

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

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

**Requirements for In-band OAM**  
**draft-brockners-inband-oam-requirements-01**

Abstract

This document discusses the motivation and requirements for including specific operational and telemetry information into data packets while the data packet traverses a path between two points in the network. This method is referred to as "in-band" Operations, Administration, and Maintenance (OAM), given that the OAM information is carried with the data packets as opposed to in "out-of-band" packets dedicated to OAM. In-band OAM complements other OAM mechanisms which use dedicated probe packets to convey OAM information.

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](http://trustee.ietf.org/license-info) 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>	<a href="#">Introduction</a>	<a href="#">3</a>
<a href="#">2.</a>	<a href="#">Conventions</a>	<a href="#">4</a>
<a href="#">3.</a>	<a href="#">Motivation for In-band OAM</a>	<a href="#">4</a>
<a href="#">3.1.</a>	<a href="#">Path Congruency Issues with Dedicated OAM Packets</a>	<a href="#">5</a>
<a href="#">3.2.</a>	<a href="#">Results Sent to a System Other Than the Sender</a>	<a href="#">5</a>
<a href="#">3.3.</a>	<a href="#">Overlay and Underlay Correlation</a>	<a href="#">5</a>
<a href="#">3.4.</a>	<a href="#">SLA Verification</a>	<a href="#">6</a>
<a href="#">3.5.</a>	<a href="#">Analytics and Diagnostics</a>	<a href="#">6</a>
<a href="#">3.6.</a>	<a href="#">Frame Replication/Elimination Decision for Bi-casting /Active-active Networks</a>	<a href="#">7</a>
<a href="#">3.7.</a>	<a href="#">Proof of Transit</a>	<a href="#">7</a>
<a href="#">3.8.</a>	<a href="#">Use Cases</a>	<a href="#">8</a>
<a href="#">4.</a>	<a href="#">Considerations for In-band OAM</a>	<a href="#">9</a>
<a href="#">4.1.</a>	<a href="#">Type of information to be recorded</a>	<a href="#">10</a>
<a href="#">4.2.</a>	<a href="#">MTU and packet size</a>	<a href="#">10</a>
<a href="#">4.3.</a>	<a href="#">Administrative boundaries</a>	<a href="#">11</a>
<a href="#">4.4.</a>	<a href="#">Selective enablement</a>	<a href="#">11</a>
<a href="#">4.5.</a>	<a href="#">Optimization of node and interface identifiers</a>	<a href="#">12</a>
<a href="#">4.6.</a>	<a href="#">Loop communication path (IPv6-specifics)</a>	<a href="#">12</a>
<a href="#">5.</a>	<a href="#">Requirements for In-band OAM Data Types</a>	<a href="#">12</a>
<a href="#">5.1.</a>	<a href="#">Generic Requirements</a>	<a href="#">12</a>
<a href="#">5.2.</a>	<a href="#">In-band OAM Data with Per-hop Scope</a>	<a href="#">13</a>
<a href="#">5.3.</a>	<a href="#">In-band OAM with Selected Hop Scope</a>	<a href="#">14</a>
<a href="#">5.4.</a>	<a href="#">In-band OAM with End-to-end Scope</a>	<a href="#">14</a>
<a href="#">6.</a>	<a href="#">Security Considerations and Requirements</a>	<a href="#">15</a>
<a href="#">6.1.</a>	<a href="#">Proof of Transit</a>	<a href="#">15</a>
<a href="#">7.</a>	<a href="#">IANA Considerations</a>	<a href="#">16</a>
<a href="#">8.</a>	<a href="#">Acknowledgements</a>	<a href="#">16</a>
<a href="#">9.</a>	<a href="#">Informative References</a>	<a href="#">16</a>
	<a href="#">Authors' Addresses</a>	<a href="#">18</a>



## 1. Introduction

This document discusses requirements for "in-band" Operations, Administration, and Maintenance (OAM) mechanisms. "In-band" OAM means to record OAM and telemetry information within the data packet while the data packet traverses a network or a particular network domain. The term "in-band" refers to the fact that the OAM and telemetry data is carried within data packets rather than being sent within packets specifically dedicated to OAM. In-band OAM mechanisms, which are sometimes also referred to as embedded network telemetry are a current topic of discussion. In-band network telemetry has been defined for P4 [P4]. The SPUD prototype [I-D.hildebrand-spud-prototype] uses a similar logic that allows network devices on the path between endpoints to participate explicitly in the tube outside the end-to-end context. Even the IPv4 route-record option defined in [RFC0791] can be considered an in-band OAM mechanism. In-band OAM complements "out-of-band" mechanisms such as ping or traceroute, or more recent active probing mechanisms, as described in [I-D.lapukhov-dataplane-probe]. In-band OAM mechanisms can be leveraged where current out-of-band mechanisms do not apply or do not offer the desired characteristics or requirements, such as proving that a certain set of traffic takes a pre-defined path, strict congruency is desired, checking service level agreements for the live data traffic, detailed statistics on traffic distribution paths in networks that distribute traffic across multiple paths, or scenarios where probe traffic is potentially handled differently from regular data traffic by the network devices. [RFC7276] presents an overview of OAM tools.

Compared to probably the most basic example of "in-band OAM" which is IPv4 route recording [RFC0791], an in-band OAM approach has the following capabilities:

- a. A flexible data format to allow different types of information to be captured as part of an in-band OAM operation, including not only path tracing information, but additional operational and telemetry information such as timestamps, sequence numbers, or even generic data such as queue size, geo-location of the node that forwarded the packet, etc.
- b. A data format to express node as well as link identifiers to record the path a packet takes with a fixed amount of added data.
- c. The ability to detect whether any nodes were skipped while recording in-band OAM information (i.e., in-band OAM is not supported or not enabled on those nodes).



- d. The ability to actively process information in the packet, for example to prove in a cryptographically secure way that a packet really took a pre-defined path using some traffic steering method such as service chaining or traffic engineering.
- e. The ability to include OAM data beyond simple path information, such as timestamps or even generic data of a particular use case.
- f. The ability to include OAM data in various different transport protocols.

## **2. Conventions**

Abbreviations used in this document:

ECMP:	Equal Cost Multi-Path
MTU:	Maximum Transmit Unit
NFV:	Network Function Virtualization
OAM:	Operations, Administration, and Maintenance
PMTU:	Path MTU
SLA:	Service Level Agreement
SFC:	Service Function Chain
SR:	Segment Routing

This document defines in-band Operations, Administration, and Maintenance (in-band OAM), as the subset in which OAM information is carried along with data packets. This is as opposed to "out-of-band OAM", where specific packets are dedicated to carrying OAM information.

## **3. Motivation for In-band OAM**

In several scenarios it is beneficial to make information about which path a packet took through the network available to the operator. This includes not only tasks like debugging, troubleshooting, as well as network planning and network optimization but also policy or service level agreement compliance checks. This section discusses the motivation to introduce new methods for enhanced in-band network diagnostics.



### **3.1. Path Congruency Issues with Dedicated OAM Packets**

Mechanisms which add tracing information to the regular data traffic, sometimes also referred to as "in-band" or "passive OAM" can complement active, probe-based mechanisms such as ping or traceroute, which are sometimes considered as "out-of-band", because the messages are transported independently from regular data traffic. "In-band" mechanisms do not require extra packets to be sent and hence don't change the packet traffic mix within the network. Traceroute and ping for example use ICMP messages: New packets are injected to get tracing information. Those add to the number of messages in a network, which already might be highly loaded or suffering performance issues for a particular path or traffic type.

Packet scheduling algorithms, especially for balancing traffic across equal cost paths or links, often leverage information contained within the packet, such as protocol number, IP-address or MAC-address. Probe packets would thus either need to be sent from the exact same endpoints with the exact same parameters, or probe packets would need to be artificially constructed as "fake" packets and inserted along the path. Both approaches are often not feasible from an operational perspective, be it that access to the end-system is not feasible, or that the diversity of parameters and associated probe packets to be created is simply too large. An in-band mechanism is an alternative in those cases.

In-band mechanisms also don't suffer from implementations, where probe traffic is handled differently (and potentially forwarded differently) by a router than regular data traffic.

### **3.2. Results Sent to a System Other Than the Sender**

Traditional ping and traceroute tools return the OAM results to the sender of the probe. Even when the ICMP messages that are used with these tools are enhanced, and additional telemetry is collected (e.g., ICMP Multi-Part [[RFC4884](#)] supporting MPLS information [[RFC4950](#)], Interface and Next-Hop Identification [[RFC5837](#)], etc.), it would be advantageous to separate the sending of an OAM probe from the receiving of the telemetry data. In this context, it is desired to not assume there is a bidirectional working path.

### **3.3. Overlay and Underlay Correlation**

Several network deployments leverage tunneling mechanisms to create overlay or service-layer networks. Examples include VXLAN-GPE, GRE, or LISP. One often observed attribute of overlay networks is that they do not offer the user of the overlay any insight into the underlay network. This means that the path that a particular





tunneled packet takes, nor other operational details such as the per-hop delay/jitter in the underlay are visible to the user of the overlay network, giving rise to diagnosis and debugging challenges in case of connectivity or performance issues. The scope of OAM tools like ping or traceroute is limited to either the overlay or the underlay which means that the user of the overlay has typically no access to OAM in the underlay, unless specific operational procedures are put in place. With in-band OAM the operator of the underlay can offer details of the connectivity in the underlay to the user of the overlay. The operator of the egress tunnel router could choose to share the recorded information about the path with the user of the overlay.

Coupled with mechanisms such as Segment Routing (SR) [[I-D.ietf-spring-segment-routing](#)], overlay network and underlay network can be more tightly coupled: The user of the overlay has detailed diagnostic information available in case of failure conditions. The user of the overlay can also use the path recording information as input to traffic steering or traffic engineering mechanisms, to for example achieve path symmetry for the traffic between two endpoints. [[I-D.brockners-lisp-sr](#)] is an example for how these methods can be applied to LISP.

### **[3.4.](#) SLA Verification**

In-band OAM can help users of an overlay-service to verify that negotiated SLAs for the real traffic are met by the underlay network provider. Different from solutions which rely on active probes to test an SLA, in-band OAM based mechanisms avoid wrong interpretations and "cheating", which can happen if the probe traffic that is used to perform SLA-check is prioritized by the network provider of the underlay.

### **[3.5.](#) Analytics and Diagnostics**

Network planners and operators benefit from knowledge of the actual traffic distribution in the network. When deriving an overall network connectivity traffic matrix one typically needs to correlate data gathered from each individual devices in the network. If the path of a packet is recorded while the packet is forwarded, the entire path that a packet took through the network is available to the egress system. This obviates the need to retrieve individual traffic statistics from every device in the network and correlate those statistics, or employ other mechanisms such as leveraging traffic engineering with null-bandwidth tunnels just to retrieve the appropriate statistics to generate the traffic matrix.



In addition, with individual path tracing, information is available at packet level granularity, rather than only at aggregate level - as is usually the case with IPFIX-style methods which employ flow-filters at the network elements. Data-center networks which use equal-cost multipath (ECMP) forwarding are one example where detailed statistics on flow distribution in the network are highly desired. If a network supports ECMP, one can create detailed statistics for the different paths packets take through the network at the egress system, without a need to correlate/aggregate statistics from every router in the system. Transit devices are off-loaded from the task of gathering packet statistics.

### **3.6. Frame Replication/Elimination Decision for Bi-casting/Active-active Networks**

Bandwidth- and power-constrained, time-sensitive, or loss-intolerant networks (e.g., networks for industry automation/control, health care) require efficient OAM methods to decide when to replicate packets to a secondary path in order to keep the loss/error-rate for the receiver at a tolerable level - and also when to stop replication and eliminate the redundant flow. Many IoT networks are time sensitive and cannot leverage automatic retransmission requests (ARQ) to cope with transmission errors or lost packets. Transmitting the data over multiple disparate paths (often called bi-casting or live-live) is a method used to reduce the error rate observed by the receiver. TSN receive a lot of attention from the manufacturing industry as shown by a various standardization activities and industry forums being formed (see e.g., IETF 6TiSCH, IEEE P802.1CB, AVnu).

### **3.7. Proof of Transit**

Several deployments use traffic engineering, policy routing, segment routing or Service Function Chaining (SFC) [[RFC7665](#)] to steer packets through a specific set of nodes. In certain cases regulatory obligations or a compliance policy require to prove that all packets that are supposed to follow a specific path are indeed being forwarded across the exact set of nodes specified. If a packet flow is supposed to go through a series of service functions or network nodes, it has to be proven that all packets of the flow actually went through the service chain or collection of nodes specified by the policy. In case the packets of a flow weren't appropriately processed, a verification device would be required to identify the policy violation and take corresponding actions (e.g., drop or redirect the packet, send an alert etc.) corresponding to the policy. In today's deployments, the proof that a packet traversed a particular service chain is typically delivered in an indirect way: Service appliances and network forwarding are in different trust



domains. Physical hand-off-points are defined between these trust domains (i.e., physical interfaces). Or in other terms, in the "network forwarding domain" things are wired up in a way that traffic is delivered to the ingress interface of a service appliance and received back from an egress interface of a service appliance. This "wiring" is verified and trusted. The evolution to Network Function Virtualization (NFV) and modern service chaining concepts (using technologies such as LISP, NSH, Segment Routing, etc.) blurs the line between the different trust domains, because the hand-off-points are no longer clearly defined physical interfaces, but are virtual interfaces. Because of that very reason, networks operators require that different trust layers not to be mixed in the same device. For an NFV scenario a different proof is required. Offering a proof that a packet traversed a specific set of service functions would allow network operators to move away from the above described indirect methods of proving that a service chain is in place for a particular application.

Deployed service chains without the presence of a "proof of transit" mechanism are typically operated as fail-open system: The packets that arrive at the end of a service chain are processed. Adding "proof of transit" capabilities to a service chain allows an operator to turn a fail-open system into a fail-close system, i.e. packets that did not properly traverse the service chain can be blocked.

A solution approach could be based on OAM data which is added to every packet for achieving Proof Of Transit. The OAM data is updated at every hop and is used to verify whether a packet traversed all required nodes. When the verifier receives each packet, it can validate whether the packet traversed the service chain correctly. The detailed mechanisms used for path verification along with the procedures applied to the OAM data carried in the packet for path verification are beyond the scope of this document. Details are addressed in [[draft-brockners-proof-of-transit](#)]. In this document the term "proof" refers to a discrete set of bits that represents an integer or string carried as OAM data. The OAM data is used to verify whether a packet traversed the nodes it is supposed to traverse.

### **3.8. Use Cases**

In-band OAM could be leveraged for several use cases, including:

- o Traffic Matrix: Derive the network traffic matrix: Traffic for a given time interval between any two edge nodes of a given domain. Could be performed for all traffic or per QoS-class.



- o Flow Debugging: Discover which path(s) a particular set of traffic (identified by an n-tuple) takes in the network. Such a procedure is particularly useful in case traffic is balanced across multiple paths, like with link aggregation (LACP) or equal cost multi-pathing (ECMP).
- o Loss Statistics per Path: Retrieve loss statistics per flow and path in the network.
- o Path Heat Maps: Discover highly utilized links in the network.
- o Trend Analysis on Traffic Patterns: Analyze if (and if so how) the forwarding path for a specific set of traffic changes over time (can give hints to routing issues, unstable links etc.).
- o Network Delay Distribution: Show delay distribution across network by node or links. If enabled per application or for a specific flow then display the path taken along with the delay incurred at every hop.
- o SLA Verification: Verify that a negotiated service level agreement (SLA), e.g., for packet drop rates or delay/jitter is conformed to by the actual traffic.
- o Low-power Networks: Include application level OAM information (e.g., battery charge level, cache or buffer fill level) into data traffic to avoid sending extra OAM traffic which incur an extra cost on the devices. Using the battery charge level as example, one could avoid sending extra OAM packets just to communicate battery health, and as such would save battery on sensors.
- o Path Verification or Service Function Path Verification: Proof and verification of packets traversing check points in the network, where check points can be nodes in the network or service functions.
- o Geo-location Policy: Network policy implemented based on which path packets took. Example: Only if packets originated and stayed within the trading-floor department, access to specific applications or servers is granted.

#### **4. Considerations for In-band OAM**

The implementation of an in-band OAM mechanism needs to take several considerations into account, including administrative boundaries, how information is recorded, Maximum Transfer Unit (MTU), Path MTU discovery and packet size, etc.





#### **4.1. Type of information to be recorded**

The information gathered for in-band OAM can be categorized into three main categories: Information with a per-hop scope, such as path tracing; information which applies to a specific set of nodes, such as path or service chain verification; information which only applies to the edges of a domain, such as sequence numbers.

- o "edge to edge": Information that needs to be shared between network edges (the "edge" of a network could either be a host or a domain edge device): Edge to edge data e.g., packet and octet count of data entering a well-defined domain and leaving it is helpful in building traffic matrix, sequence number (also called "path packet counters") is useful for the flow to detect packet loss.
- o "selected hops": Information that applies to a specific set of nodes only. In case of path verification, only the nodes which are "check points" are required to interpret and update the information in the packet.
- o "per hop": Information that is gathered at every hop along the path a packet traverses within an administrative domain:
  - \* Hop by Hop information e.g., Nodes visited for path tracing, Timestamps at each hop to find delays along the path
  - \* Stats collection at each hop to optimize communication in resource constrained networks e.g., Battery, CPU, memory status of each node piggy backed in a data packet is useful in low power lossy networks where network nodes are mostly asleep and communication is expensive

#### **4.2. MTU and packet size**

The recorded data at every hop may lead to packet size exceeding the Maximum Transmit Unit (MTU). Based on the transport protocol used MTU is discovered as a configuration parameter or Path MTU (PMTU) is discovered dynamically. Example: IPv6 recommends PMTU discovery before data packets are sent to prevent packet fragmentation. It specifies 1280 octets as the default PDU to be carried in a IPv6 datagram. A detailed discussion of the implications of oversized IPv6 header chains is found in [\[RFC7112\]](#).

The Path MTU restricts the amount of data that can be recorded for purpose of OAM within a data packet. The total size of data to be recorded needs to be preset to avoid packet size exceeding the MTU.



It is recommended to pre-calculate and configures network devices to limit the in-band OAM data that is attached to a packet.

#### **4.3. Administrative boundaries**

There are several challenges in enabling in-band OAM in the public Internet as well as in corporate/enterprise networks across administrative domains, which include but are not limited to:

- o Deployment dependent, the data fields that in-band OAM requires as part of a specific transport protocol may not be supported across administrative boundaries.
- o Current OAM implementations are often done in the slow path, i.e., OAM packets are punted to router's CPU for processing. This leads to performance and scaling issues and opens up routers for attacks such as Denial of Service (DoS) attacks.
- o Discovery of network topology and details of the network devices across administrative boundaries may open up attack vectors compromising network security.
- o Specifically on IPv6: At the administrative boundaries IPv6 packets with extension headers are dropped for several reasons described in [[RFC7872](#)].

The following considerations will be discussed in a future version of this document: If the packet is dropped due to the presence of the in-band OAM; If the policy failure is treated as feature disablement and any further recording is stopped but the packet itself is not dropped, it may lead to every node in the path to make this policy decision.

#### **4.4. Selective enablement**

Deployment dependent, in-band OAM could either be used for all, or only a subset of the overall traffic. While it might be desirable to apply in-band OAM to all traffic and then selectively use the data gathered in case needed, it might not always be feasible. Depending on the forwarding infrastructure used, in-band OAM can have an impact on forwarding performance. The SPUD prototype for example uses the notion of "pipes" to describe the portion of the traffic that could be subject to in-path inspection. Mechanisms to decide which traffic would be subject to in-band OAM are outside the scope of this document.



#### **4.5. Optimization of node and interface identifiers**

Since packets have a finite maximum size, the data recording or carrying capacity of one packet in which the in-band OAM meta data is present is limited. In-band OAM should use its own dedicated namespace (confined to the domain in-band OAM operates in) to represent node and interface IDs to save space in the header. Generic representations of node and interface identifiers which are globally unique (such as a UUID) would consume significantly more bits of in-band OAM data.

#### **4.6. Loop communication path (IPv6-specifics)**

When recorded data is required to be analyzed on a source node that issues a packet and inserts in-band OAM data, the recorded data needs to be carried back to the source node.

One way to carry the in-band OAM data back to the source is to utilize an ICMP Echo Request/Reply (ping) or ICMPv6 Echo Request/Reply (ping6) mechanism. In order to run the in-band OAM mechanism appropriately on the ping/ping6 mechanism, the following two operations should be implemented by the ping/ping6 target node:

1. All of the in-band OAM fields would be copied from an Echo Request message to an Echo Reply message.
2. The Hop Limit field of the IPv6 header of these messages would be copied as a continuous sequence. Further considerations are addressed in a future version of this document.

### **5. Requirements for In-band OAM Data Types**

The above discussed use cases require different types of in-band OAM data. This section details requirements for in-band OAM derived from the discussion above.

#### **5.1. Generic Requirements**

- REQ-G1: Classification: It should be possible to enable in-band OAM on a selected set of traffic. The selected set of traffic can also be all traffic.
- REQ-G2: Scope: If in-band OAM is used only within a specific domain, provisions need to be put in place to ensure that in-band OAM data stays within the specific domain only.
- REQ-G3: Transport independence: Data formats for in-band OAM shall be defined in a transport independent way. In-band OAM



applies to a variety of transport protocols. Encapsulations should be defined how the generic data formats are carried by a specific protocol.

- REQ-G4: Layering: It should be possible to have in-band OAM information for different transport protocol layers be present in several fields within a single packet. This could for example be the case when tunnels are employed and in-band OAM information is to be gathered for both the underlay as well as the overlay network.
- REQ-G5: MTU size: With in-band OAM information added, packets should not become larger than the path MTU.
- REQ-G6: Data Structure Reusability: The data types and data formats defined and used for in-band OAM ought to be reusable for out-of-band OAM telemetry as well.

## **5.2. In-band OAM Data with Per-hop Scope**

- REQ-H1: Missing nodes detection: Data shall be present that allows a node to detect whether all nodes that should participate in in-band OAM operations have indeed participated.
- REQ-H2: Node, instance or device identifier: Data shall be present that allows to retrieve the identity of the entity reporting telemetry information. The entity can be a device, or a subsystem/component within a device. The latter will allow for packet tracing within a device in much the same way as between devices.
- REQ-H3: Ingress interface identifier: Data shall be present that allows the identification of the interface a particular packet was received from. The interface can be a logical or physical entity.
- REQ-H4: Egress interface identifier: Data shall be present that allows the identification of the interface a particular packet was forwarded to. Interface can be a logical or physical entity.
- REQ-H5: Time-related requirements
- REQ-H5.1: Delay: Data shall be present that allows to retrieve the delay between two or more points of interest within the system. Those points can be within the same device or on different devices.





REQ-H5.2: Jitter: Data shall be present that allows to retrieve the jitter between two or more points of interest within the system. Those points can be within the same device or on different devices.

REQ-H5.3: Wall-clock time: Data shall be present that allows to retrieve the wall-clock time visited a particular point of interest in the system.

REQ-H5.4: Time precision: The precision of the time related data should be configurable. Use-case dependent, the required precision could e.g., be nano-seconds, micro-seconds, milli-seconds, or seconds.

REQ-H6: Generic data records (like e.g., GPS/Geo-location information): It should be possible to add user-defined OAM data at select hops to the packet. The semantics of the data are defined by the user.

### **5.3. In-band OAM with Selected Hop Scope**

REQ-S1: Proof of transit: Data shall be present which allows to securely prove that a packet has visited or ore several particular points of interest (i.e., a particular set of nodes).

REQ-S1.1: In case "Shamir's secret sharing scheme" is used for proof of transit, two data records, "random" and "cumulative" shall be present. The number of bits used for "random" and "cumulative" data records can vary between deployments and should thus be configurable.

REQ-S1.2: Enable a fail-open service chaining system to be converted into a fail-closed service chaining system.

### **5.4. In-band OAM with End-to-end Scope**

REQ-E1: Sequence numbering:

REQ-E1.1: Reordering detection: It should be possible to detect whether packets have been reordered while traversing an in-band OAM domain.



REQ-E1.2: Duplicates detection: It should be possible to detect whether packets have been duplicated while traversing an in-band OAM domain.

REQ-E1.3: Detection of packet drops: It should be possible to detect whether packets have been dropped while traversing an in-band OAM domain.

## **6. Security Considerations and Requirements**

General Security considerations will be addressed in a later version of this document. Security considerations for Proof of Transit alone are discussed below.

### **6.1. Proof of Transit**

Threat Model: Attacks on the deployments could be due to malicious administrators or accidental misconfigurations resulting in bypassing of certain nodes. The solution approach should meet the following requirements:

REQ-SEC1: Sound Proof of Transit: A valid and verifiable proof that the packet definitively traversed through all the nodes as expected. Probabilistic methods to achieve this should be avoided, as the same could be exploited by an attacker.

REQ-SEC2: Tampering of meta data: An active attacker should not be able to insert or modify or delete meta data in whole or in parts and bypass few (or all) nodes. Any deviation from the expected path should be accurately determined.

REQ-SEC3: Replay Attacks: A attacker (active/passive) should not be able to reuse the proof of transit bits in the packet by observing the OAM data in the packet, packet characteristics (like IP addresses, octets transferred, timestamps) or even the proof bits themselves. The solution approach should consider usage of these parameters for deriving any secrets cautiously. Mitigating replay attacks beyond a window of longer duration could be intractable to achieve with fixed number of bits allocated for proof.

REQ-SEC4: Recycle Secrets: Any configuration of the secrets (like cryptographic keys, initialisation vectors etc.) either in the controller or service functions should be reconfigurable. Solution approach should enable controls, API calls etc. needed in order to perform such recycling. It is desirable to provide recommendations on the duration



of rotation cycles needed for the secure functioning of the overall system.

REQ-SEC5: Secret storage and distribution: Secrets should be shared with the devices over secure channels. Methods should be put in place so that secrets cannot be retrieved by non authorized personnel from the devices.

## **7. IANA Considerations**

[RFC Editor: please remove this section prior to publication.]

This document has no IANA actions.

## **8. Acknowledgements**

The authors would like to thank Eric Vyncke, Nalini Elkins, Srihari Raghavan, Ranganathan T S, Karthik Babu Harichandra Babu, Akshaya Nadahalli, and Andrew Yourtchenko for the comments and advice. 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.

## **9. Informative References**

[[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.

[I-D.brockners-lisp-sr]

Brockners, F., Bhandari, S., Maino, F., and D. Lewis, "LISP Extensions for Segment Routing", [draft-brockners-lisp-sr-01](#) (work in progress), February 2014.

[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-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.lapukhov-dataplane-probe]  
Lapukhov, P. and r. remy@barefootnetworks.com, "Data-plane probe for in-band telemetry collection", [draft-lapukhov-dataplane-probe-01](#) (work in progress), June 2016.
- [P4] Kim, , "P4: In-band Network Telemetry (INT)", September 2015.
- [RFC0791] Postel, J., "Internet Protocol", STD 5, [RFC 791](#), DOI 10.17487/RFC0791, September 1981, <<http://www.rfc-editor.org/info/rfc791>>.
- [RFC4884] Bonica, R., Gan, D., Tappan, D., and C. Pignataro, "Extended ICMP to Support Multi-Part Messages", [RFC 4884](#), DOI 10.17487/RFC4884, April 2007, <<http://www.rfc-editor.org/info/rfc4884>>.
- [RFC4950] Bonica, R., Gan, D., Tappan, D., and C. Pignataro, "ICMP Extensions for Multiprotocol Label Switching", [RFC 4950](#), DOI 10.17487/RFC4950, August 2007, <<http://www.rfc-editor.org/info/rfc4950>>.
- [RFC5837] Atlas, A., Ed., Bonica, R., Ed., Pignataro, C., Ed., Shen, N., and JR. Rivers, "Extending ICMP for Interface and Next-Hop Identification", [RFC 5837](#), DOI 10.17487/RFC5837, April 2010, <<http://www.rfc-editor.org/info/rfc5837>>.
- [RFC7112] Gont, F., Manral, V., and R. Bonica, "Implications of Oversized IPv6 Header Chains", [RFC 7112](#), DOI 10.17487/RFC7112, January 2014, <<http://www.rfc-editor.org/info/rfc7112>>.
- [RFC7276] Mizrahi, T., Sprecher, N., Bellagamba, E., and Y. Weingarten, "An Overview of Operations, Administration, and Maintenance (OAM) Tools", [RFC 7276](#), DOI 10.17487/RFC7276, June 2014, <<http://www.rfc-editor.org/info/rfc7276>>.
- [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>>.
- [RFC7872] Gont, F., Linkova, J., Chown, T., and W. Liu, "Observations on the Dropping of Packets with IPv6 Extension Headers in the Real World", [RFC 7872](#), DOI 10.17487/RFC7872, June 2016, <<http://www.rfc-editor.org/info/rfc7872>>.





## Authors' Addresses

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

Email: 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

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

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