Internet Engineering Task Force
Internet-Draft
Intended status: Experimental
Expires: 16 May 2026

J. Parikh
Arista Networks
C. Pignataro
Blue Fern Consulting
A. Clemm
Sympotech
12 November 2025

Variable Length Node Data Field Option for In-situ Operations, Administration, and Maintenance (IOAM)

draft-parikh-ioam-variable-length-field-00

Abstract

The purpose of this memo is to describe a new IOAM Node Data Field type, called Flex Field, for In-Situ Operations, Administration, and Maintenance (IOAM). This option type, under IOAM Trace Option-Types will allow one to append variable length node data in an IOAM packet, along a network path.

Status of This Memo

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. [RFC9197] defines the data fields and associated data types for IOAM. Currently, every node data field (but one) defined in there is of fixed length, i.e. 4-octet or 8-octet, the only variable length field in that document is for the Opaque State Snapshot (Section 4.4.2.13. of [RFC9197]).

What if a network operator wants to add important variable length node data, from the nodes along a network path, in an IOAM packet? This memo proposes a new option type for In-Situ Operations, Administration, and Maintenance (IOAM) [RFC9197], to address that.

1.1. 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.

2. Background and Motivation

As we had discussed in the previous section, the [RFC9197] talks about IOAM data fields and different IOAM nodes types.

The term "in situ", in IOAM, refers to the fact that the OAM data is added to the data packets rather than being sent within packets specifically dedicated to OAM. IOAM (which uses the data plane) is used to complement mechanisms, such as Ping or Traceroute (which use the control plane).

Now, many applications, interested in telemetry data across a path are not only interested in the entire path’s comprehensive telemetry, (to paint a more holistic picture), but also in each individual node’s telemetry.
For example:

Individual packet latency of each node, across the path.
(This might look same like traceroute, but unlke traceroute, we could, with this, be getting the data-plane latency)

Per-interface performance metrics:
窶｢ Microbursts
窶｢ Packet Drops
窶｢ Traffic Rates
窶｢ Congestion
窶｢ Timestamps
窶｢ Path consistencies, etc.

Sustainable Networking Applications:
窶｢ Carbon-Intensity of each node along the path (And using that as an input to applications that attempt to minimize pollution attributable to specific networking traffic)
窶｢ Power Utilization of each node along the path, etc.

And so, as discussed above, using IOAM, which traverses in the data plane, we would get an accurate idea of how the above mentioned telemetry metrics would look, for actual data plane packets/traffic.

Now, due to the nature of this field, and how it is designed, it will have an upper limit when it comes to "how many nodes (data) can this field accomodate" and we have decided to keep it at 255 (with variable length). This should idealy be able to address an administrative domain’s needs. And it is and will be both, backwards compatible with existing IOAM data field definitions and forward compatible for if any new data fields are introduced in the future.

3. Transport Options for the Flex Field

So, as discussed above, we will be using XX (undefined) [---preferrably the 21st bit, so it’s next to the opaque state snapshot, for easy processing---] bit for, the IOAM-Trace-Type field present in the "Pre-allocated and Incremental Trace-Option headers" (Section 4.4.1. of [RFC9197]), to represent the new Flex Field defined in this document.

And, due to the nature of this field, when present, the Node Length (Section 4.4.1. of [RFC9197]) present in the IOAM Header, will exclude the length of this field. With this, logic processor will know that the actual bits 0-11 are done and can start on the XX bit (i.e. this document) with it’s own length.

4. IOAM Flex Field Option Type

This section defines a new node data field for IOAM Flex Field Option Type. And like other IOAM Option Types [RFC9197], this field can be transported by a variety of transport protocols, including NSH, Segment Routing, Geneve, BIER, IPv6, etc. [RFC9378]

Flex Field Node Data Field Option Type

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Length                    |   Hop Count   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Namespace-ID           | Flags |       Reserved        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|                                                               |
|                            Data                               |
~                                                               ~
.                                                               .
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: Flex Field Option Type

Length:
A 24-bit field, which will represent the length of the Flex Field. The length field MUST be present and populated with the "number of 4 octect words" in the entire Flex Field, including the header octet word (this will allow us to have significant amount of data in this field, and also have a uniformity in the length field processing across the packet, since the Opaque State Snapshot also will have a similar length calculation). This field will also let

the processing node know till what size does the Flex Field goes, so that it won’t accidentally consider the Opaque State Snapshot Data Field as a part of the Flex Field.

Hop Count:

A 8-bit fields to represent the hop count to record the number of nodes along the path that successfully processed the Flex Field. The encapsulating node MUST set the value to 1, and each subsequent node (transit nodes, as well as the decapsulating node prior to performing decapsulation) MUST increment its value by 1. If the Hop Count at a node exceeds 255, that node MUST set the Hop Count to 0 and set Flag 1 ("Node Hop Count Exceeded") to prevent further processing of the Flex Field.

Namespace-ID:

As discussed in [RFC9197], this is a 16-bit identifier of an IOAM-Namespace. The Namespace-ID value of 0x0000 is defined as the "Default-Namespace-ID" (Section 4.3. of [RFC9197]) and MUST be known to all the nodes implementing IOAM.
The Namespace-ID is populated by the encapsulating node and MUST NOT be changed by any of the intermediate nodes. For any other Namespace-ID value that does not match any Namespace-ID the node is configured to operate on, the node MUST NOT change the contents of the IOAM-Data-Fields except for the Namespace Flag (see below)

Flags:

Flags are 4-bit in length, which are used to indicate errors, which are encountered during the process of populating the data in the Flex Field. Once a flag is set, no further aggregarion occurs along the path. The encapsulating node MUST set the value of Flags to 0 upon transmission. When an intermediate node encounters receives a packet in which any of operations on that packet; instead, the IOAM data MUST be the Flags are non-zero, the node MUST NOT perform further IOAM forwarded as-is unchanged.
Here’s how the flags are defined:
窶「 Flag 0: 0x0000: Default Flag value.
窶「 Flag 1: 0x0001: Node Hop Count Exceeded.
窶「 Flag 2: 0x0010: Length Error.
窶「 Flag 16: 0x1111: Other error.

Reserved:

A 12-bit field, reserved for future usage. The IOAM encapsulating node MUST set the value to zero upon transmission. And IOAM transit nodes MUST ignore the received value (unless this field us used in a future memo)

Data:
This field is the variable length field which will contain the actual data that the operator would like to append.
Now, what gets into this space is dependent on the Namespace specific definition that the operator chooses to have inside their domain. As discussed in Section 4.3. of [RFC9197], the significance of IOAM-Data-Fields is always within a particular IOAM-Namespace and so, this gives the operators flexibility to append different types of variable length data for different Namespaces.
Example: IOAM-Namespaces can be used to identify different sets of devices (e.g., different types of devices) in a deployment; if an operator wants to insert different IOAM-Data-Fields based on the device, the devices could be grouped into multiple IOAM-Namespaces.

5. Sample Use Case

TBD

6. Secutiry Considerations

As discussed in [RFC9378], IOAM is focused on "limited domains", as defined in [RFC8799]. IOAM is not targeted for a deployment on the global Internet, which would incur additional considerations such as the crossing of Trust Boundaries, authentication of IOAM data, or the desire to obfuscate domain internals to outside parties. The part of the network that employs IOAM is referred to as the "IOAM-Domain".

7. IANA Considerations

TBD

8. References

8.1. 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>.

[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>.

[RFC9378] Brockners, F., Ed., Bhandari, S., Ed., Bernier, D., and T. Mizrahi, Ed., "In Situ Operations, Administration, and Maintenance (IOAM) Deployment", RFC 9378, DOI 10.17487/RFC9378, April 2023, <https://www.rfc-editor.org/info/rfc9378>.

8.2. Informative References

Acknowledgements

TBD

Authors' Addresses

Jainam Parikh
Arista Networks
5453 Great America Parkway
Santa Clara, 95054
United States of America
Email: jainam@parikhgroup.net

Carlos Pignataro
Blue Fern Consulting
United States of America
Email: carlos@bluefern.consulting

Alexander Clemm
Sympotech
United States of America
Email: ludwig@clemm.org