

MPLS Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: April 24, 2014

N. Kumar  
G. Swallow  
C. Pignataro  
N. Akiya  
Cisco Systems, Inc.  
M. Chen  
Huawei  
October 21, 2013

Label Switched Path (LSP) Ping/Trace for Segment Routing Networks Using  
MPLS Dataplane  
[draft-kumar-mpls-spring-lsp-ping-00](#)

## Abstract

Segment Routing architecture leverages the source routing and tunneling paradigm and can be directly applied to MPLS dataplane. A node steers a packet through a controlled set of instructions called segments, by prepending the packet with Segment Routing header.

The segment assignment and forwarding semantic nature of Segment Routing raises additional consideration for connectivity verification and fault isolation in LSP with SPRING architecture. This document illustrates the problem and describe a mechanism to perform LSP Ping and Traceroute on Segment Routing network over MPLS dataplane.

## 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 April 24, 2014.

## Copyright Notice

Copyright (c) 2013 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

<a href="#">1.</a>	Introduction . . . . .	<a href="#">2</a>
<a href="#">2.</a>	Requirements notation . . . . .	<a href="#">3</a>
<a href="#">3.</a>	Path validation in Segment Routing networks . . . . .	<a href="#">3</a>
<a href="#">4.</a>	Segment Routing Sub-TLV Format . . . . .	<a href="#">4</a>
<a href="#">4.1.</a>	IPv4 Prefix Node Segment ID . . . . .	<a href="#">4</a>
<a href="#">4.2.</a>	IPv6 Prefix Node Segment ID . . . . .	<a href="#">6</a>
<a href="#">4.3.</a>	IGP Adjacency Segment ID . . . . .	<a href="#">7</a>
<a href="#">5.</a>	Extension to Downstream Mapping TLV . . . . .	<a href="#">8</a>
<a href="#">6.</a>	Procedures . . . . .	<a href="#">8</a>
<a href="#">6.1.</a>	FECs in Target FEC Stack TLV . . . . .	<a href="#">8</a>
<a href="#">6.2.</a>	FEC Stack Change TLV . . . . .	<a href="#">9</a>
<a href="#">6.3.</a>	PHP, Adjacency SID Pop, Implicit NULL . . . . .	<a href="#">9</a>
<a href="#">6.4.</a>	Segment Protocol Check . . . . .	<a href="#">10</a>
<a href="#">7.</a>	IANA Considerations . . . . .	<a href="#">10</a>
<a href="#">8.</a>	Security Considerations . . . . .	<a href="#">11</a>
<a href="#">9.</a>	Acknowledgement . . . . .	<a href="#">11</a>
<a href="#">10.</a>	Contributing Authors . . . . .	<a href="#">11</a>
<a href="#">11.</a>	References . . . . .	<a href="#">11</a>
<a href="#">11.1.</a>	Normative References . . . . .	<a href="#">11</a>
<a href="#">11.2.</a>	Informative References . . . . .	<a href="#">11</a>
	Authors' Addresses . . . . .	<a href="#">12</a>

## [1.](#) Introduction

[I-D.filsfils-rtgwg-segment-routing] introduces and explains Segment Routing architecture that leverages the source routing and tunneling paradigm. A node steers a packet through a controlled set of instructions called segments, by prepending the packet with SR header. A detailed definition about Segment Routing architecture is available in [draft-filsfils-rtgwg-segment-routing](#) and different use-cases are discussed in [draft-filsfils-rtgwg-segment-routing-use-cases](#).

The Segment Routing architecture can be directly applied to MPLS dataplane in a way that, the segment will be of 20-bits size and SR header is the label stack.



```
500x --> Node Segment ID for Router X
        (Ex: 5006 is node segment ID for R6)
9axy --> Adj Segment ID from Router X to Y over link a
        (Ex: 9136 is Adj segment ID from R3 to R6 via link 1)
```



The forwarding semantic of Adjacency segment is to pop the segment and send the packet to a specific neighbor over a specific link. A malfunctioning node may forward packets using Adjacency segment to incorrect neighbor or over incorrect link. Exposed segment (after incorrectly forwarded Adjacency segment) might still allow such packet to traverse to intended destination, yet intended strict traversal has been broken.

Assume in above topology, R1 sends traffic with segment stack as {9124, 5007, 9378} so that the path taken will be R1-R2-R4-R5-R7-R8. If the adjacency segment 9124 is misprogrammed in R2 to send the packet to R1 or R3, it will still be delivered to R8 but is not via the expected path.

MPLS traceroute may help with detecting such deviation in above mentioned scenario. However, it may not be helpful if R3, due to misprogramming, forwards packet with adjacency segment 9236 via link L1 while it is expected to be forwarded over Link L2.

This document defines Target FEC Stack sub-TLVs and explains how they can be used to tackle above challenges.

#### 4. Segment Routing Sub-TLV Format

The format of the following FEC Sub-TLVs follows the philosophy of Target FEC Stack TLV carrying FECs corresponding to each label in the label stack. When operated with the procedures defined in [\[RFC4379\]](#), this allows LSP ping/traceroute operations to function when Target FEC Stack TLV contains more FECs than received label stack at responder nodes.

Type	Sub-Type	Value Field
----	-----	-----
1	TBD1	IPv4 Prefix Node Segment ID
	TBD2	IPv6 Prefix Node Segment ID
	TBD3	Adjacency Segment ID

Service Segments and FRR will be considered in future version.

##### 4.1. IPv4 Prefix Node Segment ID

The format is as below:

```

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

```



```

+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               IPv4 Prefix                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|Prefix Length |      Resv      |      Protocol  |  SID Length  |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Node Segment ID                          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Reserved                                |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Advertising Node Identifier              |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

#### IPv4 Prefix

This field carries the IPv4 prefix to which the Node Segment ID is assigned.

#### Prefix Length

One octet of prefix length in bits.

#### Protocol

Set to 1 if the IGP protocol is OSPF and 2 if IGP protocol is ISIS.

#### SID Length

Set to 3 or 4 depending on the Segment ID.

#### Node Segment ID

This field carries the Node segment ID. If the SID Length is 3, then the 20 rightmost bits represent the segment. If length is 4, then the value represent a 32 bits Segment ID.

#### Advertising Node Identifier

Specifies the advertising node identifier. When Protocol is set to 1, then the 32 rightmost bits represent OSPF Router ID and if protocol is set to 2, this field carries 48 bit ISIS System ID.





This field carries the Node segment ID. If the SID Length is 3, then the 20 rightmost bits represent the segment. If length is 4, then the value represent a 32 bits Segment ID.



Specifies the advertising node identifier. When Protocol is set to 1, then the 32 rightmost bits represent OSPF Router ID and if protocol is set to 2, this field carries 48 bit ISIS System ID.

The format is as below:

[illegible]

An identifier that is assigned by local LSR for a link on which Adjacency SID is binded. If the Adj-SID represents parallel adjacencies (Section 3.3.2.1 of [\[I-D.filesfiles-rtgwg-segment-routing\]](#)) this field MUST be set to zero.

An identifier that is assigned by remote LSR for a link on which Adjacency SID is binded. If the Adj-SID represents parallel adjacencies (Section 3.3.2.1 of [\[I-D.filsfils-rtgwg-segment-routing\]](#)) this field MUST be set to zero.

Set to 1 if the IGP protocol is OSPF and 2 if IGP protocol is ISIS

## SID Length



Set to 3 or 4 depending on the Segment ID.

#### Advertising Node Identifier

Specifies the advertising node identifier. When Protocol is set to 1, then the 32 rightmost bits represent OSPF Router ID and if protocol is set to 2, this field carries 48 bit ISIS System ID.

#### IGP Adjacency Segment ID

This field carries the adjacency segment ID.

## 5. Extension to Downstream Mapping TLV

In an echo reply, the Downstream Mapping TLV [[RFC4379](#)] is used to report for each interface over which a FEC could be forwarded. For an FEC, there are multiple protocols that may be used to distribute label mapping. The "Protocol" field of the Downstream Mapping TLV is used to return the protocol that is used to distribute a specific a label. The following protocols are defined in [section 3.2 of \[RFC4379\]](#):

Protocol #	Signaling Protocol
-----	-----
0	Unknown
1	Static
2	BGP
3	LDP
4	RSVP-TE

With segment routing, OSPF or ISIS can be used for label distribution, this document adds two new protocols as follows:

Protocol #	Signaling Protocol
-----	-----
5	OSPF
6	ISIS

## 6. Procedures

This section describes aspects of LSP ping/traceroute operations that require further considerations beyond [[RFC4379](#)].

### 6.1. FECs in Target FEC Stack TLV



When LSP echo request packets are generated by an initiator, FECs carried in Target FEC Stack TLV may need to or desire to have deviating contents. This document outlines expected Target FEC Stack TLV construction mechanics by initiator for known scenarios.

#### Ping

Initiator MUST include FEC(s) corresponding to the destination segment.

Initiator MAY include FECs corresponding to some or all of segments imposed in the label stack by the initiator to communicate the segments traversed.

#### Traceroute

Initiator MUST initially include FECs corresponding to all of segments imposed in the label stack.

When a received echo reply contains FEC Stack Change TLV with one or more of original segment(s) being popped, initiator MAY remove corresponding FEC(s) from Target FEC Stack TLV in the next (TTL+1) traceroute request.

When a received echo reply does not contain FEC Stack Change TLV, initiator MUST NOT attempt to remove FEC(s) from Target FEC Stack TLV in the next (TTL+1) traceroute request. Note that Downstream Label field of DSMAP/DDMAP contains hints on how initiator may be able to update the contents of next Target FEC Stack TLV. However, such hints are ambiguous without full understanding of PHP capabilities.

### **6.2. FEC Stack Change TLV**

The network node which advertised the node segment ID is responsible for generating FEC Stack Change TLV of &pop& operation for node segment ID, regardless of if PHP is enabled or not.

The network node that is immediate downstream of the node which advertised the adjacency segment ID is responsible for generating FEC Stack Change TLV of &pop& operation for adjacency segment ID.

### **6.3. PHP, Adjacency SID Pop, Implicit NULL**

Forwarding behavior of node segment ID PHP is equivalent to usage of implicit Null in MPLS protocols that embraces downstream label allocation scheme. Adjacency segment ID is also similar in a sense that it can be thought as nexthop destined locally allocated segment





that has PHP enabled. Procedures described in [Section 4.4 of \[RFC4379\]](#) relies on Stack-D and Stack-R explicitly having Implicit Null value. It may simplify implementations to reuse Implicit Null for node segment ID PHP and adjacency segment ID cases. However, it is technically incorrect for Implicit Null value to externally appear. Therefore, implicit Null MUST NOT be placed in Stack-D and Interface and Label Stack TLV for node segment ID PHP and adjacency segment ID cases.

#### **6.4. Segment Protocol Check**

If the Target FEC Sub-TLV at FEC-stack-depth is TBD1 (IPv4 Prefix Node Segment ID), set Best return code to (error code TBD) if any below conditions fail:

- \* Validate that Advertising Node Identifier of Protocol is a local node.
- \* Validate that Node Segment ID is advertised for IPv4 Prefix by the Protocol with Advertising Node Identifier.
- \* Validate that Node Segment ID is advertisement of PHP bit.

If the Target FEC Sub-TLV at FEC-stack-depth is TBD2 (IPv6 Prefix Node Segment ID), set Best return code to (error code TBD) if any below conditions fail:

- \* Validate that Advertising Node Identifier of Protocol is a local node.
- \* Validate that Node Segment ID is advertised for IPv6 Prefix by the Protocol with Advertising Node Identifier.
- \* Validate that Node Segment ID is advertised of PHP bit.

If the Target FEC sub-TLV at FEC-stack-depth is TBD3 (Adjacency Segment ID), set Best return code to (error code TBD) if any below conditions fail:

- \* Validate that Remote Interface ID matches the local identifier of the interface on which the packet was received.
- \* Validate that IGP Adjacency SID is advertised by Advertising Node Identifier of Protocol in local IGP database.

#### **7. IANA Considerations**

To be Updated.



## **8. Security Considerations**

To be Updated.

## **9. Acknowledgement**

The authors would like to thank Stefano Previdi for his review and comments.

## **10. Contributing Authors**

Tarek Saad  
Cisco Systems  
Email: tsaad@cisco.com

Siva Sivabalan  
Cisco Systems  
Email: msiva@cisco.com

## **11. References**

### **11.1. Normative References**

- [I-D.filsfils-rtgwg-segment-routing]  
Filsfils, C., Previdi, S., Bashandy, A., Decraene, B., Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R., Ytti, S., Henderickx, W., Tantsura, J., and E. Crabbe, "Segment Routing Architecture", [draft-filsfils-rtgwg-segment-routing-01](#) (work in progress), October 2013.
- [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.
- [RFC6424] Bahadur, N., Kompella, K., and G. Swallow, "Mechanism for Performing Label Switched Path Ping (LSP Ping) over MPLS Tunnels", [RFC 6424](#), November 2011.

### **11.2. Informative References**

- [RFC6291] Andersson, L., van Helvoort, H., Bonica, R., Romascanu, D., and S. Mansfield, "Guidelines for the Use of the "OAM" Acronym in the IETF", [BCP 161](#), [RFC 6291](#), June 2011.



[RFC6425] Saxena, S., Swallow, G., Ali, Z., Farrel, A., Yasukawa, S., and T. Nadeau, "Detecting Data-Plane Failures in Point-to-Multipoint MPLS - Extensions to LSP Ping", [RFC 6425](#), November 2011.

#### Authors' Addresses

Nagendra Kumar  
Cisco Systems, Inc.  
7200 Kit Creek Road  
Research Triangle Park, NC 27709  
US

Email: [naikumar@cisco.com](mailto:naikumar@cisco.com)

George Swallow  
Cisco Systems, Inc.  
1414 Massachusetts Ave  
Boxborough, MA 01719  
US

Email: [swallow@cisco.com](mailto:swallow@cisco.com)

Carlos Pignataro  
Cisco Systems, Inc.  
7200 Kit Creek Road  
Research Triangle Park, NC 27709-4987  
US

Email: [cpignata@cisco.com](mailto:cpignata@cisco.com)

Nobo Akiya  
Cisco Systems, Inc.  
2000 Innovation Drive  
Kanata, ON K2K 3E8  
Canada

Email: [nobo@cisco.com](mailto:nobo@cisco.com)

Mach(Guoyi) Chen  
Huawei

Email: [mach.chen@huawei.com](mailto:mach.chen@huawei.com)

