IPv6 Encapsulation for SFC and IFIT
draft-li-6man-ipv6-sfc-ifit-02

Abstract

   Service Function Chaining (SFC) and In-situ Flow Information
   Telemetry (IFIT) are important path services along with the packets.
   In order to support these services, several encapsulations have been
   defined. The document analyzes the problems of these encapsulations
   in the IPv6 scenario and proposes the possible optimized
   encapsulation for IPv6.

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].

Table of Contents

1. Introduction

Service Function Chaining (SFC) [RFC7665] and In-situ Flow
Information Telemetry (IFIT) [I-D.song-opsawg-ifit-framework] are
important path services along with the packets. In order to support
these services, several encapsulations have been defined. Network
Service Header (NSH) is defined in [RFC8300] as the encapsulation for
SFC. For IFIT encapsulations, In-situ OAM (iOAM) Header is defined
in [I-D.ietf-ippm-ioam-data] and Postcard-Based Telemetry (PBT)
Header is defined in [I-D.song-ippm-postcard-based-telemetry].
Inband Flow Analyzer (IFA) is also defined in [I-D.kumar-ippm-ifa] to
record flow specific information from an end station and/or switches
across a network. In the application scenario of IPv6, these
encapsulations propose challenges for the data plane. The document
analyzes the problems and proposes the possible optimized
encapsulation for IPv6.

2. Terminology

SFC: Service Function Chaining

IFIT: In-situ Flow Information Telemetry

IOAM: In-situ OAM

PBT: Postcard-Based Telemetry

IFA: Inband Flow Analyzer

SRH: Segment Routing Header

3. Problem Statement

The problems posed by the current encapsulations for SFC and IFIT in the application scenarios of IPv6 and SRv6 include:

1. According to the encapsulation order recommended in [RFC8200], if the IOAM is encapsulated in the IPv6 Hop-by-Hop options header, in the incremental trace mode of IOAM as the number of nodes traversed by the IPv6 packets increases, the recorded IOAM information will increase accordingly. This will increase the length of the Hop-by-Hop options header and cause increasing difficulties in reading the subsequent Segment Routing Extension Header (SRH) [I-D.ietf-6man-segment-routing-header] and thereby reduce the forwarding performance of the data plane greatly.

2. With the introduction of SRv6 network programming [I-D.ietf-spring-srv6-network-programming], the path services along with the IPv6 packets can be processed at all the IPv6 network nodes or only at the SRv6 enabled network nodes along the path. It is necessary to distinguish the encapsulations for the specific path service which should be processed by the IPv6 path or the SRv6 path.

3. Both NSH and IOAM need the Metadata field to record metadata information. However currently these metadata has to be recorded separately which may generate redundant metadata information or increase the cost of process.

4. There is unnecessary inconsistency in the current encapsulations for IOAM, IFA and PBT in the IPv6 scenario. Especially it seems unnecessary to define a new specific IPv6 header for IFA, i.e. IFA header.

4. Design Consideration

To solve the problems stated above, in the application scenarios of
IPv6 and SRv6, the encapsulations of SFC and IFIT can be optimized
with the following design considerations:

o To separate the SFC/IFIT path service into two parts, i.e.
   instruction and recording parts. The instruction part (normally
   with fixed length) can be placed in the front IPv6 extension
   headers including Hop-by-Hop options header, Destination options
   header, Routing header, etc. while the recording part can be
   placed in the back IPv6 extension headers such as being placed
   after IPv6 Routing Header. In this way the path service
   instruction in the IPv6 extension headers can be fixed as much as
   possible to facilitate hardware process to keep forwarding
   performance while the SFC/IFIT metadata recording part is placed
   afterwards which enables to stop recording when too much recording
   information has to be carried to reach the limitation of hardware
   process.

o To define SFC/IFIT path service instructions as IPv6 options
   uniformly whichs can be placed either in the Hop-by-hop options
   which indicates the path service processed by all IPv6 enabled
   nodes along the path or in the SRH option TLVs which indicates the
   path service processed only by the SRv6 nodes along the SRv6 path
   indicated by the Segment List in the SRH.

o To define a unified IPv6 metadata header which can be used as a
   container to record the service metadata of SFC, IFIT and other
   possible path services.

According to the above design optimization consideration, in the
application scenarios of IPv6 and SRv6 the encapsulations for SFC and
IFIT can be defined as below.

4.1. Service Options

1. NSH Service Option

```
 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
                             +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                             | Option Type  | Opt Data Len  |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |Ver|O|U|    TTL    |    Length  |U|U|U|U|MD Type| Next Protocol |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |          Service Path Identifier           | Service Index |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

           Figure 1. IPv6 Options with NSH instructions

Option Type: TBD_0

Opt Data Len: 8 octets.

Other fields: refer to [RFC8300].

2.   IOAM Service Option

```
 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
                             +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                             | Option Type  | Opt Data Len  |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |       Namespace-ID        |NodeLen  | Flags | RemainingLen|
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |            IOAM-Trace-Type            |    Reserved     |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

           Figure 2. IPv6 Options with IOAM instructions

Option Type: TBD_1

Opt Data Len: 8 octets.

Other fields: refer to [I-D.ietf-ippm-ioam-data].

3.   PBT Service Option

```
 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
                                +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                                | Option Type   | Opt Data Len  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Next Header   | TIH Length    |    Reserved   |   Hop Count   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Flow ID                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Flow ID                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Sequence Number                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Data Set ID                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

           Figure 3. IPv6 Options with PBT instructions

Option Type: TBD_2

Opt Data Len: 20 octets.

Other fields: refer to [I-D.song-ippm-postcard-based-telemetry].

4. IFA Service Option

```
 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
                                +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                                | Option Type   | Opt Data Len  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Ver=2.0|  GNS  |NextHdr = IP_xx|R|R|R|M|T|I|T|C|   Max Length  |
|       |       |               | | | |F|S| |A| |               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
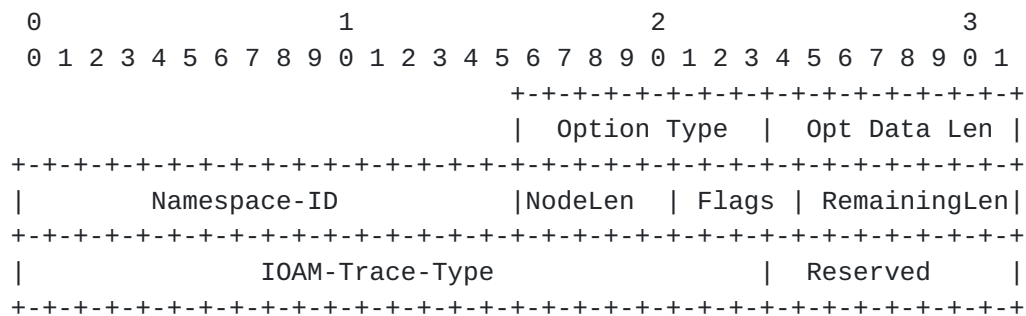   Figure 4. IPv6 Options with IFA instructions

Option Type: TBD_3

Opt Data Len: 4 octets.
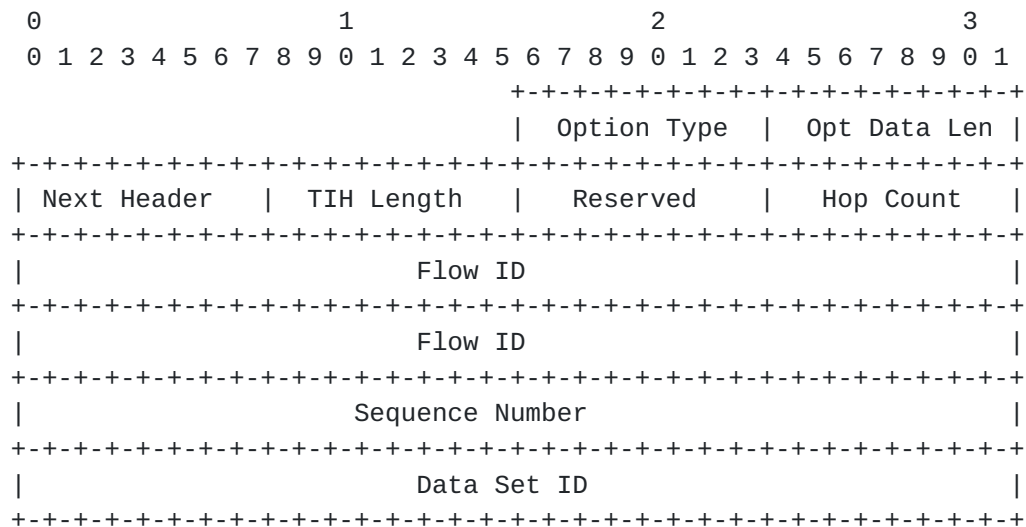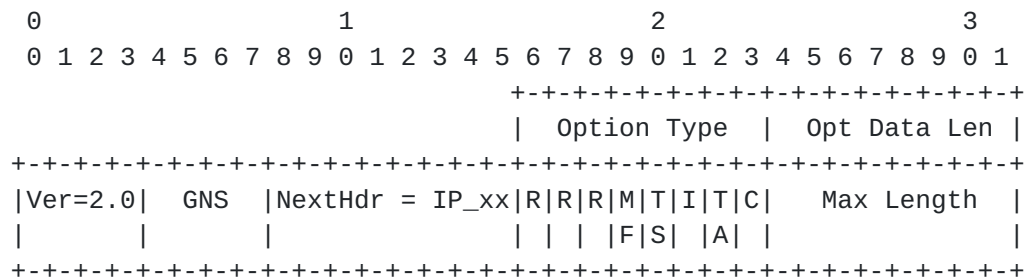
Other fields: refer to [I-D.kumar-ippm-ifa].

These options can be put in the IPv6 Hop-by-Hop Options Header or SRH
TLV.

4.2. IPv6 Service Metadata Options

   As introduced in [I-D.li-6man-enhanced-extension-header], IPv6
   Metadata Header is defined as a new type of IPv6 extension header.
   The metadata is the information recorded by each hop for specific
   path services, and carried in corresponding service metadata options.
   The length of the metadata is variable.

4.2.1. SFC Service Metadata Option

   For the SFC service, the corresponding SFC service metadata option is
   defined as shown in Figure 5.

```
 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 Type    |    Length     |            Reserved           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      SFC Metadata Class        |     Type      |U|   Length   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Variable-Length Metadata                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

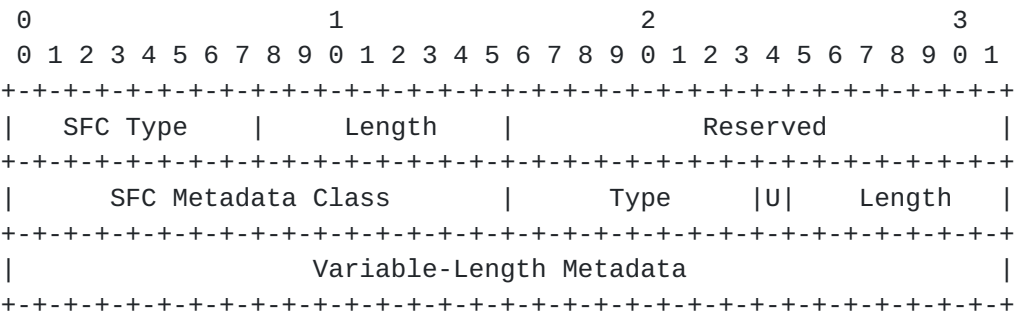   Figure 5. SFC Service Metadata

      SFC Type   8-bit identifier of the service type, i.e. SFC.
                        The value is TBD-4.

      Length   8-bit unsigned integer. Length of the
                        Service Metadata field, in octets.

      Metadata Class   Defines the scope of the Type field to
                        provide a hierarchical namespace. IANA has
                        set up the "NSH MD Class" registry, which
                        contains 16-bit values [RFC8300].

      Type   Indicates the explicit type of metadata
                        being carried. The definition of the Type is
                        the responsibility of the MD Class owner.

      Unassigned bit   One unassigned bit is available for future use.
                        This bit MUST NOT be set, and it MUST be
                        ignored on receipt.

      Length   Indicates the length of the variable-length
                        metadata, in bytes. Detailed specification
                        in [RFC8300].

4.2.2. **IOAM Service Metadata Option**

   For the IOAM service, the corresponding IOAM service metadata option
   is defined as shown in Figure 6.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   IOAM Type   |    Length     |           Reserved            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               ~
|            IOAM Service Metadata Options (variable)           |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

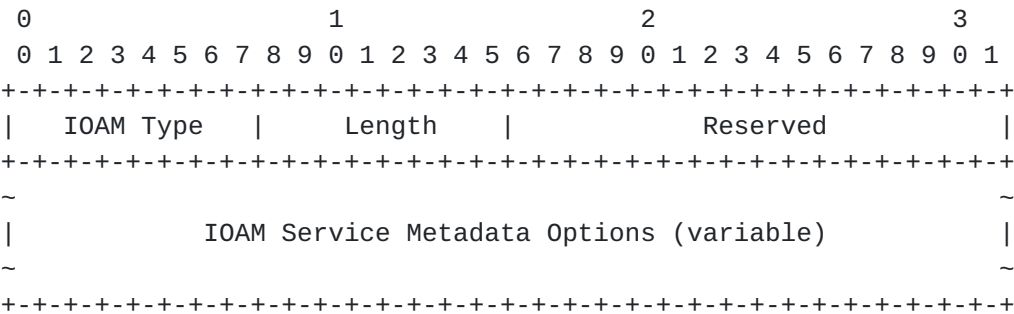   Figure 6. IOAM Service Metadata

      IOAM Type   8-bit identifier of the IOAM Service Metadata
                          type. The value is TBD-5.

      Length   8-bit unsigned integer. Length of the
                          IOAM Service Metadata field, in octets.

      RESERVED   8-bit reserved field MUST be set to zero upon
                          transmission and ignored upon receipt.

      IOAM Service   IOAM option data is present as specified by the
      Metadata Options   IOAM Type field, and is defined in Section 4 of
                          [I-D.ietf-ippm-ioam-data].

   All the IOAM IPv6 options require 4n alignment. This ensures that 4
   octet fields specified in [I-D.ietf-ippm-ioam-data] such as transit
   delay are aligned at a multiple-of-4 offset from the start of the
   IPv6 Metadata header.

   In addition, to maintain IPv6 extension header 8-octet alignment and
   avoid the need to add or remove padding at every hop, the Trace-Type
   for Incremental Tracing Option in IPv6 MUST be selected such that the
   IOAM node data length is a multiple of 8-octets.

4.2.3. **IFA Service Metadata Option**

   For the IOAM service, the corresponding IOAM service metadata option
   is defined as shown in Figure 6.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   IFA Type    |    Length     |             Reserved          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               ~
|             IFA Service Metadata Options (variable)           |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
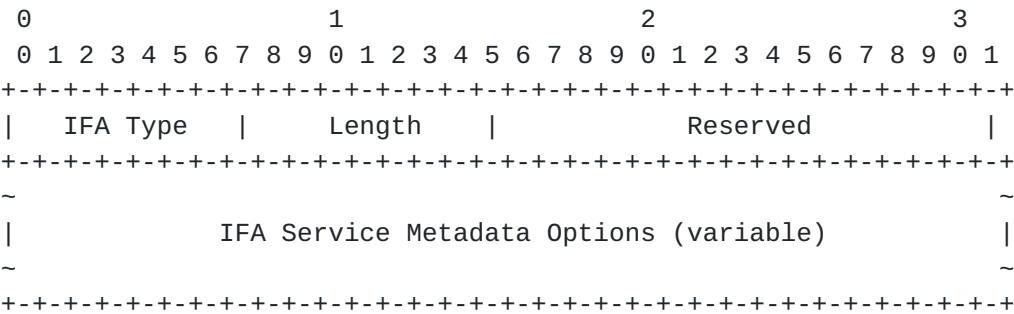
                   Figure 6. IFA Service Metadata

     IFA Type   8-bit identifier of the IFA Service Metadata
                         type. The value is TBD-6.

     Length   8-bit unsigned integer. Length of the
                         IOAM Service Metadata field, in octets.

     RESERVED   8-bit reserved field MUST be set to zero upon
                         transmission and ignored upon receipt.

     IFA Service   IFA option data is present as specified by the
     Metadata Options   IFA Type field.

5. IANA Considerations

     Value   Description   Reference

     TBD_0   NSH Service Option   [This draft]
     TBD_1   IOAM Service Option   [This draft]
     TBD_2   PBT Service Option   [This draft]
     TBD_3   IFA Service Option   [This draft]
     TBD_4   SFC Service Metadata Type   [This draft]
     TBD_5   IOAM Service Metadata Type   [This draft]
     TBD_6   IFA Service Metadata Type   [This draft]

6. Security Considerations

     TBD.

7. References

7.1. Normative References

   [I-D.guichard-spring-nsh-sr]
              Guichard, J., Song, H., Tantsura, J., Halpern, J.,
              Henderickx, W., Boucadair, M., and S. Hassan, "NSH and
              Segment Routing Integration for Service Function Chaining
              (SFC)", draft-guichard-spring-nsh-sr-01 (work in
              progress), March 2019.

   [I-D.ietf-6man-segment-routing-header]
              Filsfils, C., Dukes, D., Previdi, S., Leddy, J.,
              Matsushima, S., and d. daniel.voyer@bell.ca, "IPv6 Segment
              Routing Header (SRH)", draft-ietf-6man-segment-routing-
              header-22 (work in progress), August 2019.

   [I-D.ietf-ippm-ioam-data]
              Brockners, F., Bhandari, S., Pignataro, C., Gredler, H.,
              Leddy, J., Youell, S., Mizrahi, T., Mozes, D., Lapukhov,
              P., Chang, R., daniel.bernier@bell.ca, d., and J. Lemon,
              "Data Fields for In-situ OAM", draft-ietf-ippm-ioam-
              data-06 (work in progress), July 2019.

   [I-D.ietf-spring-srv6-network-programming]
              Filsfils, C., Camarillo, P., Leddy, J.,
              daniel.voyer@bell.ca, d., Matsushima, S., and Z. Li, "SRv6
              Network Programming", draft-ietf-spring-srv6-network-
              programming-01 (work in progress), July 2019.

   [I-D.kumar-ippm-ifa]
              Kumar, J., Anubolu, S., Lemon, J., Manur, R., Holbrook,
              H., Ghanwani, A., Cai, D., Ou, H., and L. Yizhou, "Inband
              Flow Analyzer", draft-kumar-ippm-ifa-01 (work in
              progress), February 2019.

   [I-D.song-ippm-postcard-based-telemetry]
              Song, H., Zhou, T., Li, Z., Shin, J., and K. Lee,
              "Postcard-based On-Path Flow Data Telemetry", draft-song-
              ippm-postcard-based-telemetry-04 (work in progress), June
              2019.

   [I-D.song-opsawg-ifit-framework]
              Song, H., Li, Z., Zhou, T., Qin, F., Shin, J., and J. Jin,
              "In-situ Flow Information Telemetry Framework", draft-
              song-opsawg-ifit-framework-04 (work in progress),
              September 2019.

   [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>.

   [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>.

   [RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6
              (IPv6) Specification", STD 86, RFC 8200,
              DOI 10.17487/RFC8200, July 2017,
              <https://www.rfc-editor.org/info/rfc8200>.

   [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>.

7.2. Informative References

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

Authors' Addresses

   Zhenbin Li
   Huawei Technologies
   Huawei Bld., No.156 Beiqing Rd.
   Beijing 100095
   China

   Email: lizhenbin@huawei.com

   Shuping Peng
   Huawei Technologies
   Huawei Bld., No.156 Beiqing Rd.
   Beijing 100095
   China

   Email: pengshuping@huawei.com

Kihoon LEE
LG U+
71, Magokjungang 8-ro, Gangseo-gu
Seoul
Republic of Korea

Email: soho8416@lguplus.co.kr