPv6. The Destination IP Address is set to the value to be used in the MPLS echo request packet. If the Address Type is IPv4, an address from the range 127/8. If the Address Type is IPv6, an address from the range 0:0:0:0:0:FFFF:7F00:0/104.

The Reply mode and Global Flags of the Proxy Echo Parameters object are set to the values to be used in the MPLS echo request message header. The Source UDP Port is set to the value to be used in the MPLS echo request packet. The TTL is set to the value to be used in the outgoing MPLS label stack. See section 5.2.2.2 for further details.

If the previous hop address and/or a downstream mapping object from the proxy LSR is needed the corresponding flag in the Flags field of the Proxy Echo Parameters object is set.

A list of Next Hop IP Addresses MAY be included to limit the next hops towards which the MPLS echo request message will be sent. These are encoded as Next Hop sub-objects and included in the Proxy Echo

Parameters object.

Any of following objects MAY be included; these objects will be copied into the MPLS echo request messages:

- Pad
- Vendor Enterprise Number
- Reply TOS Byte
- P2MP Egress Identifier [McstPing]
- Echo Jitter TLV [McstPing]
- Vendor Private TLVs

Downstream Mapping objects MAY be included. These objects will be matched to the next hop address for inclusion in those particular MPLS echo request messages.

The message is then encapsulated in a UDP packet. The source UDP port is chosen by the sender; the destination UDP port is set to 3503. The IP header is set as follows: the source IP address is a routable address of the sender; the destination IP address is a routable address of the midpoint. The packet is then sent with the IP TTL is set to 255.

3.2. Procedures for the proxy LSR

A proxy LSR that receives an MPLS proxy ping request message, parses the packet to ensure that it is a well-formed packet. It checks that the TLVs that are not marked "Ignore" are understood. If not, it sets the Return Code set to "Malformed echo request received" or "TLV not understood" (as appropriate), and the Subcode set to zero. If the Reply Mode of the message header is not 1, an MPLS proxy ping reply message SHOULD be sent as described below. In the latter case, the misunderstood TLVs (only) are included in an Errored TLVs object.

The header fields Sender's Handle and Sequence Number are not examined, but are saved to be included in the MPLS proxy ping reply and MPLS echo request messages.

The proxy LSR validates that it has a label mapping for the specified FEC, it then determines if it is an egress, transit or bud node and sets the Return Code as appropriate.

The proxy LSR then determines if it is authorized to send the specified MPLS echo request on behalf of the initiator. An LSR MUST be capable of filtering addresses to validate initiators. Other filters on FECs or MPLS echo request contents MAY be applied. If a filter has been invoked (i.e. configured) and an address does not pass the

filter, then an MPLS echo request message MUST NOT be sent, and the event SHOULD be logged. An MPLS proxy ping reply message MAY be sent with a Return Code of <tba>, "Remote Ping not authorized".

The destination address specified in the Proxy Echo Parameters object is checked to ensure that it conforms to the address allowed IPv4 or IPv6 address range. If not, it sets the Return Code set to "Malformed echo request received" and the Subcode set to zero. If the Reply Mode of the message header is not 1, an MPLS proxy ping reply message SHOULD be sent as described below.

If the "Request for Previous Hop" flag is set, a Previous Hop Address Object is formatted for inclusion in the MPLS proxy ping reply. If the previous HOP is unknown or ambiguous the Address Type is set to "No Address Supplied".

If there are Next Hop sub-objexts in the Proxy Echo Parameters object, each address is examined to determine if it is a next hop for this FEC. If any are not, those sub-objects are from the Proxy Echo Parameters object. The updated object is included in the MPLS proxy ping reply.

If the "Request for Downstream Mapping" flag is set the LSR formats a Downstream Mapping object for each interface over which the MPLS echo request will be sent.

If the Reply Mode of the message header is 1 or is 5 and no errors or modificatons have occurred no MPLS proxy ping reply is sent. Otherwise an MPLS proxy ping reply message SHOULD be sent as described below.

3.2.1. Sending an MPLS proxy ping reply

The Reply mode, Sender's Handle and Sequence Number fields are copied from the proxy ping request message. The objects specified above are included. The message is encapsulated in a UDP packet. The source IP address is a routable address of the proxy LSR; the source port is the well-known UDP port for LSP ping. The destination IP address and UDP port are copied from the source IP address and UDP port of the echo request. The IP TTL is set to 255.

3.2.2. Sending the MPLS echo requests

A base MPLS echo request is formed as decribed in the next section. The section below that describes how the base MPLS echo request is sent on each interface.

3.2.2.1. Forming the base MPLS echo request

A Next_Hop_List is created as follows. If Next Hop sub-objects were included in the received Proxy Parameters object, the Next_Hop_List created from the address in those sub-ojects as adjusted above. Otherwise, the list is set to all the next hops to which the FEC would be forwarded.

The proxy LSR then formats an MPLS echo request message. The Global Flags and Reply Mode are copied from the Proxy Echo Parameters object. The Return Code and Return Subcode are set to zero.

The Sender's Handle and Sequence Number are copied from the remote echo request message.

The TimeStamp Sent is set to the time-of-day (in seconds and microseconds) that the echo request is sent. The TimeStamp Received is set to zero.

A Reply-to Address object containing the initiator's address is included.

The following objects are copied from the MPLS proxy ping request message. Note that of these, only the Target FEC Stack is REQUIRED to appear in the MPLS proxy ping request message.

Target FEC Stack
Pad
Vendor Enterprise Number
Reply TOS Byte
P2MP Egress Identifier [McstPing]
Echo Jitter TLV [McstPing]
Vendor Private TLVs

The message is then encapsulated in a UDP packet. The source UDP port is copied from the Proxy Echo Parameters object. The destination port copied from the proxy ping request message.

The source IP address is set to a routable address of the proxy LSR. Per usual the TTL of the IP packet is set to 1.

If the Explicit DSCP flag is set, the Requested DSCP byte is examined. If the setting is permitted then the DSCP byte of the IP header of the MPLS Echo Request message is set to that value. Otherwise the DSCP byte is set to a default value. In this case the MPLS Proxy Echo Parameters with the Explicit DSCP flag cleared MUST be included in any MPLS proxy ping reply message. The return code MUST be set to <tba>, "Proxy ping parameters modified". The DSCP field of the MPLS Proxy Echo Parameters SHOULD be set to the actual value used.

3.2.2.2. Per interface sending procedures

The proxy LSR now iterates through the Next_Hop_List modifying the base MPLS echo request to form the MPLS echo request packet which is then sent on that particular interface.

For each next hop address, the outgoing label stack is determained. The TTL for the label corresponding to the FEC specified in the FEC stack is set such that the TTL on the wire will be one less than the TTL specified in the Proxy Echo Parameters. If any additional labels are pushed onto the stack, their TTLs are set to 255.

If the MPLS proxy ping request message contained Downstream Mapping objects, they are examined. If the Downstream IP Address matches the next hop address that Downstream Mapping object is included in the MPLS echo request.

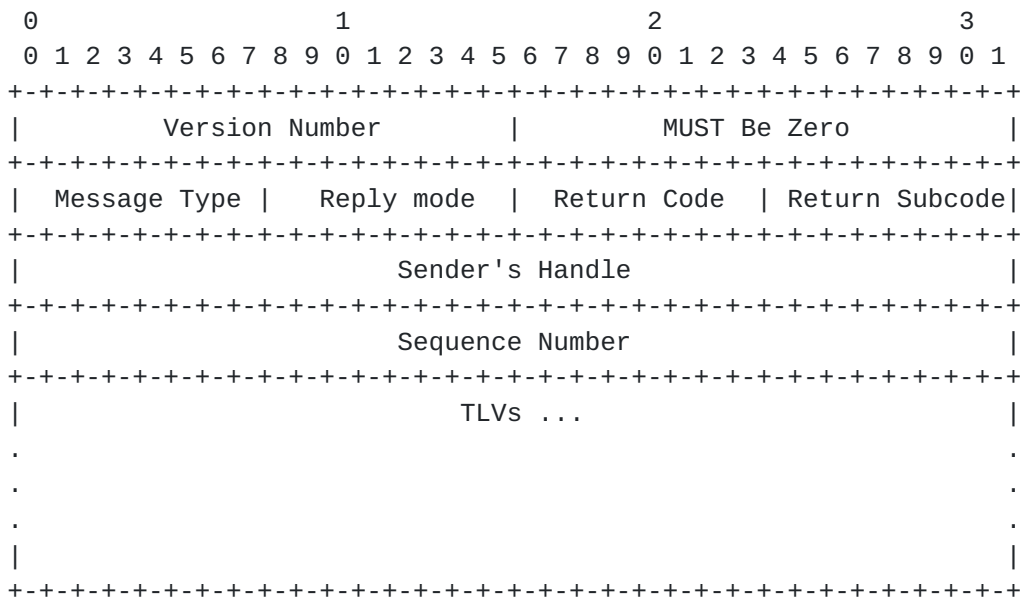
The packet is then transmitted on this interface.

4. Proxy Ping Request / Reply Messages

This document defines two new LSP Ping messages, the MPLS proxy ping request and the MPLS proxy ping reply.

4.1. Proxy Ping Request / Reply Message formats

Except where noted, the definitions of all fields in the messages are identical to those found in [RFC4379]. The messages have the following format:



Version Number

The Version Number is currently 1. (Note: the Version Number is to be incremented whenever a change is made that affects the ability of an implementation to correctly parse or process an MPLS echo request/reply. These changes include any syntactic or semantic changes made to any of the fixed fields, or to any TLV or sub-TLV assignment or format that is defined at a certain version number. The Version Number may not need to be changed if an optional TLV or sub-TLV is added.)

Message Type

Type	Message
5	MPLS proxy ping request
6	MPLS proxy ping reply

Reply mode

The reply modes are the same as [RFC4379] with the addtion of value 5. For completeness, the full list of reply modes follows:

Value	Meaning
1	Do not reply
2	Reply via an IPv4/IPv6 UDP packet
3	Reply via an IPv4/IPv6 UDP packet with Router Alert
4	Reply via application level control channel
5	Reply via an IPv4/IPv6 UDP packet only if the proxy request is not fulfilled or is modified

4.2. Proxy Ping Request Message contents

The MPLS proxy ping request message MAY contain the following objects:

Type	Object
1	Target FEC Stack
2	Downstream Mapping
3	Pad
5	Vendor Enterprise Number
10	Reply TOS Byte
tba	Proxy Echo Parameters
tba	P2MP Egress Identifier [McstPing]
tba	Echo Jitter TLV [McstPing]
	Vendor Private TLVs

4.3. Proxy Ping Reply Message Contents

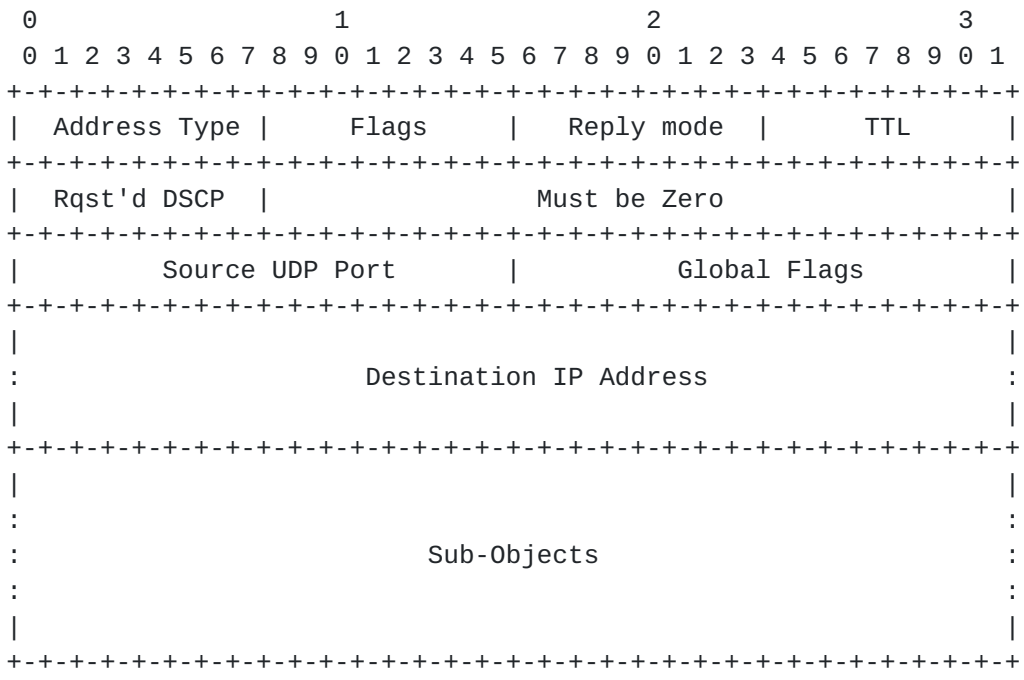
The MPLS proxy ping reply message MAY contain the following objects:

Type	Object
1	Target FEC Stack
2	Downstream Mapping
5	Vendor Enterprise Number
9	Errored TLVs
tba	Proxy Echo Parameters
tba	Previous Hop Address
	Vendor Private objects

5. Object formats

5.1. Proxy Echo Parameters Object

The Proxy Echo Parameters object is a TLV that MUST be included in an MPLS Proxy Echo Request message. The length of the TLV is 12 + K + S, where K is the length of the Destination IP Address field and S is the total length of the sub-objects.



Address Type

The type and length of the address found in the in the Destination IP Address and Next Hop IP Addresses fields. The type codes appear in the table below:

Address Family	Type	Length
IPv4	1	4
IPv6	3	16

Flags

Request for Previous Hop 0x01

When set this requests that the proxy LSR supply the previous hop address in the MPLS proxy ping reply message

Request for Downstream Mapping 0x02

When set this requests that the proxy LSR supply a Downstream Mapping object in the MPLS proxy ping reply message

Explicit DSCP Request 0x04

When set this requests that the proxy LSR supply a use the supplied DSCP byte in the echo request message

Reply mode

The reply mode to be sent in the MPLS Echo Request message; the values are as specified in [RFC4379]

TTL

The TTL to be used in the label stack entry corresponding to the topmost FEC in the in the MPLS Echo Request packet

Requested DSCP

This field is valid only if the Explicit DSCP flag is set. If not set, the field MUST be zero on transmission and ignored on receipt. When the flag is set this field contains the DSCP value to be used in the MPLS echo request packet IP header.

Source UDP Port

The source UDP port to be sent in the MPLS Echo Request packet

Global Flags

The Global Flags to be sent in the MPLS Echo Request messge

Destination IP Address

If the Address Type is IPv4, an address from the range 127/8;
If the Address Type is IPv6, an address from the range
0:0:0:0:0:FFFF:7F00:0/104

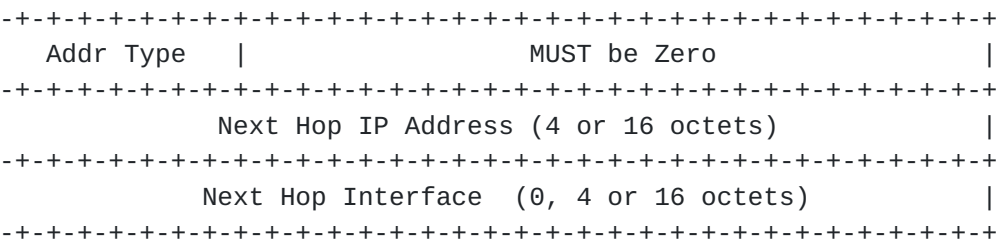
Sub-Objects

A TLV encoded list of sub-objects. Currently one is defined.

Sub-Type	Length	Value Field
1	8+	Next Hop

5.1.1. Next Hop sub-Object

This sub-object is used to describe a particular next hop towards which the Echo Request packet should be sent. If the topmost FEC in the FEC-stack is a multipoint LSP, this sub-object may appear multiple times.



Address Type

Type	Type of Next Hop	Addr Length	IF Length
1	IPv4 Numbered	4	4
2	IPv4 Unnumbered	4	4
3	IPv6 Numbered	16	16
4	IPv6 Unnumbered	16	4
5	IPv4 Protocol Adj	4	0
6	IPv6 Protocol Adj	16	0

Note: Types 1-4 correspond to the types in the DS Mapping object. They are expected to populated with information obtained through a previously returned DS Mapping object. Types 5 and 6 are intended to be populated from the local address information obtained from a previously returned Previous Hop Address Object.

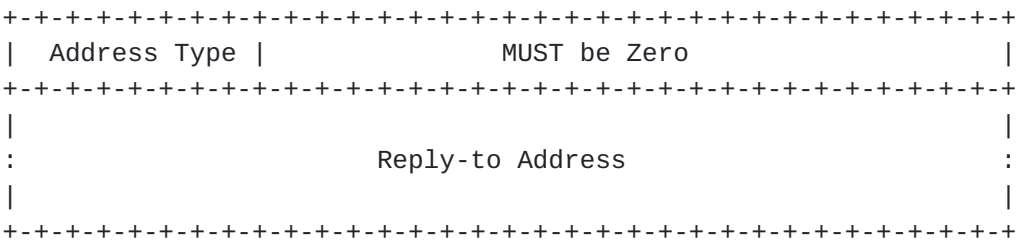
Next Hop IP Address

A next hop address that the echo request message is to be sent towards

Next Hop Interface

Identifier of the interface through which the echo request message is to be sent

5.2. Reply-to Address Object



Address Type

A type code as specified in the table below:

Type	Type of Address
1	IPv4
3	IPv6

5.3. Previous Hop Address Object

PHOP Addr Type; Local Addr Type

These two fields determine the type and length of the respective addresses. The codes are specified in the table below:

Type	Type of Address	Length
0	No Address Supplied	0
1	IPv4	4
3	IPv6	16

Previous Hop Address

The address of the immediate upstream neighbor for the topmost FEC in the FEC stack. If protocol adjacency exists by which the label for this FEC was exchanged, this address MUST be the address used in that protocol exchange.

Local Address

The local address used in the protocol adjacency exists by which the label for this FEC was exchanged.

6. Security Considerations

The mechanisms described in this document are intended to be used within a Service Provider network and to be initiated only under the authority of that administration.

If such a network also carries internet traffic, or permits IP access from other administrations, MPLS proxy ping message SHOULD be discarded at those points. This can be accomplished by filtering on source address or by filtering all MPLS ping messages on UDP port.

Any node which acts as a proxy node SHOULD validate requests against a set of valid source addresses. An implementation MUST provide such filtering capabilities.

MPLS proxy ping request messages are IP addressed directly to the Proxy node. If a node which receives an MPLS proxy ping message via TTL expiration, it MUST NOT be acted upon.

MPLS proxy ping requests are limited to making their request via the specification of a FEC. This ensures that only valid MPLS echo request messages can be created. No label spoofing attacks are

possible.

7. IANA Considerations

This document makes the following assigments (pending IANA action):

LSP Ping Message Types

Type	Value Field
tba	MPLS proxy ping request message
tba	MPLS proxy ping reply

Objects and Sub-Objects

Type	Sub-Type	Value Field
tba		Proxy Echo Parameters
	1	Next Hop
tba		Reply-to Address
tba		Previous Hop Address

Return Codes

Value	Meaning
tba	Remote ping not authorized
tba	Proxy ping parameters modified

Reply Modes [pending IANA assignment]

Value	Meaning
5	Reply via an IPv4/IPv6 UDP packet only if the proxy request is not fulfilled or is modified

8. References

8.1. Normative References

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

[KeyWords] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[McstPing] Farrel, A. et al, "Detecting Data Plane Failures in Point-to-Multipoint MPLS Traffic Engineering - Extensions to LSP Ping", draft-ietf-mpls-p2mp-lsp-ping-07.txt, September 2008.

8.2. Informative References

[P2MP-TE] Aggarwal, R., et al., "Extensions to RSVP-TE for Point-to-Multipoint TE LSPs", RFC 4875, May 2007.

[mLDP] Minei, I., et. al., "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths" draft-ietf-mpls-ldp-p2mp-05.txt, May 2008.

9. Authors' Addresses

George Swallow
Cisco Systems, Inc.
1414 Massachusetts Ave
Boxborough, MA 01719

Email: swallow@cisco.com

Vanson Lim
Cisco Systems, Inc.
1414 Massachusetts Ave
Boxborough, MA 01719

Email: vlim@cisco.com

Intellectual Property

Full Copyright Notice