IPPM Working Group
Internet-Draft
Intended status: Standards Track
Expires: 19 April 2026

X. Min
ZTE Corp.
Y. Liu
China Mobile
C. Lin
New H3C Technologies

16 October 2025

Extensions to IOAM Trace Option for Carrying Fixed-Size Data
draft-xiao-ippm-ioam-trace-extensions-02

Abstract

In situ Operations, Administration, and Maintenance (IOAM) Trace-Option data defined in RFC 9197 is a variable-length data, the length of this kind of data varies with the number of transited IOAM-capable nodes and the selection of data fields processed by each IOAM-capable node. This document extends the IOAM Trace Option to carry a fixed-size data, the length of this kind of data is fixed once the selection of data fields processed by each IOAM-capable node is determined, and doesn’t vary with the number of transited IOAM-capable nodes.

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 19 April 2026.

Copyright Notice

Table of Contents

1. Introduction

[RFC9197] defines In situ Operations, Administration, and Maintenance (IOAM), which collects operational and telemetry information in the packet while the packet traverses a path between two points in the network. As specified in Section 4.1 of [RFC9197], IOAM tracing is defined as two separate options: Pre-allocated Trace Option and Incremental Trace Option. The two IOAM Trace Options share the same format of IOAM Trace-Option header and data. The IOAM Trace-Option data is composed of a set of node data lists; among them each node data list is populated by a node along the forwarding path of IOAM packet.

The IOAM Trace-Option data defined in [RFC9197] is a variable-length data, the length of this kind of data varies with the number of transited IOAM-capable nodes and the selection of data fields processed by each IOAM-capable node. This document extends the IOAM Trace Option to carry a fixed-size data, the length of this kind of data is fixed once the selection of data fields processed by each IOAM-capable node is determined, and doesn't vary with the number of transited IOAM-capable nodes.

Note that the difference between the fixed-size IOAM tracing data defined in this document and the pre-allocated IOAM tracing data defined in [RFC9197], is that the fixed-size IOAM tracing data is processed in a compare-and-replace manner by each node along the forwarding path, which makes the IOAM data size always fixed no matter how many nodes get traversed.

2. Conventions Used in This Document

2.1. Abbreviations

ABW: Available Bandwidth

BU: Buffer Utilization

CSIG: Congestion Signaling

HPCC++: Enhanced High Precision Congestion Control

IOAM: In situ Operations, Administration, and Maintenance

LU: Link Utilization

PD: Per-hop delay

TTL: Time to Live

2.2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Problem with the current IOAM Trace Option

As specified in Section 4.4 of [RFC9197], the IOAM Trace Option (including IOAM Pre-allocated Trace Option and IOAM Incremental Trace Option) has a format as shown in Figure 1.

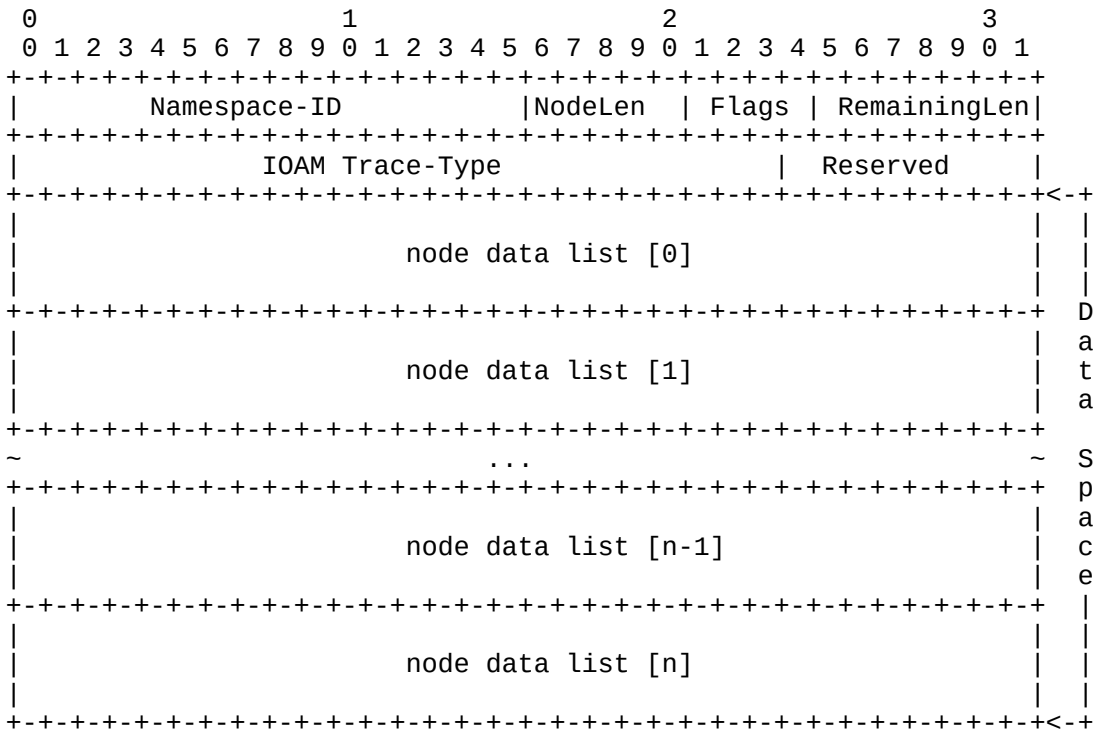

Figure 1: IOAM Trace Option defined in RFC 9197

The IOAM Trace Option defined in [RFC9197] is used to stack up multiple metadata from multiple IOAM transit nodes. Specifically, along the forwarding path, each IOAM-capable node adds its own operational and telemetry information into the IOAM Trace Option carried by a data packet in a stacking manner.

As described in [I-D.miao-ccwg-hpcc-info], the IOAM Trace Option Data can be used as a congestion control signal for Enhanced High Precision Congestion Control (HPCC++) congestion control mechanism, and at the same time, Congestion Signaling (CSIG) Data can also be used as a congestion control signal for HPCC++. As specified in [I-D.ravi-ippm-csig], a CSIG tag carries a fixed-size aggregate metric computed over the hop devices. Specifically, along the

forwarding path, each CSIG-capable node optionally inputs its own congestion information into the CSIG tag carried by a data packet in a compare-and-replace manner.

In different application scenarios, either the IOAM Trace Option defined in [RFC9197] or the CSIG tag defined in [I-D.ravi-ippm-csig] is applicable, or even both of them can be used concurrently. In other words, they're complementary to each other. Then the question goes to that whether it's possible to integrate them together. That is also the problem this document intends to address.

4. Extension to the IOAM Trace Option's Flags

In order to integrate the variable-size superimposed data and fixed-size aggregate data into IOAM Trace Option, this document defines the reserved Bit 3 of the Flags field of the IOAM Trace Option as shown in Figure 2.

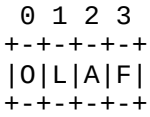

Figure 2: IOAM Trace Option's Flags format

Bit 0 "Overflow" (O-bit): When set, the Overflow flag indicates that there are not enough octets left to record the node data, as defined in Section 4.4.1 of [RFC9197].

Bit 1 "Loopback" (L-bit): When set, the Loopback flag triggers the sending of a copy of a packet back towards the source, as defined in Section 3 of [RFC9322].

Bit 2 "Active" (A-bit): When set, the Active flag indicates that a packet is an active measurement packet rather than a data packet, as defined in Section 3 of [RFC9322].

Bit 3 "Fixed-size" (F-bit): When set, the Fixed-size flag indicates that a fixed-size aggregate data rather than a variable-size superimposed data is carried, as defined in Section 5 of this document. As defined in Section 4.4 of [RFC9197], there are two types of IOAM Trace Options, Pre-allocated Trace-Option and Incremental Trace-Option, the F-bit defined in this document can only be used for Pre-allocated Trace-Option. In other words, if the F-bit is set while the IOAM Trace Option-Type indicates it's an Incremental Trace-Option, then the F-bit MUST be ignored by the receiver.

5. Fixed-size aggregate data in the IOAM Trace Option

When the IOAM Trace Option is used to carry fixed-size aggregate data, the format of the IOAM Trace Option is shown in Figure 3.

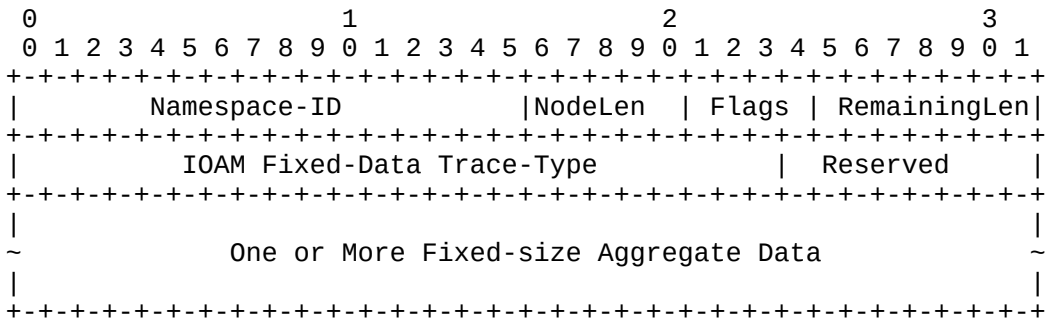

Figure 3: IOAM Trace Option carrying fixed-size aggregate data

Namespace-ID: The same as defined in [RFC9197].

NodeLen: The same as defined in [RFC9197]. MUST be set to 0.

Flags: The same as defined in [RFC9197] and [RFC9322]. This document allocates a single flag as follows:

Bit 3 "Fixed-size" (F-bit) (least significant bit). As defined in Section 4 of this document. When set, a fixed-size data space as shown in Figure 3 is carried in the IOAM Trace Option.

RemainingLen: The same as defined in [RFC9197]. MUST be set to 0.

IOAM Fixed-Data Trace-Type: The similar as the IOAM Trace-Type defined in Section 4.4.1 of [RFC9197]. The only difference with the IOAM Trace-Type is that each bit of IOAM Fixed-Data Trace-Type indicates the presence of a fixed-size 8-octet data space called Fixed-size Aggregate Data, whose length is independent of the number of transited IOAM-capable nodes.

One or More Fixed-size Aggregate Data: Each Fixed-size Aggregate Data is a 8-octet data space. When only one bit within the bitmap of IOAM-Fixed-Data Trace-Type is set, the length of this field is eight octets. When more than one bit within the bitmap of IOAM Fixed-Data Trace-Type is set, the length of this field is equal to the number of set bits multiplied by eight octets.

This document defines the format of Fixed-size Aggregate Data, as shown in Figure 4.

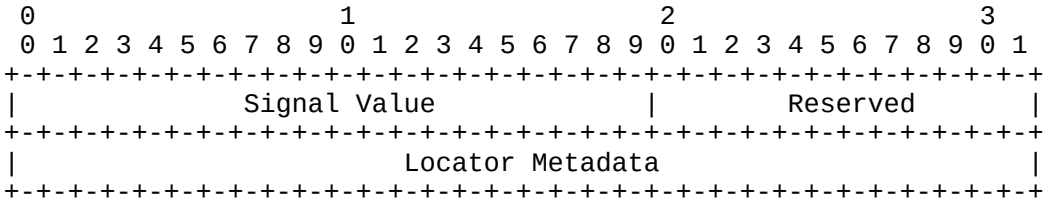

Figure 4: Fixed-size Aggregate Data Format

The fields are defined as follows:

* Signal Value: A 20-bit field. Its meaning depends on the position of set bit in IOAM Fixed-Data Trace-Type.

* Locator Metadata: A 4-octet field. This field indicates the relevant metadata about the bottleneck device and port, specifically, it contains 12-bit Port ID, 12-bit Device ID, and 8-bit TTL (also called Hop Limit for IPv6).

* Reserved: A 12-bit field that MUST be zeroed on transmission and ignored on receipt.

The 24-bit identifier of IOAM Fixed-Data Trace-Type and corresponding Signal Values are defined as follows:

Bit 0 Most significant bit. When set, indicates the Signal Value in the data space is Minimum Available Bandwidth - min(ABW). This field indicates the minimum absolute available bandwidth (in Mbps) across all the nodes' egress interfaces from where the packet is forwarded out. Each node along the forwarding path would compare its available bandwidth of the egress interface with the received field value. If a node finds that its available bandwidth of the egress interface is smaller, then the node would populate this field with a value of its available bandwidth; If a node finds that its available bandwidth of the egress interface is not smaller, then the node would not change this field.

Bit 1 When set, indicates the Signal Value in the data space is Maximum Link Utilization - max(LU). This field value devided by 0xFFFFF equals the link utilization rate of the egress interface from where the packet is forwarded out. Each node

along the forwarding path would compare its link utilization rate of the egress interface, i.e. utilized bandwidth devided by the whole capacity, with the calculated link utilization rate by the received field value. If a node finds that its link utilization rate of the egress interface is bigger, then the node would populate this field with a value of 0xFFFFF multiplied by its link utilization rate; If a node finds that its link utilization rate of the egress interface is not bigger, then the node would not change this field.

Bit 2 When set, indicates the Signal Value in the data space is Maximum Buffer Utilization - max(BU). This field value devided by 0xFFFFF equals the buffer utilization rate of the egress interface from where the packet is forwarded out. Each node along the forwarding path would compare its buffer utilization rate of the egress interface, i.e. occupied buffer devided by the whole buffer, with the calculated buffer utilization rate by the received field value. If a node finds that its buffer utilization rate of the egress interface is bigger, then the node would populate this field with a value of 0xFFFFF multiplied by its buffer utilization rate; If a node finds that its buffer utilization rate of the egress interface is not bigger, then the node would not change this field.

Bit 3-22 Undefined. These values are available for future assignment in the IOAM Fixed-Data Trace-Type Registry. An IOAM encapsulating node MUST set the value of each undefined bit to 0. If an IOAM transit node receives a packet with one or more of these bits set to 1, it MUST ignore these bits and SHOULD log and/or report an error.

Bit 23 Reserved; MUST be set to zero upon transmission and be ignored upon receipt. This bit is reserved to allow for future extensions of the IOAM Fixed-Data Trace-Type bit field.

6. Security Considerations

As discussed in [RFC9322], IOAM is assumed to be deployed in a restricted administrative domain, thus limiting the scope of the threats above and their effect. However, even given this limited scope, security threats should still be considered and mitigated.

Security issues discussed in [RFC9197] apply to this document.

7. IANA Considerations

7.1. IOAM Trace-Flags Registry

IANA is requested to allocate the following bit in the "IOAM Trace-Flags" registry as follows:

Bit 3 "Fixed-size" (F-bit)

This document is specified as the "Reference" in the registry for this bit.

Note that bit 0 is the most significant bit in the "IOAM Trace-Flags" registry. This bit was allocated by [RFC9197] as the ’Overflow’ bit.

7.2. IOAM Fixed-Data Trace-Type Registry

IANA is requested to create a new registry "IOAM Fixed-Data Trace-Type" within the defined registry group "In Situ OAM (IOAM)". This registry defines code points for each bit in the 24-bit IOAM Fixed-Data Trace-Type field for the Pre-allocated Trace Option-Type When the F-bit is set. Bits 0-2 are defined in Section 5 of this document:

Bit 0 Minimum Available Bandwidth - min(ABW) and Locator Metadata

Bit 1 Maximum Link Utilization - max(LU) and Locator Metadata

Bit 2 Maximum Buffer Utilization - max(BU) and Locator Metadata

Bit 23 reserved

Bits 3-22 are available for assignment via the "IETF Review" process, as per [RFC8126].

New registration requests MUST use the following template:

Bit: desired bit to be allocated in the 24-bit IOAM Fixed-Data Trace-Type field for the Pre-allocated Trace Option-Type when the F-bit is set

Description: brief description of the newly registered bit

Reference: reference to the document that defines the new bit

8. Acknowledgements

The authors would like to acknowledge Wei Duan and Jun Feng for their very helpful comments.

9. References

9.1. 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, <https://www.rfc-editor.org/info/rfc2119>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[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., Ed., Bhandari, S., Ed., and T. Mizrahi, Ed., "Data Fields for In Situ Operations, Administration, and Maintenance (IOAM)", RFC 9197, DOI 10.17487/RFC9197, May 2022, <https://www.rfc-editor.org/info/rfc9197>.

[RFC9322] Mizrahi, T., Brockners, F., Bhandari, S., Gafni, B., and M. Spiegel, "In Situ Operations, Administration, and Maintenance (IOAM) Loopback and Active Flags", RFC 9322, DOI 10.17487/RFC9322, November 2022, <https://www.rfc-editor.org/info/rfc9322>.

9.2. Informative References

[I-D.miao-ccwg-hpcc-info] Miao, R., Anubolu, S., Pan, R., Lee, J., Gafni, B., Tantsura, J., Alemania, A., and Y. Shpigelman, "Inband Telemetry for HPCC++", Work in Progress, Internet-Draft, draft-miao-ccwg-hpcc-info-04, 6 January 2025, <https://datatracker.ietf.org/doc/html/draft-miao-ccwg-hpcc-info-04>.

[I-D.ravi-ippm-csig] Ravi, A., Dukkipati, N., Mehta, N., and J. Kumar, "Congestion Signaling (CSIG)", Work in Progress, Internet-Draft, draft-ravi-ippm-csig-01, 2 February 2024, <https://datatracker.ietf.org/doc/html/draft-ravi-ippm-csig-01>.

Authors' Addresses

Xiao Min
ZTE Corp.
Nanjing
China
Phone: +86 18061680168
Email: xiao.min2@zte.com.cn

Yisong Liu
China Mobile
Beijing
China
Email: liuyisong@chinamobile.com

Changwang Lin
New H3C Technologies
Beijing
China
Email: linchangwang.04414@h3c.com