Internet Engineering Task Force (IETF)
Request for Comments: 7737
Updates: 7110
Category: Standards Track
ISSN: 2070-1721

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

January 2016

Label Switched Path (LSP) Ping and Traceroute Reply Mode Simplification

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 procedures for the "Reply via Specified Path" Reply Mode. The value of this Reply Mode is 5. The update creates a simple way to indicate that the reverse LSP should be used as the return path. This document also adds an optional TLV that can carry an ordered list of Reply Mode values.

This document updates RFC 7110.

Status of This Memo

This is an Internet Standards Track document.

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

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7737.

Table of Contents

1. Introduction

Multiprotocol Label Switching (MPLS) Label Switched Path (LSP) Ping, as described in [RFC4379], allows an initiator Label Switching Router (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) that can instruct the responder LSR to use a 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 operations. 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 of the operators.

This document describes the following:

- In Section 2, further description of the problems;
- In Section 3, a solution to minimize false failures while accommodating operator preferences;
- In Section 4, relationships to other LSP Ping and Traceroute features;
- 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 via Specified Path" (value=5) Reply Mode without including the Reply Path TLV. The update creates a simple way to indicate that the reverse LSP should be used as the return path.

1.1. Requirements Language

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

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 that contribute to this complication.

- 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.

- 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 the case of a bidirectional LSP, available return path(s) on transit LSRs can also depend on whether the LSP is completely co-routed, partially co-routed, or associated (i.e., the LSPs in the two directions are not co-routed).

- MPLS echo request packets may incorrectly terminate on an unintended target that can have different available return path(s) than the intended target.

- The MPLS LSP Ping operation is expected to terminate on an egress LSR. However, MPLS LSP Ping operations with specific TTL values and MPLS LSP Traceroute operations 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" (value=2) Reply Mode value in all cases. However, some operators prefer the control channel and a reverse LSP as the default return path if they are both available, although this 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 that cannot be accommodated due to unavailability, then the responder LSR often drops such packets. This failure mode results in the initiator LSR not receiving the intended MPLS LSP echo reply packets. The scenario described here is a potentially acceptable result in some failure cases, like a broken LSP, where the MPLS echo request terminated on an unintended target. However, if the initiator LSR does not receive an MPLS echo reply even after the

responder LSR receives the MPLS echo request and is able to verify the request, information is still sent back to the user(s); this is considered a false failure.

Many operators prefer particular return path(s) over other return path(s) for specific LSP types. To accommodate operator-preferred paths, implementations may default to operator-preferred return paths for particular operations or allow a default return path to be configured. It would not be considered beneficial to use a preferred return path for an intended target LSR if there is previous knowledge at the initiator LSR that the return path is not available. Using an unavailable preferred return path would undesirably result in the initiator LSR not receiving the MPLS echo return packets. It would be considered beneficial, for given operations, if the sender of the MPLS echo request would be able to determine return path availability before the operation is initiated.

This document (1) updates the procedures for the "Reply via Specified Path" Reply Mode to easily indicate the reverse LSP and (2) adds one optional TLV to describe an ordered list of Reply Modes. Based on operational needs, the TLV can list 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. This new mode of operation would result in simplified implementations across the various vendors and improve both usability and operational needs.

3. Solution

This document updates the procedures for the "Reply via Specified Path" Reply Mode to easily indicate the reverse LSP. This document also adds an optional TLV that can carry an ordered list of Reply Modes.

3.1. "Reply via Specified Path" Reply Mode Update

Some LSP types are capable of having a related LSP in the reverse direction, through signaling or other association mechanisms. Examples of such LSP types are bidirectional Resource Reservation Protocol-Traffic Engineering (RSVP-TE) LSPs [RFC3473] and MPLS Transport Profile (MPLS-TP) LSPs [RFC5960]. This document uses the term "reverse LSP" to refer to the LSP in the reverse direction of such LSP types. Note that this document restricts the scope of "reverse LSP" applicability to those reverse LSPs that are capable and allowed to carry the IP encapsulated MPLS echo reply.

[RFC7110] has defined the Reply Mode "Reply via Specified Path", 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] specifies that if the "Reply via Specified Path" Reply Mode is used the Reply Path TLV MUST be present.

This document updates the procedures for the "Reply via Specified Path" Reply Mode as follows:

- The "Reply via Specified Path" Reply Mode MAY be used without including a Reply Path TLV.

- The usage of the "Reply via Specified Path" Reply Mode without the inclusion of a Reply Path TLV implies the reverse LSP. In other words, the usage of the "Reply via Specified Path" Reply Mode without the inclusion of a Reply Path TLV has the same semantics as the usage of the "Reply via Specified Path" Reply Mode with the inclusion of a Reply Path TLV with the B bit set in the Flags field.

This document updates the first sentence of Section 5.1 of [RFC7110] as follows:

- When sending an echo request, in addition to the rules and procedures defined in Section 4.3 of [RFC4379], the Reply Mode of the echo request MUST be set to "Reply via Specified Path", and a Reply Path TLV SHOULD be carried in the echo request message correspondingly; if the Reply Path TLV is not carried in the message, then it indicates the reverse LSP as the reply path.

Note that the reverse LSP is in relation to the last Forwarding Equivalence Class (FEC) specified in the Target FEC Stack TLV.

3.2. Reply Mode Order TLV

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

1. The Reply Mode Order TLV MUST NOT be included in any 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 logs, with appropriate log dampening, to record this error case.

2. The Reply Mode Order TLV MAY be included in MPLS echo requests.

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 an MPLS echo request to a value that corresponds to a return path that is 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 the conditions described in items 6, 7, 8, and 9 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 logs, 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 specified in the TLV and MUST NOT use the value specified in the Reply Mode field of the received MPLS echo request. In other words, a valid Reply Mode Order TLV overrides the value specified in the Reply Mode field of the received MPLS echo request.

6. The Reply Mode Order TLV MUST contain at least one Reply Mode value.

7. A Reply Mode value, except for Reply Mode value 5 (Reply via Specified Path), MUST NOT be repeated (i.e., MUST NOT appear multiple times) in the Reply Mode Order TLV.

8. Reply Mode value 5 (Reply via Specified Path) MAY be included more than once in the Reply Mode Order TLV. However, in such a case, a Reply Path TLV MUST be included for all instances of Reply Mode value 5 that are included in the Reply Mode Order TLV. In other words, three instances of Reply Mode value 5 in the Reply Mode Order TLV will each require a Reply Path TLV.

9. The Reply Mode value 1 (Do not reply) MUST NOT be used in the Reply Mode Order TLV.

The responder LSR SHOULD select the first available return path in this TLV. The Reply Mode value corresponding to the selected return path MUST be set in the Reply Mode field of the MPLS echo reply to communicate back to the initiator LSR which return path was chosen.

The format of the TLV is as follows:

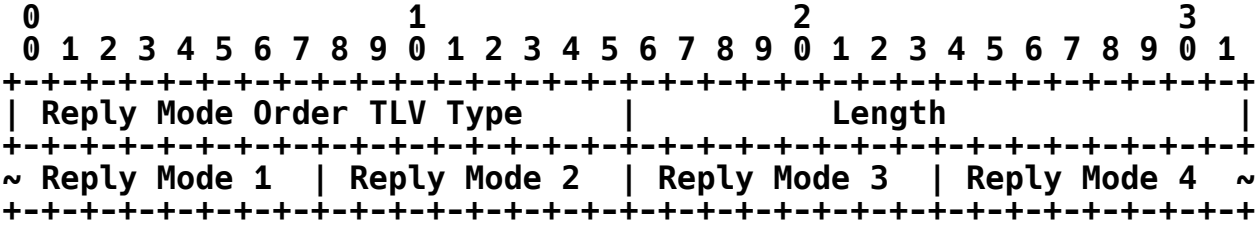


Figure 1: Reply Mode Order TLV

This is a variable-length optional TLV. The Reply Mode Order TLV Type is 32770.

The Length field is 2 octets in length. It defines the length, in octets, of the list of Reply Mode values.

Each Reply Mode field is 1 octet, and there is no padding.

4. Relationships to Other LSP Ping and Traceroute Features

4.1. Backwards Compatibility with "Reply via Specified Path" Reply Mode

[RFC7110] introduces the "Reply via Specified Path" (value=5) Reply Mode. [RFC7110] also specifies that if this Reply Mode is used the Reply Path TLV MUST be included. This document relaxes the semantics and specifies that this Reply Mode MAY be used without the Reply Path TLV. This MAY be done to indicate that the reverse LSP SHALL be used as the return path.

If an initiator LSR that sent an MPLS echo request message with the "Reply via Specified Path" Reply Mode but without including the Reply Path TLV receives back an MPLS echo reply message with a return code of "Malformed echo request received", then the initiator LSR SHOULD assume that the responder LSR does not support the mechanism defined in this document.

4.2. Reply Path TLV

A Reply Path TLV [RFC7110] is defined to identify a single return path. When the initiator LSR wants to use the Reply Mode Order TLV to specify multiple return paths, then the initiator LSR SHOULD

include multiple "Reply via Specified Path" (value=5) Reply Mode values and multiple corresponding Reply Path TLV objects (one Reply Path TLV corresponding to a "Reply via Specified Path" Reply Mode and one Reply Path TLV identifying a return path).

As described in Section 3.1, it is valid to use the "Reply via Specified Path" Reply Mode without inclusion in a Reply Path TLV. For the Reply Mode Order TLV, it is also valid to include a "Reply via Specified Path" Reply Mode value without a corresponding Reply Path TLV; this implies that the reverse LSP is the preferred return path. When multiple consecutive "Reply via Specified Path" Reply Mode values are included but fewer corresponding Reply Path TLV objects exist, the responder LSR SHOULD think that the former "Reply via Specified Path" Reply Mode values have corresponding Reply Path TLVs. The latter "Reply via Specified Path" Reply Mode values have no corresponding Reply Path TLVs. For example, if the Reply Mode Order TLV carries Reply Modes {5, 5, 5} and only two Reply Path TLVs carry FEC X and FEC Y, respectively, then the reply path order is as follows:

1. Reply via Specified Path (FEC X)

2. Reply via Specified Path (FEC Y)

3. Reply via Specified Path (reverse LSP)

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

If the initiator LSR was interested in encoding the 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:

- The Reply Mode Order TLV carrying Reply Modes {4, 5, 5, 2}

- One Reply Path TLV carrying FEC X

- One Reply Path TLV carrying FEC Y

The encoding specified by the Reply Mode Order TLV and the Reply Path TLV in the MPLS echo request message will cause 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)".

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

If the initiator LSR was interested in encoding the 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:

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

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

4.3. Proxy LSP Ping

The mechanism defined in this document will work with Proxy LSP Ping as defined by [RFC7555]. 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 Reply Mode 2 (Reply via an IPv4/IPv6 UDP packet) be used in the Reply Mode field of the MPLS proxy ping request message.

4.3.1. Proxy LSR Sending an MPLS Echo Request

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

- The Reply Mode field.
- The Reply Mode Order TLV.
- The Reply Path TLV(s). If there is more than one Reply Path TLV, then the ordering of the TLVs MUST be preserved when copying.

4.3.2. 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

The security considerations specified in [RFC4379] and [RFC7110] also apply for this document.

In addition, this document introduces the Reply Mode Order TLV. It provides a new way for an unauthorized source to gather more network information, especially information regarding the potential return path(s) of an LSP. To protect against unauthorized sources using MPLS echo request messages with the Reply Mode Order TLV to obtain network information, as also specified in [RFC4379], it is RECOMMENDED that implementations provide a means of checking the source addresses of MPLS echo request messages against an access list before accepting the message.

Another potential security issue is that the MPLS echo request and reply messages are not encrypted; the contents of the MPLS echo request and reply messages may therefore be potentially exposed. Although the exposure is within the MPLS domain, if such exposure is a concern, some encryption mechanisms [MPLS-OPP-ENCR] may be employed.

6. Manageability Considerations

Section 2 described problems that increase complexity with respect to operations and implementations. In order to simplify operations and to allow LSP Ping and Traceroute to function efficiently whilst preserving code simplicity, it is RECOMMENDED that implementations allow devices to have configuration options to set operator-preferred Reply Modes. For example:

- For those operators who are more interested in MPLS echo reply packets reaching the initiator LSR:
  1. Reply via an IPv4/IPv6 UDP packet (2)

2. Reply via application level control channel (4)

3. Reply via Specified Path (5)

- For those operators who are more interested in MPLS echo reply packets testing the paths related to the forward LSP:

  1. Reply via Specified Path (5)

  2. Reply via application level control channel (4)

  3. Reply via an IPv4/IPv6 UDP packet (2)

7. IANA Considerations

7.1. New Reply Mode Order TLV

IANA has assigned a new TLV type value from the "TLVs" sub-registry within the "Multiprotocol Label Switching Architecture (MPLS) Label Switched Paths (LSPs) Ping Parameters" registry, for the Reply Mode Order TLV.

The new TLV Type value has been assigned from the range 32768-49161, as specified in Sections 3 and 7.2 of [RFC4379]; this range is for optional TLVs that can be silently dropped if not recognized.

Type	Meaning	Reference
32770	Reply Mode Order TLV	RFC 7737

8. References

8.1. Normative References

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

[RFC4379] Kompella, K. and G. Swallow, "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", RFC 4379, DOI 10.17487/RFC4379, February 2006, <http://www.rfc-editor.org/info/rfc4379>.

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

8.2. Informative References

[MPLS-OPP-ENCR]
Farrel, A. and S. Farrell, "Opportunistic Security in MPLS Networks", Work in Progress, draft-ietf-mpls-opportunistic-encrypt-00, July 2015.

[RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, DOI 10.17487/RFC3473, January 2003, <http://www.rfc-editor.org/info/rfc3473>.

[RFC5960] Frost, D., Ed., Bryant, S., Ed., and M. Bocci, Ed., "MPLS Transport Profile Data Plane Architecture", RFC 5960, DOI 10.17487/RFC5960, August 2010, <http://www.rfc-editor.org/info/rfc5960>.

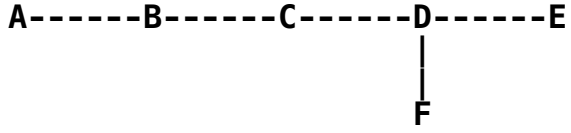
[RFC7555] Swallow, G., Lim, V., and S. Aldrin, "Proxy MPLS Echo Request", RFC 7555, DOI 10.17487/RFC7555, June 2015, <http://www.rfc-editor.org/info/rfc7555>.

Appendix A. Reply Mode Order TLV Beneficial Scenarios

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

A.1. Incorrect Forwarding Scenario

As shown in Figure 2, a network has an LSP with the forwarding path A-B-C-D-E. The LSP has a control channel.



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

Figure 2: Incorrect Forwarding

If D is incorrectly label switching to F (instead of E), then LSP Traceroute with "Reply via application level control channel (4)" will result in the following:

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

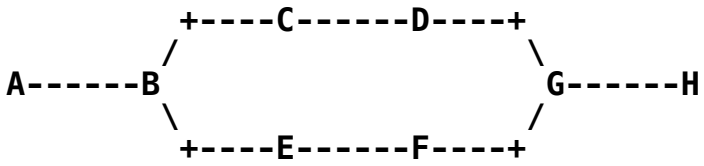
This is because F does not have a control channel on which to send the MPLS echo reply message. With the extensions described in this document, the same procedures can be performed with the Reply Mode Order TLV carrying {4, 2}. When LSP Traceroute is issued, then the following output may be displayed 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

As shown in Figure 3, a network has a bidirectional LSP where the forward LSP and the reverse LSP are not fully co-routed.



Forward Path: A-B-C-D-G-H (upper path)
Reverse Path: H-G-F-E-B-A (lower path)

Figure 3: Non-Co-Routed Bidirectional LSP

Some operators may prefer and configure "Reply via Specified Path" as the default Reply Mode but without including the Reply Path TLV, to indicate that the reverse LSP is used as the return path when MPLS echo request messages are sent on bidirectional LSPs. Without the extensions described in this document, the following behaviors will be seen:

- When LSP Ping is issued from A, the reply will come back on the reverse LSP from H.
- When LSP Traceroute is issued from A, the replies 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 LSPs on those nodes.
- When LSP Ping with a 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 the MPLS echo request terminates on a node that has a reverse LSP).

One can argue that the initiator LSR can automatically generate the same MPLS echo request with a different Reply Mode value to those nodes that time out. However, such a mechanism will result in an 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 extensions described in this document, the same procedures can be performed with the Reply Mode Order TLV carrying {5, 2}. When LSP Traceroute is issued, then the 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
```

Acknowledgements

The authors especially thank Tom Taylor, who passed away close to the time of publication of this RFC. Tom did a careful review of the document and provided useful comments.

The authors would also like to thank Santiago Alvarez and Faisal Iqbal for discussions that motivated the creation of this document; Sam Aldrin, Curtis Villamizar, Ross Callon, Jeffrey Zhang, Jeremy Whittaker, Mustapha Alissaoui, Qin Wu, Jie Dong, and Adrian Farrel for providing valuable comments that influenced the content of the document; and Dan Frost, Victor Kuarsingh, and Deborah Brungard for reviewing the document and providing useful comments.

Contributors

Shaleen Saxena
Brocade

Email: ssaxena@brocade.com

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