SFC
Internet-Draft
Intended status: Standards Track
Expires: November 19, 2022

F. Brockners, Ed.
Cisco
S. Bhandari, Ed.
Thoughtspot

May 18, 2022

Network Service Header (NSH) Encapsulation for In-situ OAM (IOAM) Data

draft-ietf-sfc-ioam-nsh-10

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 outlines how IOAM data fields are encapsulated with the Network Service Header (NSH).

Commenté [BMI1]: I wonder whether we really need to have this mention here. I'm thinking about services where packet replication may be involved. Such services have more than "two" point.

I suggest we simply go for: "while the packet traverses a network"

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

In-situ Operations, Administration, and Maintenance ~~OAM~~ (IOAM), as defined in [I-D.ietf-ippm-ioam-data], is used to record and collect OAM information while the packet traverses a particular network domain. The term "in-situ" refers to the fact that the OAM data is added to the data packets rather than is being sent within packets specifically dedicated to OAM.

This document defines how IOAM data fields are transported as part of the Network Service Header (NSH) [RFC8300] encapsulation for the Service Function Chaining (SFC) [RFC7665]. The IOAM-Data-Fields are defined in [I-D.ietf-ippm-ioam-data].

Commenté [BMI2]: Start a new para.

Considerations that motivated the design in the document are elaborated in Appendix A.

MTU-related considerations are similar to those already discussed in [I-D.ietf-ippm-ioam-data] and Section 5 of [RFC8300]. These considerations are not reiterated in this document.

An implementation of IOAM ~~which~~ that leverages the NSH to carry the IOAM data is available from the FD.io open source software project [FD.io].

Commenté [BMI3]: Start a new para.

2. Conventions

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.

The following ~~Abbreviations~~ abbreviations are used in this document:

IOAM: In-situ Operations, Administration, and Maintenance

NSH: Network Service Header

OAM: Operations, Administration, and Maintenance
SFC: Service Function Chaining

TLV: Type, Length, Value

3. IOAM ~~encapsulation~~ Encapsulation with the NSH

The NSH is defined in [RFC8300]. IOAM-Data-Fields are carried as NSH payload using a next protocol header which follows the NSH Base Header, Service Path Header, and optional Context Headers~~headers~~ (Section 2.1 of [RFC8300]).

An IOAM header is added containing the different IOAM-Data-Fields.

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]). In an administrative domain where IOAM is used, insertion of the IOAM header in the NSH is enabled at the NSH ~~tunnel~~ ingress endpoints, which also serve as IOAM encapsulating/decapsulating nodes by means of configuration. There can be multiple IOAM headers added by encapsulating nodes as configured. The IOAM transit nodes (e.g., ~~ana~~ Service Function Forwarder (SFF)) MUST process all the IOAM headers that are relevant based on its local configuration. See [I-D.ietf-ippm-ioam-deployment] for a discussion of deployment related aspects of IOAM-Data-fields.

```
  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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+<-+
 |Ver|O|U|    TTL    |   Length  |U|U|U|U|MD Type| NP = TBD_IOAM |  |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+  N
 |          Service Path Identifier              | Service Index |  S
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+  H
 |                            ...                                |  |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+<-+
 |  IOAM-Type    | IOAM HDR len  |    Reserved   | Next Protocol |  |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+  I
 !                                                               |  O
 !                                                               |  A
 ~                 IOAM Option and Optional Data Space           ~  M
 |                                                               |  |
 |                                                               |  |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+<-+
 |                                                               |
 |                                                               |
 |                 Payload + Padding (L2/L3/ESP/...)             |
 |                                                               |
 |                                                               |
 |                                                               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure x: IOAM over NSH Format

The NSH ~~header~~ and fields are defined in Section 2 of [RFC8300]. The O-bit MUST be handled following the rules in [I-D.ietf-sfc-oam-packet]. The NSH "~~NSH~~ Next Protocol" value (referred to as "NP" in ~~the diagram above~~Figure x) is TBD_IOAM.

The ~~IOAM~~ IOAM-related fields in the NSH are defined as follows:

IOAM-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 Len: ~~8 bit~~8-bit Length field that contains the length of the IOAM

Commenté [BMI4]: There is no such a notion of "NSH tunnel endpoint' in 8300 and 7665.

Commenté [BMI5]: Please check if this is the intent.

Commenté [BMI6]: Which padding?

Commenté [BMI7]: Redundant with « H » of NSH.

Commenté [BMI8]: To be consistent with RFC8300

Commenté [BMI9]: Please add a note for the RFC Editor to replace all occurrences of TBD_IOAM with the value assigned by IANA.

header in 4-octet units.

Reserved bits: Reserved bits are ~~present~~ for future use. The reserved bits MUST be set to 0x0 upon transmission and ignored upon receipt.

Next Protocol: 8-bit unsigned integer that determines the type of header following IOAM. The semantics of this field are identical to the Next Protocol field in Section 2.2 of [RFC8300].

IOAM Option and Data Space: IOAM-Data-Fields as specified by the IOAM-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]).

Multiple IOAM-Option-Types MAY be included within the NSH encapsulation. For example, if a NSH encapsulation contains two IOAM-Option-Types before a data payload, the Next Protocol field of the first IOAM option will contain the value of TBD_IOAM, while the Next Protocol field of the second IOAM-Option-Type will contain the NSH "~~NSH~~ Next Protocol" number indicating the type of the data payload.

The applicability of the IOAM Active and Loopback flags [I-D.ietf-ippm-ioam-flags] is outside the scope of this document and may be specified in the future.

When a packet with IOAM is received at an NSH based forwarding node such as an ~~Service Function Forwarder (~~SFF~~)~~ that does not understand IOAM header, it SHOULD drop the packet. The mechanism to maintain and notify of such events are local policies that are outside the scope of this document.

4. IANA Considerations

IANA is requested to allocate a protocol number~~s for the following~~ from the "NSH Next Protocol" registry available at <https://www.iana.org/assignments/nsh/nsh.xhtml#next-protocol>~~related to IOAM~~:

Next Protocol	Description	Reference
TBD_IOAM~~x~~	~~–~~IOAM~~TBD_IOAM~~	This document

5. Security Considerations

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. For additional IOAM

Commenté [BMI10]: I wonder whether we can indicate that these bits MUST NOT be changed when the packet crosses an SFP. This would ensure consistent behavior when traversing an SFC-enabled domain.

Commenté [BMI11]: Start a new para

Commenté [BMI12]: Start a new para

Commenté [BMI13]: To maintain what? the header?

related security considerations, see Section 10 in [I-D.ietf-ippm-ioam-data].

For additional OAM and NSH related security considerations see Section 5 of [I-D.ietf-sfc-oam-packet].

6. Acknowledgements

The authors would like to thank Eric Vyncke, Nalini Elkins, Srihari Raghavan, Ranganathan T S, Karthik Babu Harichandra Babu, Akshaya Nadahalli, Stefano Previdi, Hemant Singh, Erik Nordmark, LJ Wobker, Andrew Yourtchenko, Greg Mirsky and Mohamed Boucadair for the comments and advice.

7. Contributors

In addition to editors listed on the title page, the following people have contributed to this document:

Vengada Prasad Govindan
Cisco Systems, Inc.
Email: venggovi@cisco.com

Carlos Pignataro
Cisco Systems, Inc.
7200-11 Kit Creek Road
Research Triangle Park, NC 27709
United States
Email: cpignata@cisco.com

Hannes Gredler
RtBrick Inc.
Email: hannes@rtbrick.com
John Leddy
Email: john@leddy.net

Stephen Youell
JP Morgan Chase
25 Bank Street
London E14 5JP
United Kingdom
Email: stephen.youell@jpmorgan.com

Tal Mizrahi
Huawei Network.IO Innovation Lab
Israel
Email: tal.mizrahi.phd@gmail.com

David Mozes
Email: mosesster@gmail.com

Petr Lapukhov
Facebook
1 Hacker Way
Menlo Park, CA 94025
US
Email: petr@fb.com

Remy Chang
Barefoot Networks
2185 Park Boulevard
Palo Alto, CA 94306
US

8. References

8.1. Normative References

[I-D.ietf-ippm-ioam-data] Brockners, F., Bhandari, S., and T. Mizrahi, "Data Fields for In-situ OAM", draft-ietf-ippm-ioam-data-17 (work in progress), December 2021.

[I-D.ietf-sfc-oam-packet] Boucadair, M., "OAM Packet and Behavior in the Network Service Header (NSH)", draft-ietf-sfc-oam-packet-01 (work in progress), April 2022.

[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>.

[RFC8300] Quinn, P., Ed., Elzur, U., Ed., and C. Pignataro, Ed., "Network Service Header (NSH)", RFC 8300, DOI 10.17487/RFC8300, January 2018, <https://www.rfc-editor.org/info/rfc8300>.

8.2. Informative References

[FD.io] "Fast Data Project: FD.io", <https://fd.io/>.

[I-D.ietf-ippm-ioam-deployment] Brockners, F., Bhandari, S., Bernier, D., and T. Mizrahi, "In-situ OAM Deployment", draft-ietf-ippm-ioam-deployment-01 (work in progress), April 2022.

[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", draft-ietf-ippm-ioam-direct-export-07 (work in progress), October 2021.

[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", draft-ietf-ippm-ioam-flags-07 (work in progress), October 2021.

[RFC7665] Halpern, J., Ed. and C. Pignataro, Ed., "Service Function Chaining (SFC) Architecture", RFC 7665, DOI 10.17487/RFC7665, October 2015,

<https://www.rfc-editor.org/info/rfc7665>.

Appendix A. Discussion of the IOAM ~~encapsulation~~ Encapsulation ~~approach~~Approach

This ~~section~~ appendix lists several approaches considered for encapsulating IOAM with the NSH and presents the rationale for the approach chosen in this document.

An encapsulation of IOAM-Data-Fields in the NSH should be friendly to an implementation in both hardware as well as software forwarders and support a wide range of deployment cases, including large networks that desire to leverage multiple IOAM-Data-Fields at the same time.

Hardware and software friendly implementation: Hardware forwarders benefit from an encapsulation that minimizes iterative look-ups of fields within the packet: Any operation which looks up the value of a field within the packet, based on which another lookup is performed, consumes additional gates and time in an implementation - both of which are desired to be kept to a minimum. This means that flat TLV structures are to be preferred over nested TLV structures. IOAM-Data-Fields are grouped into several categories, including trace, proof-of-transit, and edge-to-edge. Each of these options defines a TLV structure. A hardware-friendly encapsulation approach avoids grouping these three option categories into yet another TLV structure, but would rather carry the options as a serial sequence.

Total length of the IOAM-Data-Fields: The total length of IOAM-Data-Fields can grow quite large in case multiple different IOAM-Data-Fields are used and large path-lengths need to be considered. If, for Example, an operator ~~would~~ considers using the IOAM Trace Option-Type and capture node-id, app_data, egress/ingress interface-id, timestamp seconds, timestamps nanoseconds at every hop, then a total of 20 octets would be added to the packet at every hop. In case this particular deployment would have a maximum path length of 15 hops in the IOAM domain, then a maximum of 300 octets were to be encapsulated in the packet.

Commenté [BMI14]: Where those are defined? Consider adding a pointer.

Different approaches for encapsulating IOAM-Data-Fields in the NSH could be considered:

1. Encapsulation of IOAM-Data-Fields as "NSH MD Type 2" (see [RFC8300], Section 2.5). Each IOAM-Option-Type (e.g., trace, proof-of-transit, and edge-to-edge) would be specified by a type, with the different IOAM-Data-Fields being TLVs within this the particular option type. NSH MD Type 2 offers support for variable length meta-data. The length field is 6-bits, resulting in a maximum of 256 (2^6 x 4) octets.

2. Encapsulation of IOAM-Data-Fields using the "Next Protocol" field. Each IOAM-Option-Type (e.g., trace, proof-of-transit, and edge-to-edge) would be specified by its own "next protocol".

3. Encapsulation of IOAM-Data-Fields using the "Next Protocol"

field. A single NSH protocol type code point would be allocated for IOAM. A "sub-type" field would then specify what IOAM options type (trace, proof-of-transit, edge-to-edge) is carried.

The third option has been chosen here. This option avoids the additional layer of TLV nesting that the use of NSH MD Type 2 would result in. In addition, this option does not constrain IOAM data to a maximum of 256 octets, thus allowing support for very large deployments.

Authors' Addresses

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

Email: fbrockne@cisco.com

Shwetha Bhandari (editor)
Thoughtspot
3rd Floor, Indiqube Orion, 24th Main Rd, Garden Layout, HSR Layout
Bangalore, KARNATAKA 560 102
India

Email: shwetha.bhandari@thoughtspot.com