

Network Working Group  
IETF Internet Draft  
Proposed Status: Standards Track  
Expires: October 2005

Seisho Yasukawa (NTT)  
Adrian Farrel (Olddog Consulting)  
Zafar Ali (Cisco Systems)  
Bill Fenner (AT&T Research)

April 2005

## **Detecting Data Plane Failures in Point-to-Multipoint MPLS Traffic Engineering - Extensions to LSP Ping**

[draft-yasukawa-mpls-p2mp-lsp-ping-02.txt](#)

Status of this Memo

This document is an Internet-Draft and is subject to all provisions of [Section 3 of RFC 3667](#). 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/lid-abstracts.txt>.

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

Copyright Notice

Copyright (C) The Internet Society (2005). All Rights Reserved.

Abstract

Recent proposals have extended the scope of Multi-Protocol Label Switching (MPLS) traffic engineered Label Switched Paths (TE LSPs) to encompass point-to-multipoint (P2MP) TE 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 recognised 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 TE LSPs. This documents does not replace any of the mechanism of LSP Ping, but clarifies their applicability to P2MP MPLS TE LSPs, and extends the techniques and mechanisms of LSP Ping to the P2MP TE environment.

## 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

<a href="#">1.</a>	<a href="#">Introduction</a>	<a href="#">3</a>
<a href="#">1.1</a>	<a href="#">Design Considerations</a>	<a href="#">3</a>
<a href="#">2.</a>	<a href="#">Notes on Motivation</a>	<a href="#">4</a>
<a href="#">2.1</a>	<a href="#">Basic Motivations for LSP Ping</a>	<a href="#">4</a>
<a href="#">2.2</a>	<a href="#">Motivations for LSP Ping for P2MP TE LSPs</a>	<a href="#">5</a>
<a href="#">3.</a>	<a href="#">Operation of LSP Ping for a P2MP TE LSP</a>	<a href="#">6</a>
<a href="#">3.1</a>	<a href="#">Identifying the LSP Under Test</a>	<a href="#">6</a>
<a href="#">3.1.1</a>	<a href="#">RSVP P2MP IPv4 Session Sub-TLV</a>	<a href="#">6</a>
<a href="#">3.1.2</a>	<a href="#">RSVP P2MP IPv6 Session Sub-TLV</a>	<a href="#">7</a>
<a href="#">3.2</a>	<a href="#">Ping Mode Operation</a>	<a href="#">7</a>
<a href="#">3.2.1</a>	<a href="#">Controlling Responses to LSP Pings</a>	<a href="#">7</a>
<a href="#">3.2.2</a>	<a href="#">P2MP Egress Identifier sub-TLVs</a>	<a href="#">9</a>
<a href="#">3.2.3</a>	<a href="#">Echo Jitter TLV</a>	<a href="#">9</a>
<a href="#">3.3</a>	<a href="#">Traceroute Mode Operation</a>	<a href="#">10</a>
<a href="#">3.3.1</a>	<a href="#">Traceroute Responses at Non-Branch Nodes</a>	<a href="#">10</a>
<a href="#">3.3.2</a>	<a href="#">Traceroute Responses at Branch Nodes</a>	<a href="#">11</a>
<a href="#">3.3.3</a>	<a href="#">Traceroute Responses at Bud Nodes</a>	<a href="#">12</a>
<a href="#">3.3.4</a>	<a href="#">Non-Response to Traceroute Echo Requests</a>	<a href="#">12</a>
<a href="#">3.3.5</a>	<a href="#">Modifications to the Downstream Mapping TLV</a>	<a href="#">12</a>
<a href="#">3.3.6</a>	<a href="#">Additions to Downstream Mapping Multipath Information</a>	<a href="#">13</a>
<a href="#">4.</a>	<a href="#">Non-compliant Routers</a>	<a href="#">14</a>
<a href="#">5.</a>	<a href="#">OAM Considerations</a>	<a href="#">15</a>
<a href="#">6.</a>	<a href="#">IANA Considerations</a>	<a href="#">15</a>
<a href="#">6.1</a>	<a href="#">New Sub TLV Types</a>	<a href="#">15</a>
<a href="#">6.2</a>	<a href="#">New Multipath Type</a>	<a href="#">16</a>
<a href="#">7.</a>	<a href="#">Security Considerations</a>	<a href="#">16</a>
<a href="#">8.</a>	<a href="#">Acknowledgements</a>	<a href="#">16</a>
<a href="#">9.</a>	<a href="#">Intellectual Property Considerations</a>	<a href="#">16</a>
<a href="#">10.</a>	<a href="#">Normative References</a>	<a href="#">17</a>
<a href="#">11.</a>	<a href="#">Informational References</a>	<a href="#">17</a>
<a href="#">12.</a>	<a href="#">Authors' Addresses</a>	<a href="#">17</a>
<a href="#">13.</a>	<a href="#">Full Copyright Statement</a>	<a href="#">18</a>



## **1. Introduction**

Simple and efficient mechanisms that can be used to detect data plane failures in point-to-point (P2P) MPLS LSP are described in [[LSP-PING](#)]. 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 reply channels for echo request messages as described in [[LSP-PING](#)] enables more robust fault isolation. This collection of mechanisms is commonly referred to as "LSP Ping".

The requirement for point-to-multipoint (P2MP) MPLS traffic engineered (TE) LSPs is stated in [[P2MP-REQ](#)]. [[P2MP-RSVP](#)] specifies a signaling solution for establishing P2MP MPLS TE LSPs. P2MP MPLS TE LSPs are at least as vulnerable to data plane faults or to discrepancies between the control and data planes as their P2P counterparts. LSP Ping Mechanisms are, therefore, also desirable to detect such data plane faults in P2MP MPLS TE LSPs.

This document extends the techniques described in [[LSP-PING](#)] such that they may be applied to P2MP MPLS TE LSPs. This document stresses the reuse of existing LSP Ping mechanisms used for P2P LSPs, and applies them to P2MP MPLS TE LSPs in order to simplify implementation and network operation.

### **1.1 Design Considerations**

As mentioned earlier, an important consideration for designing LSP Ping for P2MP MPLS TE 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 would traverse. However, it can be seen this notion needs to be extended for P2MP MPLS TE 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 against the control plane. As pointed out in [[LSP-PING](#)], mechanisms to check the liveness, function and consistency of the control plane are valuable, but such mechanisms are not covered in this document.

As is described in [[LSP-PING](#)], 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.

Yasukawa, Farrel and Ali.

[Page 3]

## **2. Notes on Motivation**

### **2.1. Basic Motivations for LSP Ping**

The motivations listed in [[LSP-PING](#)] 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 would enable users to detect such traffic "black holes" or misrouting within a reasonable period of time; and a mechanism to isolate faults.

[[LSP-PING](#)] 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. [[LSP-PING](#)] specifies a "ping mode" and a "traceroute" mode for testing MPLS LSPs.

The basic idea as expressed in [[LSP-PING](#)] 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. [[LSP-PING](#)] 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 thus should be used with caution.





## **2.2. Motivations for LSP Ping for P2MP TE LSPs**

P2MP MPLS TE LSPs may be viewed as MPLS tunnels with a single ingress and multiple egresses. MPLS packets inserted at the ingress are delivered equally (barring faults) to all egresses. There is no concept or applicability of an FEC in the context of a P2MP MPLS TE LSP.

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 MPLS TE 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 a specific egress.

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 TE LSP that it is testing. The echo request is forwarded just as any other packet using that LSP. 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 LSR, which then verifies that it is indeed an egress (leaf) of the P2MP MPLS TE LSP. An echo response message is sent by the 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 TE 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 TE 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 TE 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 the ability for the initiator to limit the scope of an LSP Ping echo request (ping or traceroute mode) to one specific intended egress of the P2MP MPLS TE LSP, or to target all egresses. Further, in the event that the initiator wishes to use ping or traceroute to a large number of leaves simultaneously, this document provides a procedure that allows the responders to randomly delay or



jitter their responses so that the chances of swamping the ingress are reduced.

LSP Ping can be used to periodically ping a P2MP MPLS TE 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.

### **3. Operation of LSP Ping for a P2MP TE LSP**

This section describes how LSP Ping is applied to P2MP MPLS TE 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 LSRs and receivers (egresses).

#### **3.1. Identifying the LSP Under Test**

[LSP-PING] 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 or an RSVP P2MP IPv6 Session sub-TLV. These sub-TLVs carry the various fields from the RSVP-TE P2MP Session and Sender-Template objects [[P2MP-RSVP](#)] and so provide sufficient information to uniquely identify the LSP.

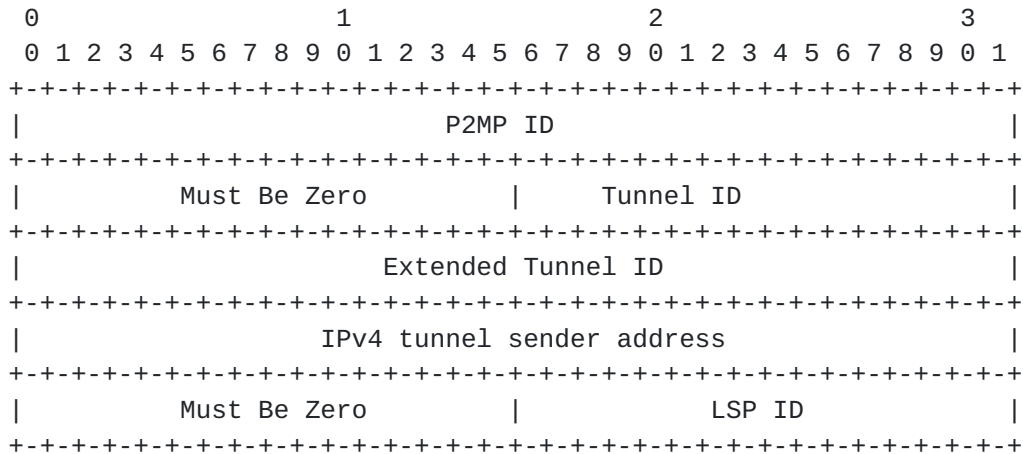
The new sub-TLVs are assigned sub-type identifiers as follows, and are described in the following sections.

Sub-Type #	Length	Value Field
-----	-----	-----
TBD	20	RSVP P2MP IPv4 Session
TBD	56	RSVP P2MP IPv6 Session

##### **3.1.1. RSVP P2MP IPv4 Session Sub-TLV**

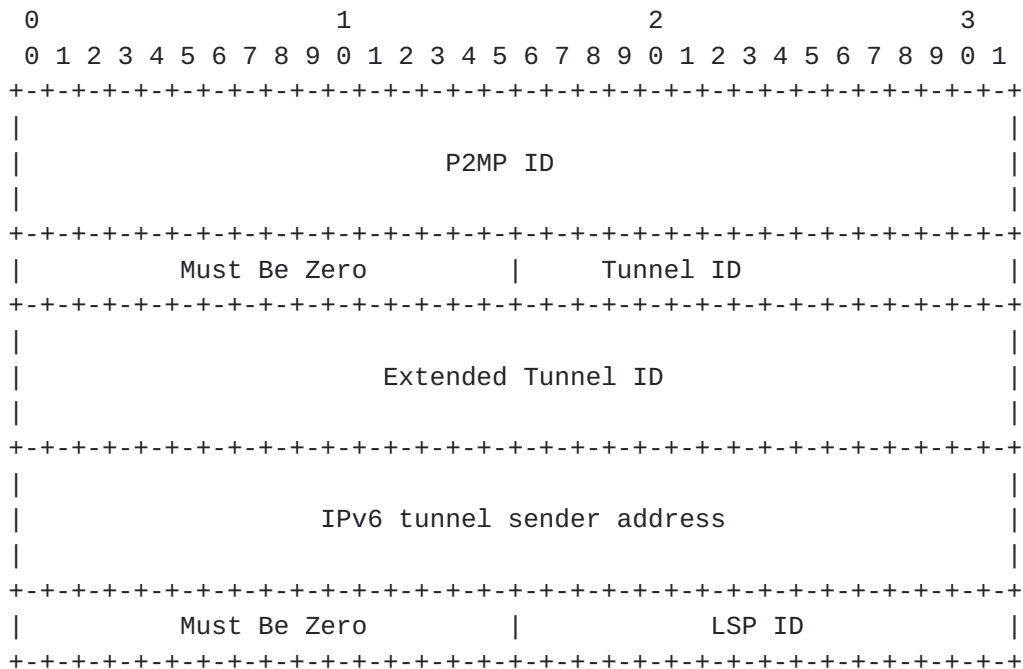
The format of the RSVP P2MP IPv4 Session Sub-TLV value field is specified in the following figure. The value fields are taken from the definitions of the P2MP IPv4 LSP Session Object, and the P2MP IPv4 Sender-Template Object in [[P2MP-RSVP](#)]. Note that the Sub-Group ID of the Sender-Template is not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





### 3.1.2. 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 [P2MP-RSVP]. Note that the Sub-Group ID of the Sender-Template is not required.



### 3.2. Ping Mode Operation

### 3.2.1. Controlling Responses to LSP Pings

As described above, 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

Yasukawa, Farrel and Ali.

[Page 7]

propagation of echo requests, and they will automatically be forwarded to all (reachable) egresses.

However, the intended egress under test is identified in the FEC Stack TLV by the inclusion of an IPv4 P2MP Egress Identifier sub-TLV or an IPv6 P2MP Egress Identifier sub-TLV. Such TLVs, if used, MUST be placed after the RSVP P2MP IPv4/6 Session sub-TLV.

An initiator may indicate that it wishes all egresses to respond to an echo request by omitting all P2MP Egress Identifier sub-TLVs.

An egress LSR that receives an echo request carrying an RSVP P2MP IPv4/6 Session 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 [[LSP-PING](#)].

If the egress that receives an echo request is an intended egress, the LSR MUST check to see whether it is an intended Ping recipient. If a P2MP Egress Identifier sub-TLV is present and contains an address that indicates any address that is local to the egress LSR, it MUST respond according to the setting of the Response Type field in the echo message following the rules defined in [[LSP-PING](#)]. If the P2MP Egress Identifier sub-TLV is present, but does not identify the egress LSR, it MUST NOT respond to the echo request. If the P2MP Egress identifier is not present, but the egress that received the echo request is an intended egress, it MUST respond according to the setting of the Response Type field in the echo message following the rules defined in [[LSP-PING](#)].

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.

Yasukawa, Farrel and Ali.

[Page 8]



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.

### **3.2.2. P2MP Egress Identifier sub-TLVs**

Two new sub-TLVs are defined for inclusion in the Target FEC Stack TLV (type 1) carried on the echo request message. These are:

Sub-Type #	Length	Value Field
-----	-----	-----
(TBD)	4	IPv4 P2MP Egress Identifier
(TBD)	16	IPv6 P2MP Egress Identifier

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

The value of an IPv6 P2MP Egress Identifier consists of sixteen octets of an IPv6 address. The IPv6 address is in network byte order.

### **3.2.3. 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
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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      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 egress to determine how long to wait before sending an echo response. An egress 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.

### **3.3. Traceroute Mode Operation**

The traceroute mode of operation is described in [[LSP-PING](#)]. 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 LSR which responds according to the Response Type in the message. A responding LSR fills in the fields of the Downstream Mapping TLV to indicate the downstream interfaces and labels used by the reported LSP from the responding LSR. 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 an LSR 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 a single destination, or to all destinations of the P2MP tree just as in the ping mode. That is, the IPv4/6 P2MP Egress Identifier sub-TLVs may be used to identify a specific egress for which traceroute information is requested. In the absence of an IPv4/6 P2MP Egress Identifier sub-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 and the message is passed to the control plane, an echo response MUST only be returned if the responding LSR lies on the path to the egress identified by the IPv4/6 P2MP Egress Identifier carried on the request, or if no such sub-TLV is present.



The echo response identifies the next hop of the path in the data plane by including a Downstream Mapping TLV as described in [\[LSP-PING\]](#).

When traceroute is being simultaneously applied to multiple 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 egress in turn, but this may be inefficient or undesirable. Therefore, the echo response contains additional information in the Multipath Information field of the Downstream Mapping TLV that identifies to which egress/egresses the echo response applies. This information **MUST** be present when the echo request applies to all egresses, and is **RECOMMENDED** to be present even when the echo request is limited to a single egress.

The format of the information in the Downstream Mapping TLV for P2MP MPLS TE LSPs is described in [section 3.3.5](#) and 3.3.6.

### **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 egresses that are being traced lie on different branches. This will always be the case for any branch node if all egresses are being traced.

[\[LSP-PING\]](#) 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.

Just as with non-branches, it is important that the echo responses provide correlation information that will allow the ingress to work out to which branch of the LSP the response applies. Further, when multiple downstream interfaces are identified, it is necessary to indicate which egresses are reached through which branches. This is achieved exactly as for non-branch nodes: that is, by including a list of egresses as part of the Multipath Information field of the appropriate Downstream Mapping TLV.

Note also that 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 LSR is not a branch node.

To report on the fact that an LSR is or is not a branch node for the P2MP MPLS TE LSP a new B-flag is added to the Downstream Mapping TLV. The flag is set to zero to indicate that the reporting LSR is not a

branch for this LSP, and is set to one to indicate that it is a

Yasukawa, Farrel and Ali.

[Page 11]

branch. The flag is placed in the fourth byte of the TLV that was previously reserved.

The format of the information in the Downstream Mapping TLV for P2MP MPLS TE LSPs is described in [section 3.3.5](#) and 3.3.6.

### **[3.3.3. Traceroute Responses at Bud Nodes](#)**

Some nodes on an P2MP MPLS TE LSP may be egresses, but also have downstream LSRs. Such LSRs are known as bud nodes.

A bud node will respond to a traceroute echo request just as a branch node would, but it is also important that it indicates to the ingress that it is an egress in its own right. This is achieved through the use of a new E-flag in the Downstream Mapping TLV that indicates that the reporting LSR is not a bud for this LSP (set to zero) or is a bud (set to one). A normal egress is not required to set this flag.

The flag is placed in the fourth byte of the TLV that was previously reserved.

### **[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 LSRs that must not reply to the request because, although they lie on the P2MP tree, they do not lie on the path to the egress that is being traced.

Thus, an LSR on a P2MP MPLS TE 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
- There is an IPv4/6 P2MP Egress Identifier present on the echo request, but the address does not identify an egress that is reached through this LSR for this particular P2MP MPLS TE LSP.

### **[3.3.5. Modifications to the Downstream Mapping TLV](#)**

A new B-flag is added to the Downstream Mapping TLV to indicate that the reporting LSR is not a branch for this LSP (set to zero) or is a branch (set to one).

A new E-flag is added to the Downstream Mapping TLV to indicate that the reporting LSR is not a bud node for this LSP (set to zero) or is a bud node (set to one).

The flags are placed in the fourth byte of the TLV that was previously reserved as shown below. All other fields are unchanged from their definitions in [[LSP-PING](#)] except for the additional information that can be carried in the Multipath Information.





```

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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               MTU                               | Address Type | Reserved |E|B|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Downstream IP Address (4 or 16 octets)                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Downstream Interface Address (4 or 16 octets)                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Hash Key Type | Depth Limit |                               Multipath Length                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
.
.                               (Multipath Information)                               .
.
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Downstream Label                               | Protocol |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
.
.
.
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Downstream Label                               | Protocol |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

### 3.3.6. Additions to Downstream Mapping Multipath Information

A new value for the Multipath Type is defined to indicate that the reported Multipath Information applies to an P2MP MPLS TE LSP and may contain a list of egress identifiers that indicate the egress nodes that can be reached through the reported interface.

Type #	Address Type	Multipath Information
---	-----	-----
TBD	P2MP egresses	List of P2MP egresses

Note that a list of egresses 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 [[LSP-PING](#)] (that is not including the downstream labels), and the downstream label (or potential stack thereof) is also handled just as in [[LSP-PING](#)]. The format of the Multipath Information for a Multipath Type of P2MP Egresses is 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	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
+--+--+--+--+--+--+--+--+--+										+--+--+--+--+--+--+--+--+--+										+--+--+--+--+--+--+--+--+--+										+--+--+--+--+--+--+--+--+--+									
Address Type										Egress Address (4 or 16 octets)																													
+--+--+--+--+--+--+--+--+--+										+--+--+--+--+--+--+--+--+--+										+--+--+--+--+--+--+--+--+--+										+--+--+--+--+--+--+--+--+--+									
(continued)																														:									
+--+--+--+--+--+--+--+--+--+																														:									
:										Further Address Types and Egress Addresses										:																			
:																														:									
+--+--+--+--+--+--+--+--+--+										+--+--+--+--+--+--+--+--+--+										+--+--+--+--+--+--+--+--+--+										+--+--+--+--+--+--+--+--+--+									

## Address Type

This field indicates whether the egress 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.

Type #	Address Type
-----	-----
1	IPv4
3	IPv6

## Egress Address

An egress 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.

#### 4. 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 LSR does not support LSP ping, then no reply will be forthcoming from that LSR for some TTL, say n. The LSR originating the echo request SHOULD continue to send echo requests with TTL=n+1, n+2, ..., n+k in the hope that some transit LSR further downstream may support MPLS echo requests and reply. In such a case, the echo request for TTL > n MUST NOT have Downstream Mapping TLVs, until a reply is received with a Downstream Mapping.



## **5. OAM Considerations**

This draft clearly facilitates OAM procedures for P2MP MPLS TE LSPs.

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 TE 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 TE 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.

## **6. IANA Considerations**

### **6.1. New Sub TLV Types**

Four new sub-TLV types are defined for inclusion within the Target FEC Stack TLV (TLV type 1).

IANA is requested to assign sub-type values to the following sub-TLVs.

- RSVP P2MP IPv4 Session (see [section 3.1](#))
- RSVP P2MP IPv6 Session (see [section 3.1](#))
- IPv4 P2MP Egress Identifier (see [section 3.2.2](#))
- IPv6 P2MP Egress Identifier (see [section 3.2.2](#))



## **6.2. New Multipath Type**

A new value for the Multipath Type is defined to indicate that the reported Multipath Information applies to an P2MP MPLS TE LSP.

IANA is requested to assign a new value as follows.

List of P2MP egresses (see [section 3.3.6](#))

## **7. Security Considerations**

This document does not introduce security concerns over and above those described in [[LSP-PING](#)]. Note that because of the scalability implications of many egresses to P2MP MPLS TE 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.

## **8. Acknowledgements**

The authors would like to acknowledge the authors of [[LSP-PING](#)] 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 Dan King for his review, and to Yakov Rekhter for useful discussions.

## **9. 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](mailto:ietf-ipr@ietf.org).





## **10. Normative References**

- [RFC2119]    Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC3667]    Bradner, S., "IETF Rights in Contributions", [BCP 78](#), [RFC 3667](#), February 2004.
- [RFC3668]    Bradner, S., Ed., "Intellectual Property Rights in IETF Technology", [BCP 79](#), [RFC 3668](#), February 2004.
- [LSP-PING]    Kompella, K., and Swallow, G., (Editors), "Detecting MPLS Data Plane Failures", [draft-ietf-mpls-lsp-ping](#), work in progress.

## **11. Informational References**

- [RFC2434]    Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP: 26, [RFC 2434](#), October 1998.
- [RFC3209]    Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", [RFC 3209](#), December 2001.
- [RFC3552]    Rescorla E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP: 72, [RFC 3552](#), July 2003.
- [RFC792]    Postel, J., "Internet Control Message Protocol", [RFC 792](#).
- [P2MP-REQ]    S. Yasukawa, et. al., "Signaling Requirements for Point to Multipoint Traffic Engineered MPLS LSPs", [draft-ietf-mpls-p2mp-sig-requirement](#), work in progress.
- [P2MP-RSVP]    R. Aggarwal, et. al., "Extensions to RSVP-TE for Point to Multipoint TE LSPs", [draft-ietf-mpls-rsvp-te-p2mp](#), work in progress.

## **12. Authors' Addresses**

Seisho Yasukawa  
NTT Corporation  
9-11, Midori-Cho 3-Chome  
Musashino-Shi, Tokyo 180-8585,  
Japan  
Phone: +81 422 59 4769  
Email: [yasukawa.seisho@lab.ntt.co.jp](mailto:yasukawa.seisho@lab.ntt.co.jp)



Adrian Farrel  
Old Dog Consulting  
EMail: [adrian@olddog.co.uk](mailto: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](mailto:zali@cisco.com)

Bill Fenner  
AT&T Labs -- Research  
75 Willow Rd.  
Menlo Park, CA 94025  
United States  
Email: [fenner@research.att.com](mailto:fenner@research.att.com)

### **13. Full Copyright Statement**

Copyright (C) The Internet Society (2005). 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 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.

### **14. Change History**

This section to be removed before publication as an RFC

#### **14.1. Changes from [draft-yasukawa-mpls-p2mp-lsp-ping 01](#) to 02**

- Add Bill Fenner as co-author.
- Add echo jitter response processing.

