

sfc  
Internet-Draft  
Intended status: Standards Track  
Expires: December 2, 2018

F. Brockners  
S. Bhandari  
V. Govindan  
C. Pignataro  
Cisco  
H. Gredler  
RtBrick Inc.  
J. Leddy  
Comcast  
S. Youell  
JMPC  
T. Mizrahi  
Marvell  
D. Mozes  
  
P. Lapukhov  
Facebook  
R. Chang  
Barefoot Networks  
May 31, 2018

**NSH Encapsulation for In-situ OAM Data**  
**draft-ietf-sfc-ioam-nsh-00**

**Abstract**

In-situ Operations, Administration, and Maintenance (OAM) records operational and telemetry information in the packet while the packet traverses a path between two points in the network. This document outlines how IOAM data fields are encapsulated in the Network Service Header (NSH).

**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 <http://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 December 2, 2018.

## Copyright Notice

Copyright (c) 2018 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 (<http://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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">2</a>
<a href="#">2.</a>	Conventions . . . . .	<a href="#">3</a>
<a href="#">3.</a>	IOAM data fields encapsulation in NSH . . . . .	<a href="#">3</a>
<a href="#">4.</a>	Considerations . . . . .	<a href="#">5</a>
<a href="#">4.1.</a>	Discussion of the encapsulation approach . . . . .	<a href="#">5</a>
<a href="#">4.2.</a>	IOAM and the use of the NSH 0-bit . . . . .	<a href="#">6</a>
<a href="#">5.</a>	IANA Considerations . . . . .	<a href="#">6</a>
<a href="#">6.</a>	Security Considerations . . . . .	<a href="#">7</a>
<a href="#">7.</a>	Acknowledgements . . . . .	<a href="#">7</a>
<a href="#">8.</a>	References . . . . .	<a href="#">7</a>
<a href="#">8.1.</a>	Normative References . . . . .	<a href="#">7</a>
<a href="#">8.2.</a>	Informative References . . . . .	<a href="#">8</a>
	Authors' Addresses . . . . .	<a href="#">8</a>

## [1.](#) Introduction

In-situ OAM (IOAM) records OAM information within the packet 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. The IOAM data fields are defined in [[I-D.ietf-ippm-ioam-data](#)]. An implementation of IOAM which leverages NSH to carry the IOAM data is available from the FD.io open source software project [[FD.io](#)].



## 2. Conventions

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 [\[RFC2119\]](#).

Abbreviations used in this document:

IOAM: In-situ Operations, Administration, and Maintenance

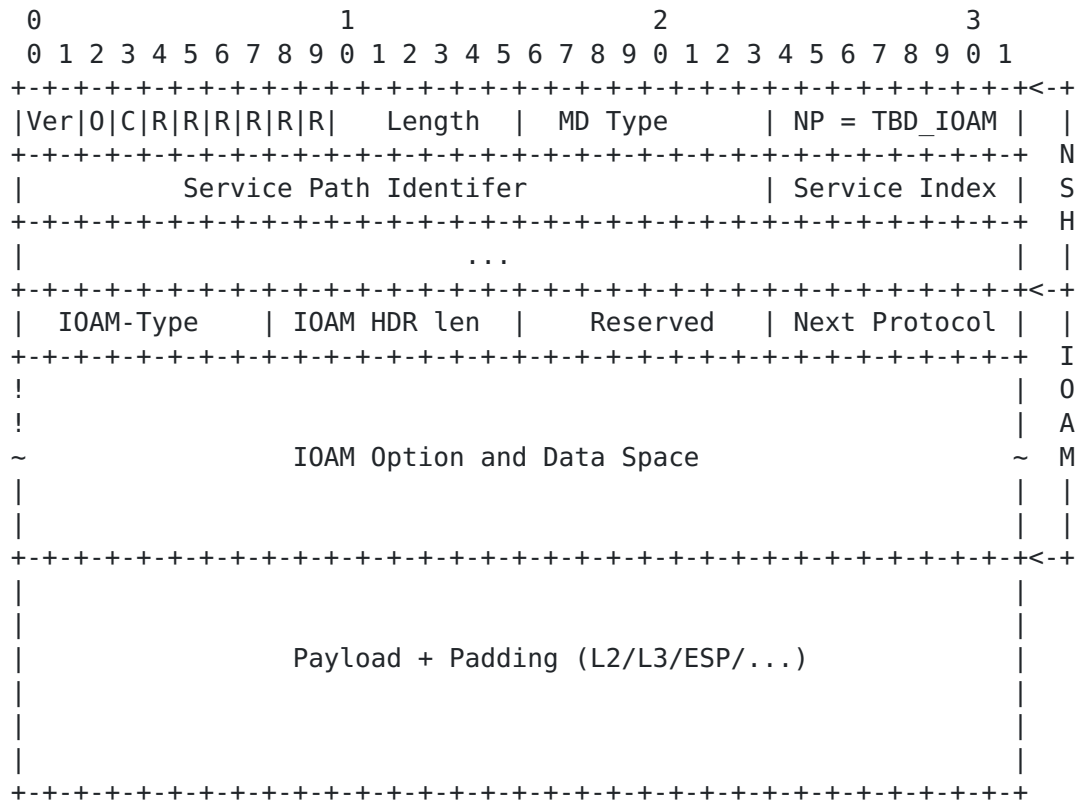
NSH: Network Service Header

OAM: Operations, Administration, and Maintenance

TLV: Type, Length, Value

## 3. IOAM data fields encapsulation in NSH

NSH is defined in [\[RFC8300\]](#). IOAM data fields are carried in NSH using a next protocol header which follows the NSH MDx metadata TLVs. An IOAM header is added containing the different IOAM data fields defined in [\[I-D.ietf-ippm-ioam-data\]](#). In an administrative domain where IOAM is used, insertion of the IOAM header in NSH is enabled at the NSH tunnel endpoints, which also serve as IOAM encapsulating/decapsulating nodes by means of configuration.



The NSH header and fields are defined in [\[RFC8300\]](#). The "NSH Next Protocol" value (referred to as "NP" in the diagram above) is TBD IOAM.

The IOAM related fields in NSH are defined as follows:

IOAM-Type: 8-bit field defining the IOAM Option type, as defined in Section 7.2 of [\[I-D.ietf-ippm-ioam-data\]](#).

IOAM HDR Len: 8 bit Length field contains the length of the IOAM 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 protocol.

IOAM Option and Data Space: IOAM option header and data is present as specified by the IOAM-Type field, and is defined in Section 4 of [\[I-D.ietf-ippm-ioam-data\]](#).



Multiple IOAM options MAY be included within the NSH encapsulation. For example, if a NSH encapsulation contains two IOAM options 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 will contain the "NSH Next Protocol" number indicating the type of the data payload.

#### **4. Considerations**

This section summarizes a set of considerations on the overall approach taken for IOAM data encapsulation in NSH, as well as deployment considerations.

##### **4.1. Discussion of the encapsulation approach**

This section is to support the working group discussion in selecting the most appropriate approach for encapsulating IOAM data fields in NSH.

An encapsulation of IOAM data fields in 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 three option categories: Trace, proof-of-transit, and edge-to-edge. Each of these three 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 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 consider using the IOAM trace option 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 of IOAM data were to be encapsulated in the packet.





Different approaches for encapsulating IOAM data fields in NSH could be considered:

1. Encapsulation of IOAM data fields as "NSH MD Type 2" (see [\[RFC8300\], section 2.5](#)). Each IOAM data field option (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 ( $2^6 \times 4$ ) octets.
2. Encapsulation of IOAM data fields using the "Next Protocol" field. Each IOAM data field option (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.

#### [4.2.](#) IOAM and the use of the NSH 0-bit

[RFC8300] defines an "0 bit" for OAM packets. Per [\[RFC8300\]](#) the 0 bit must be set for OAM packets and must not be set for non-OAM packets. Packets with IOAM data included MUST follow this definition, i.e. the 0 bit MUST NOT be set for regular customer traffic which also carries IOAM data and the 0 bit MUST be set for OAM packets which carry only IOAM data without any regular data payload.

#### [5.](#) IANA Considerations

IANA is requested to allocate protocol numbers for the following "NSH Next Protocol" related to IOAM:

+-----+-----+-----+
Next Protocol   Description   Reference
+-----+-----+-----+
x   TBD_IOAM   This document
+-----+-----+-----+



## 6. 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.

## 7. 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, and Andrew Yourtchenko for the comments and advice.

## 8. References

### 8.1. Normative References

- [I-D.ietf-ippm-ioam-data]  
Brockners, F., Bhandari, S., Pignataro, C., Gredler, H., Leddy, J., Youell, S., Mizrahi, T., Mozes, D., Lapukhov, P., Chang, R., daniel.bernier@bell.ca, d., and J. Lemon, "Data Fields for In-situ OAM", [draft-ietf-ippm-ioam-data-02](#) (work in progress), March 2018.
- [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>>.
- [RFC2784] Farinacci, D., Li, T., Hanks, S., Meyer, D., and P. Traina, "Generic Routing Encapsulation (GRE)", [RFC 2784](#), DOI 10.17487/RFC2784, March 2000, <<https://www.rfc-editor.org/info/rfc2784>>.
- [RFC3232] Reynolds, J., Ed., "Assigned Numbers: [RFC 1700](#) is Replaced by an On-line Database", [RFC 3232](#), DOI 10.17487/RFC3232, January 2002, <<https://www.rfc-editor.org/info/rfc3232>>.
- [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/>>.
- [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>>.

### Authors' Addresses

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

Email: [fbrockne@cisco.com](mailto:fbrockne@cisco.com)

Shwetha Bhandari  
Cisco Systems, Inc.  
Cessna Business Park, Sarjapura Marathalli Outer Ring Road  
Bangalore, KARNATAKA 560 087  
India

Email: [shwethab@cisco.com](mailto:shwethab@cisco.com)

Vengada Prasad Govindan  
Cisco Systems, Inc.

Email: [venggovi@cisco.com](mailto:venggovi@cisco.com)

Carlos Pignataro  
Cisco Systems, Inc.  
7200-11 Kit Creek Road  
Research Triangle Park, NC 27709  
United States

Email: [cpignata@cisco.com](mailto:cpignata@cisco.com)

Hannes Gredler  
RtBrick Inc.

Email: [hannes@rtbrick.com](mailto:hannes@rtbrick.com)

John Leddy  
Comcast

Email: John\_Leddy@cable.comcast.com

Stephen Youell  
JP Morgan Chase  
25 Bank Street  
London E14 5JP  
United Kingdom

Email: stephen.youell@jpmorgan.com

Tal Mizrahi  
Marvell  
6 Hamada St.  
Yokneam 20692  
Israel

Email: talmi@marvell.com

David Mozes

Email: mozesster@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