

Network Working Group  
Internet-Draft  
Intended status: Informational  
Expires: January 19, 2017

F. Brockners  
S. Bhandari  
C. Pignataro  
Cisco  
H. Gredler  
RtBrick Inc.  
J. Leddy  
Comcast  
S. Youell  
JMPC  
July 18, 2016

**Encapsulations for In-band OAM Data  
draft-brockners-inband-oam-transport-01**

Abstract

In-band operation, administration and maintenance (OAM) records operational and telemetry information in the packet while the packet traverses a path between two points in the network. In-band OAM is to complement current out-of-band OAM mechanisms based on ICMP or other types of probe packets. This document outlines how in-band OAM data records can be transported in protocols such as NSH, Segment Routing, VXLAN-GPE, native IPv6 (via extension header), and IPv4. Transport options are currently investigated as part of an implementation study. This document is intended to only serve informational purposes.

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 January 19, 2017.

Copyright Notice

Copyright (c) 2016 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

- [1.](#) Introduction . . . . . [2](#)
- [2.](#) Conventions . . . . . [3](#)
- [3.](#) In-Band OAM Metadata Transport in IPv6 . . . . . [4](#)
  - [3.1.](#) In-band OAM in IPv6 Hop by Hop Extension Header . . . . . [4](#)
    - [3.1.1.](#) In-band OAM Hop by Hop Options . . . . . [4](#)
    - [3.1.2.](#) Procedure at the Ingress Edge to Insert the In-band OAM Header . . . . . [6](#)
    - [3.1.3.](#) Procedure at Intermediate Nodes . . . . . [7](#)
    - [3.1.4.](#) Procedure at the Egress Edge to Remove the In-band OAM Header . . . . . [7](#)
- [4.](#) In-band OAM Metadata Transport in VXLAN-GPE . . . . . [7](#)
- [5.](#) In-band OAM Metadata Transport in NSH . . . . . [9](#)
- [6.](#) In-band OAM Metadata Transport in Segment Routing . . . . . [11](#)
  - [6.1.](#) In-band OAM in SR with IPv6 Transport . . . . . [11](#)
  - [6.2.](#) In-band OAM in SR with MPLS Transport . . . . . [11](#)
- [7.](#) IANA Considerations . . . . . [12](#)
- [8.](#) Manageability Considerations . . . . . [12](#)
- [9.](#) Security Considerations . . . . . [12](#)
- [10.](#) Acknowledgements . . . . . [12](#)
- [11.](#) References . . . . . [12](#)
  - [11.1.](#) Normative References . . . . . [12](#)
  - [11.2.](#) Informative References . . . . . [12](#)
- Authors' Addresses . . . . . [14](#)

**1. Introduction**

This document discusses transport mechanisms for "in-band" operation, administration, and maintenance (OAM) data records. In-band OAM records OAM information within the packet while the packet traverses a particular network domain. The term "in-band" 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. A discussion of the motivation and requirements for in-band OAM can be found in [[draft-brockners-inband-oam-requirements](#)]. Data types and data formats for in-band OAM are defined in [[draft-brockners-inband-oam-data](#)].

This document outlines transport encapsulations for the in-band OAM data defined in [[draft-brockners-inband-oam-data](#)]. This document is to serve informational purposes only. As part of an in-band OAM implementation study different protocol encapsulations for in-band OAM data are being explored. Once data formats and encapsulation approaches are settled, protocol specific specifications for in-band OAM data transport will address the standardization aspect.

The data for in-band OAM defined in [[draft-brockners-inband-oam-data](#)] can be carried in a variety of protocols based on the deployment needs. This document discusses transport of in-band OAM data for the following protocols:

- o IPv6
- o VXLAN-GPE
- o NSH
- o Segment Routing (IPv6 and MPLS)

This list is non-exhaustive, as it is possible to carry the in-band OAM data in several other protocols and transports.

A feasibility study of in-band OAM is currently underway as part of the FD.io project [[FD.io](#)]. The in-band OAM implementation study should be considered as a "tool box" to showcase how "in-band" OAM can complement probe-packet based OAM mechanisms for different deployments and packet transport formats. For details, see the open source code in the FD.io [[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:

MTU: Maximum Transmit Unit

OAM: Operations, Administration, and Maintenance



SR:	Segment Routing
SID:	Segment Identifier
NSH:	Network Service Header
POT:	Proof of Transit
SFC:	Service Function Chain
VXLAN-GPE:	Virtual eXtensible Local Area Network, Generic Protocol Extension

### **3. In-Band OAM Metadata Transport in IPv6**

This mechanisms of in-band OAM in IPv6 complement others proposed to enhance diagnostics of IPv6 networks, such as the IPv6 Performance and Diagnostic Metrics Destination Option described in [[I-D.ietf-ippm-6man-pdm-option](#)]. The IP Performance and Diagnostic Metrics Destination Option is destination focused and specific to IPv6, whereas in-band OAM is performed between end-points of the network or a network domain where it is enabled and used.

A historical note: The idea of IPv6 route recording was originally introduced by [[draft-kitamura-ipv6-record-route](#)] back in year 2000. With IPv6 now being generally deployed and new concepts such as Segment Routing [[I-D.ietf-spring-segment-routing](#)] being introduced, it is imperative to further mature the operations, administration, and maintenance mechanisms available to IPv6 networks.

The in-band OAM options translate into options for an IPv6 extension header. The extension header would be inserted by either a host source of the packet, or by a transit/domain-edge node.

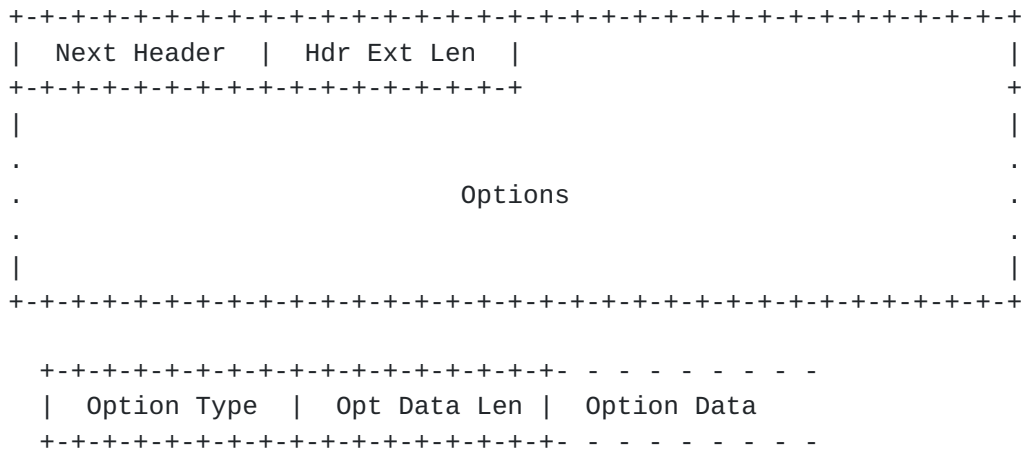
#### **3.1. In-band OAM in IPv6 Hop by Hop Extension Header**

This section defines in-band OAM for IPv6 transport. In-band OAM data is transported as an IPv6 hop-by-hop extension header.

##### **3.1.1. In-band OAM Hop by Hop Options**

Brief recap of the IPv6 hop-by-hop header as well as the options used for carrying in-band OAM data:





With 2 highest order bits of Option Type indicating the following:

- 00 - skip over this option and continue processing the header.
- 01 - discard the packet.
- 10 - discard the packet and, regardless of whether or not the packet's Destination Address was a multicast address, send an ICMP Parameter Problem, Code 2, message to the packet's Source Address, pointing to the unrecognized Option Type.
- 11 - discard the packet and, only if the packet's Destination Address was not a multicast address, send an ICMP Parameter Problem, Code 2, message to the packet's Source Address, pointing to the unrecognized Option Type.

3rd highest bit:

- 0 - Option Data does not change en-route
- 1 - Option Data may change en-route

In-band OAM data records are inserted as options in an IPv6 hop-by-hop extension header:

1. Tracing Option: The in-band OAM Tracing option defined in [\[draft-brockners-inband-oam-data\]](#) is represented as a IPv6 option in hop by hop extension header by allocating following type:

Option Type: 001xxxxxx 8-bit identifier of the type of option.  
 xxxxxx=TBD\_IANA\_TRACE\_OPTION\_IPV6.





2. Proof of Transit Option: The in-band OAM POT option defined in [\[draft-brockners-inband-oam-data\]](#) is represented as a IPv6 option in hop by hop extension header by allocating following type:

Option Type: 001xxxxxx 8-bit identifier of the type of option.  
 xxxxxx=TBD\_IANA\_POT\_OPTION\_IPV6.

3. Edge to Edge Option: The in-band OAM E2E option defined in [\[draft-brockners-inband-oam-data\]](#) is represented as a IPv6 option in hop by hop extension header by allocating following type:

Option Type: 000xxxxxx 8-bit identifier of the type of option.  
 xxxxxx=TBD\_IANA\_E2E\_OPTION\_IPV6.

### **3.1.2. Procedure at the Ingress Edge to Insert the In-band OAM Header**

In an administrative domain where in-band OAM is used, insertion of the in-band OAM header is enabled at the required edge nodes by means of configuration.

Such a config SHOULD allow selective enablement of in-band OAM header insertion for a subset of traffic (e.g., one or several "pipes").

Further the ingress edge node should be aware of maximum size of the header that can be inserted. Details on how the maximum size/size of the in-band OAM domain are retrieved are outside the scope of this document.

Let  $n$  = max number of nodes to be allocated;  
 (Based on PMTU advertised in the domain)

Let  $k$  = number of node data that can be allocated by this node  
 Let  $node\_data\_size$  = size of each  $node\_data$  based on in-band OAM type

```

if (packet matches traffic for which in-band OAM is enabled) {
  Create in-band OAM hbyh ext header with  $k$  node data preallocated
  Increment payload length in IPv6 header :
      with size of in-band OAM hbyh ext header
  Populate node data at :
      (size of in-band OAM hbyh header = 8) +  $k * node\_data\_size$ 
  from the beginning of the header
  Set segments left to :  $k - 1$ 
}

```



### **3.1.3. Procedure at Intermediate Nodes**

If a network node receives a packet with an in-band OAM header and it is enabled to process in-band OAM data it performs the following:

```
k = number of node data that this node can allocate
if (in-band OAM ext hbyh header is present) {
    if (Segments Left > 0) {
        populate node data at :
            node_data_start[Segments Left]
        Segments Left = Segments Left - 1
    }
}
```

### **3.1.4. Procedure at the Egress Edge to Remove the In-band OAM Header**

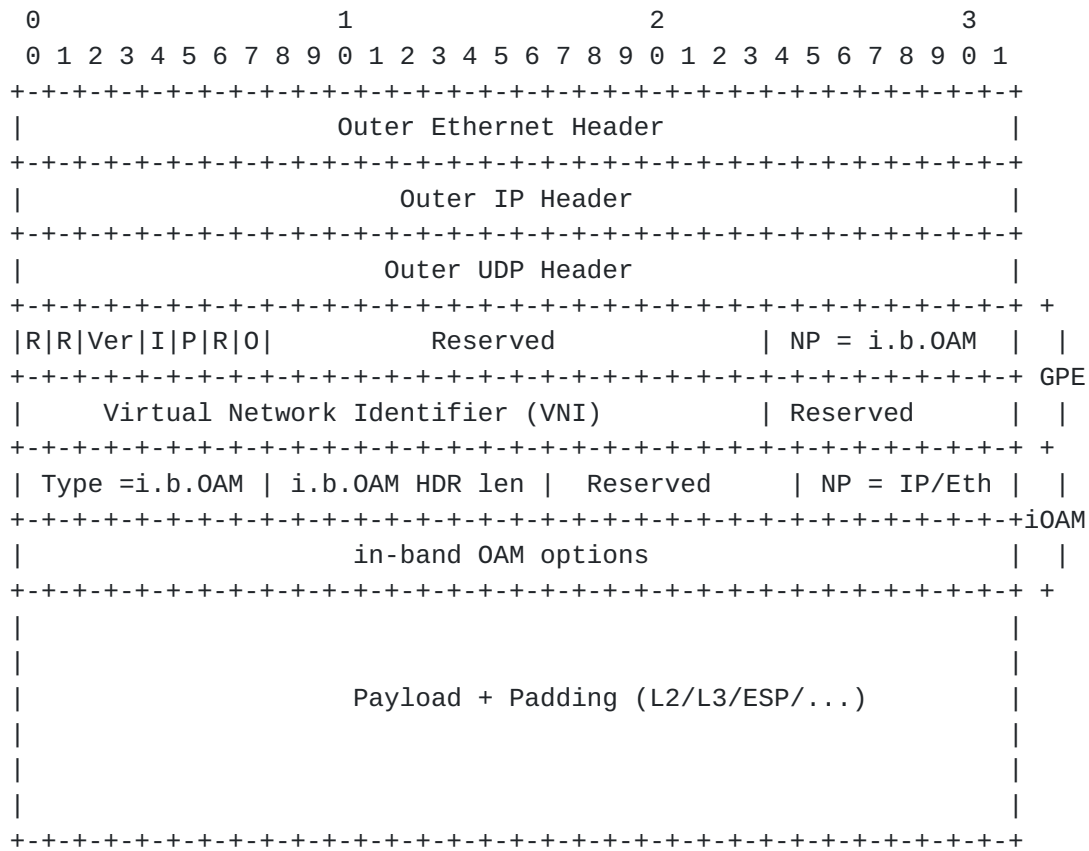
```
egress_edge = list of interfaces where in-band OAM hbyh ext
                header is to be stripped
Before forwarding packet out of interfaces in egress_edge list:
if (in-band OAM hbyh ext header is present) {
    remove the in-band OAM hbyh ext header,
    possibly store the record along with additional
    fields for analysis and export
    Decrement Payload Length in IPv6 header
    by size of in-band OAM ext header
}
```

## **4. In-band OAM Metadata Transport in VXLAN-GPE**

VXLAN-GPE [[I-D.ietf-nvo3-vxlan-gpe](#)] encapsulation is somewhat similar to IPv6 extension headers in that a series of headers can be contained in the header as a linked list. The different in-band OAM types are added as options within a new in-band OAM protocol header in VXLAN GPE.



In-band OAM header in VXLAN GPE header:



The VXLAN-GPE header and fields are defined in [\[I-D.ietf-nvo3-vxlan-gpe\]](#). in-band OAM specific fields and header are defined here:

- Type: 8-bit unsigned integer defining in-band OAM header type
- in-band OAM HDR len: 8-bit unsigned integer. Length of the in-band OAM HDR in 8-octet units
- in-band OAM options: Variable-length field, of length such that the complete in-band OAM header is an integer multiple of 8 octets long. Contains one or more TLV-encoded options of the format:

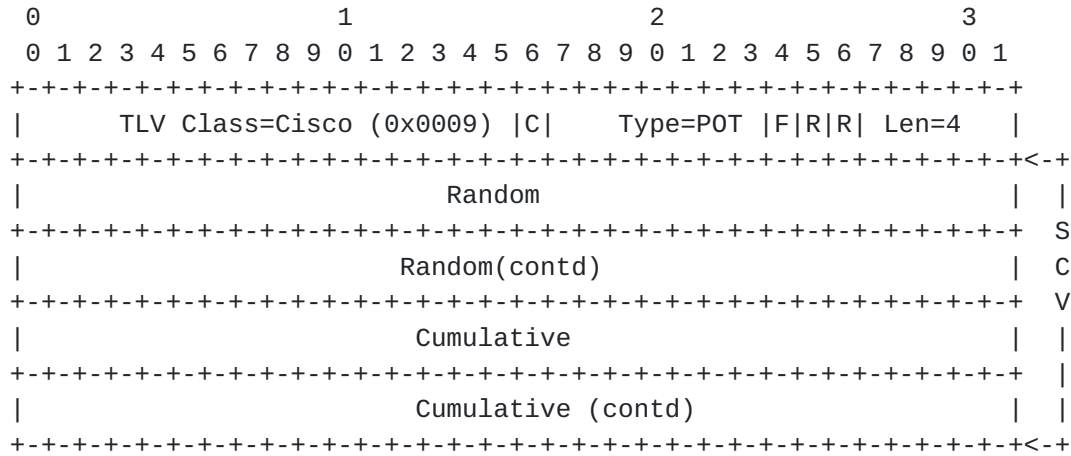








chain. The "Proof of Transit" capabilities (see [draft-brockners-inband-oam-requirements] and [draft-brockners-proof-of-transit]) of in-band OAM can be leveraged within NSH. Proof of transit in-band OAM data is added as NSH Type 2 metadata:



TLV Class: Describes the scope of the "Type" field. In some cases, the TLV Class will identify a specific vendor, in others, the TLV Class will identify specific standards body allocated types. POT is currently defined using the Cisco (0x0009) TLV class.

Type: The specific type of information being carried, within the scope of a given TLV Class. Value allocation is the responsibility of the TLV Class owner. Currently a type value of 0x94 is used for proof of transit

Reserved bits: Two reserved bit are present for future use. The reserved bits MUST be set to 0x0.

F: One bit. Indicates which POT-profile is active. 0 means the even POT-profile is active, 1 means the odd POT-profile is active.

Length: Length of the variable metadata, in 4-octet words. Here the length is 4.

Random: 64-bit Per packet Random number.

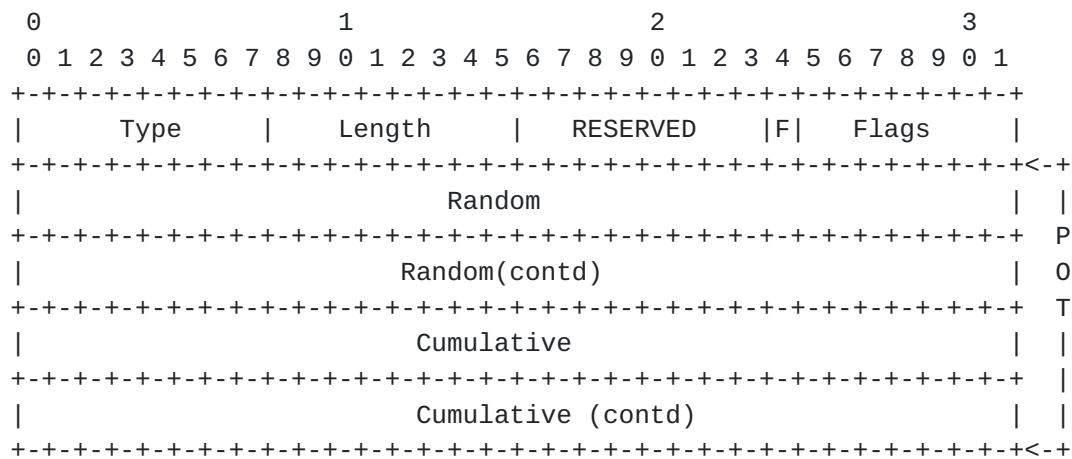
Cumulative: 64-bit Cumulative that is updated by the Service Functions.



### 6. In-band OAM Metadata Transport in Segment Routing

#### 6.1. In-band OAM in SR with IPv6 Transport

Similar to NSH, a service chain or path defined using Segment Routing for IPv6 can be verified using the in-band OAM "Proof of Transit" approach. The Segment Routing Header (SRH) for IPv6 offers the ability to transport TLV structured data, similar to what NSH does (see [[I-D.ietf-6man-segment-routing-header](#)]). A new "POT TLV" is defined for the SRH which is to carry proof of transit in-band OAM data.



Type: To be assigned by IANA.

Length: 18.

RESERVED: 8 bits. SHOULD be unset on transmission and MUST be ignored on receipt.

F: 1 bit. Indicates which POT-profile is active. 0 means the even POT-profile is active, 1 means the odd POT-profile is active.

Flags: 8 bits. No flags are defined in this document.

Random: 64-bit per packet random number.

Cumulative: 64-bit cumulative value that is updated at specific nodes that form the service path to be verified.

#### 6.2. In-band OAM in SR with MPLS Transport

In-band OAM "Proof of Transit" data can also be carried as part of the MPLS label stack. Details will be addressed in a future version of this document.



## **7. IANA Considerations**

IANA considerations will be added in a future version of this document.

## **8. Manageability Considerations**

Manageability considerations will be addressed in a later version of this document..

## **9. Security Considerations**

Security considerations will be addressed in a later version of this document. For a discussion of security requirements of in-band OAM, please refer to [[draft-brockners-inband-oam-requirements](#)].

## **10. Acknowledgements**

The authors would like to thank Eric Vyncke, Nalini Elkins, Srihari Raghavan, Ranganathan T S, Karthik Babu Harichandra Babu, Akshaya Nadahalli, Stefano Previdi and Andrew Yourtchenko for the comments and advice. For the IPv6 encapsulation, this document leverages and builds on top of several concepts described in [[draft-kitamura-ipv6-record-route](#)]. The authors would like to acknowledge the work done by the author Hiroshi Kitamura and people involved in writing it.

## **11. References**

### **11.1. Normative References**

[[draft-brockners-inband-oam-requirements](#)]

Brockners, F., Bhandari, S., and S. Dara, "Requirements for in-band OAM", July 2016.

### **11.2. Informative References**

[[draft-brockners-inband-oam-data](#)]

Brockners, F., Bhandari, S., Pignataro, C., and H. Gredler, "Data Formats for in-band OAM", July 2016.

[[draft-brockners-proof-of-transit](#)]

Brockners, F., Bhandari, S., and S. Dara, "Proof of transit", July 2016.

[[draft-kitamura-ipv6-record-route](#)]

Kitamura, H., "Record Route for IPv6 (PR6), Hop-by-Hop Option Extension", November 2000.



- [FD.io] "Fast Data Project: FD.io", <<https://fd.io/>>.
- [I-D.hildebrand-spud-prototype]  
Hildebrand, J. and B. Trammell, "Substrate Protocol for User Datagrams (SPUD) Prototype", [draft-hildebrand-spud-prototype-03](#) (work in progress), March 2015.
- [I-D.ietf-6man-segment-routing-header]  
Previdi, S., Filsfils, C., Field, B., Leung, I., Linkova, J., Aries, E., Kosugi, T., Vyncke, E., and D. Lebrun, "IPv6 Segment Routing Header (SRH)", [draft-ietf-6man-segment-routing-header-01](#) (work in progress), March 2016.
- [I-D.ietf-ippm-6man-pdm-option]  
Elkins, N., Hamilton, R., and m. mackermann@bcbsm.com, "IPv6 Performance and Diagnostic Metrics (PDM) Destination Option", [draft-ietf-ippm-6man-pdm-option-03](#) (work in progress), June 2016.
- [I-D.ietf-nvo3-vxlan-gpe]  
Kreeger, L. and U. Elzur, "Generic Protocol Extension for VXLAN", [draft-ietf-nvo3-vxlan-gpe-02](#) (work in progress), April 2016.
- [I-D.ietf-sfc-nsh]  
Quinn, P. and U. Elzur, "Network Service Header", [draft-ietf-sfc-nsh-05](#) (work in progress), May 2016.
- [I-D.ietf-spring-segment-routing]  
Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", [draft-ietf-spring-segment-routing-09](#) (work in progress), July 2016.
- [I-D.penno-sfc-trace]  
Penno, R., Quinn, P., Pignataro, C., and D. Zhou, "Services Function Chaining Traceroute", [draft-penno-sfc-trace-03](#) (work in progress), September 2015.
- [P4] Kim, , "P4: In-band Network Telemetry (INT)", September 2015.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.





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

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](mailto:John_Leddy@cable.comcast.com)



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

Email: [stephen.youell@jpmorgan.com](mailto:stephen.youell@jpmorgan.com)