Detecting Data Plane Failures in Point-to-Multipoint Multiprotocol Label Switching (MPLS) - Extensions to LSP Ping

draft-ietf-mpls-p2mp-lsp-ping-07.txt

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

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

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

Abstract

Recent proposals have extended the scope of Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) to encompass point-to-multipoint (P2MP) LSPs.

The requirement for a simple and efficient mechanism that can be used to detect data plane failures in point-to-point (P2P) MPLS LSPs has been recognized and has led to the development of techniques for fault detection and isolation commonly referred to as "LSP Ping".

The scope of this document is fault detection and isolation for P2MP MPLS LSPs. This document does not replace any of the mechanisms of LSP Ping, but clarifies their applicability to MPLS P2MP LSPs, and extends the techniques and mechanisms of LSP Ping to the MPLS P2MP environment.

Yasukawa and Farrel
Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Contents

1. Introduction ................................................... 4
   1.1 Design Considerations ...................................... 5
2. Notes on Motivation ............................................ 6
   2.1 Basic Motivations for LSP Ping ............................. 6
   2.2 Motivations for LSP Ping for P2MP LSPs .................... 8
3. Operation of LSP Ping for a P2MP LSP .......................... 9
   3.1 Identifying the LSP Under Test ............................ 9
   3.1.1 Identifying a P2MP MPLS TE LSP .......................... 9
   3.1.1.1 RSVP P2MP IPv4 Session Sub-TLV ....................... 9
   3.1.1.2 RSVP P2MP IPv6 Session Sub-TLV ....................... 10
   3.1.2 Identifying a Multicast LDP LSP .......................... 10
   3.1.2.1 Multicast LDP FEC Stack Sub-TLV ...................... 11
   3.2 Ping Mode Operation ....................................... 12
   3.2.1 Controlling Responses to LSP Pings ........................ 12
   3.2.2 Ping Mode Egress Procedures ............................ 12
   3.2.3 Jittered Responses .................................... 13
   3.2.4 P2MP Responder Identifier TLV and Sub-TLVs ................ 14
   3.2.5 Echo Jitter TLV ....................................... 15
   3.2.6 Echo Response Reporting ................................ 15
   3.3 Traceroute Mode Operation .................................. 16
   3.3.1 Traceroute Responses at Non-Branch Nodes ................ 17
   3.3.1.1 Correlating Traceroute Responses .................... 17
   3.3.2 Traceroute Responses at Branch Nodes ................... 18
   3.3.2.1 Node Properties TLV ................................ 18
   3.3.2.2 Branching Properties Sub-TLV ........................ 19
   3.3.2.3 Egress Address Sub-TLV .............................. 20
   3.3.2.4 Correlating Traceroute Responses .................... 21
   3.3.3 Traceroute Responses at Bud Nodes ..................... 21
   3.3.4 Non-Response to Traceroute Echo Requests ................ 22
   3.3.5 Additions to Downstream Mapping Multipath Information .... 22
   3.3.6 Echo Response Reporting ................................ 24
   3.3.6.1 Reporting Multiple Conditions Using The DDM TLV .... 24
4. Operation of LSP Ping for Bootstrapping Other OAM Mechanisms .. 25
5. Non-compliant Routers ........................................... 26
6. OAM Considerations ............................................. 26
7. IANA Considerations ............................................ 27
   7.1 New Sub-TLV Types ....................................... 27
   7.2 New Multipath Type ....................................... 27
   7.3 New TLVs ............................................... 28
   7.4 New Return Code ....................................... 28
   7.5 New Sub-TLV Value for the Downstream Detailed Mapping TLV ... 28
8. Security Considerations ........................................ 29
9. Acknowledgements .................................................. 29
10. Intellectual Property Considerations .............................. 29
11. Normative References ............................................. 30
12. Informative References .......................................... 30
13. Authors’ Addresses ................................................ 31
14. Full Copyright Statement ......................................... 32

0. Change Log

This section to be removed before publication as an RFC.

0.1 Changes from 00 to 01

- Update references.
- Fix boilerplate.

0.2 Changes from 01 to 02

- Update entire document so that it is not specific to MPLS-TE, but also includes multicast LDP LSPs.
- Move the egress identifier sub-TLVs from the FEC Stack TLV to a new egress identifier TLV.
- Include Multicast LDP FEC Stack sub-TLV definition from [MCAST-CV].
- Add brief section on use of LSP Ping for bootstrapping.
- Add new references to References section.
- Add details of two new authors.

0.3 Changes from 02 to 03

- Update references.
- Update boilerplate.
- Fix typos.
- Clarify in 3.2.2 that a recipient of an echo request must reply only once it has applied incoming rate limiting.
- Tidy references to bootstrapping for [MCAST-CV] in 1.1.
- Allow multiple sub-TLVs in the P2MP Egress Identifier TLV in sections 3.2.1, 3.2.2, 3.2.4, 3.3.1, and 3.3.4.
- Clarify how to handle a P2MP Egress Identifier TLV with no sub-TLVs in sections 3.2.1 and 3.2.2.

0.4 Changes from 03 to 04

- Revert to previous text in sections 3.2.1, 3.2.2, 3.2.4, 3.3.1, and 3.3.4 with respect to multiple sub-TLVs in the P2MP Egress Identifier TLV.
1. Introduction

Simple and efficient mechanisms that can be used to detect data plane failures in point-to-point (P2P) Multiprotocol Label Switching (MPLS) Label Switched Paths (LSP) are described in [RFC4379]. The techniques involve information carried in an MPLS "echo request" and "echo reply", and mechanisms for transporting the echo reply. The echo request and reply messages provide sufficient information to check correct operation of the data plane, as well as a mechanism to verify the data plane against the control plane, and thereby localize faults. The use of reliable channels for echo reply messages as
described in [RFC4379] enables more robust fault isolation. This collection of mechanisms is commonly referred to as "LSP Ping".

The requirements for point-to-multipoint (P2MP) MPLS traffic engineered (TE) LSPs are stated in [RFC4461]. [RFC4875] specifies a signaling solution for establishing P2MP MPLS TE LSPs.

The requirements for point-to-multipoint extensions to the Label Distribution Protocol (LDP) are stated in [P2MP-LDP-REQ]. [P2MP-LDP] specifies extensions to LDP for P2MP MPLS.

P2MP MPLS LSPs are at least as vulnerable to data plane faults or to discrepancies between the control and data planes as their P2P counterparts. Mechanisms are, therefore, desirable to detect such data plane faults in P2MP MPLS LSPs as described in [RFC4687].

This document extends the techniques described in [RFC4379] such that they may be applied to P2MP MPLS LSPs and so that they can be used to bootstrap other Operations and Management (OAM) procedures such as [MCAST-CV]. This document stresses the reuse of existing LSP Ping mechanisms used for P2P LSPs, and applies them to P2MP MPLS LSPs in order to simplify implementation and network operation.

1.1 Design Considerations

An important consideration for designing LSP Ping for P2MP MPLS LSPs is that every attempt is made to use or extend existing mechanisms rather than invent new mechanisms.

As for P2P LSPs, a critical requirement is that the echo request messages follow the same data path that normal MPLS packets traverse. However, it can be seen this notion needs to be extended for P2MP MPLS LSPs, as in this case an MPLS packet is replicated so that it arrives at each egress (or leaf) of the P2MP tree.

MPLS echo requests are meant primarily to validate the data plane, and they can then be used to validate data plane state against the control plane. They may also be used to bootstrap other OAM procedures such as [MPLS-BFD] and [MCAST-CV]. As pointed out in [RFC4379], mechanisms to check the liveness, function, and consistency of the control plane are valuable, but such mechanisms are not a feature of LSP Ping and are not covered in this document.

As is described in [RFC4379], to avoid potential Denial of Service attacks, it is RECOMMENDED to regulate the LSP Ping traffic passed to the control plane. A rate limiter should be applied to the well-known UDP port defined for use by LSP Ping traffic.
2. Notes on Motivation

2.1. Basic Motivations for LSP Ping

The motivations listed in [RFC4379] are reproduced here for completeness.

When an LSP fails to deliver user traffic, the failure cannot always be detected by the MPLS control plane. There is a need to provide a tool that enables users to detect such traffic "black holes" or misrouting within a reasonable period of time. A mechanism to isolate faults is also required.

[RFC4379] describes a mechanism that accomplishes these goals. This mechanism is modeled after the ping/traceroute paradigm: ping (ICMP echo request [RFC792]) is used for connectivity checks, and traceroute is used for hop-by-hop fault localization as well as path tracing. [RFC4379] specifies a "ping mode" and a "traceroute" mode for testing MPLS LSPs.

The basic idea as expressed in [RFC4379] is to test that the packets that belong to a particular Forwarding Equivalence Class (FEC) actually end their MPLS path on an LSR that is an egress for that FEC. [RFC4379] achieves this test by sending a packet (called an "MPLS echo request") along the same data path as other packets belonging to this FEC. An MPLS echo request also carries information about the FEC whose MPLS path is being verified. This echo request is forwarded just like any other packet belonging to that FEC. In "ping" mode (basic connectivity check), the packet should reach the end of the path, at which point it is sent to the control plane of the egress LSR, which then verifies that it is indeed an egress for the FEC. In "traceroute" mode (fault isolation), the packet is sent to the control plane of each transit LSR, which performs various checks that it is indeed a transit LSR for this path; this LSR also returns further information that helps to check the control plane against the data plane, i.e., that forwarding matches what the routing protocols determined as the path.

One way these tools can be used is to periodically ping a FEC to ensure connectivity. If the ping fails, one can then initiate a traceroute to determine where the fault lies. One can also periodically traceroute FECs to verify that forwarding matches the control plane; however, this places a greater burden on transit LSRs and should be used with caution.

2.2. Motivations for LSP Ping for P2MP LSPs

As stated in [RFC4687], MPLS has been extended to encompass P2MP LSPs. As with P2P MPLS LSPs, the requirement to detect, handle, and diagnose control and data plane defects is critical. For operators...
deploying services based on P2MP MPLS LSPs, the detection and specification of how to handle those defects is important because such defects may affect the fundamentals of an MPLS network, but also because they may impact service level specification commitments for customers of their network.

P2MP LDP [P2MP-LDP] uses the Label Distribution Protocol to establish multicast LSPs. These LSPs distribute data from a single source to one or more destinations across the network according to the next hops indicated by the routing protocols. Each LSP is identified by an MPLS multicast FEC.

P2MP MPLS TE LSPs [RFC4875] may be viewed as MPLS tunnels with a single ingress and multiple egresses. The tunnels, built on P2MP LSPs, are explicitly routed through the network. There is no concept or applicability of a FEC in the context of a P2MP MPLS TE LSP.

MPLS packets inserted at the ingress of a P2MP LSP are delivered equally (barring faults) to all egresses. In consequence, the basic idea of LSP Ping for P2MP MPLS TE LSPs may be expressed as an intention to test that packets that enter (at the ingress) a particular P2MP LSP actually end their MPLS path on the LSRs that are the (intended) egresses for that LSP. The idea may be extended to check selectively that such packets reach specific egresses.

The technique in this document makes this test by sending an LSP Ping echo request message along the same data path as the MPLS packets. An echo request also carries the identification of the P2MP MPLS LSP (multicast LSP or P2MP TE LSP) that it is testing. The echo request is forwarded just as any other packet using that LSP, and so is replicated at branch points of the LSP and should be delivered to all egresses. In "ping" mode (basic connectivity check), the echo request should reach the end of the path, at which point it is sent to the control plane of the egress LSRs, which verify that they are indeed an egress (leaf) of the P2MP LSP. An echo response message is sent by an egress to the ingress to confirm the successful receipt (or announce the erroneous arrival) of the echo request.

In "traceroute" mode (fault isolation), the echo request is sent to the control plane at each transit LSR, and the control plane checks that it is indeed a transit LSR for this P2MP MPLS LSP. The transit LSR also returns information on an echo response that helps verify the control plane against the data plane. That is, the information is used by the ingress to check that the data plane forwarding matches what is signaled by the control plane.

P2MP MPLS LSPs may have many egresses, and it is not necessarily the intention of the initiator of the ping or traceroute operation to collect information about the connectivity or path to all egresses. Indeed, in the event of pinging all egresses of a large P2MP MPLS
LSP, it might be expected that a large number of echo responses would arrive at the ingress independently but at approximately the same time. Under some circumstances this might cause congestion at or around the ingress LSR. Therefore, the procedures described in this document provide a mechanism that allows the responders to randomly delay (or jitter) their responses so that the chances of swamping the ingress are reduced.

Further, the procedures in this document allow the initiator to limit the scope of an LSP Ping echo request (ping or traceroute mode) to one specific intended egress.

The scalability issues surrounding LSP Ping for P2MP MPLS LSPs may be addressed by other mechanisms such as [MCAST-CV] that utilize the LSP Ping procedures in this document to provide bootstrapping mechanisms as described in Section 2.3.

LSP Ping can be used to periodically ping a P2MP MPLS LSP to ensure connectivity to any or all of the egresses. If the ping fails, the operator or an automated process can then initiate a traceroute to determine where the fault is located within the network. A traceroute may also be used periodically to verify that data plane forwarding matches the control plane state; however, this places an increased burden on transit LSRs and should be used infrequently and with caution.

2.3 Bootstrapping Other OAM Procedures Using LSP Ping

[MPLS-BFD] describes a process where LSP Ping [RFC4379] is used to bootstrap the Bidirectional Forwarding Detection (BFD) mechanism [BFD] for use to track the liveness of an MPLS LSP. In particular BFD can be used to detect a data plane failure in the forwarding path of an MPLS LSP.

Requirements for MPLS P2MP LSPs extend to hundreds or even thousands of endpoints. If a protocol required explicit acknowledgments to each probe for connectivity verification, the response load at the root would be overwhelming.

A more scalable approach to monitoring P2MP LSP connectivity is described in [MCAST-CV]. It relies on using the MPLS echo request and echo response messages of LSP Ping [RFC4379] to bootstrap the monitoring mechanism in a manner similar to [MPLS-BFD]. The actual monitoring is done using a separate process defined in [MCAST-CV].

Note that while the approach described in [MCAST-CV] was developed in response to the multicast scalability problem, it can be applied to P2P LSPs as well.
3. Operation of LSP Ping for a P2MP LSP

This section describes how LSP Ping is applied to P2MP MPLS LSPs. It covers the mechanisms and protocol fields applicable to both ping mode and traceroute mode. It explains the responsibilities of the initiator (ingress), transit nodes, and receivers (egresses).

3.1. Identifying the LSP Under Test

3.1.1. Identifying a P2MP MPLS TE LSP

[RFC4379] defines how an MPLS TE LSP under test may be identified in an echo request. A Target FEC Stack TLV is used to carry either an RSVP IPv4 Session or an RSVP IPv6 Session sub-TLV.

In order to identify the P2MP MPLS TE LSP under test, the echo request message MUST carry a Target FEC Stack TLV, and this MUST carry exactly one of two new sub-TLVs: either an RSVP P2MP IPv4 Session sub-TLV or an RSVP P2MP IPv6 Session sub-TLV. These sub-TLVs carry fields from the RSVP-TE P2MP Session and Sender-Template objects [RFC4875] and so provide sufficient information to uniquely identify the LSP.

The new sub-TLVs are assigned sub-type identifiers as follows, and are described in the following sections.

<table>
<thead>
<tr>
<th>Sub-Type #</th>
<th>Length</th>
<th>Value Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBD</td>
<td>20</td>
<td>RSVP P2MP IPv4 Session</td>
</tr>
<tr>
<td>TBD</td>
<td>56</td>
<td>RSVP P2MP IPv6 Session</td>
</tr>
</tbody>
</table>

3.1.1.1. RSVP P2MP IPv4 Session Sub-TLV

The format of the RSVP P2MP IPv4 Session sub-TLV value field is specified in the following figure. The value fields are taken from the definitions of the P2MP IPv4 LSP Session Object and the P2MP IPv4 Sender-Template Object in [RFC4875]. Note that the Sub-Group ID of the Sender-Template is not required.


3.1.1.2. RSVP P2MP IPv6 Session Sub-TLV

The format of the RSVP P2MP IPv6 Session sub-TLV value field is specified in the following figure. The value fields are taken from the definitions of the P2MP IPv6 LSP Session Object, and the P2MP IPv6 Sender-Template Object in [RFC4875]. Note that the Sub-Group ID of the Sender-Template is not required.

3.1.2. Identifying a Multicast LDP LSP

[RFC4379] defines how a P2P LDP LSP under test may be identified in an echo request. A Target FEC Stack TLV is used to carry one or more sub-TLVs (for example, an IPv4 Prefix FEC sub-TLV) that identify the LSP.

In order to identify a multicast LDP LSP under test, the echo request
message MUST carry a Target FEC Stack TLV, and this MUST carry
exactly one new sub-TLV: the Multicast LDP FEC Stack sub-TLV. This
sub-TLV uses fields from the multicast LDP messages [P2MP-LDP] and so
provides sufficient information to uniquely identify the LSP.

The new sub-TLV is assigned a sub-type identifier as follows, and
is described in the following section.

<table>
<thead>
<tr>
<th>Sub-Type #</th>
<th>Length</th>
<th>Value Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBD</td>
<td>Variable</td>
<td>Multicast LDP FEC Stack</td>
</tr>
</tbody>
</table>

3.1.2.1. Multicast LDP FEC Stack Sub-TLV

The format of the Multicast LDP FEC Stack sub-TLV is shown below.

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Address Family          | Address Length | Root LSR Addr |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Root LSR Address (Cont.)|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Opaque Length           | Opaque Value ... |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Address Family

A two octet quantity containing a value from ADDRESS FAMILY NUMBERS in [IANA-PORT] that encodes the address family for the Root LSR Address.

Address Length

The length of the Root LSR Address in octets.

Root LSR Address

An address of the LSR at the root of the P2MP LSP encoded according to the Address Family field.
Opaque Length

The length of the Opaque Value, in octets.

Opaque Value

An opaque value elements of which uniquely identifies the P2MP LSP in the context of the Root LSR.

If the Address Family is IPv4, the Address Length MUST be 4. If the Address Family is IPv6, the Address Length MUST be 16. No other Address Family values are defined at present.

3.2. Ping Mode Operation

3.2.1. Controlling Responses to LSP Pings

As described in Section 2.2, it may be desirable to restrict the operation of LSP Ping to a single egress. Since echo requests are forwarded through the data plane without interception by the control plane (compare with traceroute mode), there is no facility to limit the propagation of echo requests, and they will automatically be forwarded to all (reachable) egresses.

However, the intended egress under test can be identified by the inclusion of a P2MP Responder Identifier TLV containing an IPv4 P2MP Responder Identifier sub-TLV or an IPv6 P2MP Responder Identifier sub-TLV. The P2MP Responder Identifier TLV SHOULD contain precisely one sub-TLV. If the TLV contains no sub-TLVs it SHOULD be processed as if the whole TLV were absent (causing all egresses to respond as described below). If the TLV contains more than one sub-TLV, the first MUST be processed as described in this document, and subsequent sub-TLVs SHOULD be ignored.

An initiator may indicate that it wishes all egresses to respond to an echo request by omitting the P2MP Responder Identifier TLV.

Note that the ingress of a multicast LDP LSP will not know the identities of the egresses of the LSP except by some external means such as running P2MP LSP Ping to all egresses.

3.2.2. Ping Mode Egress Procedures

An egress node is RECOMMENDED to rate limit its receipt of echo request messages as described in [RFC4379]. After rate limiting, an egress node that receives an echo request carrying an RSVP P2MP IPv4 Session sub-TLV, an RSVP P2MP IPv6 Session sub-TLV, or a Multicast LDP FEC Stack sub-TLV MUST determine whether it is an intended egress of the P2MP LSP in question by checking with the control plane. If it is not supposed to be an egress, it MUST respond according to the...
setting of the Response Type field in the echo message following the rules defined in [RFC4379].

If the egress node that receives an echo request and allows it through its rate limiting is an intended egress of the P2MP LSP, the node MUST check to see whether it is an intended Ping recipient. If a P2MP Responder Identifier TLV is present and contains an address that indicates any address that is local to the node, the node MUST respond according to the setting of the Response Type field in the echo message following the rules defined in [RFC4379]. If the P2MP Responder Identifier TLV is present, but does not identify the egress node, it MUST NOT respond to the echo request. If the P2MP Responder Identifier TLV is not present (or, in the error case, is present, but does not contain any sub-TLVs), but the egress node that received the echo request is an intended egress of the LSP, the node MUST respond according to the setting of the Response Type field in the echo message following the rules defined in [RFC4379].

3.2.3. Jittered Responses

The initiator (ingress) of a ping request MAY request the responding egress to introduce a random delay (or jitter) before sending the response. The randomness of the delay allows the responses from multiple egresses to be spread over a time period. Thus this technique is particularly relevant when the entire LSP tree is being pinged since it helps prevent the ingress (or nearby routers) from being swamped by responses, or from discarding responses due to rate limits that have been applied.

It is desirable for the ingress to be able to control the bounds within which the egress delays the response. If the tree size is small, only a small amount of jitter is required, but if the tree is large, greater jitter is needed. The ingress informs the egresses of the jitter bound by supplying a value in a new TLV (the Echo Jitter TLV) carried on the echo request message. If this TLV is present, the responding egress MUST delay sending a response for a random amount of time between zero seconds and the value indicated in the TLV. If the TLV is absent, the responding egress SHOULD NOT introduce any additional delay in responding to the echo request.

LSP ping SHOULD NOT be used to attempt to measure the round-trip time for data delivery. This is because the LSPs are unidirectional, and the echo response is often sent back through the control plane. The timestamp fields in the echo request/response MAY be used to deduce some information about delivery times and particularly the variance in delivery times.

The use of echo jittering does not change the processes for gaining information, but note that the responding egress MUST set the value in the Timestamp Received fields before applying any delay.
It is RECOMMENDED that echo response jittering is not used except in the case of P2MP LSPs. If the Echo Jitter TLV is present in an echo request for any other type of TLV, the responding egress MAY apply the jitter behavior described here.

### 3.2.4. P2MP Responder Identifier TLV and Sub-TLVs

A new TLV is defined for inclusion in the Echo request message.

The P2MP Responder Identifier TLV is assigned the TLV type value TBD and is encoded as follows.

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Type=TBD(P2MP Responder ID TLV)|       Length = Variable       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                          Sub-TLVs                             ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Sub-TLVs:

Zero, one or more sub-TLVs as defined below.

If no sub-TLVs are present, the TLV MUST be processed as if it were absent. If more than one sub-TLV is present the first MUST be processed as described in this document, and subsequent sub-TLVs SHOULD be ignored.

The P2MP Responder Identifier TLV only has meaning on an echo request message. If present on an echo response message, it SHOULD be ignored.

Two sub-TLVs are defined for inclusion in the P2MP Responder Identifier TLV carried on the echo request message. These are:

<table>
<thead>
<tr>
<th>Sub-Type #</th>
<th>Length</th>
<th>Value Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>IPv4 P2MP Responder Identifier</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>IPv6 P2MP Responder Identifier</td>
</tr>
</tbody>
</table>

The value of an IPv4 P2MP Responder Identifier consists of four octets of an IPv4 address. The IPv4 address is in network byte order.

The value of an IPv6 P2MP Responder Identifier consists of sixteen octets of an IPv6 address. The IPv6 address is in network byte order.
3.2.5. Echo Jitter TLV

A new TLV is defined for inclusion in the Echo request message.

The Echo Jitter TLV is assigned the TLV type value TBD and is encoded as follows.

```
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type = TBD (Jitter TLV)  |          Length = 4           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Jitter time                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Jitter time:

This field specifies the upper bound of the jitter period that should be applied by a responding node to determine how long to wait before sending an echo response. A responding node SHOULD wait a random amount of time between zero seconds and the value specified in this field.

Jitter time is specified in milliseconds.

The Echo Jitter TLV only has meaning on an echo request message. If present on an echo response message, it SHOULD be ignored.

3.2.6. Echo Response Reporting

Echo response messages carry return codes and subcodes to indicate the result of the LSP Ping (when the ping mode is being used) as described in [RFC4379].

When the responding node reports that it is an egress, it is clear that the echo response applies only to the reporting node. Similarly, when a node reports that it does not form part of the LSP described by the FEC (i.e. their is a misconnection) then the echo response applies to the reporting node.

However, it should be noted that an echo response message that reports an error from a transit node may apply to multiple egress nodes (i.e. leaves) downstream of the reporting node. In the case of the Ping mode of operation, it is not possible to correlate the reporting node to the affected egresses unless the shape of the P2MP tree is already known, and it may be necessary to use the Traceroute mode of operation (see Section 3.3) to further diagnose the LSP.

Note also that a transit node may discover an error but also determine that while it does lie on the path of the LSP under test,
it does not lie on the path to the specific egress being tested. In this case, the node SHOULD NOT generate an echo response.

A reporting node that is a branch node may need to report multiple different errors (for different downstream branches). This is discussed further in Section 3.3.6.

3.3. Traceroute Mode Operation

The traceroute mode of operation is described in [RFC4379]. Like other traceroute operations, it relies on the expiration of the TTL of the packet that carries the echo request. Echo requests may include a Downstream Mapping TLV, and when the TTL expires the echo request is passed to the control plane on the transit node which responds according to the Response Type in the message. A responding node fills in the fields of the Downstream Mapping TLV to indicate the downstream interfaces and labels used by the reported LSP from the responding node. In this way, by successively sending out echo requests with increasing TTLs, the ingress may gain a picture of the path and resources used by an LSP up to the point of failure when no response is received, or an error response is generated by a node where the control plane does not expect to be handling the LSP.

This mode of operation is equally applicable to P2MP MPLS TE LSPs as described in the following sections.

The traceroute mode can be applied to all destinations of the P2MP tree just as in the ping mode. In the case of P2MP MPLS TE LSPs, the traceroute mode can also be applied to individual traceroute targets identified by the presence of a P2MP Responder Identifier TLV. These targets may be egresses or transit nodes. However, since a transit node of a multicast LDP LSP is unable to determine whether it lies on the path to any one destination or any other transit node, the traceroute mode limited to specific nodes of such an LSP MUST NOT be used.

Note that the addresses specified in the P2MP Responder Identifier TLV need not be egresses: they could be transit nodes on the LSP. The processing rules here and in the following sections apply equally to egress and transit nodes.

In the absence of a P2MP Responder Identifier TLV, the echo request is asking for traceroute information applicable to all egresses.

The echo response jitter technique described for the ping mode is equally applicable to the traceroute mode and is not additionally described in the procedures below.
3.3.1. Traceroute Responses at Non-Branch Nodes

When the TTL for the MPLS packet carrying an echo request expires the packet MUST be passed to the control plane as specified in [RFC4379]. If the LSP under test is a multicast LDP LSP and if the echo request carries a P2MP Responder Identifier TLV the node MUST treat the echo request as malformed and MUST process it according to the rules specified in [RFC4379].

Otherwise, the node MUST NOT return an echo response unless the responding node lies on the path of the P2MP LSP to the node (egress or transit) identified by the P2MP Responder Identifier TLV carried on the request, or if no such sub-TLV is present.

If sent, the echo response MUST identify the next hop of the path of the LSP in the data plane by including a Downstream Mapping TLV as described in [RFC4379].

3.3.1.1. Correlating Traceroute Responses

When traceroute is being simultaneously applied to multiple responders (e.g., egresses), it is important that the ingress should be able to correlate the echo responses with the branches in the P2MP tree. Without this information the ingress will be unable to determine the correct ordering of transit nodes. One possibility is for the ingress to poll the path to each responder in turn, but this may be inefficient, undesirable, or (in the case of multicast LDP LSPs) illegal.

The Downstream Mapping TLV that MUST be included in the echo response indicates the next hop from each responding node, and this information supplied by a non-branch node can be pieced together by the ingress to reconstruct the P2MP tree although it may be necessary to refer to the routing information distributed by the IGP to correlate next hop addresses and node reporting addresses in subsequent echo responses.

In order to facilitate more easy correlation of echo responses, the Downstream Mapping TLV can also contain Multipath Information as described in [RFC4379] to identify to which responders (transit nodes or egresses) the echo response applies. This information:

- Cannot be present when the information is not known by the responding node. For example, for a multicast LDP LSP, the branch node will not know through normal LDP signaling which leaf nodes lie on which downstream branch.

- SHOULD be present when the information is known by the responding node. That is for P2MP MPLS TE LSPs when the echo request applies to all egresses or to a specific single transit node or egress.
The format of the information in the Downstream Mapping TLV for P2MP MPLS LSPs is described in section 3.3.5.

3.3.2. Traceroute Responses at Branch Nodes

A branch node may need to identify more than one downstream interface in a traceroute echo response if some of the nodes identified in the P2MP Responder Identifier TLV that are being traced lie on different branches. This will always be the case for any branch node if all egresses are being traced.

[RFC4379] describes how multiple Downstream Mapping TLVs should be included in an echo response, each identifying exactly one downstream interface that is applicable to the LSP.

A branch node MUST follow the procedures described in Section 3.3.1 to determine whether it should respond to an echo request. The branch node MUST add a Downstream Mapping TLV (or Downstream Detailed Mapping TLV - see Section 3.3.7) to the echo response for each outgoing branch that it reports, but it MUST NOT report branches that do not lie on the path to one of the destinations being traced. Thus a branch node may sometimes only need to respond with a single Downstream Mapping TLV; for example, consider the case where the traceroute is directed to only a single egress node. Therefore, the presence of only one Downstream Mapping TLV in an echo response does not guarantee that the reporting node is not a branch node.

To report on its branching properties on a particular LSP, the responding node MAY include an optional TLV called the Node Properties TLV. This new TLV (see Section 3.3.2.1) can carry sub-TLVs, one of which (the Branching Properties sub-TLV - see Section 3.3.2.2) allows the reporting node to describe the branching characteristics of the LSP at the reporting node.

3.3.2.1. Node Properties TLV

A new TLV has been added to the set of optional TLVs that may be carried on an echo response message.

<table>
<thead>
<tr>
<th>Type #</th>
<th>Value Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBD</td>
<td>Node properties</td>
</tr>
</tbody>
</table>

The Node Properties TLV MAY be included in an echo response message. If more than one such TLV is present, the first MUST be processed and subsequent instances SHOULD be ignored.

The Node Properties TLV is used to report characteristics of the reporting node, and the LSP at that node. This distinguishes it from...
the Downstream Mapping TLV [RFC4379] and the Downstream Detailed Mapping TLV [DDMT] used to report characteristics of specific out-segments an LSP.

The Node Properties TLV is a standard LSP Ping TLV as defined in [RFC4379]. It has the following format.

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      First Sub-TLV                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Further Sub-TLVs                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The content of the Node Properties TLV is a series of one or more sub-TLVs. The Node Properties TLV SHOULD contain one or more sub-TLVs and MUST be ignored if there are no sub-TLVs present.

Each sub-TLV consists of the following fields as per [RFC4379]:

- Two octet Type field: A value indicating the sub-TLV type.
- Two octet Length field: A value indicating the total length of the Value field.
- A Value field carrying the data of the sub-TLV. The content of the Value field is padded to a four byte boundary with zero-filled octets so that the Length field is always a multiple of 4.

### 3.3.2.2. Branching Properties Sub-TLV

This document defines the Branching Properties sub-TLV carried in the Node Properties TLV. The Branching Properties sub-TLV is optional. If more than one such sub-TLV is found in a Node Properties TLV, the first MUST be processed and subsequent instances SHOULD be ignored. The sub-TLV may be used for P2MP and P2P LSPs.

The Branching Properties sub-TLV is formed as described in Section 3.3.2.1. The Value field has the following format.

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Downstream Branch Count     |         Egress Count         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
Downstream Branch Count

This field reports the number of downstream branches from the reporting node for this LSP. The number may be zero for an egress, one for a non-branch node, and more than one for a branch node. Note that the value reported here may be greater than the number of Downstream Mapping TLVs present in the echo response message since those TLVs only report on the specific egresses queried. This value may be of use in detecting faults caused by delay introduced by the data replication mechanism at branch nodes.

Egress Count

This field reports the number of egresses local to the reporting node. Thus, for non-zero values the reporting node is either a leaf or a bud. When the value reported is non-zero, the reporting node MAY also include an Egress Address Sub-TLV for each local egress (see Section 3.3.2.3).

For example, a branch node that has two downstream next hops on the LSP and that also delivers payload data to one local egress would set the two fields to 2 and 1 respectively.

3.3.2.3. Egress Address Sub-TLV

This document defines the IPv4 and IPv6 Egress Address sub-TLVs carried in the Node Properties TLV. These TLVs are optional, and more than one instance of the sub-TLVs may legitimately be present.

The Egress Address sub-TLVs are formed as described in Section 3.3.2.1. The Value field has the following formats.

IPv4 Egress Address Sub-TLV

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       IPv4 Egress Address                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IPv6 Egress Address Sub-TLV

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       IPv6 Egress Address                     |
|                          (16 octets)                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
The Egress Address sub-TLVs are optional. They MAY be included in a Node Properties TLV when reporting node is an egress (leaf or bud) for the LSP being tested. The sub-TLV may be used for P2MP and P2P LSPs.

When one or more Egress Address sub-TLVs are present and the Branch Properties sub-TLV is also present, the value of the Egress Count field in the Branch Properties sub-TLV SHOULD be the same as the number of Egress Address sub-TLVs.

The address contained in an Egress Address sub-TLV is the egress address to which the data is delivered. If there is just one egress and if the egress address is the same as the local node address carried in the main echo response message, both the Branching Properties sub-TLV and the Egress Address sub-TLV MAY be omitted as in legacy LSP Ping implementations.

### 3.3.2.4. Correlating Traceroute Responses

Just as with non-branches, it is important that the echo responses from branch nodes provide correlation information that will allow the ingress to work out to which branch of the LSP the response applies.

The P2MP tree can be determined by the ingress using the identity of the reporting node and the next hop information from the previous echo response, just as with echo responses from non-branch nodes.

As with non-branch nodes, in order to facilitate more easy correlation of echo responses, the Downstream Mapping TLV can also contain Multipath Information as described in [RFC4379] to identify to which nodes the echo response applies. This information:

- Cannot be present when the information is not known by the responding node. For example, for a multicast LDP LSP, the branch node will not know through normal LDP signaling which leaf nodes lie on which downstream branch.

- SHOULD be present when the information is known by the responding node. That is for P2MP MPLS TE LSPs when the echo request applies to all egresses or to a specific single transit node or egress.

The format of the information in the Downstream Mapping TLV for P2MP MPLS LSPs is described in section 3.3.5.

### 3.3.3. Traceroute Responses at Bud Nodes

Some nodes on a P2MP MPLS LSP may be egresses, but also have downstream node. Such nodes are known as bud nodes [RFC4461].

A bud node MUST respond to a traceroute echo request just as a branch
node would, but it MUST also indicate to the ingress that it is an egress in its own right. The issue to be resolved here is how to indicate that the reporting node is an egress when it is also providing one or more Downstream Mapping TLVs that indicate that it has downstream neighbors.

This is achieved by the inclusion of a Node Properties TLV with a Branch Properties sub-TLV indicating the number of local egresses and the number of downstream branches. The bud node MAY also include one or more Egress Address sub-TLVs in the Node Properties TLV to report on the local egresses.

3.3.4. Non-Response to Traceroute Echo Requests

The nature of P2MP MPLS TE LSPs in the data plane means that traceroute echo requests may be delivered to the control plane of nodes that must not reply to the request because, although they lie on the P2MP tree, they do not lie on the path to the node that is being traced.

Thus, a node on a P2MP MPLS LSP MUST NOT respond to an echo request when the TTL has expired if any of the following applies:

- The Reply Type indicates that no reply is required [RFC4379]
- There is a P2MP Responder Identifier TLV present on the echo request (which means that the LSP is a P2MP MPLS TE LSP), but the address does not identify a node that is reached through this node for this particular P2MP MPLS LSP.

Note that when no response to an echo request is received by the ingress (perhaps because the transit node has failed, or perhaps because the transit node does not support LSP Ping), then as per [RFC4379] the subsequent echo request (with a larger TTL) SHOULD be sent with Downstream Mapping TLV "Downstream IP Address" field set to the ALLROUTERS multicast address until a reply is received with a Downstream Mapping TLV.

3.3.5. Additions to Downstream Mapping Multipath Information

A new value for the Multipath Type is defined to indicate that the reported Multipath Information applies to a P2MP MPLS TE LSP and may contain a list of node identifiers that indicate the egress nodes and (in the case where the P2MP Responder Identifier TLV was used on the echo request to identify non-egress nodes) transit nodes that can be reached through the reported interface. This Multipath Type MUST NOT be used for a multicast LDP LSP.
Note that a list of nodes may include IPv4 and IPv6 identifiers since these may be mixed in the P2MP MPLS TE LSP.

The Multipath Length field continues to identify the length of the Multipath Information just as in [RFC4379] (that is, not including the downstream labels), and the downstream label (or potential stack thereof) is also handled just as in [RFC4379]. The format of the Multipath Information for a Multipath Type of P2MP responders is as follows.

<table>
<thead>
<tr>
<th>Address Type</th>
<th>Responder Address (4 or 16 octets)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(continued)</td>
<td></td>
</tr>
</tbody>
</table>

Address Type

This field indicates whether the address that follows is an IPv4 or IPv6 address, and so implicitly encodes the length of the address.

Two values are defined and mirror the values used in the Address Type field of the Downstream Mapping TLV itself.

<table>
<thead>
<tr>
<th>Type #</th>
<th>Address Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IPv4</td>
</tr>
<tr>
<td>3</td>
<td>IPv6</td>
</tr>
</tbody>
</table>

Responder Address

An egress or transit node of this P2MP MPLS TE LSP that is reached through the interface indicated by the Downstream Mapping TLV and for which the traceroute echo request was enquiring.

Note that padding to ensure that the whole Multipath information is aligned to a four-octet boundary is applied only after the last responder address in the list. That is, each successive Address Type follows on immediately after the previous Responder Address.
3.3.6. Echo Response Reporting

Echo responses are generated in response to traceroute echo requests at transit, branch, and bud nodes as described in Sections 3.3.1, 3.3.2, and 3.3.3, while egress responses are as described in [RFC4379].

Note, however, that a branch or bud node may have multiple downstream branches, and a transit node may have multiple downstream egresses (reached on the same branch). It may be the case that different conditions need to be reported for different branches or egresses. The echo response message defined in [RFC4379] has space for only a single return code and subcode pair, so where more than one return condition is reported by a single node it acts as follows.

- It SHOULD use the Downstream Detailed Mapping TLV [DDMT] in place of the Downstream Mapping TLV, and encode the return code as described in Section 3.3.6.1.

- It MAY report each condition in a separate echo response in which case MUST limit the downstream mapping information on each echo response to those branches/egresses to which the response applies. The use of multiple echo response messages to report errors might cause issues for an initiator that does not know how many responses it should wait for. For that reason, multiple messages should be used with care.

3.3.6.1. Reporting Multiple Conditions Using The DDM TLV

When multiple different return codes are indicated on a single echo response message they MUST be carried in separate instances on the Downstream Detailed Mapping (DDM) TLV [DDMT]. That is, each instance of a DDM TLV carries one return code, and all information carried in that TLV MUST be limited to branches/egresses to which that return code applies. However, more than one DDM TLV on the same echo response MAY carry the same return code.

The echo response message still carries a Return Code and a Return Subcode field. In order to clearly indicate that the relevant return codes are carried in the DDM TLV, a new return code is defined to be carried in the Return Code field of the echo response message as follows:

<table>
<thead>
<tr>
<th>Value</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBD</td>
<td>See DDM TLV for more details</td>
</tr>
</tbody>
</table>

The Return Subcode for this Return Code MUST be set to zero and MUST be ignored.

Yasukawa and Farrel
The DDM TLV is defined as carrying a set of sub-TLVs. A new sub-TLV, the Return Code sub-TLV, is defined here to carry a return code and return subcode.

<table>
<thead>
<tr>
<th>Return Code</th>
<th>Return Subcode</th>
<th>Reserved</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+---------------+----------------+---------</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The length of the Return Code sub-TLV is 8.

**Return Code**
As defined for inclusion in the echo response message in [RFC4379].

**Return Subcode**
As defined for inclusion in the echo response message in [RFC4379].

**Reserved**
SHOULD be set to zero on transmission and MUST be ignored on receipt.

If the Return Code of the echo response message is not set to "See DDM TLV for more details" then any Return Code sub-TLV present in a DDM TLV SHOULD be ignored.

If the Return Code of the echo response message is set to "See DDM TLV for more details" then a Return Code sub-TLV MUST be present in each DDM TLV. Subsequent Return Code sub-TLVs present in the same DDM TLV SHOULD be ignored.

4. Operation of LSP Ping for Bootstrapping Other OAM Mechanisms

Bootstrapping of other OAM procedures can be achieved using the MPLS Echo Request/Response messages. The LSP(s) under test are identified using the RSVP P2MP IPv4 or IPv6 Session sub-TLVs (see Section 3.1.1) or the Multicast LDP FEC Stack sub-TLV (see Section 3.1.2).

Other sub-TLVs may be defined in other specifications to indicate the OAM procedures being bootstrapped, and to describe the bootstrap parameters. Further details of the bootstrapping processes and the bootstrapped OAM processes are described in other documents. For example, see [MPLS-BFD] and [MCAST-CV].
5. Non-compliant Routers

If an egress for a P2MP LSP does not support MPLS LSP ping, then no reply will be sent, resulting in a "false negative" result. There is no protection for this situation, and operators may wish to ensure that end points for P2MP LSPs are all equally capable of supporting this function. Alternatively, the traceroute option can be used to verify the LSP nearly all the way to the egress, leaving the final hop to be verified manually.

If, in "traceroute" mode, a transit node does not support LSP ping, then no reply will be forthcoming from that node for some TTL, say n. The node originating the echo request SHOULD continue to send echo request with TTL=n+1, n+2, ..., n+k to probe nodes further down the path. In such a case, the echo request for TTL > n SHOULD be sent with Downstream Mapping TLV "Downstream IP Address" field set to the ALLROUTERS multicast address as described in Section 3.3.4 until a reply is received with a Downstream Mapping TLV.

6. OAM Considerations

The procedures in this document provide OAM functions for P2MP MPLS LSPs and may be used to enable bootstrapping of other OAM procedures.

In order to be fully operational several considerations must be made.

- Scaling concerns dictate that only cautious use of LSP Ping should be made. In particular, sending an LSP Ping to all egresses of a P2MP MPLS LSP could result in congestion at or near the ingress when the responses arrive.

Further, incautious use of timers to generate LSP Ping echo requests either in ping mode or especially in traceroute may lead to significant degradation of network performance.

- Management interfaces should allow an operator full control over the operation of LSP Ping. In particular, it SHOULD provide the ability to limit the scope of an LSP Ping echo request for a P2MP MPLS LSP to a single egress.

Such an interface SHOULD also provide the ability to disable all active LSP Ping operations to provide a quick escape if the network becomes congested.

- A MIB module is required for the control and management of LSP Ping operations, and to enable the reported information to be inspected.

There is no reason to believe this should not be a simple extension of the LSP Ping MIB module used for P2P LSPs.
7. IANA Considerations

7.1. New Sub-TLV Types

Three new sub-TLV types are defined for inclusion within the LSP Ping [RFC4379] Target FEC Stack TLV (TLV type 1).

IANA is requested to assign sub-type values to the following sub-TLVs from the "Multiprotocol Label Switching Architecture (MPLS) Label Switched Paths (LSPs) Parameters - TLVs" registry, "TLVs and sub-TLVs" sub-registry.

- RSVP P2MP IPv4 Session (see Section 3.1.1). Suggested value 17.
- RSVP P2MP IPv6 Session (see Section 3.1.1). Suggested value 18.
- Multicast LDP FEC Stack (see Section 3.1.2). Suggested value 19.

7.2. New Multipath Type

Section 3.3 of [RFC4379] defines a set of values for the LSP Ping Multipath Type. These values are currently not tracked by IANA.

A new value for the LSP Ping Multipath Type is defined in Section 3.3.5 of this document to indicate that the reported Multipath Information applies to a P2MP MPLS TE LSP.

IANA is requested to create a new registry as follows:

"Multiprotocol Label Switching Architecture (MPLS) Label Switched Paths (LSPs) - Multipath Types"

<table>
<thead>
<tr>
<th>Key</th>
<th>Type</th>
<th>Multipath Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>no multipath</td>
<td>Empty (Multipath Length = 0)</td>
</tr>
<tr>
<td>2</td>
<td>IP address</td>
<td>IP addresses</td>
</tr>
<tr>
<td>4</td>
<td>IP address range</td>
<td>low/high address pairs</td>
</tr>
<tr>
<td>8</td>
<td>Bit-masked IP address set</td>
<td>IP address prefix and bit mask</td>
</tr>
<tr>
<td>9</td>
<td>Bit-masked label set</td>
<td>Label prefix and bit mask</td>
</tr>
<tr>
<td>xx</td>
<td>P2MP responder IP</td>
<td>List of P2MP responders</td>
</tr>
</tbody>
</table>

A suggested value of xx is 16.

New values from this registry are to be assigned only by Standards Action.
7.3. New TLVs

Three new LSP Ping TLV types are defined for inclusion in LSP Ping messages.

IANA is requested to assign a new value from the "Multi-Protocol Label Switching Architecture (MPLS) Label Switched Paths (LSPs) Parameters - TLVs" registry, "TLVs and sub-TLVs" sub-registry as follows using a Standards Action value.

- P2MP Responder Identifier TLV (see Section 3.2.4) is a mandatory TLV. Suggested value 11.
  - Type 1: IPv4 P2MP Responder Identifier (see Section 3.2.4)
  - Type 2: IPv6 P2MP Responder Identifier (see Section 3.2.4)

- Echo Jitter TLV (see Section 3.2.5) is a mandatory TLV. Suggested value 12.

- Node Properties TLV (see Section 3.2.2.1) is an optional TLV. Suggested value 32768.
  - Type 1: IPv4 Egress Address
  - Type 2: IPv6 Egress Address
  - Type 3: Branch Properties

7.4. New Return Code

A new Return Code is defined in Section 3.3.6.1.

IANA is requested to assign a new Return Code value for the "Multi-Protocol Label Switching (MPLS) Label Switched Paths (LSPs) Parameters" registry, "Return Codes" sub-registry as follows using a Standards Action value.

<table>
<thead>
<tr>
<th>Value</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBD</td>
<td>See DDM TLV for more details</td>
</tr>
</tbody>
</table>

Suggested value 14.

7.5. New Sub-TLV Value for the Downstream Detailed Mapping TLV

[DDMT] defines a TLV called the Downstream Detailed Mapping TLV and requests IANA to maintain a registry of sub-TLVs that it can carry.

Section 3.3.6.1 of this document defines a new sub-TLV.

IANA is requested to assign a TLV type value as follows using a Standards Action value from the range 0-32767.
8. Security Considerations

This document does not introduce security concerns over and above those described in [RFC4379]. Note that because of the scalability implications of many egresses to P2MP MPLS LSPs, there is a stronger concern to regulate the LSP Ping traffic passed to the control plane by the use of a rate limiter applied to the LSP Ping well-known UDP port. Note that this rate limiting might lead to false positives.

9. Acknowledgements

The authors would like to acknowledge the authors of [RFC4379] for their work which is substantially re-used in this document. Also thanks to the members of the MBONED working group for their review of this material, to Daniel King and Mustapha Aissaoui for their review, and to Yakov Rekhter for useful discussions.

The authors would like to thank Vanson Lim, Danny Prairie, Reshad Rahman, Ben Niven-Jenkins, Hannes Gredler, Nitin Bahadur, Tetsuya Murakami and Michael Hua for their comments and suggestions.

10. Intellectual Property Considerations

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.
11. Normative References


12. Informative References

[RFC792] Postel, J., "Internet Control Message Protocol", RFC 792.


13. Authors’ Addresses

Seisho Yasukawa
NTT Corporation
(R&D Strategy Department)
3-1, Otemachi 2-Chome Chiyodaku, Tokyo 100-8116 Japan
Phone: +81 3 5205 5341
Email: s.yasukawa@hco.ntt.co.jp

Adrian Farrel
Old Dog Consulting
EMail: adrian@olddog.co.uk

Zafar Ali
Cisco Systems Inc.
2000 Innovation Drive
Kanata, ON, K2K 3E8, Canada.
Phone: 613-889-6158
Email: zali@cisco.com

Bill Fenner
Arastra, Inc.
275 Middlefield Rd.
Suite 50
Menlo Park, CA 94025
Email: fenner@fenron.com

George Swallow
Cisco Systems, Inc.
1414 Massachusetts Ave
Boxborough, MA 01719
Email: swallow@cisco.com

Thomas D. Nadeau
British Telecom
BT Centre
81 Newgate Street
EC1A 7AJ
London
Email: tom.nadeau@bt.com
14. Full Copyright Statement

Copyright (C) The IETF Trust (2008).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.