Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 2, 2015

G. Swallow
V. Lim
Cisco Systems
S. Aldrin
Huawei Technologies

January 29, 2015

Proxy MPLS Echo Request
draft-ietf-mpls-proxy-lsp-ping-03

Abstract

This document defines a means of remotely initiating Multiprotocol Label Switched Protocol Pings on Label Switched Paths. A MPLS 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 leaf/root tracing.

Table of Contents

1. Introduction 3
1.1. Requirements Language 4
1.2. Terminology 5
2. Proxy Ping Overview 5
2.1. Initiating Proxy Ping 5
2.2. Handling at Proxy LSR 6
2.1.1. Backward Compatibility 6
3. Proxy MPLS Echo Request / Reply Procedures 6
3.1. Procedures for the initiator 7
3.2. Procedures for the proxy LSR 8
3.2.1. Proxy LSR Handling when it is Egress for FEC 10
3.2.2. Downstream Detailed/Downstream Maps in Proxy Reply . . 11
3.2.3. Sending an MPLS proxy ping reply 11
3.2.4. Sending the MPLS Echo Requests 11
3.2.4.1. Forming the base MPLS Echo Request 11
3.2.4.2. Per interface sending procedures 13
4. Proxy Ping Request / Reply Messages 13
4.1. Proxy Ping Request / Reply Message formats 13
4.2. Proxy Ping Request Message contents 14
4.3. Proxy Ping Reply Message Contents 14
5. TLV formats 15
5.1. Proxy Echo Parameters TLV 15
5.1.1. Next Hop sub-TLV 18
5.2. Reply-to Address TLV 19
5.3. Upstream Neighbor Address TLV 19
5.4. Downstream Neighbor Address TLV 20
6. Security Considerations 21
7. Acknowledgements 22
8. IANA Considerations 22
9. References 23

9.1. Normative References 23
9.2. Informative References 23
Authors' Addresses 23

1. Introduction

This document is motivated by two broad issues in connection with diagnosing Point-to-Multipoint (P2MP) Label Switched Paths (LSPs). The first is scalability due to the automatic replication of Multiprotocol Label Switching (MPLS) Echo Request Messages as they proceed down the tree. The second, which is primarily motivated by Label Distribution Protocol based Point-to-Multipoint (P2MP) and Multipoint-to-Multipoint (MP2MP) Label Switched Paths [RFC6388], is the ability to trace a sub-LSP from leaf node to root node.

It is anticipated that very large Point-to-Multipoint and Multipoint-to-Multipoint (MP2MP) Label Switched Paths will exist. Further it is anticipated that many of the applications for P2MP/MP2MP 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 Time to Live (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, say two, limiting the number of responses to the number of next-next-hops of the point where a ping is initiated.

In the case of Resource Reservation Protocol-Traffic Engineering (RSVP-TE), all setup is initiated from the root of the tree. Thus, the root of the tree has knowledge of both all the leaf nodes and usually the topology of the entire tree. Thus the above alternative can easily be initiated by the root node.

In [RFC6388] the situation is quite different. Leaf nodes initiate connectivity to the tree which is granted by the first node toward the root that is part of the tree. The root node may only be aware of the immediately adjacent (downstream) nodes of the tree. Initially the leaf node only has knowledge of the (upstream) node to which it is immediately adjacent. However this is sufficient information to initiate a trace. First the above procedure is applied by asking that

node to ping across the final link. That is, a message is sent from the leaf to the upstream node requesting it to send an MPLS Echo Request for the Forward Equivalence Class (FEC) of the tree in question on said link. The leaf node also requests the identity of the upstream neighbor's upstream neighbor for that FEC. With this information the procedure can iteratively be applied until the fault is localized or the root node is reached. In all cases the TTL for the request need only be at most 2. Thus the processing load of each request is small as only a limited number of nodes will receive the request.

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 an 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 [RFC4875]. This document also provides a means for obtaining this information for [RFC6388].

While the motivation for this document came from multicast scaling concerns, it's applicability may be wider. The procedures presented in this document are applicable to all LSP ping FEC types where the MPLS Echo Request/Reply are IP encapsulated and the MPLS Echo Reply can sent out of band of the LSP over ip. Remote pinging of LSPs that involve the use of in-band control channels is beyond the scope of this document.

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.

1.1. Requirements Language

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].

The term "Must Be Zero" (MBZ) is used in TLV 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.

1.2. Terminology

Term	Definition
LSP	Label Switched Paths
LSR	Label Switching Router
MP2MP	Multipoint to Multipoint
P2MP	Point to Multipoint
TTL	Time to Live

[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 request that second node initiate an LSP ping for the LSP on behalf of the first node. Since the second node sends the LSP Ping on behalf of the first node, it does not maintain state to be able to handle the corresponding LSP Ping response. Instead the responder to the LSP ping sends the LSP Ping response to either the first node or another node configured to handle it. 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 generally 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 nodes which receive the MPLS Echo Request message

Responder - A receiver that responds to a MPLS Proxy Ping Request or an MPLS Echo Request

We note that in some scenarios, the initiator could also be the responder, in which case the response would be internal to the node.

2.1. Initiating Proxy Ping

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 instructs the proxy LSR to either Reply with Proxy information or to send a MPLS Echo Request inband of the LSP. The initiator requests Proxy information so that it can learn additional information it needs to use to form a subsequent MPLS Proxy Ping request. For example during LSP traceroute an initiator needs the downstream map information to form an MPLS Echo Request. An initiator may also want to learn a Proxy LSR's FEC neighbor information so that it can form proxy request to various nodes along the LSP.

2.2. Handling at Proxy LSR

The proxy LSR either replies with the requested Proxy information or it 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.

If the proxy LSR has a label mapping for the FEC and all authorization checks have passed, the proxy LSR formats an MPLS Echo Request. If the source address of the MPLS Echo Request is not to be set to the Proxy Request source address, the initiator MUST include a Reply-to Address TLV containing the source address to use in the MPLS Echo Request. 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.

If the proxy LSR failed to send a MPLS Echo Request as normal because it encountered an issue while attempting to send, a MPLS proxy ping reply message is sent back with a return code indicating that the MPLS Echo Request could not be sent.

2.1.1. Backward Compatibility

As described in sec 4.4 of [RFC4379], If the packet is not well-formed, LSR X SHOULD send an MPLS Echo Reply with the Return Code set to "Malformed echo request received" and the Subcode to zero. If there are any TLVs not marked as "Ignore" that Proxy LSR does not understand, Proxy LSR SHOULD send an MPLS "TLV not understood" (as appropriate), and the Subcode set to zero.

In the case the targeted proxy LSR does not understand LSP ping Echo Request at all, like any other LSR which do not understand the messages, they MUST be dropped and no messages is set back to the initiator.

3. Proxy MPLS Echo Request / Reply Procedures

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 topmost FEC in the target FEC stack is used at the Proxy LSR to lookup the MPLS label stack that will be used to encapsulate the MPLS Echo Request packet.

The MPLS Proxy Ping request message MUST contain a Proxy Echo Parameters TLV. In that TLV, 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 is from the range 127/8. If the Address Type is IPv6, an address is from the range ::FFFF:7F00:0/104.

The Reply mode and Global Flags of the Proxy Echo Parameters TLV 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 (the source port is supplied by the Proxy Ping initiator because it or a node known to it handles the LSP ping responses). The TTL is set to the value to be used in the outgoing MPLS label stack. See Section 5.1 for further details.

If the FEC's Upstream/Downstream Neighbor address information is required, the initiator sets the "Request for FEC neighbor information" Proxy Flags in the Proxy Echo Parameters TLV.

If a Downstream Detailed or Downstream Mapping TLV is required in a MPLS Proxy Ping Reply, the initiator sets the "Request for Downstream Detailed Mapping" or "Request for Downstream Mapping" Proxy Flags in the Proxy Echo Parameters TLV. Only one of the two flags can be set.

The Proxy Request reply mode is set with one of the reply modes defined in [RFC4379] as appropriate.

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-TLVs and included in the Proxy Echo Parameters TLV.

Proxy Echo Parameter TLV MPLS payload size field may be set to request that the MPLS Echo Request (including any IP and UDP header) be zero padded to the specified size. When the payload size is non zero, if sending the MPLS Echo Request involves using an IP header, the Do not Fragment (DF) bit MUST be set to 1.

Any of following TLVs MAY be included; these TLVs will be copied into

the MPLS Echo Request messages:

- Pad
- Vendor Enterprise Number
- Reply TOS Byte
- P2MP Responder Identifier [RFC6425]
- Echo Jitter TLV [RFC6425]
- Vendor Private TLVs

Downstream Detailed Mapping (DDMAP) or Downstream Mapping (DSMAP) TLVs MAY be included. These TLVs 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 User Datagram Protocol (UDP) port for the MPLS proxy ping requests message is chosen by the initiator; 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 initiator; the destination IP address is a routable address to the Proxy LSR. 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 any part of the message is malformed, it sets the Return Code set to "Malformed echo request received". If all the TLVs are well formed and any TLVs are not understood, the return code is set to "TLV not understood". The Subcode is set to zero for both cases.

If the Reply Mode of the message header is not 1(Do not reply), an MPLS proxy ping reply message SHOULD be sent as described below.

If the Return Code is "TLV not understood", no more processing of the MPLS proxy ping request message is required. The Proxy LSR sends an MPLS Proxy ping reply message with an Errored TLVs TLV containing all the not understood TLVs (only).

The Proxy LSR checks that the MPLS proxy ping request message did not arrive via one of its exception processing paths. Packets arriving via IP TTL expiry, IP destination address set to a Martian address or label ttl expiry MUST be treated as "Unauthorized" packets. An MPLS

proxy ping reply message MAY be sent with a Return Code of <TBA-7>, "Proxy Ping not authorized".

The header fields Sender's Handle and Sequence Number are not examined, but included in the MPLS proxy ping reply or MPLS Echo Request messages, if one is sent as a direct result of the received message.

The proxy LSR validates that it has a label mapping for the specified FEC, it then determines if it is an ingress, egress, transit or bud node and sets the Return Code as appropriate. A new return code (Replying router has FEC mapping for topmost FEC) has been defined for the case where the Proxy LSR is an ingress (for example head of the TE tunnel or a transit router) because the existing RFC4379 return codes don't match the situation. For example, when a Proxy LSR is a transit router, it's not appropriate for the return code to describe how the packet would transit because the MPLS proxy ping request doesn't contain information about what input interface the an MPLS Echo Request would be switched from at the Proxy LSR.

The proxy LSR then determines if it is authorized to send the specified MPLS Echo Request on behalf of the initiator. A Proxy 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-7>, "Proxy Ping not authorized".

The destination address specified in the Proxy Echo Parameters TLV is checked to ensure that it conforms to the address allowed IPv4 or IPv6 address range. If not, 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 FEC Neighbor Address info" flag is set, a Upstream Neighbor Address TLV and/or Downstream Neighbor Address TLV(s) is/are formatted for inclusion in the MPLS proxy ping reply. If the Upstream or Downstream address is unknown they are not included in the Proxy Reply.

If there are Next Hop sub-TLVs in the Proxy Echo Parameters TLV, each address is examined to determine if it is a valid next hop for this FEC. If any are not, Proxy Echo Parameters TLV SHOULD be updated removing unrecognized Next Hop sub-TLVs. The updated Proxy Echo Parameters TLV MUST be included in the MPLS proxy ping reply.

If the "Request for Downstream Detailed Mapping" or "Request for Downstream Mapping" flag is set, the LSR formats (for inclusions in the MPLS proxy ping reply) a Downstream Detailed/Downstream Mapping TLV for each interface over which the MPLS Echo Request will be sent.

If the Proxy LSR is the egress for the FEC, the behavior of the proxy LSR vary depending on whether the node is an Egress of a P2P LSP, a P2MP LSP or MP2MP LSP. Additional details can be found in the section describing "Handling when Proxy LSR it is egress for FEC".

If the Reply Mode of the MPLS proxy ping request message header is "1 - do not reply", no MPLS proxy ping reply is sent. Otherwise an MPLS proxy ping reply message or MPLS Echo Request SHOULD be sent as described below.

3.2.1. Proxy LSR Handling when it is Egress for FEC

This sections describes the different behaviors for the Proxy LSR when it's the Egress for the FEC. In the P2MP budnode and MP2MP budnode and egress cases, different behavior is required.

When the Proxy LSR is the egress of a P2P FEC, a MPLS proxy ping reply SHOULD be sent to the initiator with the return code set to 3 (Reply router is Egress for FEC) with return Subcode set to 0.

When the Proxy LSR is the egress of a P2MP FEC, it can be either a budnode or just an Egress. If the Proxy LSR is a budnode, a MPLS proxy ping reply SHOULD be sent to the initiator with the return code set to 3 (Reply router is Egress for FEC) with return Subcode set to 0 and DS/DDMAPs only if the Proxy initiator requested information to be returned in a MPLS proxy ping reply. If the Proxy LSR is a budnode but not requested to return a MPLS proxy ping reply, the Proxy LSR SHOULD send MPLS Echo Request packet(s) to the downstream neighbors (no MPLS Echo Reply is sent to the Proxy Initiator to indicate that the Proxy LSR is an egress). If the Proxy LSR is just an egress, a MPLS proxy ping reply SHOULD be sent to the initiator with the return code set to 3 (Reply router is Egress for FEC) with return Subcode set to 0.

When the Proxy LSR is the egress of a MP2MP FEC, it can be either a budnode or just an Egress. LSP pings sent from a leaf of a MP2MP has different behavior in this case. MPLS Echo Request are sent to all upstream/downstream neighbors. The Proxy LSRs need to be consistent with this variation in behavior. If the Proxy LSR is a budnode or just an egress, a MPLS proxy ping reply SHOULD be sent to the initiator with the return code set to 3 (Reply router is Egress for FEC) with return Subcode set to 0 and DS/DDMAPs included only if the Proxy initiator requested information to be returned in a MPLS proxy

ping reply. If the Proxy LSR is not requested to return information in a MPLS proxy ping reply, the Proxy LSR SHOULD send MPLS Echo Request packets to all upstream/downstream neighbors as would be done when sourcing an LSP ping from a MP2MP leaf (no MPLS Echo Reply is sent to the Proxy initiator indicating that the Proxy LSR is an egress).

3.2.2. Downstream Detailed/Downstream Maps in Proxy Reply

When the Proxy LSR is a transit or bud node, downstream maps corresponding to how the packet is transited can not be supplied unless an ingress interface for the MPLS Echo Request is specified. Since this information is not available and all valid output paths are of interest, the Proxy LSR SHOULD include DS/DDMAP(s) to describe the entire set of paths that the packet can be replicated. This is similar to the case where an LSP ping is initiated at the Proxy LSR. For mLDP there is a DSMAP/DDMAP per upstream/downstream neighbor for MP2MP LSPs, or per downstream neighbor in the P2MP LSP case.

When the Proxy LSR is a bud node or egress in a MP2MP LSP or a budnode in a P2MP LSP, an LSP ping initiated from the Proxy LSR would source packets only to the neighbors but not itself despite the fact that the Proxy LSR is itself an egress for the FEC. In order to match the behavior as seen from LSP Ping initiated at the Proxy LSR, the Proxy Reply SHOULD contain DSMAP/DDMAPs for only the paths to the upstream/downstream neighbors, but no DSMAP/DDMAP describing its own egresses paths. The proxy LSR identifies that it's an egress for the FEC using a different Proxy Reply return code. The Proxy reply return code is either set to "Reply router has a mapping for the topmost FEC" or "Reply router is Egress for the FEC".

3.2.3. 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 TLVs 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 MPLS Proxy Ping Request. The IP TTL is set to 255.

3.2.4. Sending the MPLS Echo Requests

A MPLS Echo Request is formed as described in the next section. The section below that describes how the MPLS Echo Request is sent on each interface.

3.2.4.1. Forming the base MPLS Echo Request

A Next_Hop_List is created as follows. If Next Hop sub-TLVs were included in the received Proxy Parameters TLV, the Next_Hop_List created from the address in those sub-TLVs 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 TLV. 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 MPLS Echo Request is sent. The TimeStamp Received is set to zero.

If the reply-to address TLV is present, it is used to set the echo request source address, otherwise the echo request source address is set to the proxy request source address.

The following TLVs 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 Responder Identifier [RFC6425]

Echo Jitter TLV [RFC6425]

Vendor Private TLVs

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

The source IP address is set to a routable address specified in the reply-to-address TLV or the source address of the received proxy request. Per usual the TTL of the IP packet is set to 1.

If the Explicit Differentiated Services Code Point (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. If the Proxy LSR does not permit explicit control for the DSCP byte, the MPLS Proxy Echo Parameters with the Explicit DSCP flag cleared MUST be included in any MPLS proxy ping reply message to indicate why an MPLS Echo Request was not sent. The return code MUST be set to <TBA-8>, "Proxy ping parameters need to be modified". If the Explicit DSCP flag is not set, the Proxy LSR SHOULD set the MPLS Echo Request DSCP settings to the value normally used to source LSP ping packets..

3.2.4.2. 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 determined. 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 other TTL specified in the Proxy Echo Parameters. If any additional labels are pushed onto the stack, their TTLs are set to 255. This will ensure that the requestor will not have control over tunnels not relevant to the FEC being tested.

If the MPLS proxy ping request message contained Downstream Mapping/ Downstream Detailed Mapping TLVs, they are examined. If the Downstream IP Address matches the next hop address that Downstream Mapping TLV 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

The packet format is as defined in [RFC4379]. Two new message types, Proxy Ping Request and Reply, are being added.

Message Type

Type	Message
TBA-1	MPLS proxy ping request (Pending IANA assignment)

TBA-2	MPLS proxy ping reply (Pending IANA assignment)

4.2. Proxy Ping Request Message contents

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

Type	TLV
1	Target FEC Stack
2	Downstream Mapping
3	Pad
5	Vendor Enterprise Number
10	Reply TOS Byte
11	P2MP Responder Identifier [RFC6425]
12	Echo Jitter TLV [RFC6425]
20	Downstream Detailed Mapping
21	Reply Path [RFC7110]
22	Reply TC [RFC7110]
TBA-3	Proxy Echo Parameters (Pending IANA assignment)
TBA-4	Reply-to-Address TLV
*	Vendor Private TLVs

* TLVs types in the Vendor Private TLV Space MUST be ignored if not understood

4.3. Proxy Ping Reply Message Contents

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

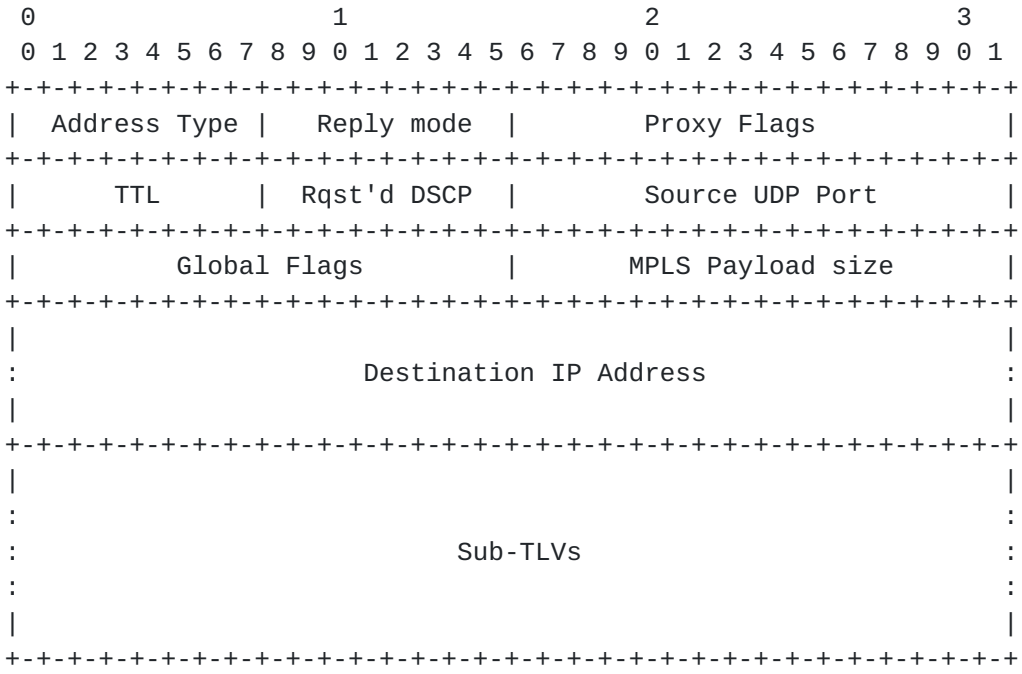
Type	TLV
1	Target FEC Stack
2	Downstream Mapping
5	Vendor Enterprise Number
9	Errored TLVs
20	Downstream Detailed Mapping
TBA-3	Proxy Echo Parameters (Pending IANA assignment)
TBA-5	Upstream Neighbor Address (Pending IANA assignment)
TBA-6	Downstream Neighbor Address (0 or more) (Pending IANA assignment)
*	Vendor Private TLVs

* TLVs types in the Vendor Private TLV Space MUST be ignored if not understood

5. TLV formats

5.1. Proxy Echo Parameters TLV

The Proxy Echo Parameters TLV is a TLV that MUST be included in an MPLS proxy ping 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-TLVs. The Proxy Echo Parameters TLV can be used to either to 1) control attributes used in Composing and Sending an MPLS Echo Request or 2) query the Proxy LSR for information about the topmost FEC in the target FEC stack but not both. In the case where the Proxy LSR is being queried (ie information needs to be returned in a Proxy Reply), no MPLS Echo Request will be sent from the Proxy LSR. The MPLS proxy ping request echo header's Reply Mode SHOULD be set to "Reply with Proxy Info".



Address Type

The type and length of the address found in the in the Destination IP Address and Next Hop IP Addresses fields. The values are shared with the Downstream Mapping Address Type Registry.

The type codes applicable in this case appear in the table below:

Address Family	Type	Length

IPv4	1	4
IPv6	3	16

Reply mode

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

Proxy Flags

The Proxy Request Initiator sets zero, one or more of these flags to request actions at the Proxy LSR.

0x01 Request for FEC Neighbor Address info

When set this requests that the proxy LSR supply the Upstream and Downstream neighbor address information in the MPLS proxy ping reply message. This flag is only applicable for the topmost FEC in the FEC stack if the FEC types corresponds with a P2MP or MP2MP LSPs. The Proxy LSR MUST respond as applicable with a Upstream Neighbor Address TLV and Downstream Neighbor Address TLV(s) in the MPLS proxy ping reply message. Upstream Neighbor Address TLV needs be included only if there is an upstream neighbor. Similarly, one Downstream Neighbor Address TLV needs to be included for each Downstream Neighbor for which the LSR learned bindings from.

Setting this flag will cause the proxy LSR to cancel sending an Echo request. Information learned with such proxy reply may be used by the proxy initiator to generate subsequent proxy requests.

0x02 Request for Downstream Mapping

When set this requests that the proxy LSR supply a Downstream Mapping TLV see [RFC4379] in the MPLS proxy ping reply message. It's not valid to have Request for Downstream Detailed Mapping flag set when this flag is set.

Setting this flag will cause the proxy LSR to cancel sending an Echo request. Information learned with such proxy reply may be used by the proxy initiator to generate subsequent proxy requests.

0x04 Request for Downstream Detailed Mapping

When set this requests that the proxy LSR supply a

Downstream Detailed Mapping TLV see [RFC6424] in the MPLS proxy ping reply message. It's not valid to have Request for Downstream Mapping flag set when this flag is set. Setting this flag will cause the proxy LSR to cancel sending an Echo request. Information learned with such proxy reply may be used by the proxy initiator to generate subsequent proxy requests.

0x08 Explicit DSCP Request

When set this requests that the proxy LSR use the supplied "Rqst'd DSCP" byte in the Echo Request message

TTL

The TTL to be used in the label stack entry corresponding to the topmost FEC in the in the MPLS Echo Request packet. Valid values are in the range [1,255]. A setting of 0 SHOULD be ignored by the Proxy LSR.

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 message

MPLS Payload Size

Used to request that the MPLS payload (IP header + UDP header + MPLS Echo Request) be padded using a zero filled Pad TLV so that the IP header, UDP header and MPLS Echo Request total the specified size. Field set to zero means no size request is being made. If the requested size is less than the minimum size required to form the MPLS Echo Request, the request will be treated as a best effort request with the Proxy LSR building the smallest possible packet (i.e. not using a Pad TLV). The IP header DF bit SHOULD be set when this field is non zero.

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
::FFFF:7F00:0/104

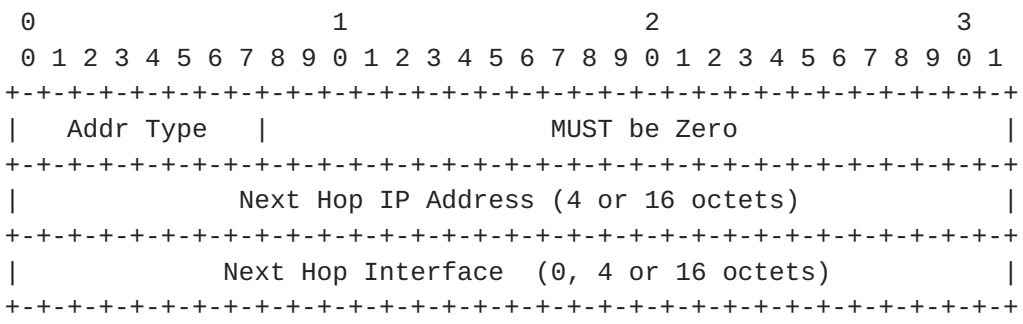
Sub-TLVs

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

Sub-Type	Length	Value Field
1	8+	Next Hop

5.1.1. Next Hop sub-TLV

This sub-TLV 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-TLV 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 TLV. They are expected to populated with information obtained through a previously returned DS Mapping TLV. Types 5 and 6 are intended to be populated from the local address information obtained from a previously returned Downstream Neighbor Address TLV(s)/Upstream Neighbor Address TLV.

Next Hop IP Address

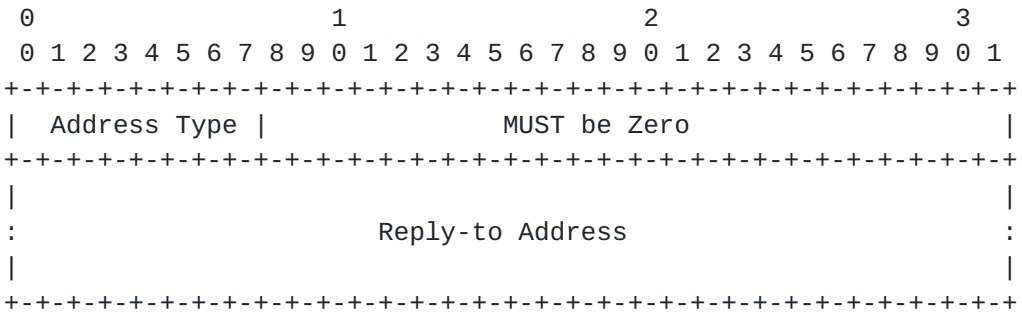
A next hop address that the echo request message is to be sent towards

Next Hop Interface

Identifier of the interface through which the echo request message is to be sent. For Addr Type 5, and 6, the Next Hop interface field isn't used and MUST be of an associated byte length of "0" octets.

5.2. Reply-to Address TLV

Used to specify the MPLS Echo Request IP source address. This address MUST be IP reachable via the Proxy LSR otherwise it will be rejected.

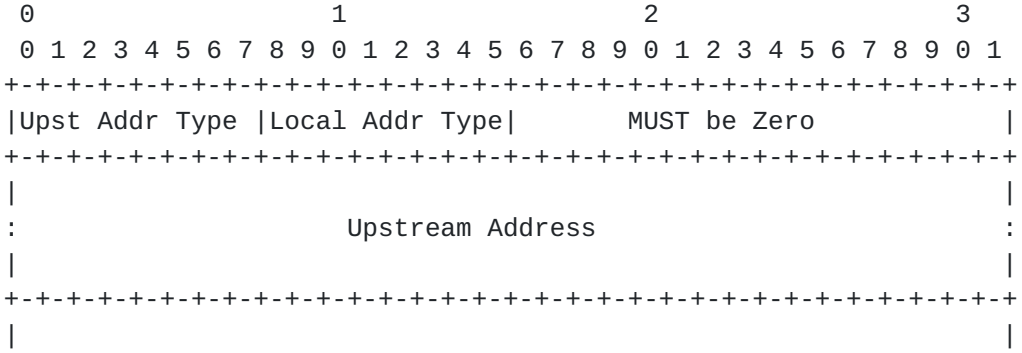


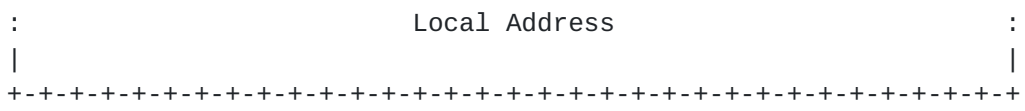
Address Type

A type code as specified in the table below:

Type	Type of Address
1	IPv4
3	IPv6

5.3. Upstream Neighbor Address TLV





Upst 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

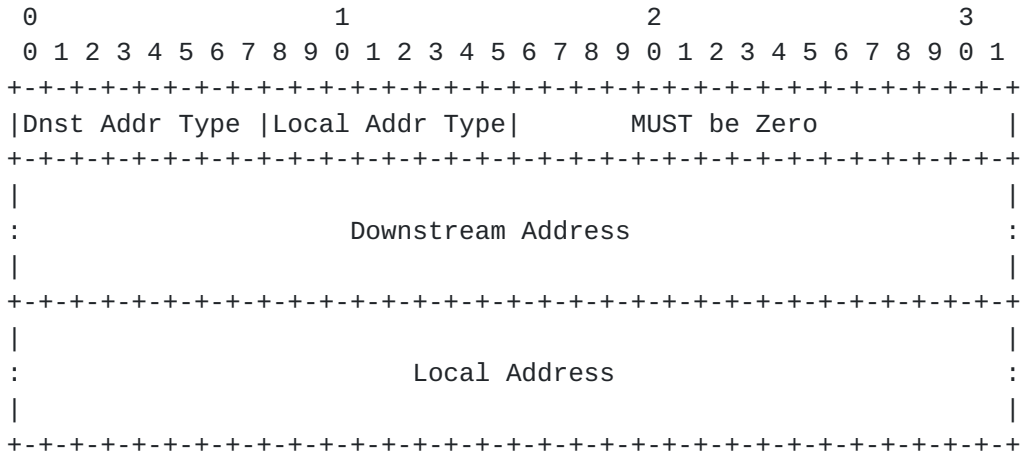
Upstream Address

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.

5.4. Downstream Neighbor Address TLV



Dnst 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

Downstream Address

The address of a immediate downstream 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 IP or Label TTL expiration, it MUST NOT be acted upon.

MPLS proxy ping request messages are IP addressed directly to the Proxy node. If a MPLS Proxy ping request IP destination address is a Martian Address, it MUST NOT be acted upon.

if a MPLS Proxy ping request IP source address is not IP reachable by the Proxy LSR, the Proxy request 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. Acknowledgements

The authors would like to thank Nobo Akiya for his detailed review and insightful comments.

8. IANA Considerations

This document makes the following assignments (pending IANA action)

LSP Ping Message Types

Type	Value Field
TBA-1	MPLS proxy ping request
TBA-2	MPLS proxy ping reply

TLVs and Sub-TLVs

Type	Sub-Type	Value Field
TBA-3		Proxy Echo Parameters
	1	Next Hop
TBA-4		Reply-to Address
TBA-5		Upstream Neighbor Address
TBA-6		Downstream Neighbor Address

Return Code [pending IANA assignment]

Value	Meaning
TBA-7	Proxy ping not authorized.
TBA-8	Proxy ping parameters need to be modified.
TBA-9	MPLS Echo Request Could not be sent.
TBA-10	Replying router has FEC mapping for topmost FEC.

Downstream Address Mapping Registry [pending IANA assignment]

Value	Meaning
TBA-11	IPv4 Protocol Adj
TBA-12	IPv6 Protocol Adj

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

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

[RFC6424] Bahadur, N., Kompella, K., and G. Swallow, "Mechanism for Performing Label Switched Path Ping (LSP Ping) over MPLS Tunnels", RFC 6424, November 2011.

[RFC6425] Saxena, S., 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, November 2011.

[RFC7110] Chen, M., Cao, W., Ning, S., Jounay, F., and Delord, S., "Return Path Specified Label Switched Path (LSP) Ping", RFC 7110, January 2014.

9.2. Informative References

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

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

Authors' Addresses

George Swallow
Cisco Systems
1414 Massachusetts Ave
Boxborough, MA 01719
USA

Email: swallow@cisco.com

Vanson Lim
Cisco Systems
1414 Massachusetts Avenue
Boxborough, MA 01719
USA

Email: vlim@cisco.com

Sam Aldrin
Huawei Technologies
2330 Central Express Way
Santa Clara, CA 95951
USA

Email: aldrin.ietf@gmail.com