Active OAM for Service Function Chains in Networks
draft-wang-sfc-multi-layer-oam-12

Abstract

   A set of requirements for active Operation, Administration and
   Maintenance (OAM) of Service Function Chains (SFCs) in networks is
   presented. Based on these requirements an encapsulation of active
   OAM message in SFC and a mechanism to detect and localize defects
   described. Also, this document updates RFC 8300 in the definition of
   O (OAM) bit in the Network Service Header (NSH) and defines how the
   active OAM message identified in SFC NSH.

Status of This Memo

   This Internet-Draft is submitted in full conformance with the
   provisions of BCP 78 and BCP 79.

   Internet-Drafts are working documents of the Internet Engineering
   Task Force (IETF). Note that other groups may also distribute
   working documents as Internet-Drafts. The list of current Internet-
   Drafts is at https://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 April 10, 2019.

Copyright Notice

Table of Contents

1. Introduction

   [RFC7665] defines components necessary to implement Service Function
   Chain (SFC). These include a classifier which performs the
   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 the 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, entirely abstract entity, which defines an
ordered set of SFs that must be applied to packets selected as a
result of classification. But SFC doesn't specify the exact mapping
between SFFs and SFs. Thus there exists another semi-abstract entity
referred to as SFP. SFP is the instantiation of the SFC in the
network and provides a level of indirection between the entirely
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 to 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 defines how active Operation, Administration and
Maintenance (OAM), per [RFC7799] definition of active OAM, identified
in Network Service Header (NSH) SFC, lists requirements to improve
the troubleshooting efficiency, and defines SFC Echo request and Echo
reply that enables on-demand Continuity Check, Connectivity
Verification among other operations over SFC in networks. Also, this
document updates Section 2.2 of [RFC8300] in part of the definition
of O bit in the (NSH).

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

NSH: Network Service Header

OAM: Operations, Administration, and Maintenance

PRNG: Pseudorandom number generator

RDI: Remote Defect Indication

RSP: Rendered Service Path

SF: Service Function

SFC: Service Function Chain

SFF: Service Function Forwarder

SFP: Service Function Path

3. Requirements for Active OAM in SFC Network

To perform the OAM task of fault management (FM) in an SFC, that
includes failure detection, defect characterization and localization,
this document defines the set of requirements for active OAM
mechanisms to be used on an SFC.

```
          +---+  +---+    +---+  +---+   +---+  +---+
          |SF1|  |SF2|    |SF3|  |SF4|   |SF5|  |SF6|
          +---+  +---+    +---+  +---+   +---+  +---+
            \    /          \   /          \   /
 +----------+     +----+         +----+        +----+
 |Classifier|-------|SFF1|---------|SFF2|--------|SFF3|
 +----------+     +----+         +----+        +----+
```
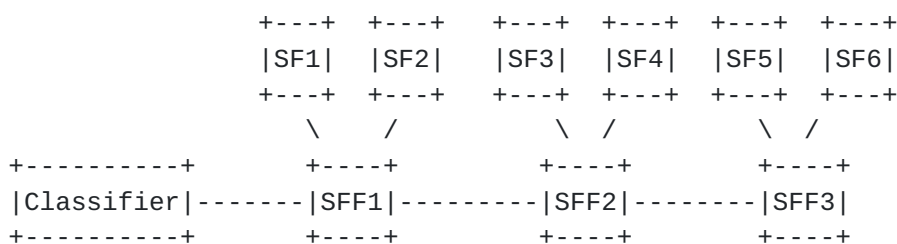
Figure 1: SFC reference model

In the example presented in Figure 1, the service SFP1 may be
realized through two RSPs, RSP1(SF1--SF3--SF5) and RSP2(SF2--SF4--
SF5). 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 follow the
   same RSP, in the forward direction from ingress toward egress
   endpoint(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 the 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 SFF1. Hence the following requirement:

REQ#3: SFC OAM MUST support Remote Defect Indication (RDI)
notification by the egress to the ingress.

REQ#4: SFC OAM MUST support connectivity verification. Definition
of the misconnection 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 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 the process of localizing the SFC failure, separating SFC OAM
layers is an 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 verified, SFF local to the
SF may be used as a source of the test packets.

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

4. Active OAM Identification in SFC NSH

The interpretation of O bit flag in the NSH header is defined in
[RFC8300] as:

O bit: Setting this bit indicates an OAM packet.

This document updates the definition of O bit as follows:

O bit: Setting this bit indicates an OAM command and/or data in
the NSH Context Header or packet payload

Active SFC OAM defined as a combination of OAM commands and/or data
included in a message that immediately follows the NSH. To identify
the active OAM message the value on the Next Protocol field MUST be

set to Active SFC OAM (TBA1) according to Section 8.1. The rules of
interpreting the values of O bit and the Next Protocol field are as
follows:

o O bit set and the Next Protocol value is not one of identifying
   active or hybrid OAM protocol (per [RFC7799] definitions), e.g.,
   defined in this specification Active SFC OAM - TLVs contain OAM
   command or data, and the type of payload determined by the Next
   Protocol field;

o O bit set and the Next Protocol value is one of identifying active
   or hybrid OAM protocol - the payload that immediately follows SFC
   NSH contains OAM command or data;

o O bit is clear - no OAM in TLV and the payload determined by the
   value of the Next Protocol field.

Several active OAM protocols will be needed to address all the
requirements listed in Section 3. Destination UDP port number may
identify protocols if IP/UDP encapsulation used. But extra IP/UDP
headers, especially in the case of IPv6, add noticeable overhead.
This document defines Active OAM Header Figure 2 to demultiplex
active OAM protocols on an SFC.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| V | Msg Type  |     Flags     |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~             SFC Active OAM Control Packet                     ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: SFC Active OAM Header

V - two bits long field indicates the current version of the SFC
active OAM header. The current value is 0.

Msg Type - six bits long field identifies OAM protocol, e.g., Echo
Request/Reply or BFD.

Flags - eight bits long field carries bit flags that define
optional capability and thus processing of the SFC active OAM
control packet, e.g., optional timestamping.

Length - two octets long field that is the length of the SFC
active OAM control packet in octets.

5. Echo Request/Echo Reply for SFC in Networks

   Echo Request/Reply is a well-known active OAM mechanism that is
   extensively used to detect inconsistencies between a state in control
   and the data planes, localize defects in the data plane. The format
   of the Echo request/Echo reply control packet is to support ping and
   traceroute functionality in SFC in networks Figure 3 resembles the
   format of MPLS LSP Ping [RFC8029] with some exceptions.

```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |        Version Number         |          Global Flags         |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    | Message Type  |  Reply mode   |  Return Code  | Return S.code |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                        Sender's Handle                        |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                        Sequence Number                        |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    ~                            TLVs                               ~
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
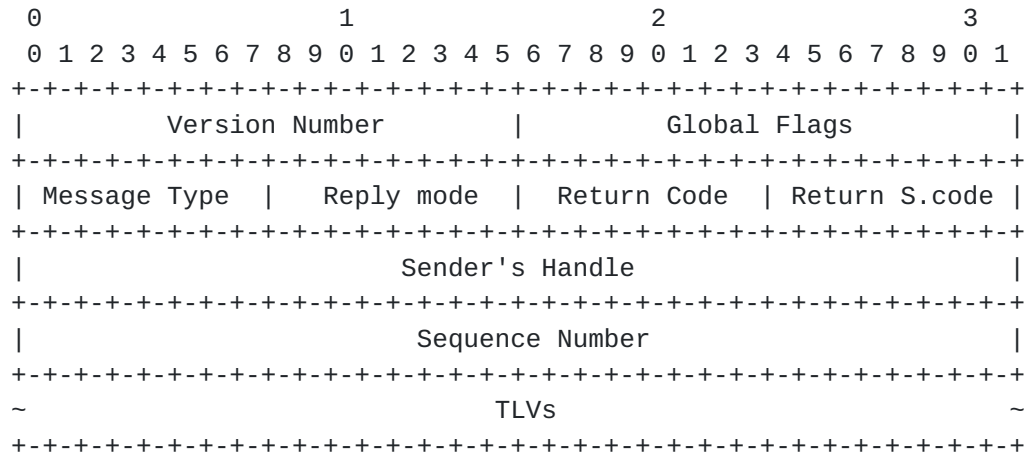
                  Figure 3: SFC Echo Request/Reply format

   The interpretation of the fields is as follows:

      The Version reflects the current version. The version number is
      to be incremented whenever a change is made that affects the
      ability of an implementation to parse or process control packet
      correctly.

      The Global Flags is a bit vector field.

      The Message Type filed reflects the type of the packet. Value
      TBA3 identifies echo request and TBA4 - echo reply

      The Reply Mode defines the type of the return path requested by
      the sender of the echo request.

      Return Codes and Subcodes can be used to inform the sender about
      the result of processing its request.

      The Sender's Handle is filled in by the sender and returned
      unchanged by the receiver in the echo reply. The sender MAY use a
      pseudo-random number generator (PRNG) to set the value of the
      Sender's Handle field. The value of the Sender's Handle field
      SHOULD NOT be changed in the course of the test session.

   The Sequence Number is assigned by the sender and can be (for
   example) used to detect missed replies. The value of the Sequence
   Number field SHOULD be monotonically increasing in the course of
   the test session.

   TLVs (Type-Length-Value tuples) have the two octets long Type
   field, two octets long Length field that is the length of the
   Value field in octets.

5.1. SFC Echo Request Transmission

   SFC echo request control packet MUST use the appropriate
   encapsulation of the monitored SFP. If Network Service Header (NSH)
   is used, echo request MUST set O bit, as defined in [RFC8300]. SFC
   NSH MUST be immediately followed by the SFC Active OAM Header defined
   in Section 4. Message Type field in the SFC Active OAM Header MUST
   be set to SFC Echo Request/Echo Reply value (TBA2) per Section 8.2.

   Value of the Reply Mode field MAY be set to:

   o Do Not Reply (TBA5) if one-way monitoring is desired. If the echo
      request is used to measure synthetic packet loss; the receiver may
      report loss measurement results to a remote node.

   o Reply via an IPv4/IPv6 UDP Packet (TBA6) value likely will be the
      most used.

   o Reply via Application Level Control Channel (TBA7) value if the
      SFP may have bi-directional paths.

   o Reply via Specified Path (TBA7) value to enforce the use of the
      particular return path specified in the included TLV to verify bi-
      directional continuity and also increase the robustness of the
      monitoring by selecting a more stable path.

5.2. SFC Echo Request Reception

5.3. SFC Echo Reply Transmission

   The Reply Mode field directs whether and how the echo reply message
   should be sent. The sender of the echo request MAY use TLVs to
   request that the corresponding echo reply is transmitted over the
   specified path. Value TBA3 is referred to as "Do not reply" mode and
   suppresses transmission of echo reply packet. The default value
   (TBA6) for the Reply mode field requests the responder to send the
   echo reply packet out-of-band as IPv4 or IPv6 UDP packet.

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.
Because SFC NSH does not identify the ingress of the SFP the echo
request, the source ID MUST be included in the message and used as
the IP destination address for IP/UDP encapsulation of the SFC echo
reply. The sender of the SFC echo request MUST include SFC Source
TLV Figure 4.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |   SFC OAM Source ID Type    |             Length              |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                             Value                             |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
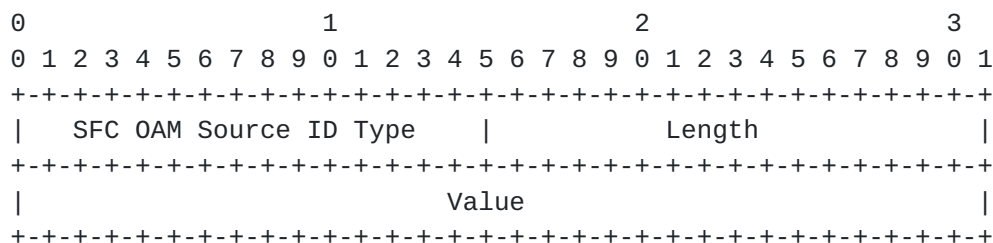
Figure 4: SFC Source TLV

where

   SFC OAM Source Id Type is two octets in length and has the value
   of TBA9 Section 8.6.

   Length is two octets long field, and the value equals the length
   of the Value field in octets.

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

The UDP destination port for SFC Echo Reply TBA10 will be allocated
by IANA Section 8.7.

5.4. Overlay Echo Reply Reception

6. Security Considerations

Overlay Echo Request/Reply operates within the domain of the overlay
network and thus inherits any security considerations that apply to
the use of that overlay technology and, consequently, underlay data
plane. Also, the security needs for SFC echo request/reply are
similar to those of ICMP ping [RFC0792], [RFC4443] and MPLS LSP ping
[RFC8029].

There are at least three approaches of attacking a node in the
overlay network using the mechanisms defined in the document. One is
a Denial-of-Service attack, by sending SFC ping to overload an
element of the SFC. The second may use spoofing, hijacking,

replying, or otherwise tampering with SFC echo requests and/or
replies to misrepresent, alter operator's view of the state of the
SFC. The third is an unauthorized source using an SFC echo request/
reply to obtain information about the SFC and/or its elements, e.g.
SFF or SF.

It is RECOMMENDED that implementations throttle the SFC ping traffic
going to the control plane to mitigate potential Denial-of-Service
attacks.

Reply and spoofing attacks involving faking or replying SFC echo
reply messages would have to match the Sender's Handle and Sequence
Number of an outstanding SFC echo request message which is highly
unlikely. Thus the non-matching reply would be discarded.

To protect against unauthorized sources trying to obtain information
about the overlay and/or underlay an implementation MAY check that
the source of the echo request is indeed part of the SFP.

7. Acknowledgments

Authors greatly appreciate thorough review and the most helpful
comments from Dan Wing.

8. IANA Considerations

8.1. SFC Active OAM Protocol

IANA is requested to assign a new type from the SFC Next Protocol
registry as follows:

```
                +-------+---------------+---------------+
                | Value |  Description  | Reference     |
                +-------+---------------+---------------+
                | TBA1  | SFC Active OAM | This document |
                +-------+---------------+---------------+
```

                     Table 1: SFC Active OAM Protocol

8.2. SFC Active OAM Message Type

IANA is requested to create a new registry called "SFC Active OAM
Message Type". All code points in the range 1 through 32767 in this
registry shall be allocated according to the "IETF Review" procedure
as specified in [RFC8126]. Remaining code points to be allocated
according to the table Table 2:

```
      +---------------+------------+------------------------+
      | Value         | Description | Reference             |
      +---------------+------------+------------------------+
      | 0             |   Reserved  |                        |
      | 1 - 32767     |   Reserved  | IETF Consensus         |
      | 32768 - 65530 |   Reserved  | First Come First Served |
      | 65531 - 65534 |   Reserved  | Private Use            |
      | 65535         |   Reserved  |                        |
      +---------------+------------+------------------------+
```

                 Table 2: SFC Active OAM Message Type

   IANA is requested to assign new type from the SFC Active OAM Message
   Type registry as follows:

```
      +-------+-----------------------------+---------------+
      | Value |          Description         | Reference     |
      +-------+-----------------------------+---------------+
      | TBA2  | SFC Echo Request/Echo Reply | This document |
      +-------+-----------------------------+---------------+
```

              Table 3: SFC Echo Request/Echo Reply Type

8.3. SFC Echo Request/Echo Reply Parameters

   IANA is requested to create new SFC Echo Request/Echo Reply
   Parameters registry.

8.4. SFC Echo Request/Echo Reply Message Types

   IANA is requested to create in the SFC Echo Request/Echo Reply
   Parameters registry the new sub-registry Message 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 [RFC8126]
   and assign values as follows:

```
+------------+-----------------+------------------------+
| Value      | Description     | Reference              |
+------------+-----------------+------------------------+
| 0          |    Reserved     |                        |
| TBA3       | SFC Echo Request | This document         |
| TBA4       |  SFC Echo Reply | This document          |
| TBA4+1-191 |    Unassigned   | IETF Review            |
| 192-251    |    Unassigned   | First Come First Served |
| 252-254    |    Unassigned   | Private Use            |
| 255        |    Reserved     |                        |
+------------+-----------------+------------------------+
```

             Table 4: SFC Echo Request/Echo Reply Message Types

8.5. SFC Echo Reply Modes

   IANA is requested to create in the SFC Echo Request/Echo Reply
   Parameters registry the new sub-registry Reply Modes All code points
   in the range 1 through 191 in this registry shall be allocated
   according to the "IETF Review" procedure as specified in [RFC8126]
   and assign values as follows:

```
+------------+--------------------------------+-------------------+
| Value      |           Description          | Reference         |
+------------+--------------------------------+-------------------+
| 0          |            Reserved            |                   |
| TBA5       |          Do Not Reply          | This document     |
| TBA6       |    Reply via an IPv4/IPv6 UDP   | This document     |
|            |             Packet             |                   |
| TBA7       |    Reply via Application Level  | This document     |
|            |         Control Channel        |                   |
| TBA8       |    Reply via Specified Path     | This document     |
| TBA8+1-191 |           Unassigned           | IETF Review       |
| 192-251    |           Unassigned           | First Come First  |
|            |                                | Served            |
| 252-254    |           Unassigned           | Private Use       |
| 255        |            Reserved            |                   |
+------------+--------------------------------+-------------------+
```

                     Table 5: SFC Echo Reply Modes

8.6. 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
   [RFC8126]. 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 [RFC8126]. Remaining code points
are allocated according to the Table 6:

```
+---------------+-------------+-------------------------+
| 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 6: SFC TLV Type Registry

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

```
+-------+-------------------+---------------+
| Value |    Description    | Reference     |
+-------+-------------------+---------------+
| TBA9  | Source IP Address | This document |
+-------+-------------------+---------------+
```

Table 7: SFC OAM Source IP Address Type

8.7. SFC OAM UDP Port

IANA is requested to allocate UDP port number according to

```
+---------+--------+-----------+---------+-------------+----------+
| Service | Port   | Transport | Descrip | Semantics   | Referenc |
| Name    | Number | Protocol  | tion    | Definition  | e        |
+---------+--------+-----------+---------+-------------+----------+
| SFC OAM | TBA10  | UDP       | SFC OAM | Section 5.3 | This     |
|         |        |           |         |             | document |
+---------+--------+-----------+---------+-------------+----------+
```

Table 8: SFC OAM Port

9. References

9.1. Normative References

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", BCP 14, RFC 2119,
              DOI 10.17487/RFC2119, March 1997,
              <https://www.rfc-editor.org/info/rfc2119>.

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

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

9.2. Informative References

   [RFC0792] Postel, J., "Internet Control Message Protocol", STD 5,
              RFC 792, DOI 10.17487/RFC0792, September 1981,
              <https://www.rfc-editor.org/info/rfc792>.

   [RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet
              Control Message Protocol (ICMPv6) for the Internet
              Protocol Version 6 (IPv6) Specification", STD 89,
              RFC 4443, DOI 10.17487/RFC4443, March 2006,
              <https://www.rfc-editor.org/info/rfc4443>.

   [RFC7665] Halpern, J., Ed. and C. Pignataro, Ed., "Service Function
              Chaining (SFC) Architecture", RFC 7665,
              DOI 10.17487/RFC7665, October 2015,
              <https://www.rfc-editor.org/info/rfc7665>.

   [RFC7799] Morton, A., "Active and Passive Metrics and Methods (with
              Hybrid Types In-Between)", RFC 7799, DOI 10.17487/RFC7799,
              May 2016, <https://www.rfc-editor.org/info/rfc7799>.

   [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,
              <https://www.rfc-editor.org/info/rfc8029>.

   [RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for
              Writing an IANA Considerations Section in RFCs", BCP 26,
              RFC 8126, DOI 10.17487/RFC8126, June 2017,
              <https://www.rfc-editor.org/info/rfc8126>.

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