Internet Engineering Task Force (IETF)
Request for Comments: 7555
Category: Standards Track
ISSN: 2070-1721

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

June 2015

Proxy MPLS Echo Request

Abstract

This document defines a means of remotely initiating Multiprotocol Label Switched Protocol (MPLS) Pings on Label Switched Paths. An MPLS Proxy Ping Request is sent to any Label Switching Router 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 tracing from leaf to leaf (or root).

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7555.

Table of Contents

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 LDP-based P2MP and Multipoint-to-Multipoint (MP2MP) LSPs [RFC6388], is the ability to trace a sub-LSP from leaf node to root node.

When tracing from a source to a particular leaf in a P2MP or MP2MP tree, nodes not along that path will need to process MPLS Echo Request messages that are received. The number of MPLS Echo Replies sent in response to an MPLS Echo Request quickly multiplies, as the Label Switching Routers (LSRs), which are part of the tree but not along the path of the trace, could be responding to the received MPLS Echo Request as well. This could also overwhelm the source to process all the MPLS Echo Reply messages it receives. 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 that 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 then work 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 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, since 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 P2MP and MP2MP LSPs that are set up with LDP as described in [RFC6388].

While the motivation for this document came from multicast scaling concerns, its 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 be sent out of band of the LSP over IP. Remote pinging of LSPs that involves 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. The procedures also do not allow the initiator to specify the label assigned to that FEC, nor do the procedures 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 Path
LSR	Label Switching Router
mLDP	Multipoint LDP
MP2MP	Multipoint to Multipoint
MTU	Maximum Transmission Unit
P2MP	Point to Multipoint
TTL	Time to Live

2. Proxy Ping Overview

This document defines a protocol interaction between a first LSR and another LSR that is part of an LSP in order to allow the first LSR to request that the second LSR initiate an LSP Ping for the LSP on the first LSR's behalf. Since the second LSR sends the LSP Ping on behalf of the first LSR, 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 LSR or another LSR 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 LSRs with the following terms:

Initiator - the LSR that initiates the ping operation by sending an MPLS Proxy Ping Request message

Proxy LSR - the LSR that is the destination of the MPLS Proxy Ping Request message and the potential initiator of the MPLS Echo Request

Receiver(s) - the LSR(s) that receive the MPLS Echo Request message

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

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

2.1. Initiating Proxy Ping

The initiator formats an MPLS Proxy Ping Request message and sends it to the Proxy LSR, an LSR it believes to be on the path of the LSP. This message instructs the Proxy LSR either to reply with Proxy information or to send an MPLS Echo Request in-band 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 Ping Requests to various LSRs along the LSP.

2.2. Handling at Proxy LSR

The Proxy LSR either replies with the requested Proxy information or 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 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 the MPLS Echo Request in-band 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 an MPLS Echo Request as normal because it encountered an issue while attempting to send, an MPLS Proxy Ping Reply message is sent back with a Return Code indicating that the MPLS Echo Request could not be sent.

2.2.1. Backward Compatibility

As described in Section 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 Return Subcode to zero. If there are any TLVs not marked as "Ignore" that the Proxy LSR does not understand, the Proxy LSR SHOULD send an MPLS "TLV not understood" (as appropriate), and the Return Subcode is set to zero.

In the case where the targeted Proxy LSR does not understand the LSP Ping Echo Request at all, like any other LSR that does not understand the messages, it MUST drop the message and MUST NOT send any message 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 look up 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 by the Proxy LSR to build the MPLS Echo Request packet. The MPLS Echo Request IP header destination address is as specified in [RFC4379]. If the Address Type is IPv4, it MUST be an address is from the range 127/8; if the Address Type is IPv6, MUST be an address 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 an LSR 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 Mapping TLV (or Downstream Mapping TLV, which is deprecated) is required in an MPLS Proxy Ping Reply, the initiator sets the "Request for Downstream Detailed Mapping" (or "Request for Downstream Mapping") Proxy Flag 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.

Although not explicitly spelled out in [RFC4379], LSP Ping packets can be formed to a desired size using a Pad TLV and then used to test the Maximum Transmission Unit (MTU) of an LSP. When testing an LSP's MTU, if the message is transported as an IP datagram, the IP header DF bit MUST be set to prevent IP fragmentation by the IP forwarding layer. The Proxy Echo Parameter TLV MPLS Payload Size field is defined for this purpose and may be set to request that the MPLS Echo Request (including any IP and UDP header) be zero-padded to the specified size. When a non-zero MPLS payload size is specified, the Proxy LSR introduces a Pad TLV to build the MPLS Echo Request packet, so in this case, the Proxy Ping Request MUST NOT include a Pad TLV.

Any of following TLVs MAY be included. These TLVs are used to form the MPLS Echo Request messages by the Proxy LSR:

- Pad
- Vendor Enterprise Number
- Reply TOS Byte
- P2MP Responder Identifier [RFC6425]
- Echo Jitter [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 UDP port for the MPLS Proxy Ping Request 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 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 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 Return 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 (only) the TLVs that were not understood.

The MPLS Proxy Ping Request is expected to be transported to the Proxy LSR via IP forwarding mechanisms instead of using the same techniques that are employed to inject an MPLS Echo Request packet into an LSP. The MPLS Echo Request would use IP TTL, MPLS TTL, and/or loopback addresses (IPv4 127.x.x.x or IPv6 ::ffff:7f00/104) in the IP header destination address field to trigger the packet to be handled via an LSR's forwarding exception processing path. The Proxy LSR MUST check whether or not MPLS Proxy Ping Request packets arrive via exception path. Packets arriving via IP TTL expiry, IP destination address set to a loopback 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 16, "Proxy Ping not authorized".

The header fields Sender's Handle and Sequence Number are not examined, but they are included in the MPLS Proxy Ping Reply or MPLS Echo Request message, if either is sent as a direct result of the received message.

The Proxy LSR validates that it has a label mapping for the specified FEC, determines if it is an ingress, egress, transit or bud node, and then sets the Return Code as appropriate. A new Return Code of 19, "Replying router has FEC mapping for topmost FEC", has been defined for the case where the Proxy LSR is an ingress (for example, the head of the TE tunnel or a transit router) because the existing Return Codes defined by RFC 4379 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 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 configured filter has been invoked 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 16, "Proxy Ping not authorized".

The destination address specified in the Proxy Echo Parameters TLV is checked to ensure that it conforms to the allowed IPv4 or IPv6 address range. If not, the Return Code is set to "Malformed echo request received" and the Return Subcode is 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.

The TTL specified in the Proxy Echo Parameters TLV is checked to ensure it contains a value in the range [1,255]. If not, the Return Code MUST be set to 17, "Proxy Ping parameters need to be modified". 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, the Upstream Neighbor Address and Downstream Neighbor Address TLVs are formatted for inclusion in the MPLS Proxy Ping reply. If the Upstream or Downstream address is unknown, the corresponding TLV is omitted.

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, the Proxy Echo Parameters TLV SHOULD be updated to remove 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 Proxy LSR formats (for inclusion in the MPLS Proxy Ping Reply) a DS/DDMAP 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 varies depending on whether the LSR is an egress of a P2P LSP, a P2MP LSP, or MP2MP LSP. Additional details can be found in Section 3.2.1, "Proxy LSR Handling When 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 section describes the different behaviors for the Proxy LSR when it's the egress for the FEC. In the P2MP bud node and MP2MP bud node egress cases, different behavior is required.

In the case where an MPLS Echo Request is originated by an LSR that is a bud or egress node of a P2MP/MP2MP, MPLS Echo Replies are returned from downstream/upstream LSRs and will not include an MPLS Echo Reply from the LSR that originated the MPLS Echo Request. This section describes the behavior required at a bud or egress node to return or not return information from MPLS Echo Replies in the Proxy Echo Reply so that no changes are required in implementations that are compliant with [RFC4379]. The Proxy Initiator should receive the same MPLS Echo Replies as in the case of the originator of the LSP Ping; any additional information (such as the Proxy LSR being a bud or egress node) is returned in the MPLS Proxy Ping Reply.

When the Proxy LSR is the egress of a P2P FEC, an MPLS Proxy Ping Reply SHOULD be sent to the initiator with the Return Code set to 3, "Replying router is an egress for the FEC at stack-depth", with Return Subcode set to zero.

When the Proxy LSR is the egress of a P2MP FEC, it can be either a bud node or just an egress. If the Proxy LSR is a bud node, an MPLS Proxy Ping Reply SHOULD be sent to the initiator with the return code set to 3, "Replying router is an egress for the FEC at stack-depth", and Return Subcode set to zero. DS/DDMAPs are included only if the Proxy Initiator requested information be returned in an MPLS Proxy Ping Reply. If the Proxy LSR is a bud node but there has not been a request to return an 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, an MPLS Proxy Ping Reply SHOULD be sent to the initiator with the Return Code set to 3, "Replying router is an egress for the FEC at stack-depth", and Return Subcode set to zero.

When the Proxy LSR is the egress of a MP2MP FEC, it can be either a bud node or just an egress. LSP Pings sent from a leaf of a MP2MP have different behavior in this case. MPLS Echo Requests 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

bud node or just an egress, an MPLS Proxy Ping Reply SHOULD be sent to the Proxy Initiator with the return code set to 3, "Replying router is an egress for the FEC at stack-depth", with Return Subcode set to zero and DS/DDMAPs included only if the Proxy Initiator requested information be returned in an MPLS Proxy Ping Reply. If the Proxy LSR is not requested to return information in an 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 Maps and Downstream Maps in Proxy Reply

When the Proxy LSR is a transit or bud node, downstream maps corresponding to how the packet is transited cannot 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 in which an LSP Ping is initiated at the Proxy LSR. For mLDP, there is a DS/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 an MP2MP LSP or a bud node 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 DS/DDMAPs for only the paths to the upstream/downstream neighbors, but no DS/DDMAP describing its own egress 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 19, "Replying router has FEC mapping for topmost FEC", or 3, "Replying router is an egress for the FEC at stack-depth".

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

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

3.2.4.1. Forming the Base MPLS Echo Request

If Next Hop sub-TLVs were included in the received Proxy Echo Parameters TLV, the Next_Hop_List is created from the addresses in those sub-TLVs adjusted as described in Section 3.2. 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 MPLS 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 MPLS Echo Request source address; otherwise, the MPLS 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. The Pad TLV is not copied if the Proxy Echo Parameter TLV MPLS payload size is set to a non-zero value.

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

If the Proxy Echo Parameter TLV MPLS payload size is non-zero, the Proxy LSR introduces a Pad TLV such that size of the MPLS Echo Request (including any IP and UDP header) is zero-padded to the specified MPLS payload size. The first octet in the Value part of the Pad TLV is set to 1, "Drop Pad TLV from reply", and the remaining octets of the Value part of the Pad TLV are filled with zeros. If the IP header is used to encapsulate the MPLS Echo Request, the DF bit MUST be set to one.

The message is then encapsulated in a UDP packet. The source UDP port is copied from the Proxy Echo Parameters TLV. The destination port is copied from the MPLS 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 17, "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 that is then sent on that particular interface.

The outgoing label stack is determined for each next-hop address. 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 the 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 TLVs or 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
3	MPLS Proxy Ping Request
4	MPLS Proxy Ping Reply

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 (DEPRECATED)
3	Pad
5	Vendor Enterprise Number
10	Reply TOS Byte
11	P2MP Responder Identifier [RFC6425]
12	Echo Jitter [RFC6425]
20	Downstream Detailed Mapping
21	Reply Path [RFC7110]
22	Reply TC [RFC7110]
23	Proxy Echo Parameters
24	Reply-to Address
*	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:

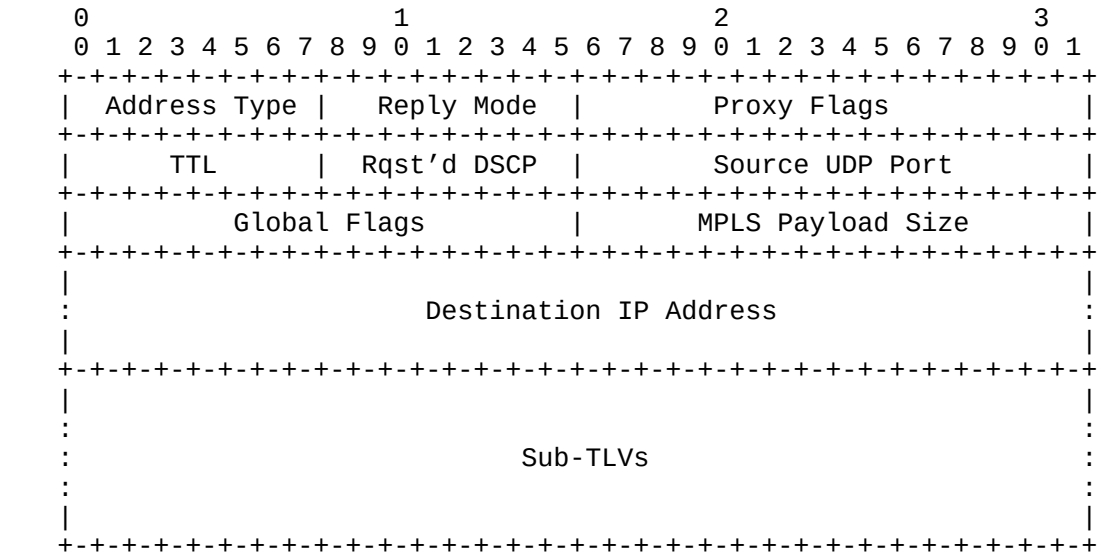
Type	TLV
1	Target FEC Stack
2	Downstream Mapping (DEPRECATED)
5	Vendor Enterprise Number
9	Errored TLVs
20	Downstream Detailed Mapping
23	Proxy Echo Parameters
25	Upstream Neighbor Address
26	Downstream Neighbor Address (0 or more)
*	Vendor Private TLVs

* 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 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 (i.e., information needs to be returned in an MPLS Proxy Ping Reply), no MPLS Echo Request will be sent from the Proxy LSR. The MPLS Proxy Ping Request Proxy Echo Parameters TLV’s Proxy Flags SHOULD be set appropriately, as described below.



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.

0x0001 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 type corresponds with a P2MP or MP2MP LSP. The Proxy LSR MUST respond (as applicable) with Upstream Neighbor Address and Downstream Neighbor Address TLV(s) in the MPLS Proxy Ping Reply message. The 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 from which the LSR learned bindings.

Setting this flag will cause the Proxy LSR to cancel sending any MPLS Echo Request. The initiator may use information learned from the MPLS Proxy Ping Reply that is sent instead to generate subsequent proxy requests.

0x0002 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. Either this flag may be set or the "Request for Downstream Detailed Mapping" flag may be set, but not both.

Setting this flag will cause the Proxy LSR to cancel sending an MPLS Echo Request. Information learned with such a Proxy Reply may be used by the Proxy Initiator to generate subsequent Proxy Requests.

0x0004 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 the "Request for Downstream Mapping" flag set when this flag is set. Setting this flag will cause the Proxy LSR to cancel sending an MPLS Echo Request. The initiator may use information learned from the MPLS Proxy Ping Reply that is sent instead to generate subsequent proxy requests.

0x0008 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 MPLS Echo Request packet. Valid values are in the range [1,255].

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. Having the 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 MUST 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

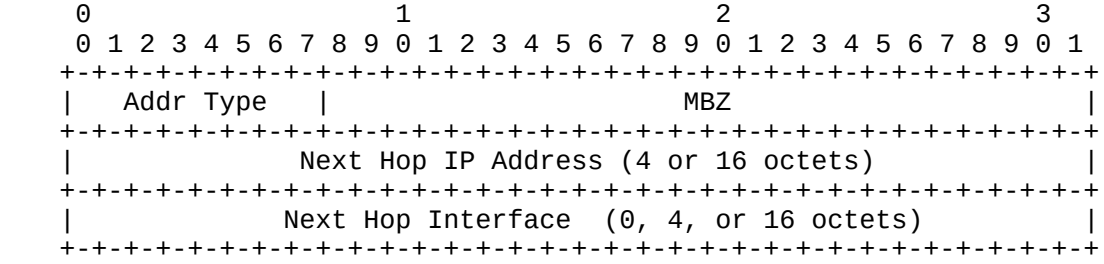
Sub-TLVs

List of TLV-encoded sub-TLVs. Currently one is defined.

Sub-Type	Length	Sub-TLV Name
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	Interface Field (IF) Length
1	IPv4 Numbered	4	4
2	IPv4 Unnumbered	4	4
3	IPv6 Numbered	16	16
4	IPv6 Unnumbered	16	4
5	Reserved		
6	IPv4 Protocol Adj	4	0
7	IPv6 Protocol Adj	16	0

Note: Types 1-4 correspond to the types in the DSMAP TLV. They are expected to be populated with information obtained through a previously returned DSMAP TLV. Types 6 and 7 are intended to be populated from the local address information obtained from a previously returned Downstream Neighbor Address TLV or Upstream Neighbor Address TLV.

Next Hop IP Address

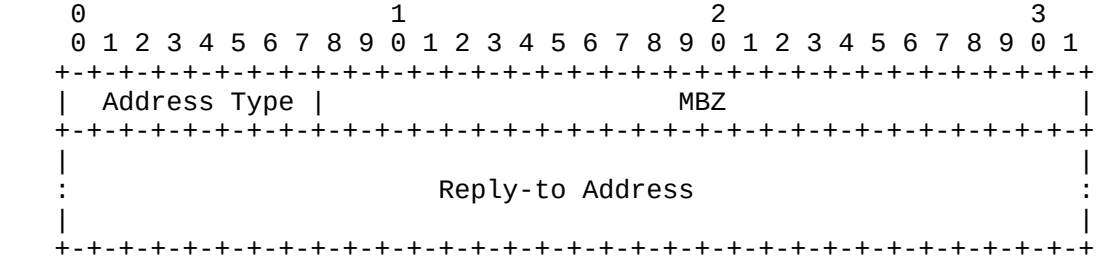
A next hop address that the Echo Request message is to be sent towards

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 zero 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.

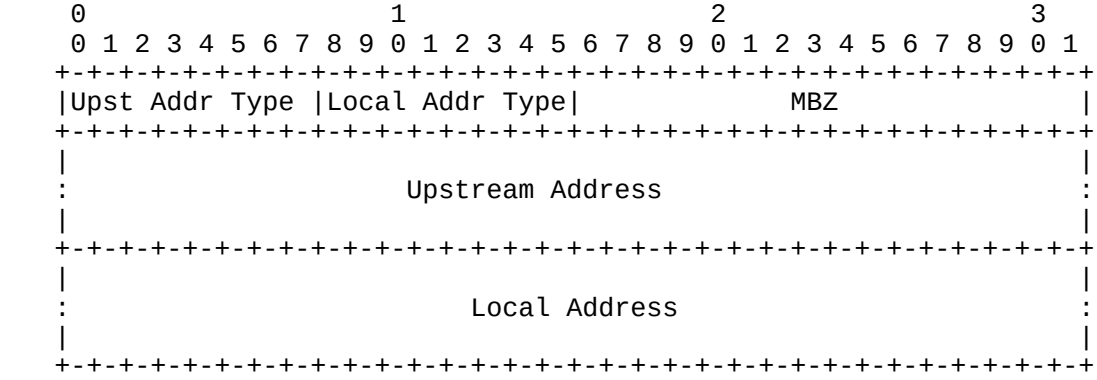


Address Type

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

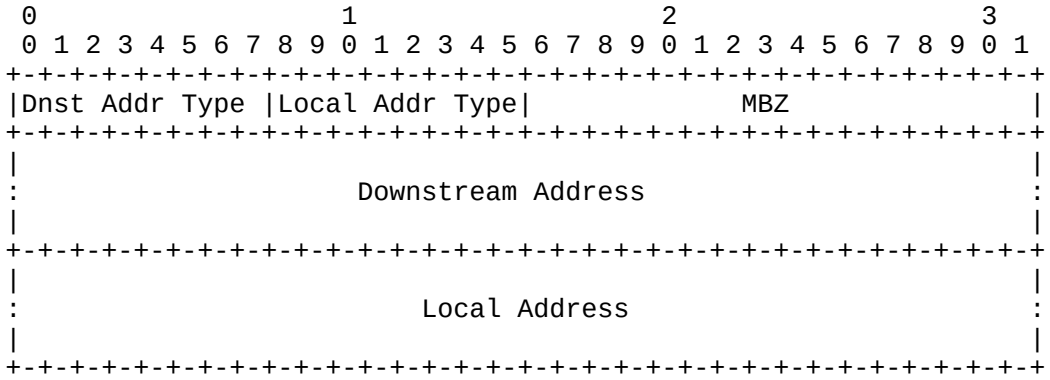
Upstream Address

The address of the immediate upstream neighbor for the topmost FEC in the FEC stack. If the 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 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 an immediate downstream neighbor for the topmost FEC in the FEC stack. If the 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 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, the MPLS Proxy Ping message SHOULD be discarded at the points where IP packets are received from other administrations. This can be accomplished by filtering on source address or by filtering all MPLS ping messages on UDP port.

Any node that acts as a Proxy LSR 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 LSR. If a Proxy LSR receives an MPLS Proxy Ping message via expiration of the IP or Label Stack Entry TTL, it MUST NOT be acted upon.

If an 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. IANA Considerations

Per this document, IANA has made the following assignments.

MPLS LSP Ping Message Types

Value	Meaning
3	MPLS Proxy Ping Request
4	MPLS Proxy Ping Reply

TLVs

Type	TLV Name
23	Proxy Echo Parameters
24	Reply-to Address
25	Upstream Neighbor Address
26	Downstream Neighbor Address

Return Codes

Value	Meaning
16	Proxy Ping not authorized
17	Proxy Ping parameters need to be modified
18	MPLS Echo Request could not be sent
19	Replying router has FEC mapping for topmost FEC

7.1. Proxy Echo Parameters Sub-TLVs

The IANA has created and maintains this new registry for Proxy Echo Parameters Sub-TLVs. Assignments will use the same rules spelled out in Section 7.2 of [RFC4379].

Sub-Type	Sub-TLV Name
0	Reserved
1	Next Hop

7.2. Proxy Flags

IANA has created and maintains a new registry for the Proxy Flags that are used with the Proxy Echo Parameters TLV. See Section 5.1 for details. The registry is in the "Multi-Protocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" registry in the "Multiprotocol Label Switching Architecture (MPLS)" name space. The registration procedure is Standards Action [RFC5226]. The initial values are as follows.

Bit Number	Name
0	Request for FEC Neighbor Address info
1	Request for Downstream Mapping
2	Request for Downstream Detailed Mapping
3	Explicit DSCP Request
4-15	Unassigned

7.3. Downstream Address Mapping Registry

This document makes the following assignments in the Downstream Address Mapping Registry. This document updates the registry defined by [RFC6426]. The registration procedure remains Standards Action and a note has been added as follows:

> When a code point is assigned that is not also assigned in the Next Hop Address Type Registry, the code point there must be marked "Reserved".

Type #	Address Type	K Octets	
6	Reserved	N/A	RFC 7555
7	Reserved	N/A	RFC 7555

7.4. Next Hop Sub-TLV Address Type Registry

IANA has created a new registry called the "Next Hop Address Type Registry". The allocation policy for this registry is Standards Action. Further, a note has been added as follows:

> When a code point is assigned that is not also assigned in the Downstream Address Mapping Registry, the code point there must be marked "Reserved".

The initial allocations are:

Type	Type of Next Hop	Addr Length	IF Length	Reference
1	IPv4 Numbered	4	4	[RFC4379]
2	IPv4 Unnumbered	4	4	[RFC4379]
3	IPv6 Numbered	16	16	[RFC4379]
4	IPv6 Unnumbered	16	4	[RFC4379]
5	Reserved			RFC 7555
6	IPv4 Protocol Adj	4	0	RFC 7555
7	IPv6 Protocol Adj	16	0	RFC 7555
8-255	Unassigned			

8. References

8.1. Normative References

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

[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>.

[RFC6424] Bahadur, N., Kompella, K., and G. Swallow, "Mechanism for Performing Label Switched Path Ping (LSP Ping) over MPLS Tunnels", RFC 6424, DOI 10.17487/RFC6424, November 2011, <http://www.rfc-editor.org/info/rfc6424>.

[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>.

[RFC7110] Chen, M., Cao, W., Ning, S., Jounay, F., and S. Delord, "Return Path Specified Label Switched Path (LSP) Ping", RFC 7110, DOI 10.17487/RFC7110, January 2014, <http://www.rfc-editor.org/info/rfc7110>.

8.2. Informative References

[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>.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, DOI 10.17487/RFC5226, May 2008, <http://www.rfc-editor.org/info/rfc5226>.

[RFC6388] Wijnands, IJ., Ed., Minei, I., Ed., Kompella, K., and B. Thomas, "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", RFC 6388, DOI 10.17487/RFC6388, November 2011, <http://www.rfc-editor.org/info/rfc6388>.

Acknowledgements

The authors would like to thank Nobo Akiya, Adrian Farrel, Tom Yu, Tom Taylor, and Warren Kumari for their detailed reviews and insightful comments.

Authors' Addresses

George Swallow
Cisco Systems
1414 Massachusetts Ave
Boxborough, MA 01719
United States

EMail: swallow@cisco.com

Vanson Lim
Cisco Systems
1414 Massachusetts Avenue
Boxborough, MA 01719
United States

EMail: vlim@cisco.com

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

EMail: aldrin.ietf@gmail.com