

Internet Engineering Task Force
Internet-Draft
Updates: 7110 (if approved)
Intended status: Standards Track
Expires: November 2, 2015

N. Akiya
Big Switch Networks
G. Swallow
C. Pignataro
Cisco Systems
L. Andersson
M. Chen
Huawei
May 1, 2015

Label Switched Path (LSP) Ping/Traceroute Reply Mode Simplification
draft-ietf-mpls-lsp-ping-reply-mode-simple-03

Abstract

The Multiprotocol Label Switching (MPLS) Label Switched Path (LSP) Ping and Traceroute use the Reply Mode field to signal the method to be used in the MPLS echo reply. This document updates the "Reply via Specified Path (5)" Reply Mode value to easily indicate the reverse LSP. This document also adds an optional TLV which can carry ordered list of Reply Mode values.

This document updates RFC7110.

Requirements Language

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

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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."

This Internet-Draft will expire on November 2, 2015.

Copyright Notice

Copyright (c) 2015 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Problem Statements	3
3. Solution	5
3.1. Reply via Specified Path Update	5
3.2. Reply Mode Order TLV	6
4. Relations to Other LSP Ping/Trace Features	7
4.1. Reply Path TLV	7
4.1.1. Example 1: Reply Mode Order TLV Usage with Reply Path TLV	8
4.1.2. Example 2: Reply Mode Order TLV Usage with Reply Path TLV	8
4.2. Proxy LSP Ping	9
4.2.1. Proxy LSR Sending an MPLS Echo Request	9
4.2.2. Proxy LSR Sending an MPLS Proxy Ping Reply	9
5. Security Considerations	10
6. IANA Considerations	10
6.1. New Reply Mode Order TLV	10
7. Acknowledgements	10
8. Contributing Authors	10
9. References	10
9.1. Normative References	10
9.2. Informative References	11
Appendix A. Reply Mode Order TLV Beneficial Scenarios	11
A.1. Incorrect Forwarding Scenario	11
A.2. Non-Co-Routed Bidirectional LSP Scenario	12
Authors' Addresses	13

1. Introduction

The MPLS LSP Ping, described in [RFC4379], allows an initiator LSR to encode instructions (Reply Mode) on how a responder LSR should send the response back to the initiator LSR. [RFC7110] also allows the initiator LSR to encode a TLV (Reply Path TLV) which can instruct the responder LSR to use specific LSP to send the response back to the initiator LSR. Both approaches are powerful as they provide the ability for the initiator LSR to control the return path.

However, it is becoming increasingly difficult for an initiator LSR to select a valid return path to encode in the MPLS LSP echo request packets. If the initiator LSR does not select a valid return path, the MPLS LSP echo reply will not get back to the initiator LSR. This results in a false failure of MPLS LSP Ping and Traceroute operation. In an effort to minimize such false failures, different implementations have chosen different default return path encoding for different LSP types and LSP operations. The problem with implementations having different default return path encoding is that the MPLS echo reply will not work in many cases, and the default value may not be the preferred choice by the operators.

This document describes:

- o In Section 2, further description of the problems;
- o In Section 3, a solution to minimize false failures while accommodating operator preferences;
- o In Section 4, relationships to other LSP Ping/Traceroute features;
- o In Appendix A, examples of scenarios where the mechanism described in this document provides benefits.

This document updates [RFC7110] by allowing the usage of the Reply Mode value 5 (Reply via Specified Path) without including the Reply Path TLV.

2. Problem Statements

It is becoming increasingly difficult for implementations to automatically supply a workable return path encoding for all MPLS LSP Ping and Traceroute operations across all LSP types. There are several factors which are contributing to this complication.

- o Some LSPs have a control-channel, and some do not. Some LSPs have a reverse LSP, and some do not. Some LSPs have IP reachability in the reverse direction, and some do not.

- o LSRs on some LSPs can have different available return path(s). Available return path(s) can depend on whether the responder LSR is a transit LSR or an egress LSR. In case of a bi-directional LSP, available return path(s) on transit LSRs can also depend on whether LSP is completely co-routed, partially co-routed or associated (i.e., LSPs in the two directions are not co-routed).
- o MPLS echo request packets may incorrectly terminate on an unintended target, which can have different available return path(s) than the intended target.
- o The MPLS LSP Ping operation is expected to terminate on egress LSR. However, the MPLS LSP Ping operation with specific TTL values and MPLS LSP Traceroute operation can terminate on both transit LSR(s) and the egress LSR.

Except for the case where the responder LSR does not have an IP route back to the initiator LSR, it is possible to use the "Reply via an IPv4/IPv6 UDP packet (2)" Reply Mode value in all cases. However, some operators are preferring control-channel and reverse LSP as default return path if they are available, which is not always the case.

When specific return path encoding is supplied by users or applications, then there are no issues in choosing the return path encoding. When specific return path encoding is not supplied by users or applications, then implementations use extra logic to compute, and sometimes guess, the default return path encodings. If a responder LSR receives an MPLS echo request containing return path instructions which cannot be accommodated due to unavailability, then the responder LSR often drops such packets. This results in the initiator LSR not receiving the MPLS LSP echo reply packets back. This consequence may be acceptable for failure cases (e.g., broken LSPs) where the MPLS echo request terminated on unintended target. However, the initiator LSR not receiving back MPLS echo reply packets, even when the intended target received and verified the requests, is not desirable as false failures will be conveyed to users.

Many operators prefer some return path(s) over others for specific LSP types. To accommodate this, implementations may default to operator preferred return path (or allow default return path to be configured) for a specific operation. However, if the sender of MPLS echo request knew that preferred return path will not be available at the intended target node, then it is not very beneficial to use a Reply Mode corresponding to preferred return path (i.e., the sender of the MPLS echo request will not receive the MPLS echo reply in the successful case). What would be beneficial, for a given operation,

is for the sender of the MPLS echo request to determine which return path(s) can and cannot be used ahead of time.

This document adds one Reply Mode value to describe the reverse LSP, and one optional TLV to describe an ordered list of reply modes. Based on operational needs, the TLV can describe multiple Reply Mode values in a preferred order to allow the responder LSR to use the first available Reply Mode from the list. This eliminates the need for the initiator LSR to compute, or sometimes guess, the default return path encoding. And that will result in simplified implementations across vendors, and result in improved usability to fit operational needs.

3. Solution

This document updates "Reply via Specified Path (5)" Reply Mode to easily indicate the reverse LSP. This document also adds an optional TLV which can carry ordered list of reply modes.

3.1. Reply via Specified Path Update

Some LSP types are capable of having related LSP in reverse direction, through signaling or other association mechanisms. Examples of such LSP types are RSVP LSPs and TP LSPs. This document uses the term "Reverse LSP" to refer to the LSP in reverse direction of such LSP types. Note that this document restricts the scope of "Reverse LSP" applicability to those reverse LSPs which are capable and allowed to carry the IP encapsulated MPLS echo reply.

[RFC7110] has defined the Reply Mode "Reply via Specified Path (5)" which allows the initiator LSR to instruct the responder LSR to send the MPLS echo reply message on the reverse LSP. However, the instruction also requires the initiator LSR to include the "Reply Path TLV" with the B bit (Bidirectional bit) set in the Flags field. Additionally, [RFC7110] defines the usage of the "Reply via Specified Path (5)" Reply Mode without inclusion of the "Reply Path TLV" to be invalid.

This document updates the "Reply via Specified Path (5)" Reply Mode as follows:

- o The usage of the "Reply via Specified Path (5)" without inclusion of a "Reply Path TLV" is no longer invalid.
- o The usage of the "Reply via Specified Path (5)" without inclusion of a "Reply Path TLV" implies the reverse LSP. In other words, the usage of the "Reply via Specified Path (5)" without inclusion of a "Reply Path TLV" has the same semantics as the usage of the

"Reply via Specified Path (5)" with inclusion of a "Reply Path TLV" with the B bit set in the Flags field.

Note that the reverse LSP is in relation to the last FEC specified in the Target FEC Stack TLV.

When a responder LSR is using this Reply Mode, transmitting MPLS echo reply packet MUST use IP destination address of 127/8 for IPv4 and 0:0:0:0:0:FFFF:7F00/104 for IPv6.

3.2. Reply Mode Order TLV

This document also introduces a new optional TLV to describe list of Reply Mode values. The new TLV will contain one or more Reply Mode value(s) in preferred order. The first Reply Mode value is the most preferred and the last Reply Mode value is the least preferred. Following rules apply when using Reply Mode Order TLV.

1. The Reply Mode Order TLV MUST NOT be included in MPLS echo reply. If the initiator LSR receives an MPLS echo reply with the Reply Mode Order TLV, the initiator LSR MUST ignore the whole Reply Mode Order TLV and MUST only use the value from the Reply Mode field of the received MPLS echo reply. It may be beneficial for implementations to provide counters and/or loggings, with appropriate log dampening, to record this error case.
2. The Reply Mode Order TLV MAY be included in MPLS echo request.
3. The Reply Mode field of an MPLS echo request MUST be set to a valid value even when supplying the Reply Mode Order TLV. The initiator LSR SHOULD set the Reply Mode field of MPLS echo request to a value that corresponds to a return path which most likely to be available, in case the responder LSR does not understand the Reply Mode Order TLV.
4. If a responder LSR understands the Reply Mode Order TLV but the TLV is not valid (due to conditions described in the items 6, 7, 8 and 9 immediately below), then the responder LSR MUST ignore the whole Reply Mode Order TLV and MUST only use the value from the Reply Mode field of the received MPLS echo request. It may be beneficial for implementations to provide counters and/or loggings, with appropriate log dampening, to record this error case.
5. If a responder LSR understands the Reply Mode Order TLV and the TLV is valid, then the responder LSR MUST consider the Reply Mode values described in the TLV and MUST NOT use the value described in the Reply Mode field of received MPLS echo request. In other

Order TLV to describe the reverse LSP and other FECs for return paths, then the initiator SHOULD include two "Reply via Specified Path (5)" Reply Mode values and two "Reply Path TLV" objects (one "Reply Path TLV" corresponding to each "Reply via Specified Path (5)").

- o The reverse LSP is described by the "Reply via Specified Path (5)" Reply Mode value and the corresponding "Reply Path TLV" with the B bit set in the Flags field. In this "Reply Path TLV", no Sub-TLVs are present.
- o Other return FECs are described by the "Reply via Specified Path (5)" Reply Mode value and the corresponding "Reply Path TLV" describing the FECs for return paths. In this "Reply Path TLV", the B bit is cleared in the Flags field.

4.1.1. Example 1: Reply Mode Order TLV Usage with Reply Path TLV

If the initiator LSR was interested in encoding following return paths:

1. Reply via application level control channel
2. FEC X
3. FEC Y
4. Reply via an IPv4/IPv6 UDP packet

Then the MPLS echo request message is to carry:

- o The Reply Mode Order TLV carrying Reply Modes {4, 5, 2}
- o The Reply Path TLV carrying {FEC X, FEC Y}

Described encoding of the Reply Mode Order TLV and the Reply Path TLV in the MPLS echo request message will result in the responder LSR to prefer "Reply via application level control channel (4)", followed by FEC X, FEC Y and then "Reply via an IPv4/IPv6 UDP packet (2)".

4.1.2. Example 2: Reply Mode Order TLV Usage with Reply Path TLV

If the initiator LSR was interested in encoding following return paths:

1. Reverse LSP
2. Reply via an IPv4/IPv6 UDP packet

3. FEC X

4. FEC Y

Then the MPLS echo request message is to carry:

- o The Reply Mode Order TLV carrying Reply Modes {5, 2, 5}
- o One Reply Path TLV with the B bit set.
- o One Reply Path TLV carrying {FEC X, FEC Y}

Described encoding of the Reply Mode Order TLV and the Reply Path TLV in the MPLS echo request message will result in the responder LSR to prefer the reverse LSP, followed by "Reply via an IPv4/IPv6 UDP packet (2)", FEC X and then FEC Y.

4.2. Proxy LSP Ping

The mechanism defined in this document will work with Proxy LSP Ping defined by [I-D.ietf-mpls-proxy-lsp-ping]. The MPLS proxy ping request message can carry a Reply Mode value in the header and one or more Reply Mode values in the Reply Mode Order TLV. It is RECOMMENDED that the Reply Mode 2 (Reply via an IPv4/IPv6 UDP packet) be used in the Reply Mode field of the MPLS proxy ping request message.

4.2.1. Proxy LSR Sending an MPLS Echo Request

If the proxy LSR is sending an MPLS echo request, then the proxy LSR MUST copy following elements from the MPLS proxy ping request message to the MPLS echo request message.

- o The Reply Mode field.
- o The Reply Mode Order TLV.
- o The Reply Path TLV(s). If there are more than one Reply Path TLVs, then then order of them MUST be preserved when copying.

4.2.2. Proxy LSR Sending an MPLS Proxy Ping Reply

If the proxy LSR is sending an MPLS proxy ping reply, then it is RECOMMENDED that the Reply Mode Order TLV be ignored and the Reply Mode field in the MPLS proxy ping request message be used.

5. Security Considerations

Beyond those specified in [RFC4379] and [RFC7110], there are no further security measures required.

6. IANA Considerations

6.1. New Reply Mode Order TLV

IANA is requested to assign a new TLV type value from the "TLVs" sub-registry within the "Multiprotocol Label Switching Architecture (MPLS)" registry, for the "Reply Mode Order TLV".

The new TLV Type value should be assigned from the range (32768-49161) specified in [RFC4379] section 3 that allows the TLV type to be silently dropped if not recognized.

Type	Meaning	Reference
----	-----	-----
TBD1	Reply Mode Order TLV	this document

7. Acknowledgements

Authors would like to thank Santiago Alvarez and Faisal Iqbal for discussions which motivated creation of this document. Authors would also like to thank Sam Aldrin, Curtis Villamizar, Ross Callon, Jeffrey Zhang, Jeremy Whittaker, Mustapha Alissaoui, Qin Wu, Jie Dong and Adrian Farrel for providing valuable comments to influence the contents of the draft.

8. Contributing Authors

Shaleen Saxena
 Brocade
 Email: ssaxena@brocade.com

9. References

9.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4379] Kompella, K. and G. Swallow, "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", RFC 4379, February 2006.

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

9.2. Informative References

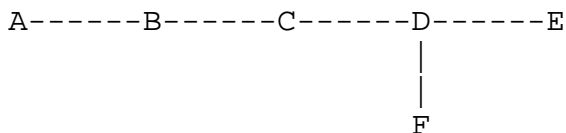
[I-D.ietf-mpls-proxy-lsp-ping]
Swallow, G., Lim, V., and S. Aldrin, "Proxy MPLS Echo Request", draft-ietf-mpls-proxy-lsp-ping-05 (work in progress), March 2015.

Appendix A. Reply Mode Order TLV Beneficial Scenarios

This section lists examples of how the Reply Mode Order TLV can benefit.

A.1. Incorrect Forwarding Scenario

A network has a following LSP, and the LSP has a control channel.



Forward Paths: A-B-C-D-E

Figure 2: Incorrect Forwarding

Imagine that D is incorrectly label switching to F (instead of E). In this scenario, LSP Traceroute with "Reply via application level control channel (4)" will result in following result.

```

Success (Reply from B)
Success (Reply from C)
Success (Reply from D)
Timeout...
Complete
  
```

This is because F does not have a control channel to send the MPLS echo reply message. With the extension described in this document, same procedures can be performed with the Reply Mode Order TLV carrying {4, 2}. When LSP Traceroute is issued, then following output may be displayed without any unnecessary timeout.

```

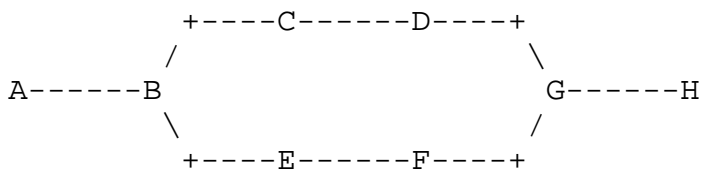
Success (Reply from B, Reply Mode: 4)
Success (Reply from C, Reply Mode: 4)
Success (Reply from D, Reply Mode: 4)
  
```

FEC Mismatch (Reply from F, Reply Mode: 2)
Complete

The result provides more diagnostic information to the initiator LSR, and without any delay (i.e. timeout from one or more downstream LSRs).

A.2. Non-Co-Routed Bidirectional LSP Scenario

A network has a following bidirectional LSP where the forward LSP and the reverse LSP are not fully co-routed.



Forward Paths: A-B-C-D-G-H (upper path)

Reverse Paths: H-G-F-E-B-A (lower path)

Figure 3: Non-Co-Routed Bidirectional LSP

Some operators may prefer and configure the system to default the Reply Mode to indicate the reverse LSP when MPLS echo request messages are sent on bidirectional LSPs. Without extensions described in this document, following behaviors will be seen:

- o When LSP Ping is issued from A, reply will come back on the reverse LSP from H.
- o When LSP Traceroute is issued from A, reply will come back on the reverse LSP from B, G and H, but will encounter a timeout from C and D as there are no reverse LSP on those nodes.
- o When LSP Ping with specific TTL value is issued from A, whether a timeout will be encountered depends on the value of the TTL used (i.e. whether or not MPLS echo request terminates on a node that has reverse LSP).

One can argue that the initiator LSR can automatically generate a same MPLS echo request with different Reply Mode value to those nodes that timeout. However, such mechanism will result in extended time for the entire operation to complete (i.e. multiple seconds to multiple minutes). This is undesirable, and perhaps unacceptable if the "user" is an application.

With the extension described in this document, same procedures can be performed with the Reply Mode Order TLV carrying {5, 2}. When LSP Traceroute is issued, then following output may be displayed without any unnecessary timeout.

```
Success (Reply Mode: 5)
Success (Reply Mode: 2)
Success (Reply Mode: 2)
Success (Reply Mode: 5)
Success (Reply Mode: 5)
Complete
```

Authors' Addresses

Nobo Akiya
Big Switch Networks

Email: nobo.akiya.dev@gmail.com

George Swallow
Cisco Systems

Email: swallow@cisco.com

Carlos Pignataro
Cisco Systems

Email: cpignata@cisco.com

Loa Andersson
Huawei

Email: loa@mail01.huawei.com

Mach(Guoyi) Chen
Huawei

Email: mach.chen@huawei.com