

Workgroup: MPLS Working Group
Internet-Draft: draft-gandhi-mpls-ioam-10
Published: 10 March 2023
Intended Status: Standards Track
Expires: 11 September 2023
Authors: R. Gandhi, Ed. F. Brockners B. Wen
 Cisco Systems, Inc. Cisco Systems, Inc. Comcast
 B. Decraene H. Song
 Orange Futurewei Technologies
 MPLS Data Plane Encapsulation for In Situ OAM Data

Abstract

In Situ Operations, Administration, and Maintenance (IOAM) is used for recording and collecting operational and telemetry information while the packet traverses a path between two points in the network. This document defines how IOAM data fields are transported with MPLS data plane encapsulation using MPLS Network Action (MNA).

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 <https://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 11 September 2023.

Copyright Notice

Copyright (c) 2023 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 (<https://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 Revised BSD License text as described in

Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

- [1. Introduction](#)
- [2. Conventions](#)
 - [2.1. Requirement Language](#)
 - [2.2. Abbreviations](#)
- [3. MPLS Extensions for IOAM Data Fields](#)
 - [3.1. Post-Stack Network Action for IOAM](#)
 - [3.2. MNA Sub-Stack for IOAM Post-Stack Network Action](#)
- [4. Edge-To-Edge IOAM Network Action](#)
 - [4.1. MNA Sub-Stack for Edge-To-Edge IOAM](#)
 - [4.2. Procedure for Edge-To-Edge IOAM Network Action](#)
- [5. Hop-By-Hop IOAM Network Action](#)
 - [5.1. MNA Sub-Stack for Hop-By-Hop IOAM](#)
 - [5.2. Procedure for Hop-By-Hop IOAM Network Action](#)
 - [5.3. Hop-By-Hop and Edge-To-Edge IOAM Network Actions](#)
- [6. Select IOAM Network Action](#)
- [7. Considerations for IOAM](#)
 - [7.1. Considerations for ECMP](#)
 - [7.2. Node Capability](#)
 - [7.3. Nested MPLS Encapsulation](#)
- [8. Security Considerations](#)
- [9. IANA Considerations](#)
- [10. References](#)
 - [10.1. Normative References](#)
 - [10.2. Informative References](#)
- [Acknowledgements](#)
- [Contributors](#)
- [Authors' Addresses](#)

1. Introduction

In Situ Operations, Administration, and Maintenance (IOAM) is used for recording and collecting operational and telemetry information while the packet traverses a path between two points in the network. 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. The IOAM data fields are defined in [RFC9197]. The IOAM data fields are further updated in [RFC9326] for direct export use-cases.

This document defines how IOAM data fields are transported with MPLS data plane encapsulations using MPLS Network Action (MNA).

[I-D.ietf-mpls-mna-hdr] defines mechanisms for carrying MNA Sub-Stack (NAS) above the Bottom of the label stack (BOS) and

[[I-D.jags-mpls-ps-mna-hdr](#)] defines mechanisms for carrying Post-Stack Network Actions in MPLS packets after the Bottom of the label stack. This document uses these two mechanisms and describes the procedures for carrying IOAM data fields in an MPLS packet.

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

The MPLS Network Action (MNA) terminology defined in [[I-D.ietf-mpls-mna-hdr](#)] are used in this document.

Abbreviations used in this document:

ECMP Equal Cost Multi-Path

E2E Edge-To-Edge

HBH Hop-By-Hop

IHS Ingress-To-Egress (I2E), Hop-By-Hop (HBH) or Select Scope

IOAM In Situ Operations, Administration, and Maintenance

MPLS Multiprotocol Label Switching

MNA MPLS Network Action

NAI Network Action Indicator

NASL Network Action Sub-Stack Length

OAM Operations, Administration, and Maintenance

POT Proof-of-Transit

3. MPLS Extensions for IOAM Data Fields

3.1. Post-Stack Network Action for IOAM

The Post-Stack Network Action for IOAM is defined using the specification from [[I-D.jags-mpls-ps-mna-hdr](#)].

The Post-Stack Network Action encoding contains IOAM Post-Stack Network Action Opcode (value TBA2 allocated by IANA), length in number of 4-Octet units, and IOAM Option-Type with IOAM-Data-Fields in ancillary data as shown in Figure 1. The IOAM-Data-Fields MUST follow the definitions corresponding to their IOAM-Option-Types (e.g., see Section 4.4 of [RFC9197] and Section 3 of [RFC9326]).

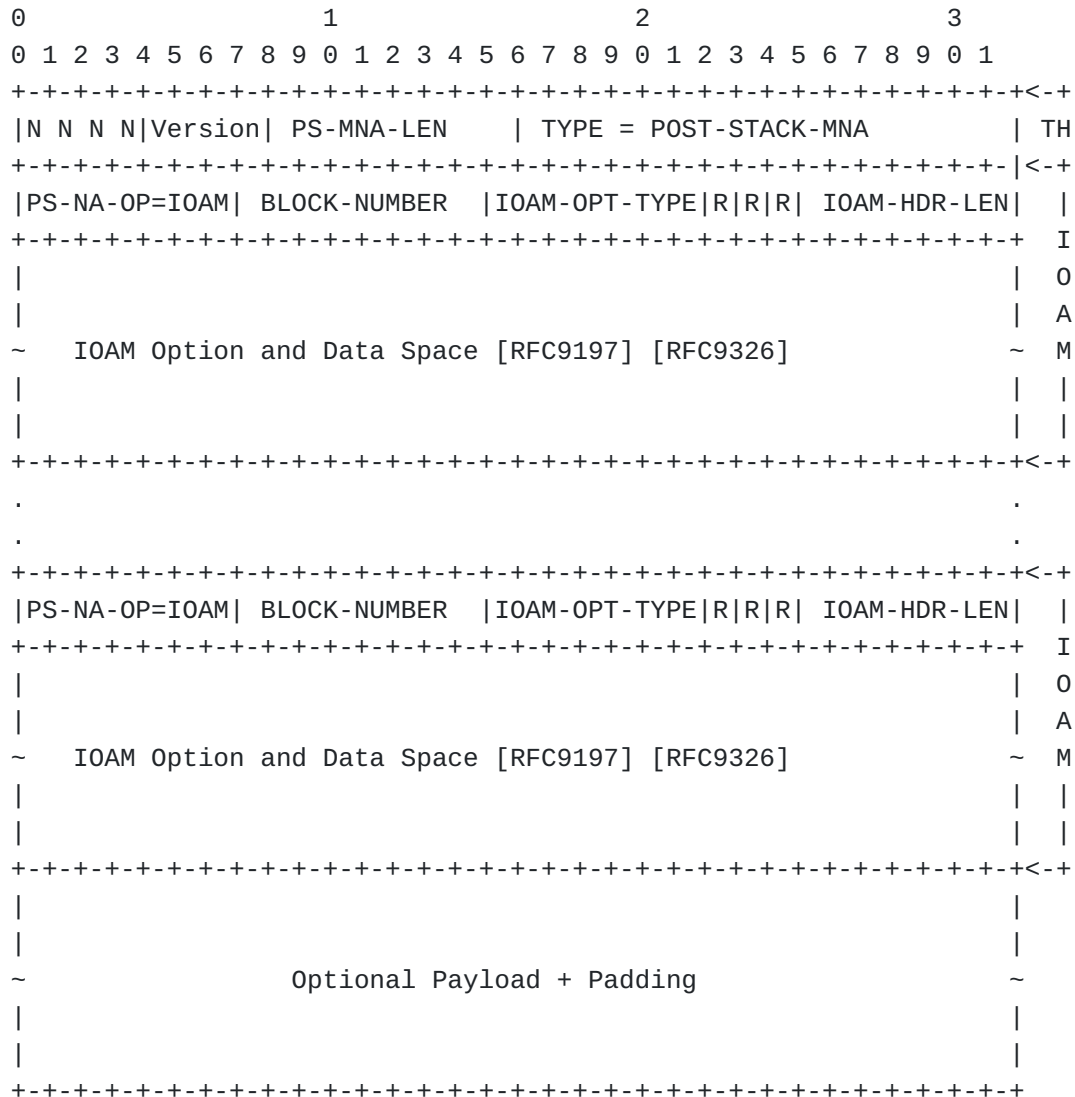


Figure 1: Post-Stack Network Action with IOAM Data Fields

The 4-Octet Top Header is added with the following fields in the the Post-Stack Network Action as defined in [[I-D. jags-mp1s-ps-mna-hdr](#)]

NNNN: The first 4-bit (value TBA3 allocated by IANA) as defined in [I-D.jags-mpls-ps-mna-hdr]. As specified in [I-D.jags-mpls-ps-mna-hdr], the NNNN value can be 0001b for Generic Associated Channel (G-ACH) [RFC5586].

PS-MNA-LEN:

Length of the Post-Stack Network Action Stack in the packet in 4-Octet units including the Top Header.

TYPE: TYPE is defined for POST-STACK-MNA (value TBA4 allocated by IANA) in [[I-D.jags-mpls-ps-mna-hdr](#)]. As specified in [[I-D.jags-mpls-ps-mna-hdr](#)], the POST-STACK-MNA Type can be a Generic Associated Channel (G-ACH) Type [RFC5586] (value TBA5 allocated by IANA).

An IOAM-Data-Field is added in the Post-Stack Network Action containing the following fields:

IOAM-OPT-TYPE: 7-bit field defining the IOAM Option-Type, as defined in the "IOAM Option-Type Registry" specified in [[RFC9197](#)] and [[RFC9326](#)]).

IOAM-HDR-LEN: 7-bit unsigned integer. Length of the IOAM Data-Fields in 4-Octet units.

IOAM Option and Data Space: IOAM-Data-Fields as specified by the IOAM-OPT-Type field. IOAM-Data-Fields are defined corresponding to the IOAM-Option-Type (e.g., see Section 4.4 of [[RFC9197](#)] and Section 3 of [[RFC9326](#)]).

BLOCK-NUMBER: The Block Number can be used to aggregate the IOAM data collected in data plane, e.g., to compute measurement metrics for each block of a data flow. It is also used to correlate the IOAM data on different nodes.

3.2. MNA Sub-Stack for IOAM Post-Stack Network Action

The P flag defined in [[I-D.jags-mpls-ps-mna-hdr](#)] is used to indicate the presence of Post-Stack Network Action and Ancillary Data.

A Post-Stack Network Action Indicator (P flag) MUST be set to "1" to indicate the presence of the Post-Stack Network Action with IOAM-Data-Fields in the Post-Stack Network Action (after the Bottom of the Stack).

The IHS scope field defined in [[I-D.ietf-mpls-mna-hdr](#)] is used to indicate that E2E or HBH or Select processing is required for the Post-Stack Network Action and Ancillary Data.

If both edge and intermediate nodes need to process the IOAM data fields then IHS scope MUST be set to "HBH, value 0x1". If only edge nodes need to process the IOAM data fields then IHS scope MUST be set to "I2E, value 0x0". The HBH scope allows to skip the IOAM data processing on the intermediate nodes i.e., avoids the need to parse all IOAM-Data-Fields to detect the HBH option type.

The MNA Label used in this document is a Base Special Purpose Label (bSPL value TBA1 to be allocated by IANA), which is also called Network Action Sub-Stack Indicator [[I-D.ietf-mpls-mna-hdr](#)].

Opcode 2 is used to carry the Flag-Based Network Action Indicators (NAIs) in the Ancillary Data field [[I-D.ietf-mpls-mna-hdr](#)]. The Flag-Based Network Action Indicators is set to 0x0 to indicate that no Network Action other than the Post-Stack Network Action(s) is required for the packet.

The U Flag for Unknown Action Handling is specified in [[I-D.ietf-mpls-mna-hdr](#)].

The Network Action Sub-Stack Length (NASL) is set to 0 if no additional Label Stack Entry (LSE) is added after the LSE carrying the P Flag.

4. Edge-To-Edge IOAM Network Action

4.1. MNA Sub-Stack for Edge-To-Edge IOAM

The IHS scope is set to "I2E, value 0x0" to indicate the scope of E2E IOAM-Data-Fields in the MPLS Sub-Stack as shown in Figure 2. Note that there may be additional Network Action (NA) LSEs added in the MNA Sub-Stack and NASL is updated based on the number of LSEs after the LSE with P flag.

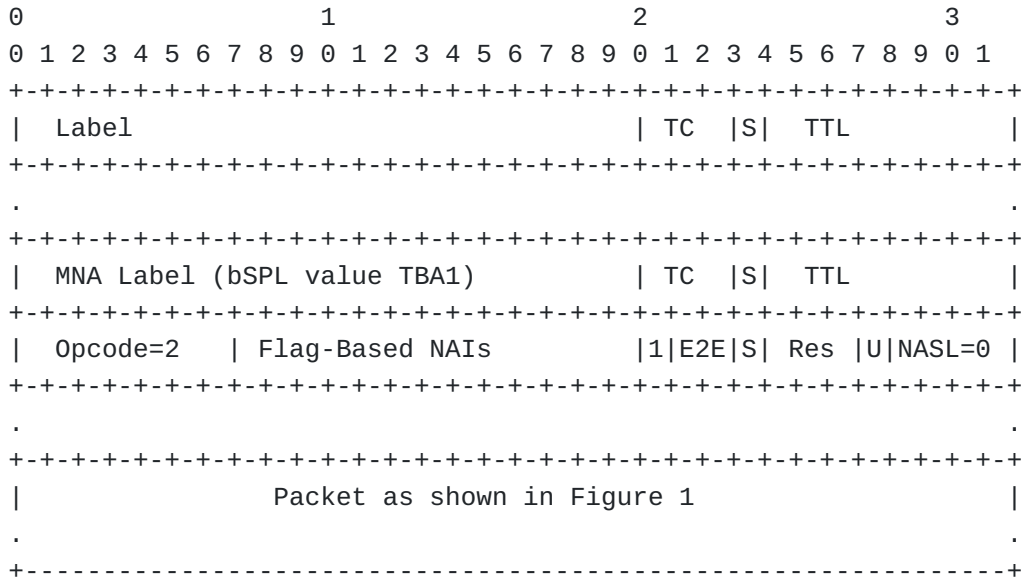


Figure 2: Example MNA Sub-Stack for E2E IOAM

The E2E IOAM-Data-Fields carry the Option-Type(s) that require processing on the encapsulating and decapsulating nodes only. The IOAM Option-Type carried can be IOAM Edge-To-Edge Option-Type (value

3) defined in [\[RFC9197\]](#) as well as Direct Export (DEX) Option-Type (value 4) defined in [\[RFC9326\]](#). The E2E IOAM-Data-Fields SHOULD NOT carry any IOAM Option-Type that require IOAM processing on the intermediate nodes as it will not be processed by them since IHS scope is set to "I2E, value 0x0".

4.2. Procedure for Edge-To-Edge IOAM Network Action

The E2E IOAM Network Action procedure is summarized as following:

- *The encapsulating node inserts an MNA Sub-Stack with the MNA Label (bSPL value TBA1) with the Post-Stack Network Action Indicator (P flag) set to "1" and IHS scope set to "I2E, value 0x0" and one or more IOAM-Data-Fields in the Post-Stack Network Action in the MPLS packet.
- *The intermediate nodes do not process IOAM-Data-Fields.
- *The penultimate node MUST NOT remove the MNA Sub-Stack from the MPLS header so that the MNA Sub-Stack is received at the decapsulating node.
- *The decapsulating node MAY punt the IOAM data from the packet with the receive timestamp to the slow path for IOAM-Data-Fields processing. The receive timestamp is required by the various E2E IOAM use-cases, including streaming telemetry. Note that the packet is not necessarily punted to the control-plane.
- *The decapsulating node processes the IOAM-Data-Fields using the procedures defined in [\[RFC9197\]](#). An example of IOAM processing is to export the IOAM-Data-Fields, send IOAM-Data-Fields via streaming telemetry, etc.
- *The decapsulating node MUST remove the Post-Stack Network Action with IOAM-Data-Fields from the received packet. The decapsulated packet is forwarded downstream or terminated locally similar to the regular data packets.

5. Hop-By-Hop IOAM Network Action

5.1. MNA Sub-Stack for Hop-By-Hop IOAM

The IHS scope is set to "HBH, value 0x1" to indicate the scope of HBH IOAM-Data-Fields in the MNA Sub-Stack as shown in Figure 3. Note that there may be additional Network Action (NA) LSEs added in the MNA Sub-Stack and NASL is updated based on the number of LSEs after the LSE with P flag.

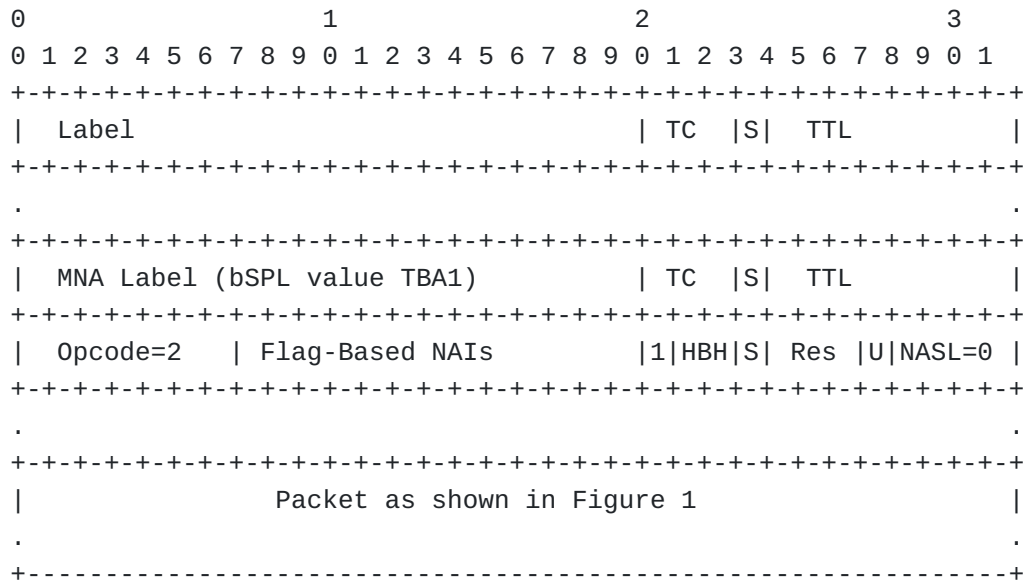


Figure 3: Example MNA Sub-Stack for HBH IOAM

The HBH IOAM-Data-Fields carry the Option-Type(s) that require processing at the intermediate and/or encapsulating and decapsulating nodes. The IOAM Option-Type carried can be IOAM Pre-allocated Trace Option-Type (value 0), IOAM Incremental Trace Option-Type (value 1) and IOAM Proof of Transit (POT) Option-Type (value 2), and Edge-To-Edge Option-Type (value 3) defined in [RFC9197] as well as Direct Export (DEX) Option-Type (value 4) defined in [RFC9326].

5.2. Procedure for Hop-By-Hop IOAM Network Action

The Hop-By-Hop IOAM Network Action procedure is summarized as following:

- *The encapsulating node inserts an MNA Sub-Stack containing MNA Label (bSPL value TBA1) with the Post-Stack Network Action Indicator (P flag) set to "1" and IHS scope set to "HBH, value 0x1" and one or more IOAM-Data-Fields in the Post-Stack Network Action in the MPLS packet.
- *The intermediate node enabled with HBH IOAM function processes the data packet including the IOAM-Data-Fields as defined in [RFC9197] and [RFC9326] when the node recognizes the HBH scope in the MNA Sub-Stack.
- *The intermediate node MAY punt the IOAM data from the packet with the receive timestamp to the slow path for IOAM-Data-Fields processing when the node recognizes the HBH scope. The receive timestamp is required by the various HBH OAM use-cases, including

streaming telemetry. Note that the packet is not necessarily punted to the control-plane.

*The intermediate node forwards the data packet downstream.

*The processing on the penultimate node is same as E2E case.

*The processing on the decapsulating node is same as E2E case.

5.3. Hop-By-Hop and Edge-To-Edge IOAM Network Actions

Both HBH and E2E IOAM NAs may be added in an MNA Sub-Stack. In this case, the HBH IOAM NAs MUST be added after the BOS and before the E2E IOAM NAs. When an intermediate node that processes the HBH IOAM NAs, encounters the first E2E IOAM NA, it will stop processing of the E2E NA.

6. Select IOAM Network Action

The procedure for carrying the IOAM NAs for Select Scope is the same as HBH Scope IOAM except the IHS scope is set to "Select, value 0x2". In this case, only the select nodes will process the IOAM-Data-Fields.

7. Considerations for IOAM

7.1. Considerations for ECMP

The encapsulating node needs to make sure the IOAM-Data-Fields do not start with a well-known IP Version Number (e.g. 0x4 for IPv4 and 0x6 for IPv6) as that can alter the hashing function for ECMP that uses the IP header. This is achieved by using the first nibble NNNN in the Top Header with a different IP Version Number (value TBA3 to be allocated by IANA or 0001b for G-ACH) after the MPLS label stack.

7.2. Node Capability

The decapsulating node that has to remove the IOAM-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.

The intermediate node that is not capable of supporting the IOAM functions defined in this document, can simply skip the IOAM processing.

The node that does not recognize the MNA Label received at the top of the label stack will drop the packet.

7.3. Nested MPLS Encapsulation

When a packet is received with MPLS Encapsulated IOAM Network Action, the nested MPLS encapsulating node that supports a different IOAM Network Action, the node MUST add a new MNA Label with Sub-Stack with the supported IOAM Post-Stack Network Action as part of the new MPLS encapsulation.

8. Security Considerations

The security considerations of IOAM in general are discussed in [RFC9197] and apply to the procedure defined in this document.

IOAM is considered a "per domain" feature, where one or several operators decide on configuring IOAM according to their needs. IOAM is intended for deployment in limited domains [RFC8799]. As such, it assumes that a node involved in IOAM operation has previously verified the integrity of the path. Still, operators need to properly secure the IOAM domain to avoid malicious configuration and use, which could include injecting malicious IOAM packets into the domain.

9. IANA Considerations

The PS-NA-OPCODE for IOAM is allocated from the "Post-Stack Network Action" registry.

Value	Name	Description	Reference
TBA2	IOAM	In Situ OAM	This document

Table 1: Post-Stack Network Action Opcode

10. References

10.1. Normative References

[I-D.ietf-mpls-mna-hdr] Rajamanickam, J., Ed., Gandhi, R., Ed., Zigler, R., Ed., Song, H., Ed., and K. Kompella, Ed., "MPLS Network Action Sub-Stack Solution", Work in Progress, Internet-Draft, draft-ietf-mpls-mna-hdr-01, March 2023, <<https://www.ietf.org/archive/id/draft-ietf-mpls-mna-hdr-01.txt>>.

[I-D.jags-mpls-ps-mna-hdr] Rajamanickam, J., Ed., Gandhi, R., Ed., Zigler, R., Li, T., and J. Dong, "Post-Stack MPLS Network Action (MNA) Solution", Work in Progress, Internet-Draft, draft-jags-mpls-ps-mna-hdr-00, March 2023, <<https://www.ietf.org/archive/id/draft-jags-mpls-ps-mna-hdr-00.txt>>.

[RFC2119]

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

[RFC8174]

Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

[RFC9197]

Brockners, F., Bhandari, S., and T. Mizrahi, "Data Fields for In Situ Operations, Administration, and Maintenance", RFC 9197, May 2022, <<https://www.rfc-editor.org/info/rfc9197>>.

[RFC9326]

Song, H., Gafni, B., Brockners, F., Bhandari, S., and T. Mizrahi, "In Situ Operations, Administration, and Maintenance OAM Direct Exporting", RFC 9326, November 2022, <<https://www.rfc-editor.org/info/rfc9326>>.

10.2. Informative References

[RFC8799]

Carpenter, B. and B. Liu, "Limited Domains and Internet Protocols", RFC 8799, DOI 10.17487/RFC8799, July 2020, <<https://www.rfc-editor.org/info/rfc8799>>.

Acknowledgements

The authors would like to thank Patrick Khordoc, Sagar Soni, Shwetha Bhandari, Clarence Filsfils, and Vengada Prasad Govindan for the discussions on IOAM. The authors would also like to thank Tarek Saad, Loa Andersson, Greg Mirsky, Stewart Bryant, Xiao Min, Jaganbabu Rajamanickam, and Cheng Li for providing many useful comments. The authors would also like to thank Mach Chen, Andrew Malis, Matthew Bocci, and Nick Delregno for the MPLS-RT reviews of the early version of the draft.

Contributors

The following people have substantially contributed to this document:

Zafar Ali
Cisco Systems, Inc.
Email: zali@cisco.com

Voitek Kozak
Comcast
Email: Voitek_Kozak@comcast.com

Authors' Addresses

Rakesh Gandhi (editor)
Cisco Systems, Inc.
Canada

Email: rgandhi@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

Bruno Decraene
Orange

Email: bruno.decraene@orange.com

Haoyu Song
Futurewei Technologies
United States of America

Email: haoyu.song@futurewei.com