Interdomain Routing Working Group
Internet-Draft
Intended status: Standards Track
Expires: July 9, 2020

F. Qin
China Mobile
H. Yuan
UnionPay
T. Zhou
M. Liu
G. Fioccola
Huawei

January 6, 2020

BGP SR Policy Extensions to Enable IFIT
draft-qin-idr-sr-policy-ifit-00

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) provides a reference framework that supports network OAM applications to apply dataplane on-path telemetry techniques acquiring data about a packet on its forwarding path. This document defines extensions to BGP to distribute SR policies carrying IFIT information. So that IFIT behavior 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) [I-D.song-opsawg-ifit-framework] provides a reference framework that supports network OAM applications to apply dataplane on-path telemetry techniques, including In-situ OAM (IOAM) [I-D.ietf-ippm-ioam-data], Postcard Based Telemetry (PBT)

[I-D.song-ippm-postcard-based-telemetry], In-band Flow Analyzer (IFA) [I-D.kumar-ippm-ifa], Enhanced Alternate Marking (EAM) [I-D.zhou-ippm-enhanced-alternate-marking], and Hybrid Two Steps (HTS) [I-D.mirsky-ippm-hybrid-two-step]. 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. The SR policy native IFIT can facilitate the closed loop control, and enable the automation of SR service.

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

2. 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.

3. SR Policy for IOAM

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.

3.1. 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     |    Length     |    Namespace ID               |
+---------------+---------------+--------------+--------+-------+
|         IOAM Trace Type                      | Flags  | Rsvd  |
+----------------------------------------------+--------+-------+
```

Fig. 1 IOAM Pre-allocated Trace Option Sub-TLV

Where:

Type: to be assigned by IANA.

Length: 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.

3.2. 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     |     Length    |     Namespace ID              |
+---------------+---------------+--------------+--------+-------+
|         IOAM Trace Type                      | Flags  | Rsvd  |
+----------------------------------------------+--------+-------+
```

Fig. 2 IOAM Incremental Trace Option Sub-TLV

Where:

Type: to be assigned by IANA.

Length: 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.

3.3. 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      |    Length     |
+-------------------------------+---------------+---------------+
|        Namespace ID           |            Flags              |
+-------------------------------+---------------+---------------+
|              IOAM Trace Type                  |     Rsvd      |
+-----------------------------------------------+---------------+
|                         Flow ID                               |
+---------------------------------------------------------------+
```

Fig. 3 IOAM Directly Export Option Sub-TLV

Where:

Type: to be assigned by IANA.

Length: 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.ioamteam-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.ioamteam-ippm-ioam-direct-export].

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

3.4. 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     |     Length    |
+-------------------------------+---------------+---------------+
|        Namespace ID           |         IOAM E2E Type         |
+-------------------------------+-------------------------------+
```

Fig. 4 IOAM Edge-to-Edge Option Sub-TLV

Where:

Type: to be assigned by IANA.

Length: 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].

4. SR Policy for Enhanced Alternate Marking

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.

The Enhanced Alternate Marking (EAM) [I-D.zhou-ippm-enhanced-alternate-marking] defines data fields for the alternate marking with enough space, in particular for Postcard-based Telemetry. More information can be considered within the alternate marking field to facilitate the efficiency and ease the deployment.

The format of 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      |    Length     |
+-------------------------------+-------+-------+-------+-------+
|           FlowMonID                   |      Period    | Rsvd  |
+---------------------------------------+---------------+-------+
```

Where:

Type: to be assigned by IANA.

Length: 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 2 of [I-D.zhou-ippm-enhanced-alternate-marking].

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.

5. IANA Considerations

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

Codepoint	Description	Reference
TBD1	IOAM Pre-allocated Trace Option Sub-TLV	This document
TBD2	IOAM Incremental Trace Option Sub-TLV	This document
TBD3	IOAM Directly Export Option Sub-TLV	This document
TBD4	IOAM Edge-to-Edge Option Sub-TLV	This document
TBD5	Enhanced Alternate Marking Sub-TLV	This document

6. Security Considerations

TBD.

7. Acknowledgements

8. References

8.1. 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>.

[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>.

[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>.

8.2. Informative References

[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-08 (work in progress), November 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., remy@barefootnetworks.com, r., daniel.bernier@bell.ca, d., and J. Lemon, "Data Fields for In-situ OAM", draft-ietf-ippm-ioam-data-08 (work in progress), October 2019.

[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-00 (work in progress), October 2019.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-06 (work in progress), December 2019.

[I-D.ioamteam-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-ioamteam-ippm-ioam-direct-export-00 (work in progress), October 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.mirsky-ippm-hybrid-two-step]
Mirsky, G., Lingqiang, W., and G. Zhui, "Hybrid Two-Step Performance Measurement Method", draft-mirsky-ippm-hybrid-two-step-04 (work in progress), October 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-06 (work in progress), October 2019.

[I-D.song-opsawg-ifit-framework]
Song, H., Qin, F., Chen, H., Jin, J., and J. Shin, "In-situ Flow Information Telemetry", draft-song-opsawg-ifit-framework-10 (work in progress), December 2019.

[I-D.zhou-ippm-enhanced-alternate-marking]
Zhou, T., Fioccola, G., Li, Z., Lee, S., and M. Cociglio, "Enhanced Alternate Marking Method", draft-zhou-ippm-enhanced-alternate-marking-04 (work in progress), October 2019.

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

Min Liu
Huawei
156 Beiqing Rd., Haidian District
Beijing
China

Email: lucy.liumin@huawei.com

Giuseppe Fioccola
Huawei
Riesstrasse, 25
Munich
Germany

Email: giuseppe.fioccola@huawei.com