

SFC WG
Internet-Draft
Intended status: Standards Track
Expires: December 31, 2017

T. Ao
ZTE Corporation
G. Mirsky
ZTE Corp.
Z. Chen
China Telecom
June 29, 2017

SFC OAM for path consistency
draft-ao-sfc-oam-path-consistency-00

Abstract

Service Function Chain(SFC) defines an ordered set of service functions(SFs) to be applied to packets and/or frames and/or flows selected as a result of classification. SFC Operation, Administration and Maintenance can monitor the continuity of the SFC, i.e., that all elements of the SFC are reachable to each other in the downstream direction. But SFC OAM must support verification that the order of traversing these SFs corresponds to the state defined by the SFC control plane or orchestrator, the metric referred in this document as the path consistency of the SFC. This document defines a new SFC OAM method to support SFC consistency, i.e. verification that all elements of the given SFC are being traversed in the expected order.

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 31, 2017.

Copyright Notice

Copyright (c) 2017 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 used in this document	3
2.1. Terminology	3
2.2. Requirements Language	3
3. Consistency OAM: Theory of Operation	3
3.1. COAM packet	4
3.2. SF Sub-TLV	4
4. Security Considerations	5
5. IANA Considerations	5
5.1. COAM Message Types	5
5.2. SFF Information Record TLV Type	6
5.3. SF Information Sub-TLV Type	6
5.4. SF Types	6
5.5. SF Identifier Types	7
6. References	8
6.1. Normative References	8
6.2. Informational References	8
Authors' Addresses	8

1. Introduction

Service Function Chain (SFC) is a chain with a series of ordered Service Functions(SFs). Service Function Path (SFP) is a path of a SFC. SFC is described in detail in the SFC architecture document [RFC7665]. The SFs in the SFC are ordered and only when traffic is processed by one SF then it should be processed by the next SF, otherwise errors may occur. Sometimes, a SF needs to use the metadata from its upstream SF process. That's why it's very important for the operator to make sure that the order of traversing the SFs is exactly as defined by the control plane or the

orchestrator. This document refers to the correspondence between the state of control plane and the SFP itself as the SFP consistency.

This document defines the method to check the path consistency of the SFP. It is an extension of the Overlay Echo-Request/Echo-reply specified in the [I-D.ooamdt-rtgwg-demand-cc-cv].

2. Conventions used in this document

2.1. Terminology

SFC(Service Function Chain): An ordered set of some abstract SFs.

SFF: Service Function Forwarder

SF: Service Function

OAM: Operation, Administration and Maintenance

SFP: Service Function Path

COAM(Consistency OAM): OAM that can be used to check path consistency.

2.2. 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.

3. Consistency OAM: Theory of Operation

Consistency OAM uses two functions: COAM Request and COAM Reply. The SFF, that is ingress of the SFP, transmits COAM Request packet. Every intermediate SFF that receives the COAM Request MUST perform the following actions:

- collect information of traversed by the COAM Request packet SFs and send it to the ingress SFF as COAM Reply packet over IP network [I-D.wang-sfc-multi-layer-oam];

- forward the COAM Request to next downstream SFF if the one exists.

As result, the ingress SFF collects information about all traversed SFFs and SFs, information of the actual path the COAM packet has traveled, so that we can verify the path consistency of the SFC. The

mechanism for the SFP consistency verification is outside the scope of this document.

3.1. COAM packet

Consistency OAM introduces two new types of messages to the OOAM Echo Request/Reply operation [I-D.ooamdt-rtgwg-demand-cc-cv] with the following values Section 5.1:

- o TBA1 - COAM Request
- o TBA2 - COAM Reply

An SFF, upon receiving the Consistency OAM Request, MUST include the corresponding SFs information, Section 3.2, into the Value field of the COAM Reply packet.

The COAM packet is displayed in Figure 1.

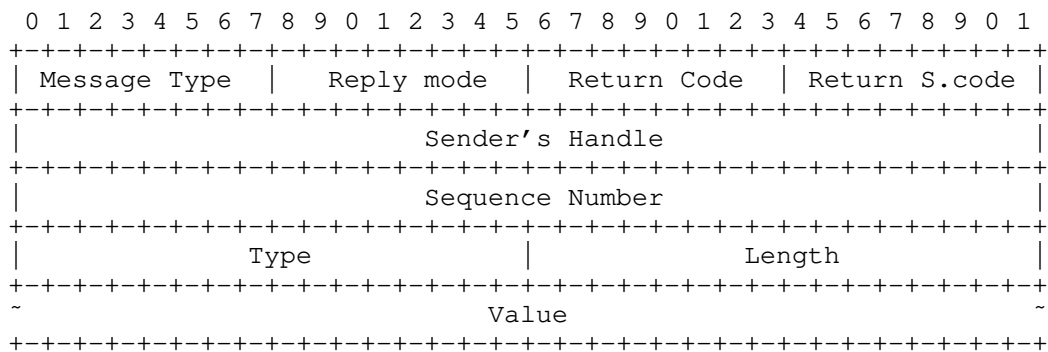


Figure 1: COAM Packet Header

3.2. SF Sub-TLV

Every SFF receiving COAM Request packet MUST include the SF characteristic data into the COAM Reply packet. The per SF data included in COAM Reply packet as SF Information sub-TLV that is displayed in Figure 2.

After the COAM traversed the SFP, all the information of the SFs on the SFP are collected in the TLVs with COAM Reply.

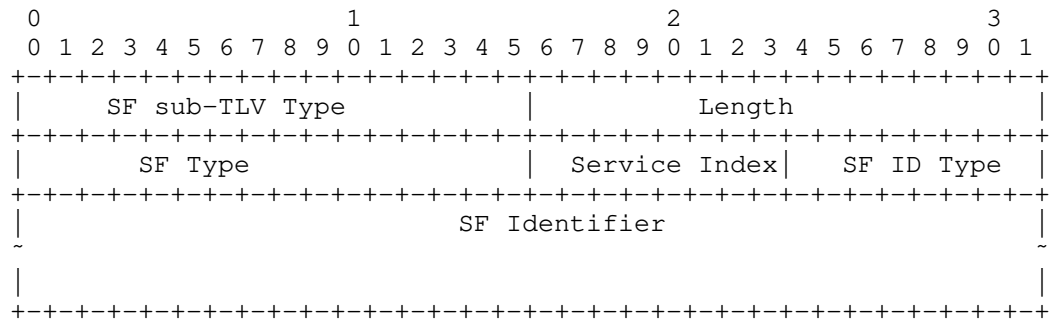


Figure 2: Service Function sub-TLV

SF TLV Type: indicate that the TLV is a SF TLV which contains the information of one SF.

SF Type: indicates the type of SF, e.g., Firewall, Deep Packet Inspection, WAN optimization controller, etc.

Service Index: indicates the SF's position on the SFP.

SF ID Type:

0x01: IPv4

0x02: IPv6

0x03: MAC address

0x04-0xFF: Reserved

SF Identifier: An identifier of the SF. The length of the SF Identifier depends on the type of the SF ID Type. For example, if the SF Identifier is its IPv4 address, the SF Identifier should be 32 bits.

4. Security Considerations

Will be added in the future updates.

5. IANA Considerations

5.1. COAM Message Types

IANA is requested to assign values from its Message Types sub-registry in Overlay Echo Request/Echo Reply Message Types registry as follows:

Value	Description	Reference
TBA1	SFP Consistency Echo Request	This document
TBA2	SFP Consistency Echo Reply	This document

Table 1: SFP Consistency Echo Request/Echo Reply Message Types

5.2. SFF Information Record TLV Type

IANA is requested to assign new type value from SFC OAM TLV Type registry as follows:

Value	Description	Reference
TBA3	SFF Information Record Type	This document

Table 2: SFF-Information Record

5.3. SF Information Sub-TLV Type

IANA is requested to assign new type value from SFC OAM TLV Type registry as follows:

Value	Description	Reference
TBA4	SF Information	This document

Table 3: SF-Information Sub-TLV Type

5.4. SF Types

IANA is requested create in the registry SF Types. All code points in the range 1 through 32759 in this registry shall be allocated according to the "IETF Review" procedure as specified in [RFC5226]. Code points in the range 32760 through 65279 in this registry shall be allocated according to the "First Come First Served" procedure as specified in [RFC5226]. Remaining code points are allocated according to the Table 4:

Value	Description	Reference
0	Reserved	This document
1- 32759	Unassigned	IETF Review
32760 - 65279	Unassigned	First Come First Served
65280 - 65519	Experimental	This document
65520 - 65534	Private Use	This document
65535	Reserved	This document

Table 4: SF Type Registry

This document defines the following new value in SF Type registry:

Value	Description	Reference
TBA5	Firewall	This document

Table 5: SF Types

5.5. SF Identifier Types

IANA is requested create in the registry SF Types the new sub-registry SF Identifier Types. All code points in the range 1 through 191 in this registry shall be allocated according to the "IETF Review" procedure as specified in [RFC5226] and assign values as follows:

Value	Description	Reference
0	Reserved	This document
TBA6	IPv4	This document
TBA7	IPv6	This document
TBA8	MAC	This document
TBA8+1-191	Unassigned	IETF Review
192-251	Unassigned	First Come First Served
252-254	Unassigned	Private Use
255	Reserved	This document

Table 6: SF Identifier Type

6. References

6.1. Normative References

- [I-D.ooamdt-rtgwg-demand-cc-cv]
Mirsky, G., Kumar, N., Kumar, D., Chen, M., Yizhou, L.,
and D. Dolson, "Echo Request and Echo Reply for Overlay
Networks", draft-ooamdt-rtgwg-demand-cc-cv-03 (work in
progress), March 2017.
- [I-D.wang-sfc-multi-layer-oam]
Mirsky, G., Meng, W., Khasnabish, B., and C. Wang, "Multi-
Layer OAM for Service Function Chains in Networks", draft-
wang-sfc-multi-layer-oam-09 (work in progress), June 2017.
- [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>>.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an
IANA Considerations Section in RFCs", RFC 5226,
DOI 10.17487/RFC5226, May 2008,
<<http://www.rfc-editor.org/info/rfc5226>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
May 2017, <<http://www.rfc-editor.org/info/rfc8174>>.

6.2. Informational References

- [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

Ting Ao
ZTE Corporation
No.889, BiBo Road
Shanghai 201203
China

Phone: +86 21 68897642
Email: ao.ting@zte.com.cn

Greg Mirsky
ZTE Corp.
1900 McCarthy Blvd. #205
Milpitas, CA 95035
USA

Email: gregimirsky@gmail.com

Zhonghua Chen
China Telecom
No.1835, South PuDong Road
Shanghai 201203
China

Phone: +86 18918588897
Email: 18918588897@189.cn

SFC WG
Internet-Draft
Intended status: Standards Track
Expires: December 31, 2017

T. Ao
ZTE Corporation
G. Mirsky
ZTE Corp.
Z. Chen
China Telecom
June 29, 2017

Controlled Return Path for Service Function Chain (SFC) OAM
draft-ao-sfc-oam-return-path-specified-00

Abstract

This document defines extensions to the Service Function Chain (SFC) Operation, Administration and Maintenance (OAM) that enable control of the Echo Reply return path by specifying it as Reverse Service Function Path. Enforcing the specific return path can be used to verify bidirectional connectivity of SFC and increase robustness of SFC OAM.

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 31, 2017.

Copyright Notice

Copyright (c) 2017 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 used in this document	3
2.1. Terminology	3
2.2. Requirements Language	3
3. Extension	3
4. SFC Reply Path TLV	4
5. Theory of Operation	5
5.1. Case of Bi-directional SFC	5
6. Security Considerations	5
7. IANA Considerations	5
7.1. SFC Return Path Type	6
7.2. New Return Codes	6
8. References	6
8.1. Normative References	6
8.2. Informative References	7
Authors' Addresses	7

1. Introduction

While Echo Request, defined in [I-D.ooamdt-rtgwg-demand-cc-cv], always traverses the Service Function Chain (SFC) it directed to, the corresponding Echo Reply is sent over IP network [I-D.wang-sfc-multi-layer-oam]. There are scenarios when it is beneficial to direct the responder to use path other than the IP network. This document defines extensions to the Service Function Chain (SFC) Operation, Administration and Maintenance (OAM) that enable control of the Echo Reply return path by specifying it as Reply Service Function Path. This document defines a new Type-Length-Value (TLV), Reply Service Function Path TLV, for Reply via Specified Path mode of Overlay Echo Reply (Section 4).

The Reply Service Function Path TLV provides efficient mechanism to test bidirectional and hybrid SFCs, as these were defined in Section 2.2 [RFC7665], that allows an operator to test both directions of the bidirectional or hybrid SFP with a single Overlay Echo Request/Echo Reply operation.

2. Conventions used in this document

2.1. Terminology

SF - Service Function

SFF - Service Function Forwarder

SFC - Service Function Chain, an ordered set of some abstract SFs.

SFP - Service Function Path

SPI - Service Path Index

OAM - Operation, Administration, and Maintenance

2.2. 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.

3. Extension

Following reply modes had been defined in [I-D.ooamdt-rtgwg-demand-cc-cv]:

- o Do Not Reply
- o Reply via an IPv4/IPv6 UDP Packet
- o Reply via Application Level Control Channel
- o Reply via Specified Path

The Reply via Specified Path mode is intended to enforce use of the particular return path specified in the included TLV. This mode may help to verify bidirectional continuity or increase robustness of the monitoring of the SFC by selecting more stable path. In case of SFC, the sender of Echo Request instructs the egress SFF to send Echo Reply message along the SFP specified in the SFC Reply Path TLV Section 4.

4. SFC Reply Path TLV

The SFC Reply Path TLV carries the information that sufficiently identifies the return SFP that the Overlay Echo Reply message is expected to follow. The format of SFC Reply Path TLV is display in Figure 1.

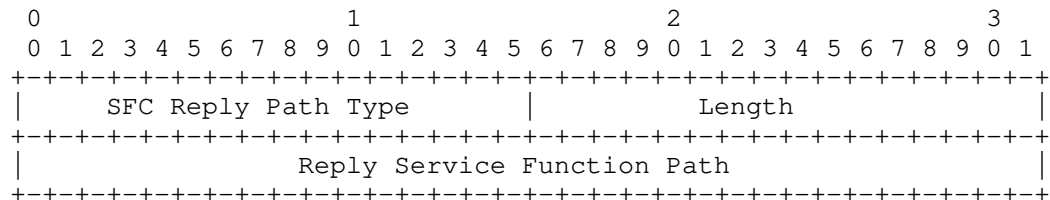


Figure 1: SFC Reply TLV Format

where:

- o Reply Path TLV Type: is 2 octets long, indicates the TLV that contains a information about the SFC Reply path.
- o Length: is 2 octets long, MUST be equal to 4
- o Reply Service Function Path is used to describe the return path that an Overlay Echo Reply is requested to follow.

The format of the Reply Service Function Path field displayed in Figure 2

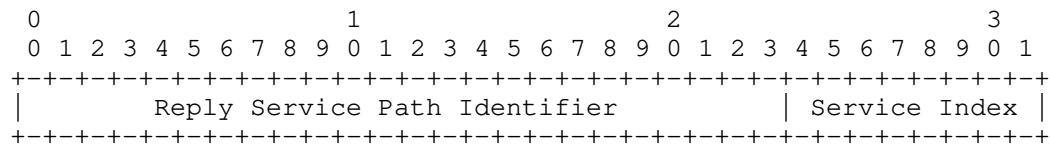


Figure 2: Reply Service Function Path Field Format

where:

- o Reply Service Path Identifier: is SFP identifier for the path that the Overlay Echo Reply message is requested to be sent over.
- o Service Index: used for forwarding in the reply SFP.

5. Theory of Operation

[RFC7110] defined mechanism to control return path for MPLS LSP Echo Reply. In case of SFC, the return path is a SFP along which Overlay Echo Reply message MUST be transmitted. Hence, the SFC Reply Path TLV included in the Overlay Echo Request message MUST sufficiently identify the SFP that the sender of the Echo Request message expects the receiver to use for the corresponding Overlay Echo Reply.

When sending an Echo Request the sender MUST set the value of Reply Mode field to "Reply via Specified Path", defined in [I-D.ooamdt-rtgwg-demand-cc-cv], and MUST include SFC Reply Path TLV. The SFC Reply Path TLV includes identifier of the reverse SFP and an appropriate Service Index.

Echo Reply is expected to be sent by the egress SFF of the SFP being tested or by the SFF at which SFC TTL expires as defined [I-D.ietf-sfc-nsh]. Processing described below equally applies in both cases and referred as responding SFF.

If the Echo Request message with SFC Reply Path TLV, received by the responding SFF, has Reply Mode value of "Reply via Specified Path" but no SFC Reply Path TLV is present, then the responding SFF MUST send Echo Reply with Return Code set to "Reply Path TLV is missing" value (TBA2). If the responding SFF cannot find requested SFP it MUST send Echo Reply with Return Code set to "Reply SFP was not found" and include the SFC Reply Path TLV from the Echo Request message.

5.1. Case of Bi-directional SFC

Ability to specify the return path to be used for Echo Reply is very useful in bi-directional SFC. For bi-directional SFC, since the last SFF of the forward SFP may not co-locate with classifier of the reverse SFP, it is assumed that last SFF doesn't know the reply path of a SFC. So even for bi-directional SFC, a reverse SFP also need to be indicated in reply path TLV in echo request message.

6. Security Considerations

Will be added in the future updates.

7. IANA Considerations

7.1. SFC Return Path Type

IANA is requested to assign from its Overlay Echo Request/Echo Reply TLV registry new type as following:

Value	Description	Reference
TBA1	SFC Reply Path Type	This document

Table 1: SFC Return Path Type

7.2. New Return Codes

IANA is requested to assign new return codes from the Overlay Echo Request/Echo Reply Return Codes registry as following:

Value	Description	Reference
TBA2	Reply Path TLV is missing	This document
TBA3	Reply SFP was not found	This document

Table 2: SFC Overlay Echo Reply Return Codes

8. References

8.1. Normative References

[I-D.ietf-sfc-nsh]

Quinn, P. and U. Elzur, "Network Service Header", draft-ietf-sfc-nsh-12 (work in progress), February 2017.

[I-D.ooamdt-rtgwg-demand-cc-cv]

Mirsky, G., Kumar, N., Kumar, D., Chen, M., Yizhou, L., and D. Dolson, "Echo Request and Echo Reply for Overlay Networks", draft-ooamdt-rtgwg-demand-cc-cv-03 (work in progress), March 2017.

[I-D.wang-sfc-multi-layer-oam]

Mirsky, G., Meng, W., Khasnabish, B., and C. Wang, "Multi-Layer OAM for Service Function Chains in Networks", draft-wang-sfc-multi-layer-oam-09 (work in progress), June 2017.

- [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>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<http://www.rfc-editor.org/info/rfc8174>>.

8.2. Informative References

- [RFC7110] Chen, M., Cao, W., Ning, S., Jounay, F., and S. Delord, "Return Path Specified Label Switched Path (LSP) Ping", RFC 7110, DOI 10.17487/RFC7110, January 2014, <<http://www.rfc-editor.org/info/rfc7110>>.
- [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

Ting Ao
ZTE Corporation
No.889, BiBo Road
Shanghai 201203
China

Phone: +86 21 68897642
Email: ao.ting@zte.com.cn

Greg Mirsky
ZTE Corp.
1900 McCarthy Blvd. #205
Milpitas, CA 95035
USA

Email: gregimirsky@gmail.com

Zhonghua Chen
China Telecom
No.1835, South PuDong Road
Shanghai 201203
China

Phone: +86 18918588897
Email: 18918588897@189.cn

SFC Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 31, 2017

A. Farrel
J. Drake
Juniper Networks
June 29, 2017

Operating the Network Service Header (NSH) with Next Protocol "None"
draft-farrel-sfc-convent-02

Abstract

This document describes the use of the Network Service Header (NSH) in a Service Function Chaining (SFC) enabled network with no payload data and carrying only metadata. This is achieved by defining a new NSH "Next Protocol" type value of "None".

This document illustrates some of the functions that may be achieved or enhanced by this mechanism, but it does not provide an exhaustive list of use cases, nor is it intended to be definitive about the functions it describes. It is expected that other documents will describe specific use cases in more detail and will define the protocol mechanics for each use case.

Requirements Language

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].

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 31, 2017.

Copyright Notice

Copyright (c) 2017 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. The Network Service Header	3
2.1. Next Protocol 'None'	4
3. Processing Rules	4
4. Backward Compatibility	5
5. Overview of Use Cases	6
5.1. Per-SFP Metadata	6
5.2. Per-Flow Metadata	6
5.3. Coordination Between SFC-Aware SFIs	6
5.4. Operations, Administration, and Maintenance (OAM)	7
5.5. Control Plane and Management Plane Uses	8
5.6. Non-Applicable Use Cases	8
6. Security Considerations	8
7. IANA Considerations	9
8. Contributors	9
9. Acknowledgements	9
10. References	9
10.1. Normative References	9
10.2. Informative References	10
Authors' Addresses	10

1. Introduction

An architecture for Service Function Chaining (SFC) is presented in [RFC7665]. That architecture enables packets to be forwarded along Service Function Paths (SFPs) to pass through various Service Functions (SFs) that act on the packets. Each packet is encapsulated with a Network Service Header (NSH) [I-D.ietf-sfc-nsh] identify the SFP that the packet travels along (by means of a Service Path Identifier - SPI) and the hop (i.e., the next SF to be executed)

along the SFP that the packet has reached (by means of a Service Index - SI). The SPI and SI are fields encoded in the NSH.

Packets are classified at the SFC ingress boundaries (section 4.4 of [RFC7665]) and have an NSH applied to them. Such packets are forwarded between Service Function Forwarders (SFFs) using tunnels across the underlay network, and each SFF may hand the packet off to one or more Service Function Instances (SFIs) according to the definition of the SFP.

The SFC classifier or any SFC-aware SFI may wish to share information (possibly state information) about the SFP, the traffic flow, or a specific packet, and they may do this by adding "metadata" to packets as part of the NSH. Metadata may be used to enhance or enable the function performed by SFC-aware SFs, may enable coordination and data exchange between SFIs, or may be used to assist a network operator in the diagnosis and monitoring of an SFP. The nature of metadata to be supplied and consumed is implementation- and deployment-specific.

This document defines a mechanism for metadata to be carried on an SFP without the need for payload data. This may enable diagnosis and monitoring of SFPs, and coordination between SFC-aware SFIs, without the need for traffic to be flowing, and without the need to rewrite data packets to insert what might be substantial amounts of metadata.

This function is achieved by defining a new value for the NSH "Next Protocol" field to indicate "None". Such packets are contained within the SFC-enabled domain.

This document illustrates some of the functions that may be achieved or enhanced by this mechanism, but it does not provide an exhaustive list of use cases, nor is it intended to be definitive about the functions it describes. It is expected that other documents will describe specific use cases in more detail and will define the protocol mechanics for each use case.

2. The Network Service Header

The NSH is defined in [I-D.ietf-sfc-nsh]. It includes a field called "Next Protocol" that is used to indicate the nature of the payload data that follows the NSH. The field can be used by any component that processes the NSH (for example, to understand how to interpret and parse the payload) and by nodes at the end of the SFP that remove the NSH and forward the payload data.

2.1. Next Protocol 'None'

This document defines a new value for the "Next Protocol" field. When set to TBD1, the field indicates that the next protocol is "None" meaning that there is no user/payload data following the NSH.

When the next protocol is "None" the rest of the NSH still has meaning and, in particular, the metadata carried in the NSH may still be present.

3. Processing Rules

An SFC-aware node wishing to send metadata without a data packet:

- o MUST create a packet carrying an NSH and the desired metadata
- o MUST set the "Next Protocol" field to TBD1
- o SHOULD ensure that there are no bytes following the end of the NSH (i.e., that there is no payload data)
- o MUST encapsulate and send the packet as normal for tunneling to the next hop on the SFP as normal for an NSH packet.

A packet with no payload data may be simply inserted at the head end of an SFP (such as a Classifier) and may be easily forwarded by an SFF or SFI on the SFP using the normal processing rules defined in [I-D.ietf-sfc-nsh].

A packet with no payload may also be generated by an SFC-aware SFI as a result of processing an incoming packet (i.e., triggered by a condition arising from processing a normal NSH packet with a payload). In such cases, the SPI/SI can be inherited from the original packet or can be set according to information supplied through the control plane or management plane. This document does not further specify the triggers to generate an NSH packet with a "Next Protocol" set to "None".

A transit node (SFF, SFI, or classifier) receiving a packet with "Next Protocol" indicating "None" MUST NOT attempt to parse or process beyond the end of the NSH, but can process the NSH and especially the metadata as normal.

A node that is the egress of an SFP would normally strip the NSH and forward the payload according to the setting of the "Next Protocol" field. Such nodes MUST NOT forward packets with "Next Protocol" indicating "None" even if there some bytes after the NSH.

4. Backward Compatibility

This section describes procedures for default handling on unknown "Next Protocol" field values. This material updates the procedures described in [I-D.ietf-sfc-nsh] and may be transferred to that document.

SFC-aware nodes that do not understand the meaning of a value contained in the "Next Protocol" field of the NSH are unable to parse the payload. Such nodes are not obliged to discard the packet unless they are specifically called upon to be able to examine the payload.

Thus:

- o Transit SFFs will normally not inspect the "Next Protocol" field or the packet payload and will forward the packets based solely on the SPI/SI
- o An SFC Proxy must not pass to an SFI a packet of type where it cannot indicate the packet type to the SFI
- o An SFC Proxy must not pass to an SFI a packet of type that the SFI does not support
- o An SFC Proxy should not return to the SFF a packet it has not passed to the SFI
- o An SFI should not return to the SFF a packet it hasn't processed unless local policy defines "process" for this SF to mean "do not process" in this case.
- o Reclassifiers would normally require to understand the payload packet type, but it is possible to imagine reclassifiers that take action based on unknown values of the "Next Protocol" field or that perform protocol-independent actions (such as hashing the whole packet).

All this means that legacy SFC-aware nodes that are unaware of the meaning of the "Next Protocol" value "None" will act as follows:

- o SFFs will forward the packets
- o SFC Proxies will drop the packets
- o SFIs will most likely drop the packets
- o Reclassifiers will most likely drop the packets

SFC-aware nodes at the end of an SFP possibly forward packets with no knowledge of the payload in a "pop and forward" form of processing where the NSH is removed and the packet is simply put on an interface and the payload protocol is known a priori (or assumed). It is a general processing rule for all forwarders that they SHOULD NOT attempt to send packets with zero length, and since packets with the NSH "Next Protocol" set "None" are expected to have zero payload length.

5. Overview of Use Cases

5.1. Per-SFP Metadata

Per-SFP metadata is metadata that applies to an SFP and any data packets on that SFP. It does not need to be transmitted with every packet, but can be installed at the SFIs on the SFP and applied to all packets on the SFP.

Per-SFP metadata may be sent along the path of an SFP simply by setting the correct SPI in the NSH, and setting the SI to the correct value for the hop of the SFP at which the metadata is to be introduced. Classifiers and reclassifiers will know the correct SI values to use from information supplied by the control or management plane as is the case for NSH packets with payload data.

5.2. Per-Flow Metadata

Per-flow metadata is metadata that applies to a subset of the packets on an SFP, such as packets matching a particular 5 tuple of source address, destination address, source port, destination port, and payload protocol. This metadata also does not need to be transmitted with every packet, but can be installed at the SFIs on the SFP and applied to the packets that match the flow description.

If there is just one flow on an SFP then there is no difference between per-flow metadata and per-SFP metadata as described in .

In normal processing, the flow to which per-flow metadata applies can be deduced by looking at the payload data in the context of the value of the "Next Protocol" field. However, when "Next Protocol" indicates "None" this cannot be done. In this case the identity of the flow is carried in the metadata.

5.3. Coordination Between SFC-Aware SFIs

A pair of SFC-aware SFIs (adjacent or not) on an SFP may desire to coordinate state and may do this by sending information encoded in metadata.

To do this using the mechanisms defined in this document:

- o There must be an SFP that passes through the two SFIs in the direction of sender to receiver
- o The sender must know the correct SPI to use
- o The sender must know the correct SI to use for the point at which it resides on the SFP
- o Ideally the receiver will know to remove the packet from the SFP and not forward it further as this might share metadata wider than desirable and would cause unnecessary packets in the network. Note, however, that continued forwarding of such packets would not be substantially harmful in its own right.

Note that technically (according to the SFC architecture) the process of inserting a packet into an SFP is performed by a Classifier. However, a Classifier may be co-resident with an SFI so an implementation of an SF may also be able to generate NSH packets as described in this document.

Note also that a system with SFIs that need to coordinate between each other may be configured so that there is a specific, dedicated SFP between those service functions that is used solely for this purpose. Thus, such an SFI does not need to insert NSH packets onto SFPs used to carry payload data, but can use (and know the SPI of) this special, dedicated SFP.

5.4. Operations, Administration, and Maintenance (OAM)

Requirements for Operations, Administration, and Maintenance (OAM) in SFC networks are discussed in [I-D.ietf-sfc-oam-framework]. The NSH definition in [I-D.ietf-sfc-nsh] includes an O-bit that indicates that packet contains OAM information.

Since OAM information will be carried in packets that also include payload data, that information must be carried in metadata. Therefore, the mechanism defined in this document can be used to carry OAM information independent of payload data.

Sending OAM separate from (but interleaved with) packets that carry payload data may have several advantages including:

- o Sending OAM when there is no other traffic flowing.
- o Sending OAM at predictable intervals.

- o Measuring path qualities distinct from behavior of SFIs.
- o Sending OAM without needing to rewrite payload data buffers.
- o Keeping OAM processing components separate from other processing components.

5.5. Control Plane and Management Plane Uses

As described in Section 5.3, SFPs can be established specifically to carry metadata-only packets. And as described in Section 5.1, metadata-only packets can be sent down existing SFPs. This means that metadata-only packets can be used to carry control plane and management plane messages used to control and manage the SFC network.

In effect, SFPs can be established to serve as a Data Control Network (DCN) or Management Control Network (MCN). Further details of this process are out of scope of this document, but it should be understood that, just as for OAM, an essential feature of using a control channel is that the various speakers are assigned identifiers (i.e., addresses). In this case, those identifiers could be SPI/SI pairs, or could be IP addresses as used in the normal control and management plane of the SFC network.

5.6. Non-Applicable Use Cases

Per-packet metadata is metadata that applies specifically to a payload packet. It informs an SFI how to handle the payload packet, and does not apply to any other packets.

The mechanisms described in this document are not applicable to per-packet metadata because, by definition, if the "Next Protocol" indicates "None" then there is no packet following the NSH for the metadata to be associated with.

6. Security Considerations

Metadata-only packets as enabled by this document provide a covert channel. However, this is only different from the metadata feature in the normal NSH in that it can be sent without the presence of a data flow.

Metadata may, of course, contain sensitive data and may also contain information used to control the behavior of SFIs in the network. As such, this data needs to be protected according to its value and according to the perceived vulnerabilities of the network. Protection of metadata may be achieved by using encrypted transport between SFC entities or by encrypting the metadata in its own right.

The need to protect the metadata is not modified by this document and forms part of the NSH definition found in [I-D.ietf-sfc-nsh].

The mechanism described in this document might possibly be used to introduce packets into the SFC overlay network. Therefore measures SHOULD be taken to ensure authorization of sources of such packets, and tunneling of such packets into the network SHOULD be prevented. The amount of packets with "Next Protocol" set to "None" on an SFP MAY be rate limited at any point on the SFP to provide additional security.

Further discussion of NSH security is presented in [I-D.ietf-sfc-nsh].

7. IANA Considerations

IANA has been requested to create a registry of "Next Protocol" values in [I-D.ietf-sfc-nsh]. This document requests IANA to allocate a value from that registry to indicate "None" (TBD1 in this document).

It is strongly suggested that a value of 0 (zero) be assigned.

8. Contributors

Lucy Yong
Retired

9. Acknowledgements

Thanks to the attendees at the SFC interim meeting in Westford in January 2017 for discussions that suggested the value of this document.

Thanks to Eric Rosen and Med Boucadair for valuable review comments.

10. References

10.1. Normative References

[I-D.ietf-sfc-nsh]

Quinn, P. and U. Elzur, "Network Service Header", draft-ietf-sfc-nsh-12 (work in progress), February 2017.

- [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>>.

10.2. Informative References

- [I-D.ietf-sfc-oam-framework] Aldrin, S., Krishnan, R., Akiya, N., Pignataro, C., and A. Ghanwani, "Service Function Chaining Operation, Administration and Maintenance Framework", draft-ietf-sfc-oam-framework-01 (work in progress), February 2016.
- [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

Adrian Farrel
Juniper Networks

Email: afarrel@juniper.net

John Drake
Juniper Networks

Email: jdrake@juniper.net

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: January 4, 2018

S. Aldrin
Google
C. Pignataro, Ed.
N. Kumar, Ed.
Cisco
N. Akiya
Big Switch Networks
R. Krishnan
A. Ghanwani
Dell
July 3, 2017

Service Function Chaining
Operation, Administration and Maintenance Framework
draft-ietf-sfc-oam-framework-02

Abstract

This document provides reference framework for Operations, Administration and Maintenance (OAM) for Service Function Chaining (SFC).

Requirements Language

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 RFC 2119 [RFC2119].

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 4, 2018.

Copyright Notice

Copyright (c) 2017 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	3
1.1. Document Scope	3
2. SFC Layering Model	4
3. SFC OAM Components	4
3.1. Service Function Component	5
3.1.1. Service Function Availability	5
3.1.2. Service Function Performance Measurement	6
3.2. Service Function Chain Component	6
3.2.1. Service Function Chain Availability	6
3.2.2. Service Function Chain Performance Measurement	7
3.3. Classifier Component	7
4. SFC OAM Functions	7
4.1. Connectivity Functions	7
4.2. Continuity Functions	8
4.3. Trace Functions	8
4.4. Performance Measurement Function	9
5. Gap Analysis	9
5.1. Existing OAM Functions	9
5.2. Missing OAM Functions	10
5.3. Required OAM Functions	10
6. SFC OAM Model	11
6.1. SFC OAM packet Marker	11
6.2. OAM packet processing and forwarding semantic	11
6.3. OAM Function Types	12
6.4. OAM toolset applicability	12
6.4.1. ICMP Applicability	12
6.4.2. Seamless BFD Applicability	12
6.4.3. In-Situ OAM	13
6.4.4. SFC Traceroute	13
6.5. Security Considerations	13
6.6. IANA Considerations	14

6.7. Acknowledgements	14
7. References	14
7.1. Normative References	14
7.2. Informative References	15
Authors' Addresses	16

1. Introduction

Service Function Chaining (SFC) enables the creation of composite services that consist of an ordered set of Service Functions (SF) that are to be applied to packets and/or frames selected as a result of classification. Service Function Chaining is a concept that provides for more than just the application of an ordered set of SFs to selected traffic; rather, it describes a method for deploying SFs in a way that enables dynamic ordering and topological independence of those SFs as well as the exchange of metadata between participating entities. The foundations of SFC are described in the following documents:

- o SFC Problem Statement [RFC7498]
- o SFC Architecture [RFC7665]

The reader is assumed to be familiar with the material in these documents.

This document provides reference framework for Operations, Administration and Maintenance (OAM, [RFC6291]) of SFC. Specifically, this document provides:

- o In Section 2, an SFC layering model;
- o In Section 3, aspects monitored by SFC OAM;
- o In Section 4, functional requirements for SFC OAM;
- o In Section 5, a gap analysis for SFC OAM.

1.1. Document Scope

The focus of this document is to provide an architectural framework for SFC OAM, particularly focused on the aspect of the Operations component within OAM. Actual solutions and mechanisms are outside the scope of this document.

2. SFC Layering Model

Multiple layers come into play for implementing the SFC. These include the service layer at SFC layer and the underlying Network, Transport, Link, etc., layers.

- o The service layer, refer to as the "Service Layer" in Figure 1, consists of classifiers and service functions, and uses the overlay network reach from a classifier to service functions and service functions to service functions.
- o The overlay network layer, refer to as the "Network" in Figure 1, extends in between various service functions and is mostly transparent to the service functions. It leverages various overlay network technologies interconnecting service functions and allows establishing of service function paths.
- o The underlay network layer, refer to as the "Transport" in Figure 1, is dictated by the networking technology of the PSN. It may be either based on MPLS LSPs or IP.
- o The link layer, refer to as the "Link" in Figure 1, is dependent upon the physical technology used. Ethernet is a popular choice for this layer, but other alternatives are deployed (e.g. POS, DWDM etc...).

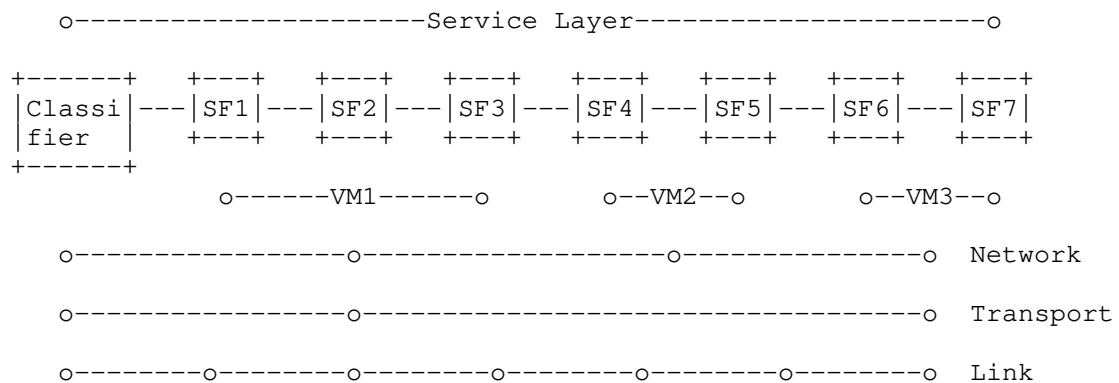


Figure 1: SFC Layering Example

3. SFC OAM Components

The SFC operates at the service layer. For the purpose of defining the OAM framework, the service layer is broken up into three distinct components.

1. Service function component: A function providing a specific service. OAM solutions for this component are to test the service functions from any SFC aware network devices (i.e. classifiers, controllers, other service nodes).
2. Service function chain component: An ordered set of service functions. OAM solution for this component are to test the service function chains and the service function paths.
3. Classifier component: A policy that describes the mapping from flows to service function chains. OAM solutions for this component are to test the validity of the classifiers.

Below figure illustrates an example where OAM for the three defined components are used within the SFC environment.

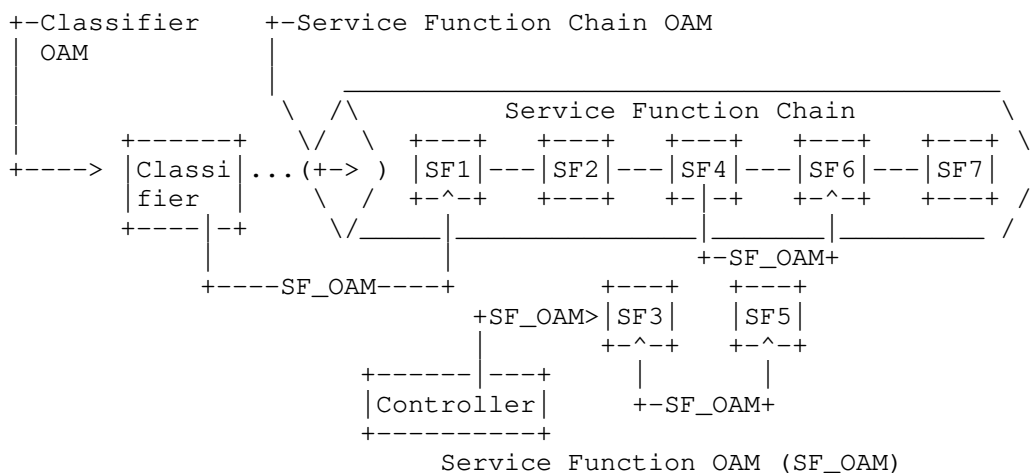


Figure 2: SFC OAM for Three Components

It is expected that multiple SFC OAM solutions will be defined, many targeting one specific component of the service layer. However, it is critical that SFC OAM solutions together provide the coverage of all three SFC OAM components: the service function component, the service function chain component and the classifier component.

3.1. Service Function Component

3.1.1. Service Function Availability

One SFC OAM requirement for the service function component is to allow an SFC aware network device to check the availability to a specific service function, located on the same or different network

devices. Service function availability is an aspect which raises an interesting question. How does one determine that a service function is available? On one end of the spectrum, one might argue that a service function is sufficiently available if the service node (physical or virtual) hosting the service function is available and is functional. On the other end of the spectrum, one might argue that the service function availability can only be concluded if the packet, after passing through the service function, was examined and verified that the packet got expected service applied.

The former approach will likely not provide sufficient confidence to the actual service function availability, i.e. a service node and a service function are two different entities. The latter approach is capable of providing an extensive verification, but comes with a cost. Some service functions make direct modifications to packets, while other service functions do not make any modifications to packets. Additionally, purpose of some service functions is to, conditionally, drop packets intentionally. In such case, packets will not be coming out from the service function. The fact is that there are many flavors of service functions available, and many more flavors of service functions will likely be introduced in future. Even a given service function may introduce a new functionality within a service function (ex: a new signature in a firewall). The cost of this approach is that verifier functions will need to be continuously modified to "keep up" with new services coming out: lack of extendibility.

This framework document provides a RECOMMENDED architectural model where generalized approach is taken to verify that a service function is sufficiently available. TBD - details will be provided in a later revision.

3.1.2. Service Function Performance Measurement

Second SFC OAM requirement for the service function component is to allow an SFC aware network device to check the loss and delay of a specific service function, located on the same or different network devices. TBD - details will be provided in a later revision.

3.2. Service Function Chain Component

3.2.1. Service Function Chain Availability

Verifying an SFC is a complicated process as the SFC could be comprised of varying SF's. Thus, SFC requires the OAM layer to perform validation and verification of SF's within an SFC Path, as well as connectivity and fault isolation.

In order to perform service connectivity verification of an SFC, the OAM could be initiated from any SFC aware network devices for end-to-end paths or partial path terminating on a specific SF within the SFC. This OAM function is to ensure the SF's chained together has connectivity as it is intended to when SFC was established. Necessary return code should be defined to be sent back in the response to OAM packet, in order to qualify the verification.

When ECMP exists at the service layer on a given SFC, there must be an ability to discover and traverse all available paths.

TBD - further details will be provided in a later revision.

3.2.2. Service Function Chain Performance Measurement

The ingress of the service function chain or an SFC aware network device must have an ability to perform loss and delay measurements over the service function chain as a unit (i.e. end-to-end) or to a specific service function through the SFC.

3.3. Classifier Component

A classifier defines a flow and maps incoming traffic to a specific SFC, and it is vital that the classifier is correctly defined and functioning. The SFC OAM must be able to test the definition of flows and the mapping functionality to expected SFCs.

4. SFC OAM Functions

Section 3 described SFC OAM operations required on each SFC component. This section explores the same from the OAM functionality point of view, which many will be applicable to multiple SFC components.

Various SFC OAM requirements provides the need for various OAM functions at different layers. Many of the OAM functions at different layers are already defined and in existence. In order to support SFC and SF's, these functions have to be enhanced to operate a single SF to multiple SF's in an SFC and also multiple SFC's.

4.1. Connectivity Functions

Connectivity is mainly an on-demand function to verify that the connectivity exists between network elements and the availability exists to service functions. Ping is a common tool used to perform this function. OAM messages SHOULD be encapsulated with necessary SFC header and with OAM markings when testing the service function chain component. OAM messages MAY be encapsulated with necessary SFC

header and with OAM markings when testing the service function component. Some of the OAM functions performed by connectivity functions are as follows:

- o Verify the MTU size from a source to the destination SF or through the SFC. This requires the ability for OAM packet to take variable length packet size.
- o Verify the packet re-ordering and corruption.
- o Verify the policy of an SFC or SF using OAM packet.
- o Verification and validating forwarding paths.
- o Proactively test alternate or protected paths to ensure reliability of network configurations.

4.2. Continuity Functions

Continuity is a model where OAM messages are sent periodically to validate or verify the reachability to a given SF or through a given SFC. This allows monitor network device to quickly detect failures like link failures, network failures, service function outages or service function chain outages. BFD is one such function which helps in detecting failures quickly. OAM functions supported by continuity check are as follows:

- o Ability to provision continuity check to a given SF or through a given SFC.
- o Notifying the failure upon failure detection for other OAM functions to take appropriate action.

4.3. Trace Functions

Tracing is an important OAM function that allows the operation to trigger an action (ex: response generation) from every transit device on the tested layer. This function is typically useful to gather information from every transit devices or to isolate the failure point towards an SF or through an SFC. Some of the OAM functions supported by trace functions are:

- o Ability to trigger action from every transit device on the tested layer towards an SF or through an SFC, using TTL or other means.
- o Ability to trigger every transit device to generate response with OAM code(s) on the tested layer towards an SF or through an SFC, using TTL or other means.

- o Ability to discover and traverse ECMP paths within an SFC.
- o Ability to skip un-supported SF's while tracing SF's in an SFC.

4.4. Performance Measurement Function

Performance management functions involve measuring of packet loss, delay, delay variance, etc. These measurements could be measured pro-actively and on-demand.

SFC OAM framework should provide the ability to perform packet loss for an SFC. In an SFC, there are various SF's chained together. Measuring packet loss is very important function. Using on-demand function, the packet loss could be measured using statistical means. Using OAM packets, the approximation of packet loss for a given SFC could be measured.

Delay within an SFC could be measured from the time it takes for a packet to traverse the SFC from ingress SF to egress SF. As the SFC's are generally unidirectional in nature, measurement of one-way delay is important. In order to measure one-way delay, the clocks have to be synchronized using NTP, GPS, etc.

Delay variance could also be measured by sending OAM packets and measuring the jitter between the packets passing through the SFC.

Some of the OAM functions supported by the performance measurement functions are:

- o Ability to measure the packet processing delay of a service function or a service function path along an SFC.
- o Ability to measure the packet loss of a service function or a service function path along an SFC.

5. Gap Analysis

This Section identifies various OAM functions available at different levels. It will also identify various gaps, if not all, existing within the existing toolset, to perform OAM function on an SFC.

5.1. Existing OAM Functions

There are various OAM tool sets available to perform OAM function and network layer, protocol layers and link layers. These OAM functions could validate some of the underlay and overlay networks. Tools like ping and trace are in existence to perform connectivity check and tracing intermediate hops in a network. These tools support

different network types like IP, MPLS, TRILL etc. There is also an effort to extend the tool set to provide connectivity and continuity checks within overlay networks. BFD is another tool which helps in detection of data forwarding failures.

Layer	Connectivity	Continuity	Trace	Performance
Underlay N/w	Ping	E-OAM, BFD	Trace	IPPM, MPLS
Overlay N/w	Ping	BFD, NVo3	Trace	IPPM
SF	None	+ None	+ None	+ None
SFC	None	+ None	+ None	+ None

Figure 3: OAM Tool GAP Analysis

Layer	Configuration	Orchestration	Topology	Notification
Underlay N/w	CLI, Netconf	CLI, Netconf	SNMP	SNMP, Syslog
Overlay N/w	CLI, Netconf	CLI, Netconf	SNMP	SNMP, Syslog
SF	CLI	+ CLI	+ None	+ None
SFC	CLI	+ CLI	+ None	+ None

Figure 4: OAM Tool GAP Analysis (contd.)

5.2. Missing OAM Functions

As shown in Figure 3, OAM functions for SFC are not standardized yet. Hence, there are no standard based tools available to verify SF and SFC.

5.3. Required OAM Functions

Primary OAM functions exist for network, transport, link and other layers. Tools like ping, trace, BFD, etc., exist in order to perform these OAM functions. Configuration, orchestration and manageability of SF and SFC could be performed using CLI, Netconf etc.

As seen in Figure 3 and 4, for configuration, manageability and orchestration, providing data and information models for SFC is very much needed. With virtualized SF and SFC, manageability of these functions has to be done programmatically.

6. SFC OAM Model

This section describes the operational aspects of SFC OAM at Service layer to perform the SFC OAM function defined in Section 4 and analyze the applicability of various existing OAM toolsets in the Service layer.

6.1. SFC OAM packet Marker

SFC OAM function described in Section 4 performed at service layer or overlay network layer must mark the packet as OAM packet that can be used by the relevant nodes to differentiate the OAM packet from data packets. The base header defined in Section 3.2 of [I-D.ietf-sfc-nsh] assigns a bit to indicate OAM packets. When NSH encapsulation is used at the service layer, the 0 bit must be set to differentiate the OAM packet. Any other overlay encapsulations used in future must have a way to mark the packet as OAM packet.

6.2. OAM packet processing and forwarding semantic

Upon receiving OAM packet, SF may choose to discard the packet if it does not support OAM functionality or if the local policy prevent it from processing OAM packet. When SF supports OAM functionality, it is desired to process the packet and respond back accordingly that helps with end-to-end verification. To avoid hitting any performance impact, SF can rate limit the number of OAM packets processed.

Service Function Forwarder (SFF) may choose not to forward the OAM packet to SF if the SF does not support OAM function or if the policy does not allow to forward OAM packet to SF. SFF may choose to skip the SF, modify the header and forward to next SFC node in the chain. How SFF detects if the connected SF supports or allowed to process OAM packet is outside the scope of this document. It could be a configuration parameter instructed by the controller or can be a dynamic negotiation between SF and SFF.

If the SFF receiving the OAM packet is the last SFF in the chain, it must send a relevant response to the initiator of the OAM packet. Depending on the type of OAM solution and tool set used, the response could be a simple response (ICMP reply or BFD reply packet) or could include additional data from the received OAM packet (like stats data consolidated along the path). The proposed solution should detail it further.

The classifier will normally be the node that initiates the OAM packet in order to validate the local classification policy or to validate the SFC or SFP. When the classifier initiates OAM packet, it must set the OAM marker in the overlay encapsulation.

6.3. OAM Function Types

As described in Section 4, there are different OAM functions that may require different OAM solution or tool sets. While the presence of OAM marker in overlay header (For ex: O bit in NSH header) indicates it as OAM packet, it is not sufficient to indicate what OAM function the packet is intended for. We can use the Next Protocol field in NSH header to indicate what OAM function is it intended to or what toolset is used.

6.4. OAM toolset applicability

As described in Section 5.1, there are different tool sets available to perform OAM functions at different layers. This section describes the applicability of some of the available tool sets in service layer.

6.4.1. ICMP Applicability

[RFC0792] and [RFC4443] describes the use of ICMP in IPv4 and IPv6 network respectively. It explains how ICMP messages can be used to test the network reachability between different end points and perform basic network diagnostics.

ICMP could be leveraged for basic OAM functions like SF availability or SFC availability. Initiator can generate ICMP echo message and control the overlay encapsulation header to get the response from relevant node. For example, a classifier initiating OAM can generate ICMP echo message can set the TTL field in NSH header to 255 to get the response from last SFF and thereby test the SFC availability. Alternately, Initiator can set the TTL to other value to get the response from specific SF and there by test the SF availability. Alternately, Initiator could send OAM packets with sequentially incrementing the TTL in NSH header to trace the Service Function Path.

It could be observed that ICMP at its current stage may not be able to perform all SFC OAM functions, but as explained above, it can be used to test the basic OAM functions.

6.4.2. Seamless BFD Applicability

[RFC5880] defines Bidirectional Forwarding Detection (BFD) mechanism for fast failure detection. [RFC5881] and [RFC5884] defines the applicability of BFD in IPv4, IPv6 and MPLS networks. [RFC7880] defines Seamless BFD (S-BFD), a simplified mechanism of using BFD. [RFC7881] explains its applicability in IPv4, IPv6 and MPLS network.

S-BFD could be leveraged to perform SF or SFC availability. Classifier or Initiator could generate BFD control packet and set the "Your Discriminator" value as last SFF in the control packet. Upon receiving the control packet, last SFF will reply back with relevant DIAG code. We could also use the TTL field in NSH header to perform the SF availability. For example, Initiator can set the "Your Discriminator" value to the SF that is intended to be tested and set the TTL field in NSH header in a way that it will be expired on the relevant SF. How the initiator gets the Discriminator value of the SF is outside the scope of this document.

6.4.3. In-Situ OAM

[I-D.brockners-proof-of-transit] defines the mechanism to perform proof of transit to securely verify if a packet traversed the relevant path or chain. While the mechanism is defined inband (i.e, it will be included in data packets), it can be used to perform various SFC OAM functions as well.

In-Situ OAM could be used with 0 bit set and perform SF availability, SFC availability of performance measurement.

6.4.4. SFC Traceroute

[I-D.penno-sfc-trace] defines a protocol that checks for path liveness and trace the service hops in any SFP. Section 3 of [I-D.penno-sfc-trace] defines the SFC trace packet format while section 4 and 5 of [I-D.penno-sfc-trace] defines the behavior of SF and SFF respectively.

Initiator can control the SIL in SFC trace packet to perform SF and SFC availability test.

6.5. Security Considerations

SFC and SF OAM must provide mechanisms for:

- o Preventing usage of OAM channel for DDOS attacks.
- o OAM packets meant for a given SFC should not get leaked beyond that SFC.
- o Prevent OAM packets to leak the information of an SFC beyond its administrative domain.

6.6. IANA Considerations

No action is required by IANA for this document.

6.7. Acknowledgements

TBD

7. References

7.1. Normative References

- [I-D.brockners-proof-of-transit]
Brockners, F., Bhandari, S., Dara, S., Pignataro, C., Leddy, J., Youell, S., Mozes, D., and T. Mizrahi, "Proof of Transit", draft-brockners-proof-of-transit-03 (work in progress), March 2017.
- [I-D.ietf-sfc-nsh]
Quinn, P. and U. Elzur, "Network Service Header", draft-ietf-sfc-nsh-13 (work in progress), June 2017.
- [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.
- [RFC0792] Postel, J., "Internet Control Message Protocol", STD 5, RFC 792, DOI 10.17487/RFC0792, September 1981, <<http://www.rfc-editor.org/info/rfc792>>.
- [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>>.
- [RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", RFC 4443, DOI 10.17487/RFC4443, March 2006, <<http://www.rfc-editor.org/info/rfc4443>>.
- [RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, DOI 10.17487/RFC5880, June 2010, <<http://www.rfc-editor.org/info/rfc5880>>.

- [RFC5881] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD) for IPv4 and IPv6 (Single Hop)", RFC 5881, DOI 10.17487/RFC5881, June 2010, <<http://www.rfc-editor.org/info/rfc5881>>.
- [RFC5884] Aggarwal, R., Kompella, K., Nadeau, T., and G. Swallow, "Bidirectional Forwarding Detection (BFD) for MPLS Label Switched Paths (LSPs)", RFC 5884, DOI 10.17487/RFC5884, June 2010, <<http://www.rfc-editor.org/info/rfc5884>>.
- [RFC7498] Quinn, P., Ed. and T. Nadeau, Ed., "Problem Statement for Service Function Chaining", RFC 7498, DOI 10.17487/RFC7498, April 2015, <<http://www.rfc-editor.org/info/rfc7498>>.
- [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>>.
- [RFC7880] Pignataro, C., Ward, D., Akiya, N., Bhatia, M., and S. Pallagatti, "Seamless Bidirectional Forwarding Detection (S-BFD)", RFC 7880, DOI 10.17487/RFC7880, July 2016, <<http://www.rfc-editor.org/info/rfc7880>>.
- [RFC7881] Pignataro, C., Ward, D., and N. Akiya, "Seamless Bidirectional Forwarding Detection (S-BFD) for IPv4, IPv6, and MPLS", RFC 7881, DOI 10.17487/RFC7881, July 2016, <<http://www.rfc-editor.org/info/rfc7881>>.
- [RFC8029] Kompella, K., Swallow, G., Pignataro, C., Ed., Kumar, N., Aldrin, S., and M. Chen, "Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures", RFC 8029, DOI 10.17487/RFC8029, March 2017, <<http://www.rfc-editor.org/info/rfc8029>>.

7.2. Informative References

- [RFC6291] Andersson, L., van Helvoort, H., Bonica, R., Romascanu, D., and S. Mansfield, "Guidelines for the Use of the "OAM" Acronym in the IETF", BCP 161, RFC 6291, DOI 10.17487/RFC6291, June 2011, <<http://www.rfc-editor.org/info/rfc6291>>.

Authors' Addresses

Sam K. Aldrin
Google

Email: aldrin.ietf@gmail.com

Carlos Pignataro (editor)
Cisco Systems, Inc.

Email: cpignata@cisco.com

Nagendra Kumar (editor)
Cisco Systems, Inc.

Email: naikumar@cisco.com

Nobo Akiya
Big Switch Networks

Email: nobo.akiya.dev@gmail.com

Ram Krishnan
Dell

Email: ramkri123@gmail.com

Anoop Ghanwani
Dell

Email: anoop@alumni.duke.edu

SFC Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 30, 2017

G. Mirsky
ZTE Corp.
G. Fioccola
Telecom Italia
T. Mizrahi
Marvell
June 28, 2017

Performance Measurement (PM) with Alternate Marking Method in Service
Function Chaining (SFC) Domain
draft-mirsky-sfc-pmamm-01

Abstract

This document describes how the alternate marking method be used as the passive performance measurement method in a Service Function Chaining (SFC) domain.

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 30, 2017.

Copyright Notice

Copyright (c) 2017 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 used in this document	2
2.1. Terminology	2
2.2. Requirements Language	3
3. Mark Field in NSH Base Header	3
4. Theory of Operation	4
4.1. Single Mark Enabled Measurement	4
4.2. Double Mark Enabled Measurement	5
5. IANA Considerations	6
5.1. Mark Field in NSH Base Header	6
6. Security Considerations	6
7. Acknowledgement	6
8. References	6
8.1. Normative References	6
8.2. Informative References	7
Authors' Addresses	7

1. Introduction

[RFC7665] introduced architecture of a Service Function Chain (SFC) in the network and defined its components as classifier, Service Function Forwarder (SFF), and Service Function (SF).

[I-D.ietf-ippm-alt-mark] describes passive performance measurement method, which can be used to measure packet loss, latency and jitter on live traffic. Because this method is based on marking consecutive batches of packets the method often referred as Alternate Marking Method (AMM).

This document defines how the alternate marking method can be used to measure packet loss and delay metrics of a service flow over e2e or any segment of the SFC.

2. Conventions used in this document

2.1. Terminology

MM: Marking Method

OAM: Operations, Administration and Maintenance

SFC: Service Function Chain

SF: Service Function

SFF: Service Function Forwarder

SFP: Service Function Path

NSH: Network Service Header

2.2. 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.

3. Mark Field in NSH Base Header

[I-D.ietf-sfc-nsh] defines format of the Network Service Header (NSH).

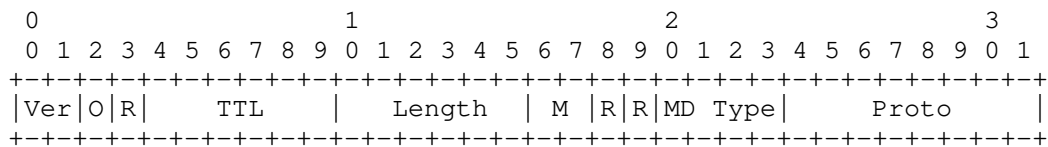


Figure 1: NSH Base format

This document defines two bit long field, referred as Mark field (M in Figure 1, as part of NSH Base and designated for the alternate marking performance measurement method [I-D.ietf-ippm-alt-mark]. The Mark field MUST NOT be used in defining forwarding and/or quality of service treatment of a SFC packet. The Mark field MUST be used only for the performance measurement of data traffic in SFC layer. Because setting of the field to any value does not affect forwarding and/or quality of service treatment of a packet, the alternate marking method in SFC layer can be viewed as true example of passive performance measurement method.

The Figure 2 displays format of the Mark field.

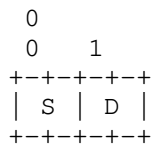


Figure 2: Mark field format

where:

- o S- Single mark method;
- o D - Double mark method.

4. Theory of Operation

The marking method can be successfully used in the SFC. Without limiting any generality consider SFC presented in Figure 3. Any combination of markings, Loss and/or Delay, can be applied to a service flow by any component of the SFC at either ingress or egress point to perform node, link, segment or end-to-end measurement to detect performance degradation defect and localize it efficiently.

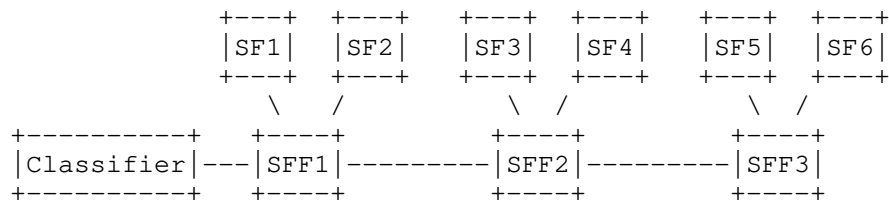


Figure 3: SFC network

Using the marking method a component of the SFC creates distinct sub-flows in the particular service traffic over SFC. Each sub-flow consists of consecutive blocks that are unambiguously recognizable by a monitoring point at any component of the SFC and can be measured to calculate packet loss and/or packet delay metrics.

4.1. Single Mark Enabled Measurement

As explained in the [I-D.ietf-ippm-alt-mark], marking can be applied to delineate blocks of packets based either on equal number of packets in a block or based on equal time interval. The latter method offers better control as it allows better account for capabilities of downstream nodes to report statistics related to batches of packets and, at the same time, time resolution that affects defect detection interval.

If the Single Mark measurement used, then the D flag MUST be set to zero on transmit and ignored by monitoring point.

The S flag is used to create alternate flows to measure the packet loss by switching value of the S flag every N-th packet or at certain

time intervals. Delay metrics MAY be calculated with the alternate flow using any of the following methods:

- o First/Last Packet Delay calculation: whenever the marking, i.e. value of S flag, changes a component of the SFC can store the timestamp of the first/last packet of the block. The timestamp can be compared with the timestamp of the packet that arrived in the same order through a monitoring point at downstream component of the SFC to compute packet delay. Because timestamps collected based on order of arrival this method is sensitive to packet loss and re-ordering of packets
- o Average Packet Delay calculation: an average delay is calculated by considering the average arrival time of the packets within a single block. A component of the SFC may collect timestamps for each packet received within a single block. Average of the timestamp is the sum of all the timestamps divided by the total number of packets received. Then difference between averages calculated at two monitoring points is the average packet delay on that segment. This method is robust to out of order packets and also to packet loss (only a small error is introduced). This method only provides single metric for the duration of the block and it doesn't give the minimum and maximum delay values. This limitation could be overcome by reducing the duration of the block by means of an highly optimized implementation of the method.

4.2. Double Mark Enabled Measurement

Double Mark method allows measurement of minimum and maximum delays for the monitored flow but it requires more nodal and network resources. If the Double Mark method used, then the S flag MUST be used to create the alternate flow, i.e. mark larger batches of packets. The D flag MUST be used to mark single packets to measure delay jitter.

The first marking (S flag alternation) is needed for packet loss and also for average delay measurement. The second marking (D flag is put to one) creates a new set of marked packets that are fully identified over the SFC, so that a component can store the timestamps of these packets; these timestamps can be compared with the timestamps of the same packets on another component of the SFC to compute packet delay values for each packet. The number of measurements can be easily increased by changing the frequency of the second marking. But the frequency of the second marking must be not too high in order to avoid out of order issues. This method is useful to have not only the average delay but also the minimum and maximum delay values and, in wider terms, to know more about the statistic distribution of delay values.

5. IANA Considerations

5.1. Mark Field in NSH Base Header

This document requests IANA to allocate Mark field as two bits-long field from NSH Base Header Reserved Bits [I-D.ietf-sfc-nsh].

This document requests IANA to register values of the Mark field of NSH as the following:

Bit Position	Marking	Description	Reference
0	S	Single Mark Measurement	This document
1	D	Double Mark Measurement	This document

Table 1: Mark field of SFC NSH

6. Security Considerations

This document lists the OAM requirement for SFC domain and does not raise any security concerns or issues in addition to ones common to networking and SFC.

7. Acknowledgement

TBD

8. References

8.1. Normative References

- [I-D.ietf-sfc-nsh] Quinn, P. and U. Elzur, "Network Service Header", draft-ietf-sfc-nsh-12 (work in progress), February 2017.
- [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>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<http://www.rfc-editor.org/info/rfc8174>>.

8.2. Informative References

- [I-D.ietf-ippm-alt-mark]
Fioccola, G., Capello, A., Cociglio, M., Castaldelli, L.,
Chen, M., Zheng, L., Mirsky, G., and T. Mizrahi,
"Alternate Marking method for passive and hybrid
performance monitoring", draft-ietf-ippm-alt-mark-05 (work
in progress), June 2017.
- [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

Greg Mirsky
ZTE Corp.

Email: gregimirsky@gmail.com

Giuseppe Fioccola
Telecom Italia

Email: giuseppe.fioccola@telecomitalia.it

Tal Mizrahi
Marvell
6 Hamada St.
Yokneam
Israel

Email: talmi@marvell.com

Network Working Group
Internet-Draft
Intended status: Informational
Expires: January 2, 2018

D. Purkayastha
A. Rahman
D. Trossen
InterDigital Communications, LLC
July 1, 2017

USE CASE FOR HANDLING DYNAMIC CHAINING AND SERVICE INDIRECTION
draft-purkayastha-sfc-service-indirection-00

Abstract

Many stringent requirements are imposed on today's network, such as low latency, high availability and reliability in order to support several use cases such as IoT, Gaming, Content distribution, Robotics etc. Networks need to be flexible and dynamic in terms of allocation of services and resources. Network Operators should be able to reconfigure the composition of a service and steer users towards new service end points as user move or resource availability changes. SFC allows network operators to easily create and reconfigure service function chains dynamically in response to changing network requirements. We discuss a use case where Service Function Chain can adapt or self-organize as demanded by the network condition without requiring SPI re-classification. This can be achieved, for example, by decoupling the service consumer and service endpoint by a new service function proposed in this draft. We describe few requirements for this service function to enable dynamic switching between consumer and end point.

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 2, 2018.

Copyright Notice

Copyright (c) 2017 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. NSH and Re-classification	3
2.1. Dynamic service chain creation using NSH	4
3. Challenges with dynamic indirection	5
4. Desired Features	8
5. Service Request Routing (SRR) Service Function	8
6. Next Steps	10
7. IANA Considerations	10
8. Security Considerations	10
9. Informative References	10
Authors' Addresses	11

1. Introduction

The requirements on today's networks are very diverse, enabling multiple use cases such as IoT, Content Distribution, Gaming, Network functions such as Cloud RAN. Every use case imposes certain requirements on the network. These requirements vary from one extreme to other and often they are in a divergent direction. Network operator and service providers are pushing many functions towards the edge of the network in order to be closer to the users. This reduces latency and backhaul traffic, as user request can be processed locally.

It becomes more challenging for the network when user mobility as well as non-deterministic availability of compute and storage resources are considered. The impact is felt most in the edge of the network because as the users move, their point of attachment changes frequently, which results in (at least partially) relocating the service as well as the service endpoint. Furthermore, network functions are pushed more and more towards the edge, where compute

and storage resources are constrained and availability is non-deterministic. Also, storage resources may need to be moved where the user concentration is more in case of content delivery applications.

Take the following video orchestration service example from ETSI MEC Requirements document [ETSI_MEC]. The proposed use case of edge video orchestration suggests a scenario where visual content can be produced and consumed at the same location close to consumers in a densely populated and clearly limited area. Such a case could be a sports event or concert where a remarkable number of consumers are using their handheld devices to access user select tailored content. The overall video experience is combined from multiple sources, such as local recording devices, which may be fixed as well as mobile, and master video from central production server. The user is given an opportunity to select tailored views from a set of local video sources.

In such a dynamic network environment, the capability to dynamically compose new services from available services as well as move a service instance in response to user mobility or resource availability is desirable. SFC allows network operators as well as service providers to compose new services by chaining individual service functions towards the composed new service. In a dynamic network environment where service functions move frequently because of user movement, load balancing or resource modification, service function chains and the service end points need to be created and recreated frequently. SFC, as defined in IETF, is capable of modifying the service chain dynamically in response to network conditions.

In this document we address this dynamicity by introducing a special Service Function, called SRR (service request routing). We describe the problems associated with today's network and Layer 3 based approach to handle dynamicity in the network. We then discuss how such new Service Function with certain capabilities can handle the dynamicity better than these conventional methods.

2. NSH and Re-classification

[RFC7498] captures the problems associated with existing service deployments that are problematic. High level problems are listed below.

- o Network topology: Network service deployment is tightly coupled with network topology thus reducing the flexibility in service delivery. It adds complexity in deploying network service when

certain traffic types may need some service and other traffic types do not need the same service.

- o Configuration complexity is the direct result of dependency on network topology.
- o Limited availability of services
- o Altering the order of a deployed chain is complex and cumbersome
- o Coupling of service functions to topology may require service functions to support many transport encapsulations or for a transport gateway function to be present.
- o In a dynamic environment like the Edge of a network service delivery, routing changes fast. It may be difficult to deliver service dynamically due to the risk and complexity of VLANs and/or routing modifications.

These factors provide motivation for a simplified and flexible service insertion model that addresses many of the current shortcomings and provides new, much needed functionality to enable service deployments in modern network environments. Service chaining accomplishes this by considering service functions as resources, with associated attributes, available for scheduled consumption. Selective traffic, subject to policy, may then be "steered" to the requisite service resources, along with any "extra" information referred to as metadata. This metadata is used for policy enforcement.

A basic form of service chaining may be realized using existing transport encapsulations. This method of chaining relies upon the tunneling of selected data between service functions. Although this form of service chaining achieves some level of abstraction from the underlying topology, it does not truly create a service plane. NSH [I-D.ietf-sfc-nsh] is a distinct identifiable plane that can be used across all transports to create a service chain and exchange metadata along the chain.

2.1. Dynamic service chain creation using NSH

We revisit the dynamic service chain creation capability of NSH. NSH defines a new service plane protocol [I-D.ietf-sfc-nsh]. A Network Service Header (NSH) contains service path information and optionally metadata that are added to a packet or frame and used to create a service plane. A control plane is required in order to exchange NSH values with participating nodes, and to provision the same nodes with requisite information such as service path ID to overlay mapping.

The Network Service Header has three parts, Base header, Service Path Header and Context Header. NSH Service Path Header is a 4-byte service path header follows the base header and defines two fields used to construct a service path:

- o Service path identifier (SPI)
- o Service index (SI)

The following figure depicts the service path header.

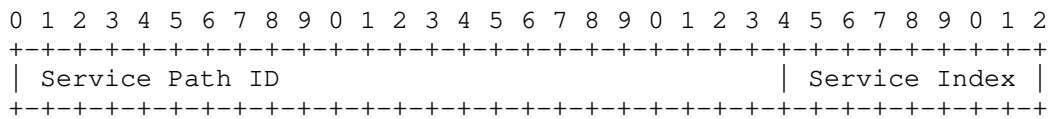


Figure 1: NSH Path Header

The service path identifier (SPI) is used to identify the service path that interconnects the needed service functions. It allows nodes to utilize the identifier to select the appropriate network transport protocol and forwarding techniques. The service index (SI) identifies the location of a packet within a service path. As packets traverse a service path, the SI is decremented post-service.

SPI represents the service path and altering the path identifier results in a change of a service path. A change in SPI value is a result of re-classification. It means a node in the service path determined, based on policy, that the initial classification was incorrect or incomplete. If the updated classification results in the necessity of a new service path, the node updates the SPI and SI fields accordingly. The new identifier is then used to select the appropriate overlay topology. This allows service functions to alter the path of a packet without having to participate in the network topology and its associated control plane(s). The method to determine that an existing classification is incorrect and how to determine the new classification is not defined.

3. Challenges with dynamic indirection

The emerging trend in today's network is to deploy network functions, services and applications at the edge of the network to support latency requirements, computational offload, traffic optimization etc. As users are moving, application or services being used by users, may need to be moved closer to the user's new location. This implies another instance of the service function may need to be instantiated close to the user's new location. It may result in re-

establishing service path from the newly instantiated service function to other service instances. It is also possible that the newly instantiated service function may be redirected to a new service end point (e.g. Application Server) for various reasons, such as incomplete content, proximity to data store, load balancing etc. In another scenario, a single instance of the service function may not handle all users. A single service function may be instantiated more than once to balance user load. As the number of instances increase and along with mobility, the complexity of service routing increases. It is anticipated that there may be a constant action of function chaining, re-chaining occurring in the network.

The challenge of dynamic indirection may be better described by analyzing the working of CDNs, which dynamically (re-)direct user-initiated requests towards the most appropriate content instance. This task becomes more difficult if granularity of the instance placement increases. For instance, in case of a CDN being realized close to end users, specifically in edge of the network, the specific content instance might need to be selected dynamically. After initial selection, the instance may change during service execution.

In a conventional network, an instance of a service is found and selected using DNS. The subsequent service request is then routed through the network between the client and the service. If the user is doing a DNS lookup to access content served by a CDN then the DNS service will maintain a list of IP addresses that can be returned for a given domain name and will try to return an IP address of a node geographically close to the client. Should the service provider want to replace an instance of their service with another one at a different IP address (and potentially a different physical location for various reasons such as load balancing, reliability etc.) then the DNS tables must be updated, i.e., the service needs to be (re-)registered quickly. This is done by updating the local authoritative DNS server which then propagates the new mapping to DNS services across the world. DNS propagation can take up to 48 hours so fast and dynamic switching from one service instance to another is not possible in conventional networks. When relying on many surrogate service endpoints to exist in the edge network, there is a clear issue of certain resources not being available in one surrogate instance while existing in another so that changes in redirection might be desirable, while also changes in local load drive the need for such change in redirection.

The other issue in conventional network lies with mobility management procedure. These procedures use an anchor point, which terminates a session at the network edge. As user moves around, traffic is redirected from the anchor point to the new point of attachment. Relying on typical mobility management approaches found in IP

networks, usually leads to inefficient 'triangular' routing of requests through this common 'anchor' point. This triangular routing increases the latency in reaching the new service function or service end points as users move.

Traffic steering is a common procedure in managed networks, particularly at the edge, due to desired subscriber-centric traffic policies (e.g., related to pricing structures), resource requirements (e.g., related to using particular paths in the network) or mobility (e.g., users moving in a cellular network). Today's methods for traffic steering include anchor-based mobility management as well as traffic classification, for instance, in packet gateways of cellular systems (using, e.g., deep packet inspection as well as port and address classification). While the former leads to inefficient 'triangular' traffic forwarding, the latter often requires additional state in the forwarders to differentiate traffic from one user to another.

The analysis of CDN network shows that dynamic indirection is a necessary requirement, which needs to be supported by the networks. The goal for this indirection is to provide user applications lowest possible latency. But as discussed above, relying on today's technique, does not help in guaranteeing same latency to user applications. On the other hand, there is a high possibility that latency may increase if we rely on Layer 3 based service redirection techniques.

SFC handles indirection through the use of SPI. A packet needs to be reclassified and the intermediate node changes the SPI. Following are the typical steps that happens in order to implement the indirection.

- o A packet arrives at a particular node
- o The node contacts the policy manager
- o Identifies the current classification is incorrect
- o Reclassifies the packet, i.e. change the SPI
- o Inserts the packet in the pipe, possibly towards the SFF

The indirection mechanism in SFC involves certain steps to process policy information and change the SPI in the packet header, making it suitable to handle dynamic indirection requirements. Our proposed SF in this document provides an additional method to handle dynamic indirection of service requests, not relying on the reclassification

mechanism. Combining these two techniques may provide flexibility and improvement over single method.

4. Desired Features

In order to route the service requests to service end points in a dynamic manner, we identify the following desirable features:

- o Fast switching from one service instance to another by not relying on DNS for service location resolution. Instead of DNS, the function should be able to identify the path, which will allow to reach the service end point.
- o Direct path mobility, where the path between the requester and the responding service can be determined as being optimal (e.g., shortest path or direct path to a selected instance), is needed to avoid the use of anchor points and reduce service-level latency
- o Indirect service requests at the network level, transparent to the requesting client and without the involvement of the DNS. End user is not aware of the decision made by the SF.
- o New methods for forwarding, such as path-based forwarding, direct path routing in mobility cases, path pinning for traffic steering and simplified service-specific peering towards the Internet.

5. Service Request Routing (SRR) Service Function

The following diagram shows the application of the new proposed SRR service function in an example of media clients connecting to media servers. There may be more than one media functions to support CDN like architecture, Surrogate servers to handle mobility and load balancing.

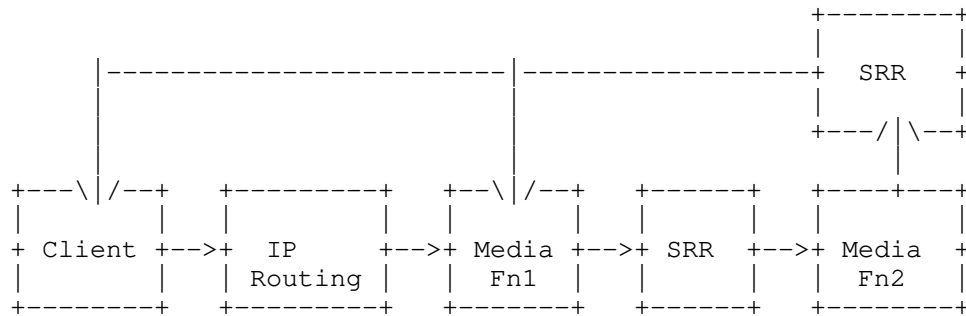


Figure 2: Use of SRR function

The clients are connected to media functions through frontend routed network, e.g., relying on standard IP routing, while media functions are chained via the new proposed service request routing (SRR) function. Alternatively, we also envision to utilize the SRR function directly between client SF and media function SF, as outlined in the figure below

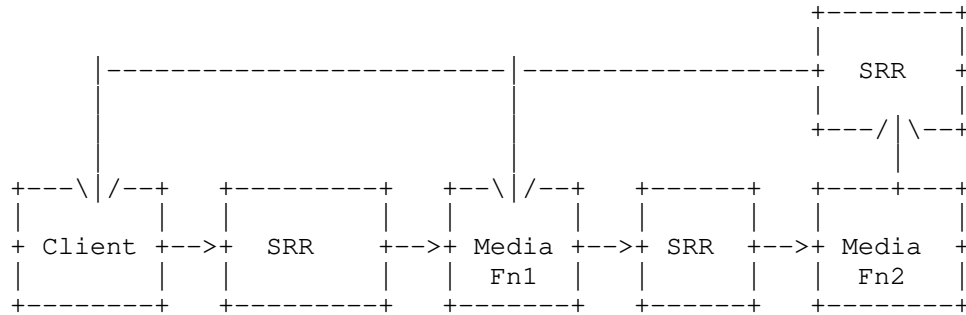


Figure 3: SRR function between Client and Media Function

The SRR service function decouples clients from media functions. This brings in flexibility in routing service requests from client to service end points. In the edge network, where users are moving and service end points may also change, having flexibility to decide and steer service requests directly helps in guaranteeing the same latency to user applications. Clearly, that is achieved by reducing the switching time from SF to another. As service end point changes, the routing functions makes instantaneous decision to route the request to the appropriate media server.

The possible improvements of using SRR within an SFC framework are listed below:

- o Fast (between 10 and 20ms) switching times from one service instance to another by not relying on the DNS for service discovery and directly routing service requests at the level of the transport network.
- o The capability to indirect service requests at the network level will help in reducing latency, when service end points change. E.g. when a service request is being sent to one surrogate instance but results in a HTTP 404 or 5xx error response, the original request is redirected to another alternative surrogate with minimal latency, i.e., right at the destination of said failed service request. Nesting these operations effectively leads to a net-level 'search' among all available surrogate instances until the search is exhausted (with a negative result) or the resource is found.
- o New methods for forwarding, such as path-based forwarding, will enable direct path routing in mobility cases, path pinning for traffic steering and simplified service-specific peering towards the Internet. Such capability would allow for localizing traffic, reduce latency and costs.

6. Next Steps

Does the WG see value in supporting the requirements for SFC to enable routing of service requests between service consumers and service endpoints in a dynamic manner as outlined in Section 4?

7. IANA Considerations

This document requests no IANA actions.

8. Security Considerations

TBD.

9. Informative References

[ETSI_MEC]

ETSI, "Mobile Edge Computing (MEC), Technical Requirements", GS MEC 002 1.1.1, March 2016, <http://www.etsi.org/deliver/etsi_gs/MEC/001_099/002/01.01.01_60/gs_MEC002v010101p.pdf>.

[I-D.ietf-sfc-nsh]

Quinn, P. and U. Elzur, "Network Service Header", draft-ietf-sfc-nsh-13 (work in progress), June 2017.

[RFC7498] Quinn, P., Ed. and T. Nadeau, Ed., "Problem Statement for Service Function Chaining", RFC 7498, DOI 10.17487/RFC7498, April 2015, <<http://www.rfc-editor.org/info/rfc7498>>.

Authors' Addresses

Debashish Purkayastha
InterDigital Communications, LLC
Conchoken
USA

Email: Debashish.Purkayastha@InterDigital.com

Akbar Rahman
InterDigital Communications, LLC
Montreal
Canada

Email: Akbar.Rahman@InterDigital.com

Dirk Trossen
InterDigital Communications, LLC
64 Great Eastern Street, 1st Floor
London EC2A 3QR
United Kingdom

Email: Dirk.Trossen@InterDigital.com
URI: <http://www.InterDigital.com/>

SFC WG
Internet-Draft
Intended status: Standards Track
Expires: December 16, 2017

G. Mirsky
ZTE Corp.
W. Meng
ZTE Corporation
B. Khasnabish
ZTE TX, Inc.
C. Wang
June 14, 2017

Multi-Layer OAM for Service Function Chains in Networks
draft-wang-sfc-multi-layer-oam-09

Abstract

A multi-layer approach to the task of Operation, Administration and Maintenance (OAM) of Service Function Chains (SFCs) in networks is presented. Based on the SFC OAM requirements, a multi-layer model is introduced. A mechanism to detect and localize defects using the multi-layer model is also described.

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 16, 2017.

Copyright Notice

Copyright (c) 2017 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
2.1. Requirements Language	3
2.2. Terminology	3
3. Multi-layer Model of SFC OAM	3
4. Requirements for SFC OAM Multi-layer Model	4
5. SFC OAM multi-layer model	5
6. Theory of Operation	6
7. Security Considerations	7
8. IANA Considerations	7
8.1. SFC TLV Type	7
8.2. SFC OAM UDP Port	8
9. References	8
9.1. Normative References	8
9.2. Informative References	9
Authors' Addresses	9

1. Introduction

[RFC7665] defines components necessary to implement Service Function Chain (SFC). These include a classifier which performs classification of incoming packets. A Service Function Forwarder (SFF) is responsible for forwarding traffic to one or more connected Service Functions (SFs) according to the information carried in the SFC encapsulation. SFF also handles traffic coming back from the SF and transports the data packets to the next SFF. And the SFF serves as termination element of the Service Function Path (SFP). SF is responsible for specific treatment of received packets.

Resulting from that SFC is constructed by a number of these components, there are different views from different levels of the SFC. One is the SFC, fully abstract entity, that defines an ordered set of SFs that must be applied to packets selected as a result of classification. But SFC doesn't define exact mapping between SFFs and SFs. Thus there exists another semi-abstract entity referred as SFP. SFP is the instantiation of the SFC in the network and provides a level of indirection between the fully abstract SFC and a fully specified ordered list of SFFs and SFs identities that the packet will visit when it traverses the SFC. The latter entity is being referred as Rendered Service Path (RSP). The main difference between

SFP and RSP is that in the former the authority to select the SFF/SF has been delegated to the network.

This document proposes the multi-layer model of SFC Operation, Administration and Maintenance (OAM) and requirements to improve the troubleshooting efficiency.

2. Conventions

2.1. 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.

2.2. Terminology

Unless explicitly specified in this document, active OAM in SFC and SFC OAM are being used interchangeably.

e2e: End-to-End

FM: Fault Management

OAM: Operations, Administration, and Maintenance

RDI: Remote Defect Indication

RSP: Rendered Service Path

SF: Service Function

SFC: Service Function Chain

SFF: Service Function Forwarder

SFP: Service Function Path

3. Multi-layer Model of SFC OAM

As described in [I-D.ietf-sfc-oam-framework], multiple layers come into play to realize the SFC, including the Service layer, the underlying Network layer, as well as the Link layer, which are depicted in Figure 1:

- o The Service layer consists of classifiers and/or service functions/SFs.
- o Network and Transport layers leverage various overlay network technologies interconnecting SFs to establish SFP.
- o The Link layer is technology specific and reflects the technology used in the underlay network.

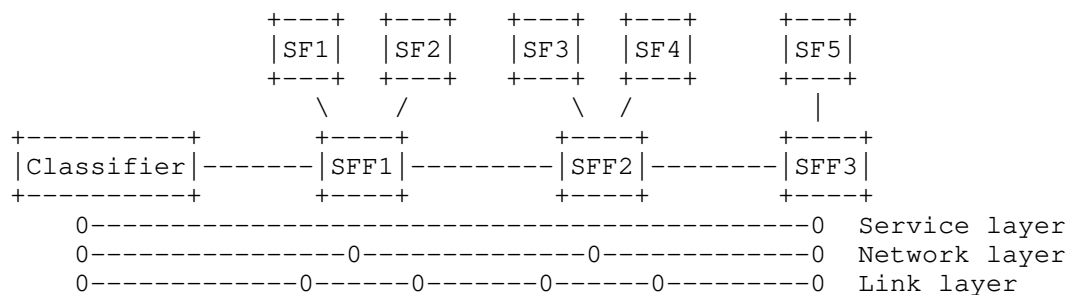


Figure 1: SFC OAM Multi-Layer model

4. Requirements for SFC OAM Multi-layer Model

To perform the OAM task of fault management (FM) in an SFC, that includes failure detection, defect characterization and localization, this document defines the multi-layer model of OAM, presented in Section 3, and set of requirements towards active OAM mechanisms to be used on an SFC.

In example presented in Figure 1 the service SFP1 may be realized through two RSPs, RSP1(SF1--SF3--SF5) and RSP2(SF2--SF4--SF6). To perform end-to-end (e2e) FM SFC OAM:

REQ#1: Packets of active OAM in SFC SHOULD be fate sharing with data traffic, i.e. in-band with the monitored traffic, i.e. follow exactly the same RSP, in forward direction, i.e. from ingress toward egress end point(s) of the OAM test.

REQ#2: SFC OAM MUST support pro-active monitoring of any element in the SFC availability.

The egress, SFF3 in example in Figure 1, is the entity that detects the failure of the SFC. It must be able to signal the new defect state to the ingress, i.e. SFF1. Hence the following requirement:

REQ#3: SFC OAM MUST support Remote Defect Indication (RDI) notification by egress to the ingress, i.e. source of continuity checking.

REQ#4: SFC OAM MUST support connectivity verification. Definition of mis-connectivity defect entry and exit criteria are outside the scope of this document.

Once the SFF1 detects the defect objective of OAM switches from failure detection to defect characterization and localization.

REQ#5: SFC OAM MUST support fault localization of Loss of Continuity check in the SFC.

REQ#6: SFC OAM MUST support tracing an SFP in order to realize the RSP.

It is practical, as presented in Figure 1, that several SFs share the same SFF. In such case SFP1 may be realized over two RSPs, RSP1(SF1--SF3--SF5) and RSP2(SF2--SF4--SF6).

REQ#7: SFC OAM MUST have the ability to discover and exercise all available RSPs in the transport network.

In process of localizing the SFC failure separating SFC OAM layers is very attractive and efficient approach. To achieve that continuity among SFFs that are part of the same SFP should be verified. Once SFFs reachability along the particular SFP has been confirmed task of defect localization may focus on SF reachability verification. Because reachability of SFFs has already been verified, SFF local to the SF may be used as source.

REQ#8: SFC OAM MUST be able to trigger on-demand FM with responses being directed towards initiator of such proxy request.

By using the multi-layer model OAM that confirms to the above listed requirements is capable to perform efficient defect localization on an SFC.

5. SFC OAM multi-layer model

Figure 2 presents a use case of applying the proposed SFC OAM multi-layer model. In this scenario operator needs to discover SFFs and SFs of the same SFC. The Layer 1 includes the SFFs that are part of the SFP. The Layer 2 - the SFs along the RSP. When trying to do SFC OAM, classifier or service nodes select and confirm which SFC OAM layering they plan to do, then encapsulate the layering information in the SFC OAM packets, and send the SFC OAM packets along the

service function paths to the destination. When receiving the SFC OAM packets, service nodes analyze the layering information and then decide whether sending these packets to next SFFs directly without being processed by SFs for Layer 1 process or sending to SFs for Layer 2 process.

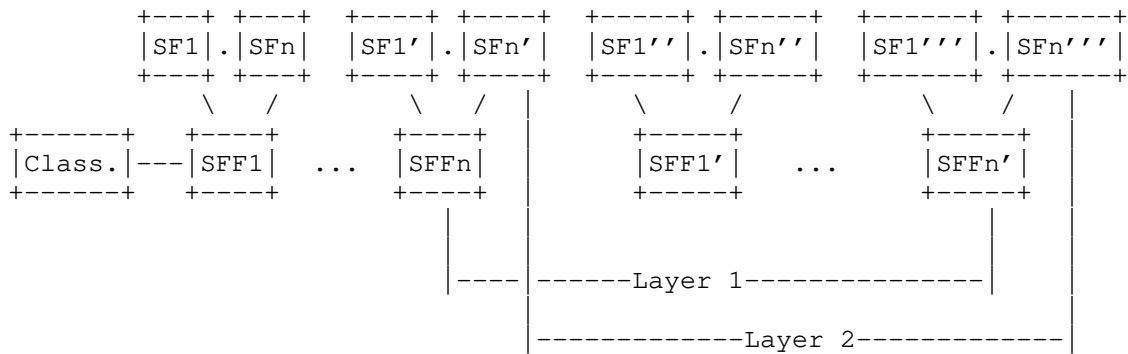


Figure 2: SFC OAM multi-layering model

6. Theory of Operation

Echo Request/Reply is well-known OAM mechanism that is extensively used to detect inconsistencies between states in control plane and data plane, localize defects in the data plane. In SFC OAM Echo Request/Reply is built as extension of Overlay Echo Request/Reply functions [I-D.ooamdt-rtgwg-demand-cc-cv].

Responder to the SFC Echo Request sends the Echo Reply over IP network if the reply mode is Reply via an IPv4/IPv6 UDP Packet [I-D.ooamdt-rtgwg-demand-cc-cv]. Because SFC NSH does not identify the ingress of the SFP the Echo Request MUST include this information that to be used as IP destination address for IP/UDP encapsulation of the SFC Echo Reply. Sender of the SFC Echo Request MUST include SFC Source TLV Figure 3.

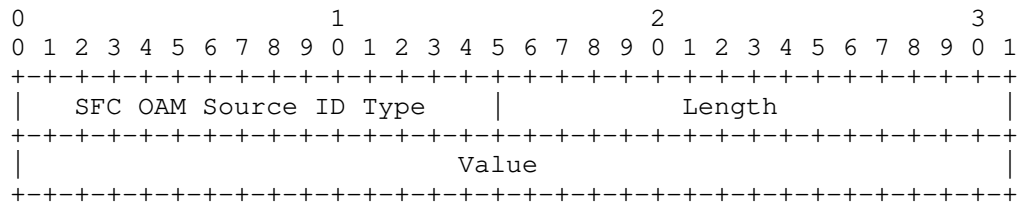


Figure 3: SFC Source TLV

where

SFC OAM Source Id Type is two octets in length and has the value of TBD1 Section 8.1.

Length is two octets long field and the valuse is equal to the length of the Value field.

Value field contains IP address of the sender of the SFC OAM control message, IPv4 or IPv6.

The UDP destination port for SFC Echo Reply TBD2 will be allocated by IANA Section 8.2.

7. Security Considerations

TBD

8. IANA Considerations

8.1. SFC TLV Type

IANA is requested to create SFC OAM TLV Type registry. All code points in the range 1 through 32759 in this registry shall be allocated according to the "IETF Review" procedure as specified in [RFC5226]. Code points in the range 32760 through 65279 in this registry shall be allocated according to the "First Come First Served" procedure as specified in [RFC5226]. Remaining code points are allocated according to the Table 1:

Value	Description	Reference
0	Reserved	This document
1- 32759	Unassigned	IETF Review
32760 - 65279	Unassigned	First Come First Served
65280 - 65519	Experimental	This document
65520 - 65534	Private Use	This document
65535	Reserved	This document

Table 1: SFC TLV Type Registry

This document defines the following new value in SFC OAM TLV Type registry:

Value	Description	Reference
TBD1	Source IP Address	This document

Table 2: SFC OAM Source IP Address Type

8.2. SFC OAM UDP Port

IANA is requested to allocate UDP port number according to

Service Name	Port Number	Transport Protocol	Description	Semantics Definition	Reference
SFC OAM	TBD2	UDP	SFC OAM	Section 6	This document

Table 3: SFC OAM Port

9. References

9.1. Normative References

[I-D.ooamdt-rtgwg-demand-cc-cv]
 Mirsky, G., Kumar, N., Kumar, D., Chen, M., Yizhou, L.,
 and D. Dolson, "Echo Request and Echo Reply for Overlay
 Networks", draft-ooamdt-rtgwg-demand-cc-cv-03 (work in
 progress), March 2017.

- [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>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<http://www.rfc-editor.org/info/rfc8174>>.

9.2. Informative References

- [I-D.ietf-sfc-oam-framework] Aldrin, S., Krishnan, R., Akiya, N., Pignataro, C., and A. Ghanwani, "Service Function Chaining Operation, Administration and Maintenance Framework", draft-ietf-sfc-oam-framework-01 (work in progress), February 2016.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, DOI 10.17487/RFC5226, May 2008, <<http://www.rfc-editor.org/info/rfc5226>>.
- [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

Greg Mirsky
ZTE Corp.

Email: gregimirsky@gmail.com

Wei Meng
ZTE Corporation
No.50 Software Avenue, Yuhuatai District
Nanjing
China

Email: meng.wei2@zte.com.cn, vally.meng@gmail.com

Bhumip Khasnabish
ZTE TX, Inc.
55 Madison Avenue, Suite 160
Morristown, New Jersey 07960
USA

Email: bhumip.khasnabish@ztetx.com

Cui Wang

Email: lindawangjoy@gmail.com