SPRING Working Group
Internet-Draft
Intended status: Standards Track
Expires: February 23, 2020

R. Gandhi, Ed.
Z. Ali
C. Filsfils
F. Brockners
Cisco Systems, Inc.
B. Wen
V. Kozak
Comcast

August 22, 2019

Segment Routing with MPLS Data Plane Encapsulation for In-situ OAM Data draft-gandhi-spring-ioam-sr-mpls-02

Abstract

In-situ Operations, Administration, and Maintenance (IOAM) records operational and telemetry information in the data packet while the packet traverses a path between two nodes in the network. Segment Routing (SR) technology leverages the source routing paradigm. This document defines how IOAM data fields are transported with the Segment Routing with MPLS data plane (SR-MPLS) encapsulation. The procedures defined are also equally applicable to all other MPLS data plane encapsulations.

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."

Table of Contents

1. Introduction

In-situ Operations, Administration, and Maintenance (IOAM) records operational and telemetry information within the packet while the packet traverses a particular network domain. The term "in-situ" refers to the fact that the IOAM data fields are added to the data packets rather than being sent within the probe packets specifically dedicated to OAM or Performance Measurement (PM). The IOAM data fields are defined in [I-D.ietf-ippm-ioam-data], and can be used for various use-cases for OAM and PM.

Segment Routing (SR) technology leverages the source routing paradigm [I-D.ietf-spring-segment-routing-mpls]. A node steers a packet through a controlled set of instructions, called segments, by pre-pending the packet with an SR header. In the MPLS data plane,

the SR header is instantiated through a label stack.

This document defines how IOAM data fields are transported with the SR with MPLS data plane (SR-MPLS) encapsulation. The procedures defined are also equally applicable to all other MPLS data plane encapsulations.

2. Conventions

2.1. Requirement 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 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2.2. Abbreviations

Abbreviations used in this document:

ECMP	Equal Cost Multi-Path
IOAM	In-situ Operations, Administration, and Maintenance
MPLS	Multiprotocol Label Switching
OAM	Operations, Administration, and Maintenance
PBT	Postcard Based Telemetry
PM	Performance Measurement
PoT	Proof-of-Transit
SR	Segment Routing
SR-MPLS	Segment Routing with MPLS Data plane

3. IOAM Data Field Encapsulation in SR-MPLS Header

SR-MPLS encapsulation is defined in [I-D.ietf-spring-segment-routing-mpls]. The IOAM data fields are defined in [I-D.ietf-ippm-ioam-data]. IOAM data fields are carried in the SR-MPLS header as shown in Figure 1 and Figure 2. More than one trace options can be present in the IOAM data fields. The Indicator Label is added at the bottom of the MPLS label stack (S

flag set to 1) to indicate the presence of the IOAM data field(s) in the MPLS header.

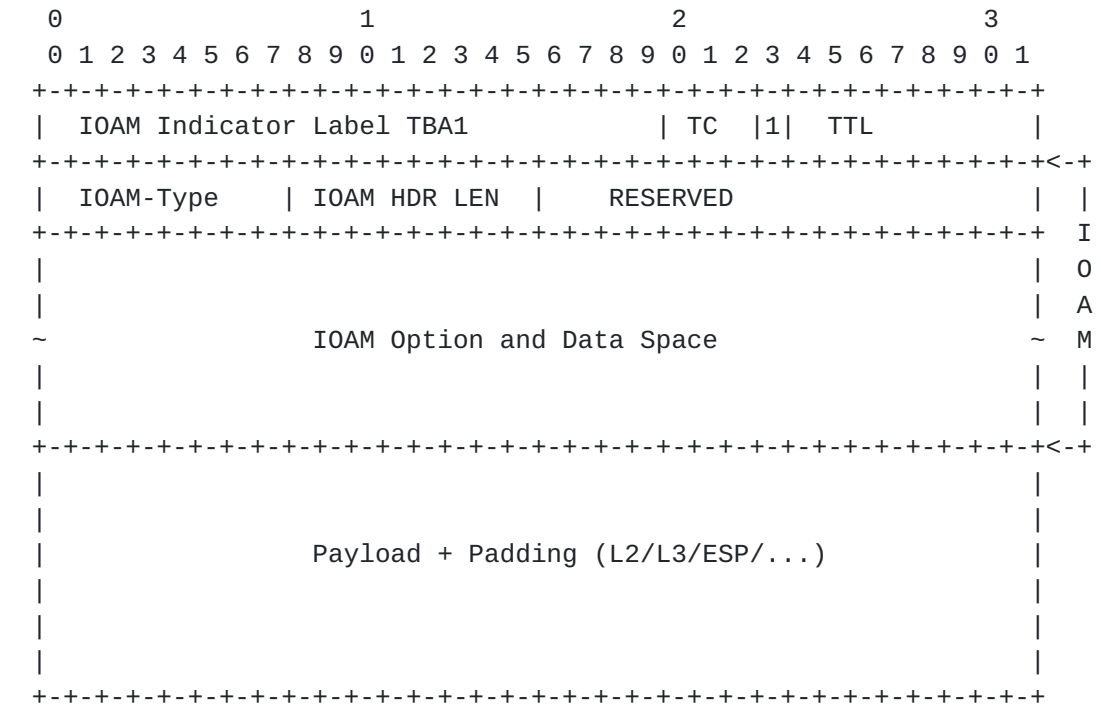
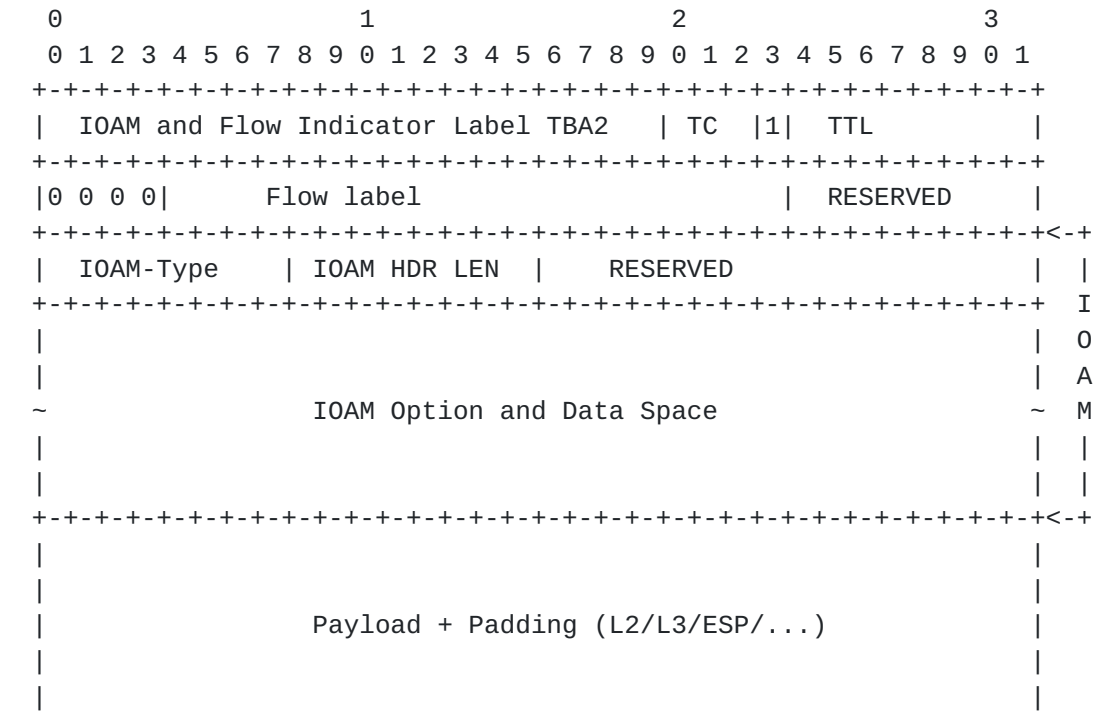


Figure 1: IOAM data encapsulation in SR-MPLS Header



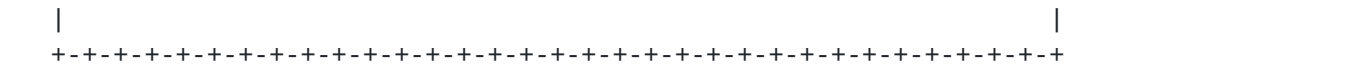


Figure 2: IOAM data encapsulation with Flow Label in SR-MPLS Header

Indicator Label and Flow Label as defined in this document.

The fields related to the encapsulation of IOAM data fields in the SR-MPLS header are defined as follows:

IOAM-Type: 8-bit field defining the IOAM Option type, as defined in Section 7.2 of [I-D.ietf-ippm-ioam-data].

IOAM HDR LEN: 8-bit unsigned integer. Length of the IOAM HDR in 4-octet units.

RESERVED: 8-bit reserved field MUST be set to zero upon transmission and ignored upon receipt.

IOAM Option and Data Space: IOAM option header and data is present as defined by the IOAM-Type field, and is defined in Section 4 of [I-D.ietf-ippm-ioam-data].

4. Procedure for Edge-to-Edge IOAM

This section summarizes the procedure for data encapsulation and decapsulation for IOAM Edge-to-Edge Option Type [I-D.ietf-ippm-ioam-data] in SR-MPLS header.

- The encapsulating node inserts the IOAM Indicator Label or IOAM Flow Indicator Label with Flow Label and one or more IOAM data field(s) in the MPLS header. The procedure to generate the Flow Label is outside the scope of this document.

- The decapsulating node "forwards and punts the timestamped copy" of the data packet including IOAM data fields when the node recognizes the IOAM Indicator Label and IOAM Flow Indicator Label. The copy of the data packet is punted to the slow path for OAM processing and is not necessarily punted to the control-plane. The receive timestamp is required by various OAM use-cases.

- The decapsulating node processes the IOAM data field(s) using the procedures defined in [I-D.ietf-ippm-ioam-data]. An example of IOAM processing may be to export the data fields, send data fields via Telemetry, etc.

o The decapsulating node also pops the Indicator Label and the IOAM data fields from the MPLS header.

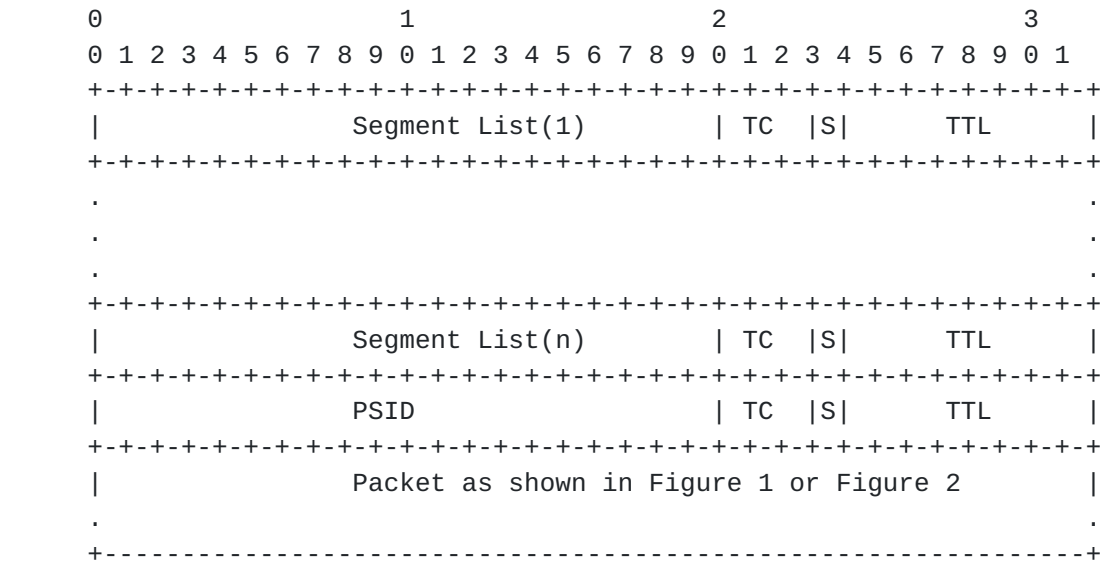


Figure 3: Data Packet over SR-MPLS Policy

4.1. Edge-to-Edge IOAM Indicator Labels

IOAM Indicator Label (value TBA1) and IOAM and Flow Indicator Label (value TBA2) are used to indicate the presence of the IOAM data field in the MPLS header.

The Indicator Label with value TBA2 is used to carry a second label underneath with protocol value 0000b and 20-bit Flow Label. The protocol value 0000b allows to avoid incorrect IP header based hashing over ECMP paths that uses the value 0x4 (for IPv4) and value 0x6 (for IPv6) [RFC4928]. The Flow Label identifies the traffic flow that can be used for IOAM purpose as well as for hashing over ECMP paths.

The IOAM Indicator Label and IOAM and Flow Indicator Label can be allocated using one of the following methods:

o Labels assigned by IANA with value TBA1 and TBA2 from the Extended Special-Purpose MPLS Values [mpls-spl-terminology].

o Labels allocated by a controller from the global table of the decapsulating node. The controller provisions the label on both encapsulating and decapsulating nodes.

o Labels allocated by the decapsulating node. The signaling

extension for this is outside the scope of this document.

5. Procedure for Hop-by-Hop IOAM

The hop-by-hop IOAM includes IOAM-Types IOAM Pre-allocated Trace Option Type, IOAM Incremental Trace Option Type and IOAM POT Option Type.

Different Indicator Labels (TBA3 and TBA4) are used for hop-by-hop IOAM.

The details for hop-by-hop IOAM will be added in a future version of the document.

6. Considerations for ECMP

The encapsulating node needs to make sure the IOAM data field does not start with a well known IP protocol value (e.g. 0x4 for IPv4 and 0x6 for IPv6) as it can alter the hashing function for ECMP that uses the IP header. This can be achieved by using the IOAM and Flow Indicator Label (value TBA2 and TBA4) that follows by protocol value 0000b. This approach is consistent with the use of utilizing 0000b as the first nibble after the MPLS label stack, as described in [RFC4928] [RFC4385].

Note that the hashing function for ECMP that uses the labels from the MPLS header may also now include the Indicator Label.

The entropy label can be used for hashing function for ECMP as defined in [RFC6790].

7. Node Capability

The decapsulating node that has to pop the Indicator Label, data fields, and perform the IOAM function may not be capable of supporting it. The encapsulating node needs to know if the decapsulating node can support the IOAM function. The signaling extension for this capability exchange is outside the scope of this document.

8. IANA Considerations

IANA maintains the "Special-Purpose Multiprotocol Label Switching (MPLS) Label Values" registry (see

<https://www.iana.org/assignments/mpls-label-values/mpls-label-values.xml>). IANA is requested to allocate IOAM Indicator Label value and IOAM and Flow Indicator value from the "Extended Special-Purpose MPLS Label Values" registry:

Value	Description	Reference
TBA1	E2E IOAM Indicator Label	This document
TBA2	E2E IOAM and Flow Indicator Label	This document
TBA3	HbH IOAM Indicator Label	This document
TBA4	HbH IOAM and Flow Indicator Label	This document

9. Security Considerations

The security considerations of SR-MPLS are discussed in [I-D.ietf-spring-segment-routing-mpls], and the security considerations of IOAM in general are discussed in [I-D.ietf-ippm-ioam-data].

IOAM is considered a "per domain" feature, where one or several operators decide on leveraging and configuring IOAM according to their needs. Still, operators need to properly secure the IOAM domain to avoid malicious configuration and use, which could include injecting malicious IOAM packets into a domain.

10. Acknowledgements

The authors would like to thank Shwetha Bhandari and Vengada Prasad Govindan for the discussions on IOAM. The authors would also like to thank Tarek Saad, Loa Andersson and Cheng Li for providing many useful comments.

11. References

11.1. 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.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", RFC 8174, May 2017.

[I-D.ietf-spring-segment-routing-mpls] Bashandy, A., Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with MPLS data plane", draft-ietf-spring-segment-routing-mpls, work in progress.

[I-D.ietf-ippm-ioam-data] Brockners, F., Bhandari, S., Pignataro, C., Gredler, H., Leddy, J., Youell, S., Mizrahi, T., Mozes, D., Lapukhov, P., Chang, R., and Bernier, D., "Data Fields for In-situ OAM", draft-ietf-ippm-ioam-data, work in progress.

11.2. Informative References

[RFC4385] Bryant, S., Swallow, G., Martini, L., and D. McPherson, "Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN", RFC 4385, February 2006.

[RFC4928] Swallow, G., Bryant, S., and L. Andersson, "Avoiding Equal Cost Multipath Treatment in MPLS Networks", BCP 128, RFC 4928, June 2007.

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, November 2012.

[mpls-spl-terminology] L. Andersson, et al. "Special Purpose Label terminology", draft-ietf-mpls-spl-terminology, work in progress.

Contributors

Sagar Soni
Cisco Systems, Inc.
Email: sagsoni@cisco.com

Patrick Khordoc
Cisco Systems, Inc.
Email: pkhordoc@cisco.com

Authors' Addresses

Rakesh Gandhi (editor)
Cisco Systems, Inc.
Canada

Email: rgandhi@cisco.com

Zafar Ali
Cisco Systems, Inc.

Email: zali@cisco.com

Clarence Filsfils
Cisco Systems, Inc.
Belgium

Email: cf@cisco.com

Frank Brockners
Cisco Systems, Inc.
Hansaallee 249, 3rd Floor
DUESSELDORF, NORDRHEIN-WESTFALEN 40549
Germany

Email: fbrockne@cisco.com

Bin Wen
Comcast

Email: Bin_Wen@cable.comcast.com

Voitek Kozak
Comcast

Email: Voitek_Kozak@comcast.com