MPLS Working Group
Internet-Draft
Updates: 5586 (if approved)
Intended status: Standards Track
Expires: 5 August 2022

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

1 February 2022

MPLS Data Plane Encapsulation for In-situ OAM Data
draft-gandhi-mpls-ioam-03

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 new Generic Associated Channel (G-ACh) and updates the RFC 5586.

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 5 August 2022.

Copyright Notice

Table of Contents

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 [I-D.ietf-ippm-ioam-data]. The IOAM data fields are further updated in [I-D.ietf-ippm-ioam-direct-export] for direct export use-cases and in [I-D.ietf-ippm-ioam-flags] for Loopback and Active flags.

This document defines how IOAM data fields are transported with MPLS data plane encapsulations using new Generic Associated Channel (G-ACh) and updates the [RFC5586].

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:

IOI IOAM Indicator

ECMP Equal Cost Multi-Path

E2E Edge-To-Edge

EL Entropy Label

ELI Entropy Label Indicator

ELC Entropy Label Control

G-ACh Generic Associated Channel

HBH Hop-By-Hop

HBI Hop-By-Hop Indicator

IOAM In-situ Operations, Administration, and Maintenance

MPLS Multiprotocol Label Switching

OAM Operations, Administration, and Maintenance

POT Proof-of-Transit

PW PseudoWire

3. MPLS Extensions for IOAM Data Fields

3.1. IOAM Generic Associated Channel

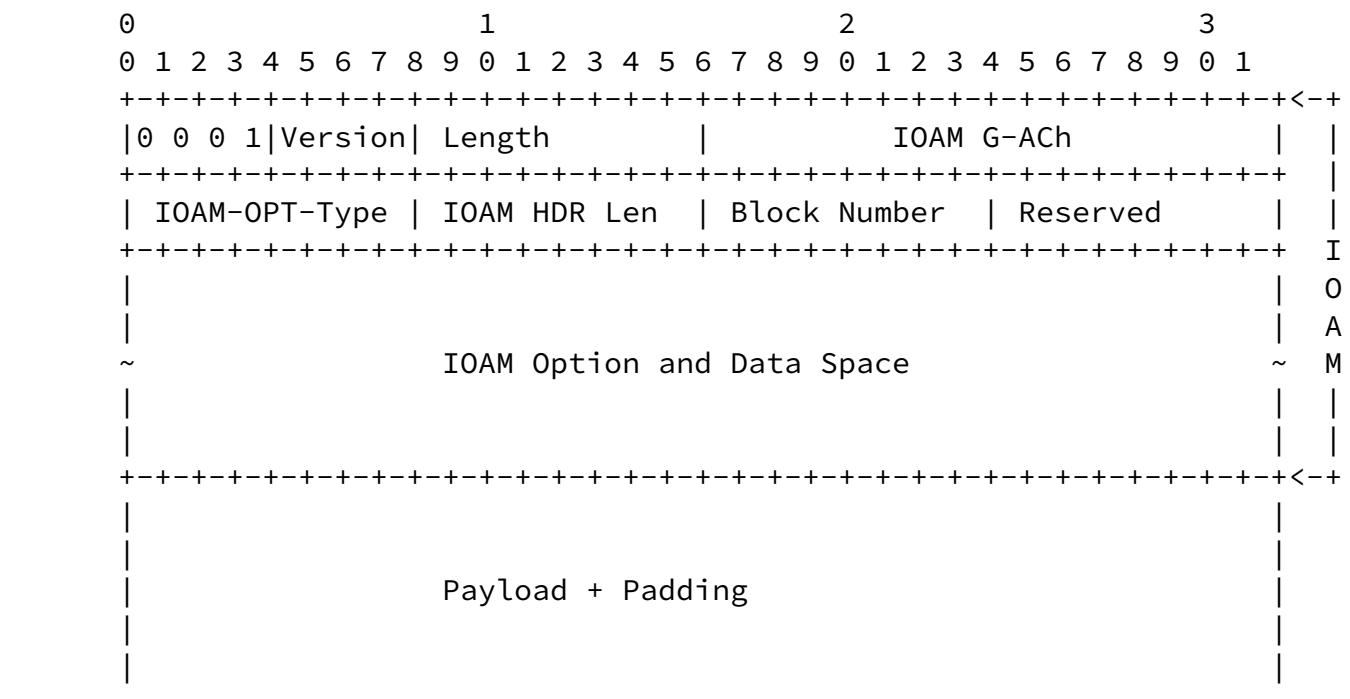
The IOAM header is added containing different IOAM-Data-Fields in the

MPLS header as shown in Figure 1. The IOAM-Data-Fields MUST follow the definitions corresponding to IOAM-Option-Types (e.g. see Section 5 of [I-D.ietf-ippm-ioam-data] and Section 3.2 of [I-D.ietf-ippm-ioam-direct-export]). More than one trace options can be present in the IOAM-Data-Fields.

G-ACh [RFC5586] provides a mechanism to transport OAM and other control messages over MPLS data plane. The IOAM G-ACh header [RFC5586] with new IOAM G-ACh type MUST be added immediately after the MPLS label stack in the MPLS header as shown in Figure 1, before the IOAM-Data-Fields. The G-ACh label (GAL) [RFC5586] MUST NOT be added in the MPLS label stack.

This document updates the following paragraph in Section 2.1 of [RFC5586]: "The G-ACh MUST NOT be used to transport user traffic" to "The G-ACh MAY be used with user traffic to transport OAM information".

Note that the G-ACh is not really used to transport the user traffic in this document but to transport the IOAM-Data-Fields with the user traffic.



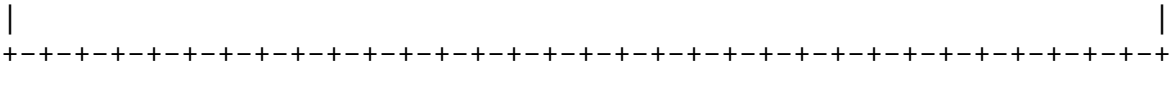


Figure 1: IOAM Generic Associated Channel with IOAM Data Fields

The IOAM-Data-Fields are encapsulated using the following fields in the MPLS header:

IP Version Number 0001b: The first four octets are IP Version Field part of a G-ACh header, as defined in [RFC5586].

Version: The Version field is set to 0, as defined in [RFC4385].

Length: Length of IOAM G-ACh data in 4-octet units. Note that this field is marked as Reserved in [RFC5586] and is updated for the new IOAM G-ACh type by this document.

IOAM G-ACh: Generic Associated Channel (G-ACh) Type (value TBA1) for IOAM [RFC5586].

Reserved: Reserved Bits MUST be set to zero upon transmission and ignored upon receipt.

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.

IOAM-OPT-Type: 8-bit field defining the IOAM Option type, as defined in the "IOAM Option-Type Registry" specified in [I-D.ietf-ippm-ioam-data].

IOAM HDR Length: 8-bit unsigned integer. Length of the IOAM Header 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 5 of [I-D.ietf-ippm-ioam-data] and Section 3.2 of

[I-D.ietf-ippm-ioam-direct-export].

3.2. IOAM Indicators

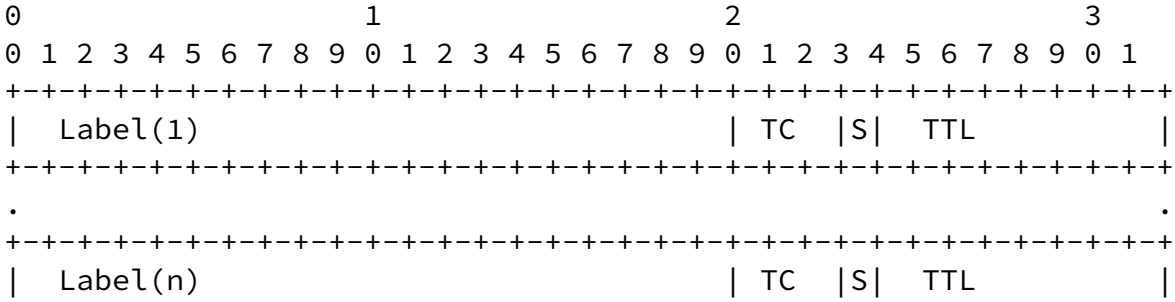
An IOAM Indicator MUST be used to indicate the presence of the IOAM-Data-Fields in the MPLS header. If both edge and intermediate nodes need to process IOAM data then both IOAM Indicator and HBH Indicator MUST be used. The HBH Indicator allows to optimize the IOAM processing on intermediate nodes and avoids the need to check all IOAM-Data-Fields.

[RFC6790] defines the MPLS Entropy Label (EL) and Entropy Label Indicator (ELI). [I-D.decraene-mpls-slid-encoded-entropy-label-id] defines Entropy Label Control (ELC) field and is carried in the TTL field of the Entropy Label. Additional flags are defined in [I-D.jags-mpls-ext-hdr-entropy-lbl]. A flag called IOI (IOAM Indicator) in the ELC is defined in this document to indicate the presence of IOAM. A flag called HBI (Hop-By-Hop Indicator) in the ELC is defined in [I-D.jags-mpls-ext-hdr-entropy-lbl] to indicate that HBH processing is required. The bit positions of these flags in the ELC field can be user-defined, consistently in the network. Alternatively, the bit positions of these flag can be allocated by IANA.

4. Edge-to-Edge IOAM

4.1. IOAM Indicator

The IOAM Indicator (Flag IOI in the Entropy Label Control field) is used to indicate the presence of the IOAM-Data-Fields in the MPLS header as shown in Figure 2.



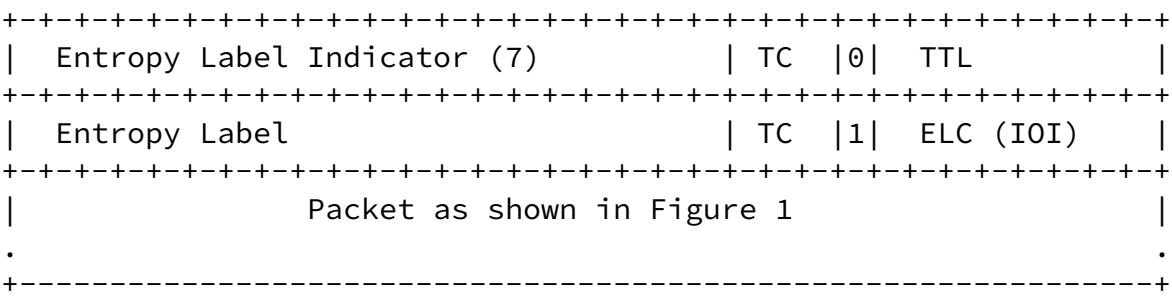


Figure 2: Example MPLS Encapsulation for IOAM

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 [I-D.ietf-ippm-ioam-data]. 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.

4.2. Procedure for Edge-to-Edge IOAM

The E2E IOM procedure is summarized as following:

* The encapsulating node inserts the ELI, EL pair with the IOAM Indicator (Flag IOI) below the label whose FEC is the end (decapsulating) node and one or more IOAM-Data-Fields in the MPLS header.

* The intermediate nodes do not process IOAM-Data-Fields.

* The decapsulating node MAY punt a copy of the packet with the receive timestamp to the slow path for IOAM-Data-Fields processing when the node recognizes the IOAM Indicator. The receive timestamp is required by the various E2E OAM 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 [I-D.ietf-ippm-ioam-data]. 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 IOAM-Data-Fields from the

received packet. The decapsulated packet is forwarded downstream or terminated locally similar to the regular IOAM-Data-Fields.

5. Hop-By-Hop IOAM

5.1. Hop-By-Hop Indicator

The IOAM Indicator (Flag IOI) along with Hop-By-Hop Indicator (Flag HBI) in the Entropy Label Control field) are used to indicate the presence of the HBH IOAM-Data-Fields in the MPLS header as shown in Figure 3.

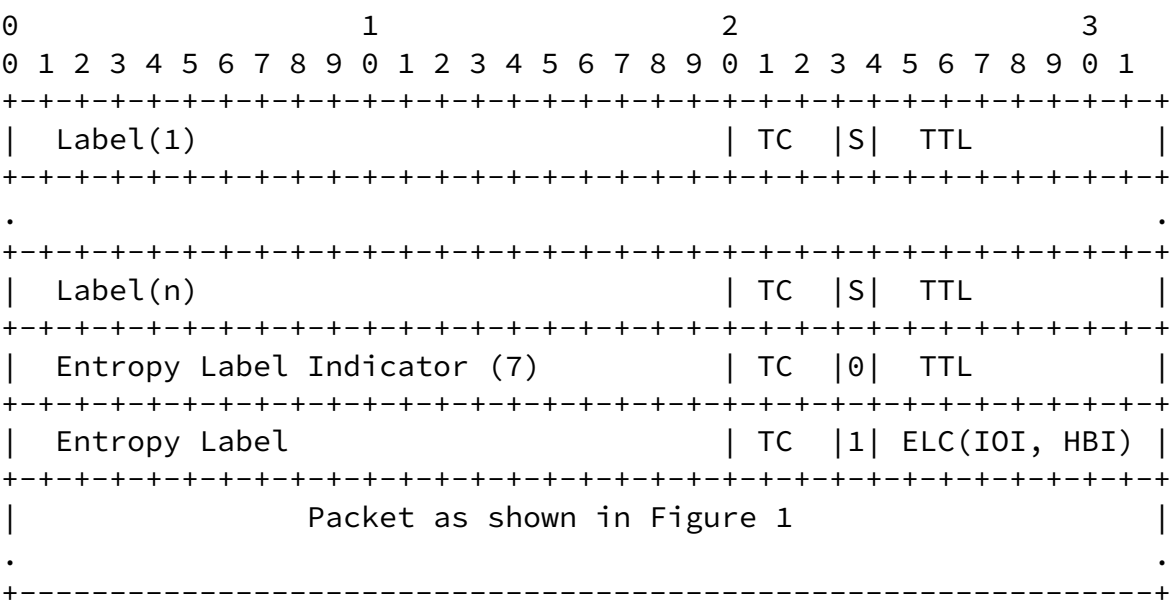


Figure 3: Example MPLS Encapsulation for HBH IOAM

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, IOAM Incremental Trace Option-Type and IOAM Proof of Transit (POT) Option-Type, as well as Edge-to-Edge Option-Type [I-D.ietf-ippm-ioam-data].

5.2. Procedure for Hop-By-Hop IOAM

The HBH IOAM procedure is summarized as following:

* The encapsulating node inserts the ELI, EL pair with the IOAM Indicator (Flag IOI) and HBH Indicator (Flag HBI) below the label whose FEC is the end (decapsulating) node and one or more IOAM-Data-Fields in the MPLS header.

* The intermediate node enabled with HBH IOAM function processes the data packet including the IOAM-Data-Fields as defined in [I-D.ietf-ippm-ioam-data] when the node recognizes the HBH Indicator in the MPLS header.

* The intermediate node MAY punt a copy of the packet with the receive timestamp to the slow path for IOAM-Data-Fields processing when the node recognizes the HBH Indicator. 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 a copy of the processed data packet downstream.

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

6. Alternate Method Using MPLS BOS Extension Header with IOAM Data Fields

In this method, BOS (Bottom of Stack) Extension Header defined in [I-D.jags-mpls-ext-hdr-entropy-lbl] is used with BOS Opcode for IOAM (value to be assigned by IANA) as shown in Figure 4 (instead of using the G-ACh channel for IOAM defined in this document).

Flag NH (Next Header Present for BOS Extension Header) and flag H (Hop-By-Hop processing required for this BOS Extension Header) as defined in [I-D.jags-mpls-ext-hdr-entropy-lbl] are also applicable to IOAM BOS Extension Header.

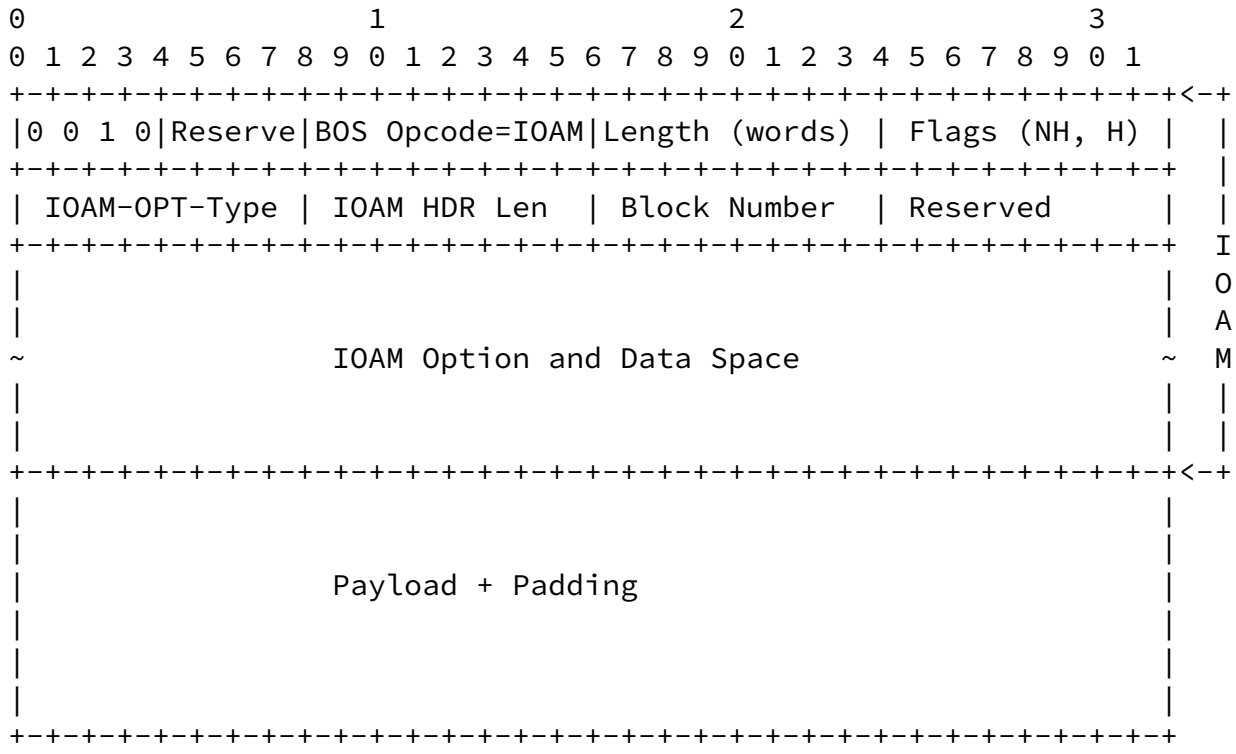
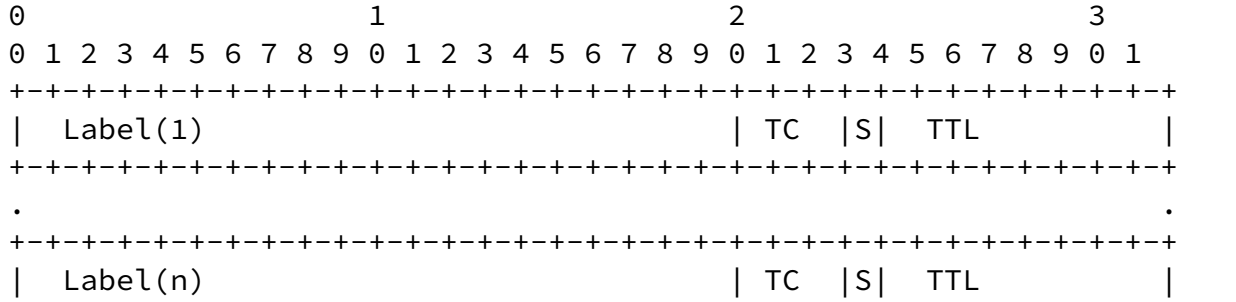


Figure 4: IOAM BOS Extended Header with IOAM Data Fields

Bottom of Stack Presence Indicator (BPI) flag in ELC as defined in [I-D.jags-mpls-ext-hdr-entropy-lbl] is set to "1" to indicate the presence of the BOS Extension Header (instead of using the IOI flag in the ELC defined in this document) as shown in Figure 5. There is no change in the usage of the HBI flag defined in [I-D.jags-mpls-ext-hdr-entropy-lbl] and further described in this document.



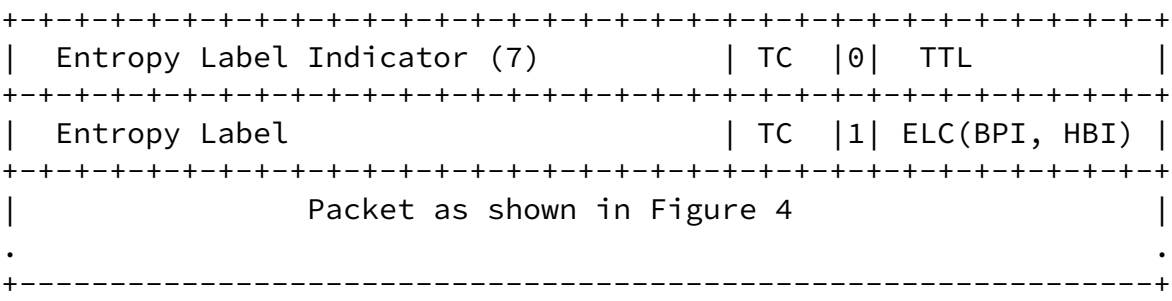


Figure 5: Example MPLS Encapsulation for IOAM using BOS Header

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 IOAM G-ACh with IP Version Number 0001b after the MPLS label stack [RFC5586].

When entropy label [RFC6790] is used for hashing function for ECMP, the procedure defined in this document does not alter the ECMP behaviour.

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.

7.3. Nested MPLS Encapsulation

The packets with IOAM-Data-Fields may carry one or more Entropy Labels with an IOAM Indicator in the MPLS header. An intermediate

node that supports IOAM SHOULD check the first Entropy Label in the label stack for the IOAM Indicator and the HBH Indicator to process the IOAM-Data-Fields.

An intermediate node that supports IOAM, SHOULD copy the ELC field from the received Entropy Label to the new Entropy Label when inserting the new Entropy Label in the MPLS header and this can be based on a local policy.

When a packet is received with IOAM, the nested MPLS encapsulating node that supports a different IOAM, the node MUST add a new Entropy Label with the supported IOAM as part of the new MPLS encapsulation.

8. Security Considerations

The security considerations of IOAM in general are discussed in [I-D.ietf-ippm-ioam-data] 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.

Routers that support G-ACh are subject to the same security considerations as defined in [RFC4385] and [RFC5586].

9. IANA Considerations

IANA maintains G-ACh Type Registry (see https://www.iana.org/assignments/g-ach-parameters/g-ach-parameters.xhtml). IANA is requested to allocate a value for IOAM G-ACh Type from "MPLS Generalized Associated Channel (G-ACh) Types (including Pseudowire Associated Channel Types)" registry.

Value	Description	Reference
TBA1	IOAM G-ACh Type	This document

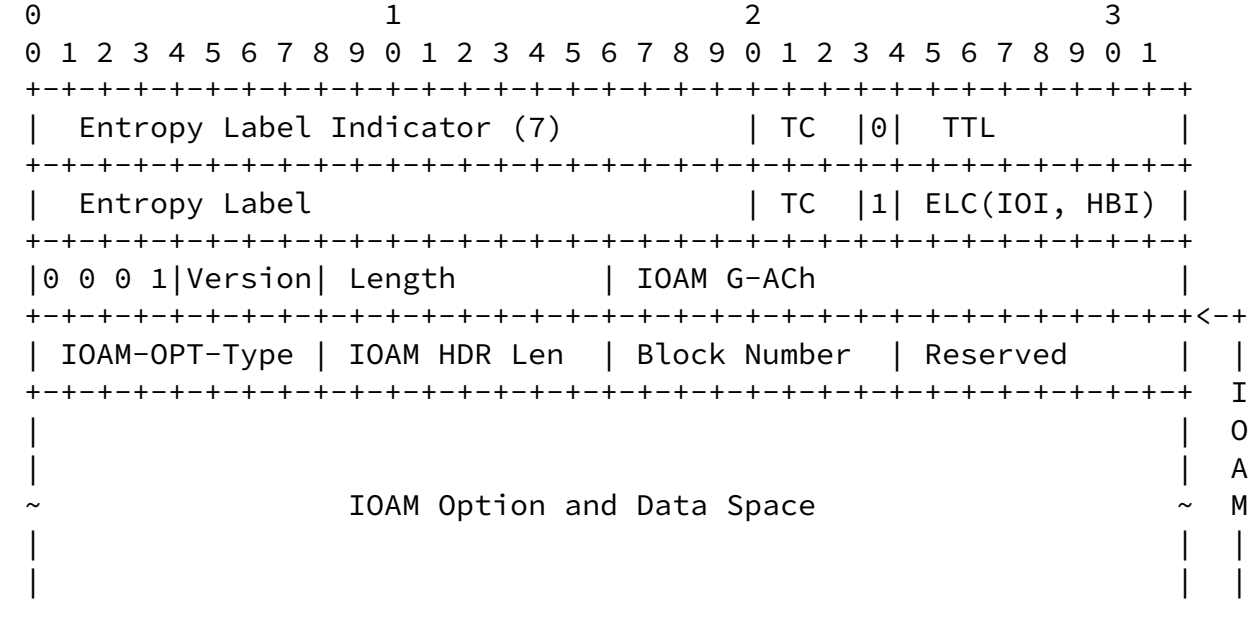
Table 1: IOAM G-ACh Type

10. Appendix

10.1. MPLS Encapsulation with Control Word and Another G-ACh for IOAM Data Fields

The IOAM-Data-Fields, including IOAM G-ACh header are added in the MPLS encapsulation immediately after the MPLS header. Any Control Word [RFC4385] or another G-ACh [RFC5586] MUST be added after the IOAM-Data-Fields in the packet as shown in the Figure 6 and Figure 7, respectively. This allows the intermediate nodes to easily access the HBH IOAM-Data-Fields located immediately after the MPLS header. The decapsulating node can remove the MPLS encapsulation including the IOAM-Data-Fields and then process the Control Word or another G-ACh following it. The subsequent G-ACh and Control Word are

located through the use of the "Length" field in the IOAM G-ACh.



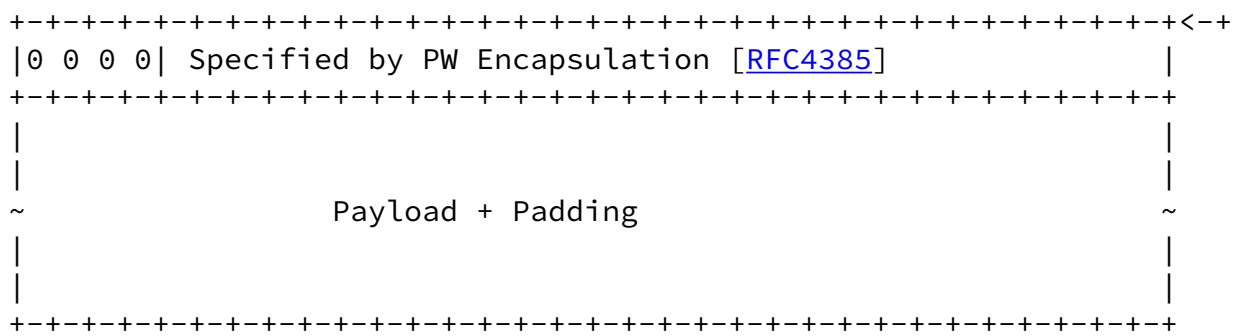
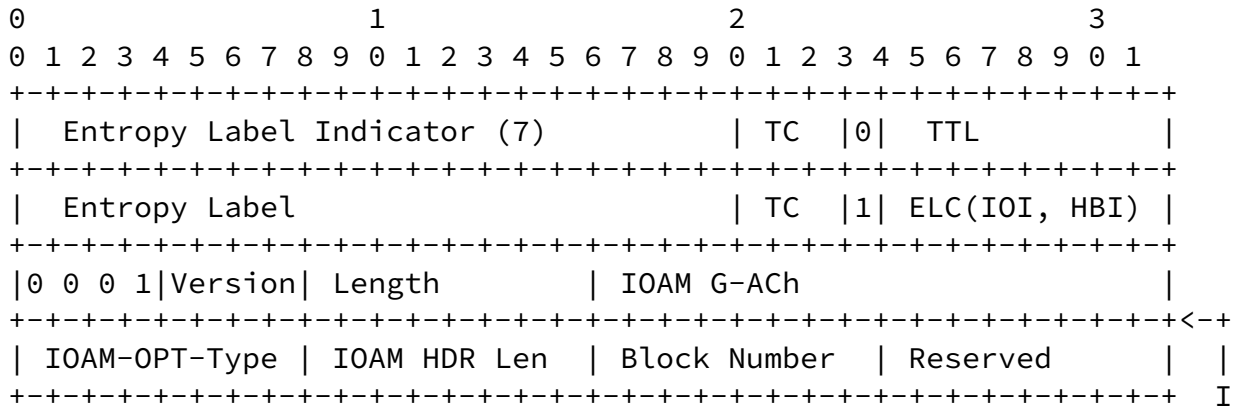


Figure 6: Example MPLS Encapsulation with Generic PW Control Word with HBH IOAM



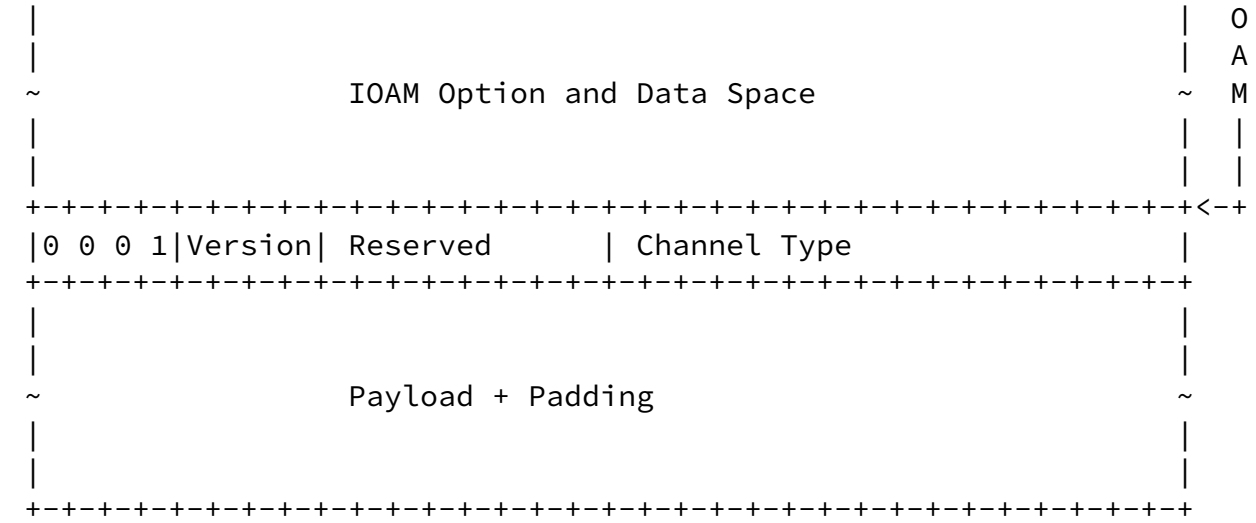


Figure 7: Example MPLS Encapsulation with Another G-ACh with HBH IOAM

11. References

11.1. Normative References

[I-D.decraene-mpls-slid-encoded-entropy-label-id]
Decraene, B., Filsfils, C., Henderickx, W., Saad, T., Beeram, V. P., and L. Jalil, "Using Entropy Label for Network Slice Identification in MPLS networks.", Work in Progress, Internet-Draft, draft-decraene-mpls-slid-encoded-entropy-label-id-02, 6 August 2021, <https://www.ietf.org/archive/id/draft-decraene-mpls-slid-encoded-entropy-label-id-02.txt>.

[I-D.ietf-ippm-ioam-data]
Brockners, F., Bhandari, S., and T. Mizrahi, "Data Fields for In-situ OAM", Work in Progress, Internet-Draft, draft-ietf-ippm-ioam-data-17, 13 December 2021, <https://www.ietf.org/archive/id/draft-ietf-ippm-ioam-data-17.txt>.

[I-D.ietf-ippm-ioam-direct-export]
Song, H., Gafni, B., Zhou, T., Li, Z., Brockners, F., Bhandari, S., Sivakolundu, R., and T. Mizrahi, "In-situ OAM Direct Exporting", Work in Progress, Internet-Draft,

draft-ietf-ippm-ioam-direct-export-07, 13 October 2021, <https://www.ietf.org/archive/id/draft-ietf-ippm-ioam-direct-export-07.txt>.

[I-D.ietf-ippm-ioam-flags]
Mizrahi, T., Brockners, F., Bhandari, S., Sivakolundu, R., Pignataro, C., Kfir, A., Gafni, B., Spiegel, M., and J. Lemon, "In-situ OAM Loopback and Active Flags", Work in Progress, Internet-Draft, draft-ietf-ippm-ioam-flags-07, 13 October 2021, <https://www.ietf.org/archive/id/draft-ietf-ippm-ioam-flags-07.txt>.

[I-D.jags-mpls-ext-hdr-entropy-lbl]
Rajamanickam, J., Gandhi, R., Bhattacharya, J., Decraene, B., and R. Zigler, "MPLS Extension Header Encodings Using Entropy Label", Work in Progress, Internet-Draft, draft-jags-mpls-ext-hdr-entropy-lbl-00, 26 January 2022, <https://www.ietf.org/archive/id/draft-jags-mpls-ext-hdr-entropy-lbl-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>.

[RFC5586] Bocci, M., Ed., Vigoureux, M., Ed., and S. Bryant, Ed., "MPLS Generic Associated Channel", RFC 5586, DOI 10.17487/RFC5586, June 2009, <https://www.rfc-editor.org/info/rfc5586>.

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, DOI 10.17487/RFC6790, November 2012, <https://www.rfc-editor.org/info/rfc6790>.

[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>.

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, DOI 10.17487/RFC4385, February 2006, <https://www.rfc-editor.org/info/rfc4385>.

[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, 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.

Authors' Addresses

Rakesh Gandhi (editor)
Cisco Systems, Inc.
Canada

Email: rgandhi@cisco.com

Zafar Ali
Cisco Systems, Inc.

Email: zali@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

Voitek Kozak
Comcast

Email: Voitek_Kozak@comcast.com