BGP SR Policy Extensions to Enable IFIT
draft-qin-idr-sr-policy-ifit-04

Abstract

   Segment Routing (SR) policy is a set of candidate SR paths consisting
   of one or more segment lists and necessary path attributes. It
   enables instantiation of an ordered list of segments with a specific
   intent for traffic steering. In-situ Flow Information Telemetry
   (IFIT) refers to network OAM data plane on-path telemetry techniques,
   in particular the most popular are In-situ OAM (IOAM) and Alternate
   Marking. This document defines extensions to BGP to distribute SR
   policies carrying IFIT information. So that IFIT methods can be
   enabled automatically when the SR policy is applied.

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

   Segment Routing (SR) policy [I-D.ietf-spring-segment-routing-policy]
   is a set of candidate SR paths consisting of one or more segment
   lists and necessary path attributes. It enables instantiation of an
   ordered list of segments with a specific intent for traffic steering.

   In-situ Flow Information Telemetry (IFIT) denotes a family of flow-
   oriented on-path telemetry techniques (e.g. IOAM, Alternate
   Marking), which can provide high-precision flow insight and real-time
   network issue notification (e.g., jitter, latency, packet loss).In

particular, IFIT refers to network OAM data plane on-path telemetry
techniques, including In-situ OAM (IOAM) [I-D.ietf-ippm-ioam-data]
and Alternate Marking [RFC8321]. It can provide flow information on
the entire forwarding path on a per-packet basis in real time.

An automatic network requires the Service Level Agreement (SLA)
monitoring on the deployed service. So that the system can quickly
detect the SLA violation or the performance degradation, hence to
change the service deployment. For this reason, the SR policy native
IFIT can facilitate the closed loop control and enable the automation
of SR service.

This document defines extensions to Border Gateway Protocol (BGP) to
distribute SR policies carrying IFIT information. So that IFIT
behavior can be enabled automatically when the SR policy is applied.

This BGP extension allows to signal the IFIT capabilities together
with the SR-policy. In this way IFIT methods are automatically
activated and running. The flexibility and dynamicity of the IFIT
applications are given by the use of additional functions on the
controller and on the network nodes, but this is out of scope here.

2. Motivation

IFIT Methods are being introduced in multiple protocols and below is
a proper picture of the relevant documents for Segment Routing.
Indeed the IFIT methods are becoming mature for Segment Routing over
the MPLS data plane (SR-MPLS) and Segment Routing over IPv6 data
plane (SRv6), that is the main focus of this draft:

   IOAM: the reference documents for the data plane are
   [I-D.ietf-ippm-ioam-ipv6-options] for SRv6 and
   [I-D.gandhi-mpls-ioam-sr] for SR-MPLS.

   Alternate Marking: the reference documents for the data plane are
   [I-D.ietf-6man-ipv6-alt-mark] for SRv6 and
   [I-D.ietf-mpls-rfc6374-sfl], [I-D.gandhi-mpls-rfc6374-sr] for SR-
   MPLS.

The definition of these data plane IFIT methods for SR-MPLS and SRv6
imply requirements for various routing protocols, such as BGP, and
this document aims to define BGP extensions to distribute SR policies
carrying IFIT information. This allows to signal the IFIT
capabilities so IFIT methods are automatically configured and ready
to run when the SR Policy candidate paths are distributed through
BGP.

It is to be noted that, for PCEP, [I-D.chen-pce-pcep-ifit] proposes
the extensions to PCEP to distribute paths carrying IFIT information
and therefore to enable IFIT methods for SR policy too.

3. IFIT methods for SR Policy

In-situ Operations, Administration, and Maintenance (IOAM)
[I-D.ietf-ippm-ioam-data] records operational and telemetry
information in the packet while the packet traverses a path between
two points in the network. In terms of the classification given in
RFC 7799 [RFC7799] IOAM could be categorized as Hybrid Type 1. IOAM
mechanisms can be leveraged where active OAM do not apply or do not
offer the desired results. When SR policy enables the IOAM, the IOAM
header will be inserted into every packet of the traffic that is
steered into the SR paths.

The Alternate Marking [RFC8321]technique is an hybrid performance
measurement method, per RFC 7799 [RFC7799] classification of
measurement methods. Because this method is based on marking
consecutive batches of packets. It can be used to measure packet
loss, latency, and jitter on live traffic.

This document aims to define the control plane. While the relevant
documents for the data plane application of IOAM and Alternate
Marking are respectively [I-D.ietf-ippm-ioam-ipv6-options] and
[I-D.ietf-6man-ipv6-alt-mark] for Segment Routing over IPv6 data
plane (SRv6).

4. IFIT Attributes in SR Policy

As defined in [I-D.ietf-idr-segment-routing-te-policy], the SR Policy
encoding structure is as follows:

```
SR Policy SAFI NLRI: <Distinguisher, Policy-Color, Endpoint>
Attributes:
   Tunnel Encaps Attribute (23)
      Tunnel Type: SR Policy
          Binding SID
          Preference
          Priority
          Policy Name
          Explicit NULL Label Policy (ENLP)
          Segment List
              Weight
              Segment
              Segment
              ...
          ...
```

A candidate path includes multiple SR paths, each of which is
specified by a segment list. IFIT can be applied to the candidate
path, so that all the SR paths can be monitored in the same way. The
new SR Policy encoding structure is expressed as below:

```
     SR Policy SAFI NLRI: <Distinguisher, Policy-Color, Endpoint>
     Attributes:
        Tunnel Encaps Attribute (23)
           Tunnel Type: SR Policy
                 Binding SID
                 Preference
                 Priority
                 Policy Name
                 Explicit NULL Label Policy (ENLP)
                 IFIT Attributes
                 Segment List
                     Weight
                     Segment
                     Segment
                     ...
              ...
```

IFIT attributes can be attached at the candidate path level as sub-
TLVs. There may be different IFIT tools. The following sections
will describe the requirement and usage of different IFIT tools, and
define the corresponding sub-TLV encoding in BGP.

Note that the IFIT attributes here described can also be generalized
and included as sub-TLVs for other SAFIs and NLRIs.

5. IFIT Attributes Sub-TLV

The format of the IFIT Attributes Sub-TLV is defined as follows:

```
 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
                     +---------------+---------------+
                     |     Type      |    Length     |
 +-------------------------------+---------------+---------------+
 |                                                               |
 //                          sub-TLVs                           //
 |                                                               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
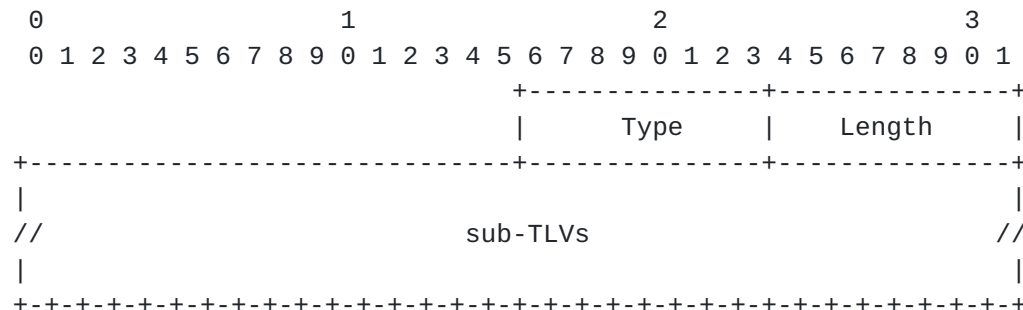
                  Fig. 1 IFIT Attributes Sub-TLV

Where:

Type: to be assigned by IANA.

Length: the total length of the value field not including Type and Length fields.

sub-TLVs currently defined:

   * IOAM Pre-allocated Trace Option Sub-TLV,

   * IOAM Incremental Trace Option Sub-TLV,

   * IOAM Directly Export Option Sub-TLV,

   * IOAM Edge-to-Edge Option Sub-TLV,

   * Enhanced Alternate Marking (EAM) sub-TLV.

The presence of the IFIT Attributes Sub-TLV implies support of IFIT methods (IOAM and/or Alternate Marking). It is worth mentioning that IOAM and Alternate Marking can be activated one at a time or can coexist; so it is possible to have only IOAM or only Alternate Marking enabled as Sub-TLVs. The sub-TLVs currently defined for IOAM and Alternate Marking are detailed in the next sections.

5.1. IOAM Pre-allocated Trace Option Sub-TLV

The IOAM tracing data is expected to be collected at every node that a packet traverses to ensure visibility into the entire path a packet takes within an IOAM domain. The preallocated tracing option will create pre-allocated space for each node to populate its information.

The format of IOAM pre-allocated trace option sub-TLV is defined as follows:

```
 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
+---------------+---------------+------------------------------+
|    Type=1     |   Length=6    |   Namespace ID               |
+---------------+---------------+-------------+--------+-------+
|          IOAM Trace Type                    | Flags  | Rsvd  |
+---------------------------------------------+--------+-------+
```
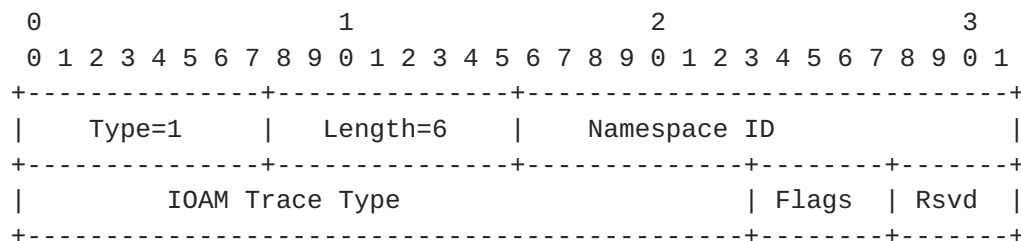
          Fig. 2 IOAM Pre-allocated Trace Option Sub-TLV

Where:

Type: 1 (to be assigned by IANA).

Length: 6, it is the total length of the value field (not including Type and Length fields).

Namespace ID: A 16-bit identifier of an IOAM-Namespace. The definition is the same as described in section 4.4 of [I-D.ietf-ippm-ioam-data].

IOAM Trace Type: A 24-bit identifier which specifies which data types are used in the node data list. The definition is the same as described in section 4.4 of [I-D.ietf-ippm-ioam-data].

Flags: A 4-bit field. The definition is the same as described in [I-D.ietf-ippm-ioam-flags] and section 4.4 of [I-D.ietf-ippm-ioam-data].

Rsvd: A 4-bit field reserved for further usage. It MUST be zero.

5.2. IOAM Incremental Trace Option Sub-TLV

The incremental tracing option contains a variable node data fields where each node allocates and pushes its node data immediately following the option header.

The format of IOAM incremental trace option sub-TLV is defined as follows:

```
 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
+---------------+---------------+-------------------------------+
|    Type=2     |   Length=6    |     Namespace ID              |
+---------------+---------------+--------------+--------+-------+
|            IOAM Trace Type                   | Flags  | Rsvd  |
+----------------------------------------------+--------+-------+
```
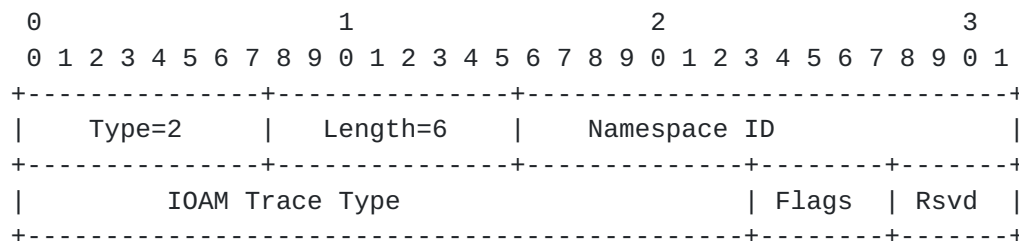
             Fig. 3 IOAM Incremental Trace Option Sub-TLV

Where:

Type: 2 (to be assigned by IANA).

Length: 6, it is the total length of the value field (not including Type and Length fields).

All the other fields definistion is the same as the pre-allocated trace option sub-TLV in section 4.1.

5.3. IOAM Directly Export Option Sub-TLV

IOAM directly export option is used as a trigger for IOAM data to be directly exported to a collector without being pushed into in-flight data packets.

The format of IOAM directly export option sub-TLV is defined as follows:

```
 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
                                +---------------+---------------+
                                |    Type=3     |   Length=12   |
+-----------------------------------------------+---------------+
|          Namespace ID          |             Flags            |
+-------------------------------+---------------+---------------+
|              IOAM Trace Type                  |     Rsvd      |
+-----------------------------------------------+---------------+
|                         Flow ID                               |
+--------------------------------------------------------------+
```
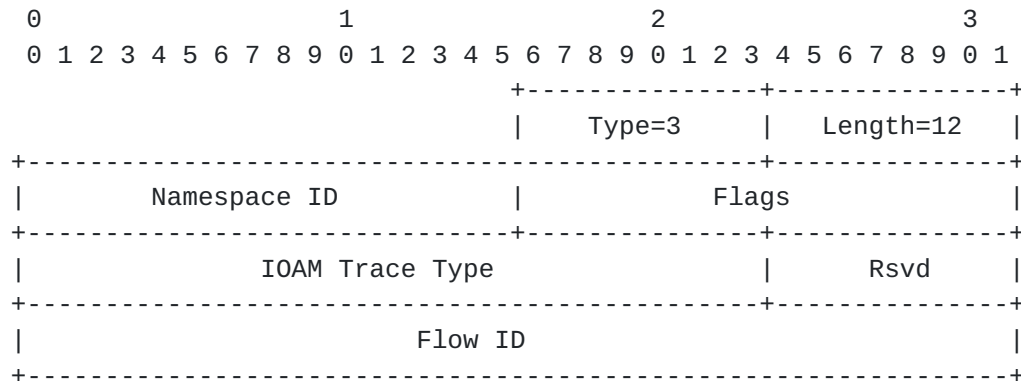
             Fig. 4 IOAM Directly Export Option Sub-TLV

Where:

Type: 3 (to be assigned by IANA).

Length: 12, it is the total length of the value field (not including Type and Length fields).

Namespace ID: A 16-bit identifier of an IOAM-Namespace. The definition is the same as described in section 4.4 of [I-D.ietf-ippm-ioam-data].

IOAM Trace Type: A 24-bit identifier which specifies which data types are used in the node data list. The definition is the same as described in section 4.4 of [I-D.ietf-ippm-ioam-data].

Flags: A 16-bit field. The definition is the same as described in section 3.2 of [I-D.ietf-ippm-ioam-direct-export].

Flow ID: A 32-bit flow identifier. The definition is the same as described in section 3.2 of [I-D.ietf-ippm-ioam-direct-export].

Rsvd: A 4-bit field reserved for further usage. It MUST be zero.

5.4. IOAM Edge-to-Edge Option Sub-TLV

The IOAM edge to edge option is to carry data that is added by the
IOAM encapsulating node and interpreted by IOAM decapsulating node.

The format of IOAM edge-to-edge option sub-TLV is defined as follows:

```
 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
                                +---------------+---------------+
                                |    Type=4     |   Length=4    |
+-----------------------------------------------+---------------+
|        Namespace ID           |        IOAM E2E Type          |
+------------------------------+-------------------------------+
```
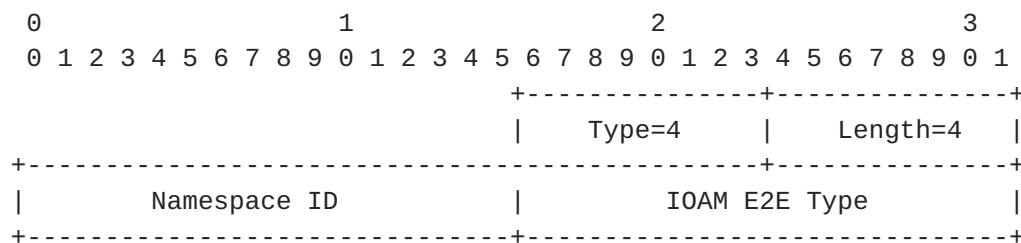
               Fig. 5 IOAM Edge-to-Edge Option Sub-TLV

Where:

Type: 4 (to be assigned by IANA).

Length: 4, it is the total length of the value field (not including
Type and Length fields).

Namespace ID: A 16-bit identifier of an IOAM-Namespace. The
definition is the same as described in section 4.6 of
[I-D.ietf-ippm-ioam-data].

IOAM E2E Type: A 16-bit identifier which specifies which data types
are used in the E2E option data. The definition is the same as
described in section 4.6 of [I-D.ietf-ippm-ioam-data].

5.5. Enhanced Alternate Marking (EAM) sub-TLV

The format of Enhanced Alternate Marking (EAM) sub-TLV is defined as
follows:

```
 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
                                +---------------+---------------+
                                |    Type=5     |   Length=4    |
+------------------------------+-------+-------+-------+-------+
|          FlowMonID           |       | Period| Rsvd  |
+------------------------------+---------------+-------+
```
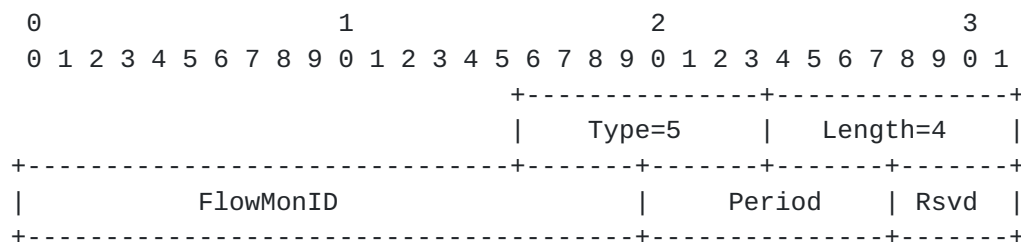
            Fig. 6 Enhanced Alternate Marking Sub-TLV

Where:

Type: 5 (to be assigned by IANA).

Length: 4, it is the total length of the value field (not including Type and Length fields).

FlowMonID: A 20-bit identifier to uniquely identify a monitored flow within the measurement domain. The definition is the same as described in section 5.3 of [I-D.ietf-6man-ipv6-alt-mark].

Period: Time interval between two alternate marking period. The unit is second.

Rsvd: A 4-bit field reserved for further usage. It MUST be zero.

6. SR Policy Operations with IFIT Attributes

The details of SR Policy installation and use are specified in [I-D.ietf-spring-segment-routing-policy]. This document complements SR Policy Operations described in [I-D.ietf-idr-segment-routing-te-policy] by adding the IFIT Attributes.

The operations described in [I-D.ietf-idr-segment-routing-te-policy] are always valid. The only difference is the addition of IFIT Attributes Sub-TLVs for the SR Policy NLRI, that can affect its acceptance by a BGP speaker, but the implementation MAY provide an option for ignoring the unrecognized or unsupported IFIT sub-TLVs. SR Policy NLRIs that have been determined acceptable, usable and valid can be evaluated for propagation, including the IFIT information.

The error handling actions are also described in [I-D.ietf-idr-segment-routing-te-policy].

The validation of the IFIT Attributes sub-TLVs introduced in this document MUST be performed to determine if they are malformed or invalid. The validation of the individual fields of the IFIT Attributes sub-TLVs are handled by the SRPM (SR Policy Module).

7. IANA Considerations

This document defines a new sub-TLV in the registry "BGP Tunnel Encapsulation Attribute sub-TLVs" to be assigned by IANA:

```
Codepoint     Description                       Reference
-------------------------------------------------------------
TBD1          IFIT Attributes Sub-TLV           This document
```

This document requests creation of a new registry called "IFIT
Attributes Sub-TLVs". The allocation policy of this registry is
"Specification Required" according to RFC 8126 [RFC8126].

Following initial Sub-TLV codepoints are assigned by this document:

Value	Description	Reference
1	IOAM Pre-allocated Trace Option Sub-TLV	This document
2	IOAM Incremental Trace Option Sub-TLV	This document
3	IOAM Directly Export Option Sub-TLV	This document
4	IOAM Edge-to-Edge Option Sub-TLV	This document
5	Enhanced Alternate Marking Sub-TLV	This document

8. Security Considerations

The security mechanisms of the base BGP security model apply to the
extensions described in this document as well. See the Security
Considerations section of [I-D.ietf-idr-segment-routing-te-policy].

SR operates within a trusted SR domain RFC 8402 [RFC8402] and its
security considerations also apply to BGP sessions when carrying SR
Policy information. The isolation of BGP SR Policy SAFI peering
sessions may be used to ensure that the SR Policy information is not
advertised outside the SR domain. Additionally, only trusted nodes
(that include both routers and controller applications) within the SR
domain must be configured to receive such information.

Implementation of IFIT methods (IOAM and Alternate Marking) are
mindful of security and privacy concerns, as explained in
[I-D.ietf-ippm-ioam-data] and RFC 8321 [RFC8321]. Anyway incorrect
IFIT parameters in the BGP extension SHOULD not have an adverse
effect on the SR Policy as well as on the network, since it affects
only the operation of the telemetry methodology.

9. Acknowledgements

The authors of this document would like to thank Ketan Talaulikar,
Joel Halpern, Jie Dong for their comments and review of this
document.

10. References

10.1. Normative References

[I-D.ietf-6man-ipv6-alt-mark]
          Fioccola, G., Zhou, T., Cociglio, M., Qin, F., and R.
          Pang, "IPv6 Application of the Alternate Marking Method",
          draft-ietf-6man-ipv6-alt-mark-01 (work in progress), June
          2020.

[I-D.ietf-idr-segment-routing-te-policy]
          Previdi, S., Filsfils, C., Talaulikar, K., Mattes, P.,
          Rosen, E., Jain, D., and S. Lin, "Advertising Segment
          Routing Policies in BGP", draft-ietf-idr-segment-routing-
          te-policy-09 (work in progress), May 2020.

[I-D.ietf-ippm-ioam-data]
          Brockners, F., Bhandari, S., and T. Mizrahi, "Data Fields
          for In-situ OAM", draft-ietf-ippm-ioam-data-10 (work in
          progress), July 2020.

[I-D.ietf-ippm-ioam-direct-export]
          Song, H., Gafni, B., Zhou, T., Li, Z., Brockners, F.,
          Bhandari, S., Sivakolundu, R., and T. Mizrahi, "In-situ
          OAM Direct Exporting", draft-ietf-ippm-ioam-direct-
          export-01 (work in progress), August 2020.

[I-D.ietf-ippm-ioam-flags]
          Mizrahi, T., Brockners, F., Bhandari, S., Sivakolundu, R.,
          Pignataro, C., Kfir, A., Gafni, B., Spiegel, M., and J.
          Lemon, "In-situ OAM Flags", draft-ietf-ippm-ioam-flags-02
          (work in progress), July 2020.

[I-D.ietf-ippm-ioam-ipv6-options]
          Bhandari, S., Brockners, F., Pignataro, C., Gredler, H.,
          Leddy, J., Youell, S., Mizrahi, T., Kfir, A., Gafni, B.,
          Lapukhov, P., Spiegel, M., Krishnan, S., Asati, R., and M.
          Smith, "In-situ OAM IPv6 Options", draft-ietf-ippm-ioam-
          ipv6-options-03 (work in progress), September 2020.

[I-D.ietf-spring-segment-routing-policy]
          Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and
          P. Mattes, "Segment Routing Policy Architecture", draft-
          ietf-spring-segment-routing-policy-08 (work in progress),
          July 2020.

   [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>.

   [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>.

   [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>.

   [RFC8321] Fioccola, G., Ed., Capello, A., Cociglio, M., Castaldelli,
              L., Chen, M., Zheng, L., Mirsky, G., and T. Mizrahi,
              "Alternate-Marking Method for Passive and Hybrid
              Performance Monitoring", RFC 8321, DOI 10.17487/RFC8321,
              January 2018, <https://www.rfc-editor.org/info/rfc8321>.

   [RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L.,
              Decraene, B., Litkowski, S., and R. Shakir, "Segment
              Routing Architecture", RFC 8402, DOI 10.17487/RFC8402,
              July 2018, <https://www.rfc-editor.org/info/rfc8402>.

10.2. Informative References

   [I-D.chen-pce-pcep-ifit]
              Chen, H., Yuan, H., Zhou, T., Li, W., Fioccola, G., and Y.
              Wang, "Path Computation Element Communication Protocol
              (PCEP) Extensions to Enable IFIT", draft-chen-pce-pcep-
              ifit-01 (work in progress), September 2020.

   [I-D.gandhi-mpls-ioam-sr]
              Gandhi, R., Ali, Z., Filsfils, C., Brockners, F., Wen, B.,
              and V. Kozak, "MPLS Data Plane Encapsulation for In-situ
              OAM Data", draft-gandhi-mpls-ioam-sr-03 (work in
              progress), September 2020.

   [I-D.gandhi-mpls-rfc6374-sr]
              Gandhi, R., Filsfils, C., Voyer, D., Salsano, S., and M.
              Chen, "Performance Measurement Using RFC 6374 for Segment
              Routing Networks with MPLS Data Plane", draft-gandhi-mpls-
              rfc6374-sr-05 (work in progress), June 2020.

   [I-D.ietf-mpls-rfc6374-sfl]
              Bryant, S., Swallow, G., Chen, M., Fioccola, G., and G.
              Mirsky, "RFC6374 Synonymous Flow Labels", draft-ietf-mpls-
              rfc6374-sfl-07 (work in progress), June 2020.

Appendix A.

Authors' Addresses

   Fengwei Qin
   China Mobile
   No. 32 Xuanwumenxi Ave., Xicheng District
   Beijing
   China

   Email: qinfengwei@chinamobile.com

   Hang Yuan
   UnionPay
   1899 Gu-Tang Rd., Pudong
   Shanghai
   China

   Email: yuanhang@unionpay.com

   Tianran Zhou
   Huawei
   156 Beiqing Rd., Haidian District
   Beijing
   China

   Email: zhoutianran@huawei.com

   Giuseppe Fioccola
   Huawei
   Riesstrasse, 25
   Munich
   Germany

   Email: giuseppe.fioccola@huawei.com

   Yali Wang
   Huawei
   156 Beiqing Rd., Haidian District
   Beijing
   China

   Email: wangyali11@huawei.com