IDR F. Qin China Mobile Internet-Draft

Intended status: Standards Track

H. Yuan Expires: 22 April 2024 UnionPay S. Yang China Telecom T. Zhou G. Fioccola Huawei

> **BGP SR Policy Extensions to Enable IFIT** draft-ietf-idr-sr-policy-ifit-07

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.

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 22 April 2024.

Copyright Notice

Copyright (c) 2023 IETF Trust and the persons identified as the document authors. All rights reserved.

20 October 2023

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (https://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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

<u>1</u> .	Intr	oduction			•	2
1.	<u>.1</u> . I	Requirements Language				<u>3</u>
<u>2</u> .	Moti	vation				3
<u>3</u> .	IFIT	methods for SR Policy				4
<u>4</u> .	IFIT	Attributes in SR Policy				4
<u>5</u> .	IFIT	Attributes Sub-TLV				<u>6</u>
<u>5</u> .	<u>. 1</u>	IOAM Pre-allocated Trace Option Sub-TLV				7
<u>5</u> .	<u>. 2</u>	IOAM Incremental Trace Option Sub-TLV				8
<u>5</u> .	<u>.3</u> .	IOAM Directly Export Option Sub-TLV				8
<u>5</u> .	<u>. 4</u>	IOAM Edge-to-Edge Option Sub-TLV				9
<u>5</u> .	<u>.5</u> . I	Enhanced Alternate Marking (EAM) sub-TLV .				<u>10</u>
<u>6</u> .	SR P	olicy Operations with IFIT Attributes				<u>11</u>
<u>7</u> .	IANA	Considerations				<u>11</u>
<u>8</u> .	Secu	rity Considerations				<u>12</u>
<u>9</u> .	Cont	ributors				<u>13</u>
<u> 10</u> .	Ackn	owledgements				<u>13</u>
<u>11</u> .	Refe	rences				<u>13</u>
<u>11</u>	<u>1.1</u> .	Normative References				<u>13</u>
<u>11</u>	1.2.	Informative References				<u>15</u>
Auth	nors'	Addresses				<u>15</u>

1. Introduction

Segment Routing (SR) policy [RFC9256] 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 floworiented 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 (Operations, Administration, and Maintenance) data plane on-path telemetry techniques, including In-situ OAM (IOAM) [RFC9197] and Alternate Marking [RFC9341]. It can provide flow information on the entire forwarding path on a perpacket 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.

IFIT is a solution focusing on network domains according to [RFC8799] that introduces the concept of specific domain solutions. A network domain consists of a set of network devices or entities within a single administration. As mentioned in [RFC8799], for a number of reasons, such as policies, options supported, style of network management and security requirements, it is suggested to limit applications including the emerging IFIT techniques to a controlled domain. Hence, the IFIT methods MUST be typically deployed in such controlled domains.

1.1. 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 BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

Motivation

IFIT Methods are being introduced in multiple protocols and in particular for Segment Routing over the MPLS data plane (SR-MPLS) and Segment Routing over IPv6 data plane (SRv6).

It is worth mentioning that, at the moment, the IFIT methods (IOAM and Alternate Marking) are more mature for SRv6 compared to SR-MPLS. The reference documents are [RFC9486] and [RFC9343] for SRv6.

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 (Path Computation Element Communication Protocol), [I-D.ietf-pce-pcep-ifit] proposes the extensions to PCEP to distribute paths carrying IFIT information and therefore to enable IFIT methods for SR policy too. These documents complement the deployment scenario described in [I-D.song-opsawg-ifit-framework].

3. IFIT methods for SR Policy

In-situ Operations, Administration, and Maintenance (IOAM) [RFC9197] 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 [RFC9341] 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 [RFC9486] and [RFC9343] 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], a new SAFI is defined (the SR Policy SAFI with codepoint 73) as well as a new NLRI. The NLRI contains the SR Policy candidate path and, according to [I-D.ietf-idr-segment-routing-te-policy], the content of the SR Policy Candidate Path is encoded in the Tunnel Encapsulation Attribute defined in [RFC9012] using a new Tunnel-Type called SR Policy Type with codepoint 15. The SR Policy encoding structure is as follows:

```
SR Policy SAFI NLRI: <Distinguisher, Policy-Color, Endpoint>
Attributes:
   Tunnel Encapsulation Attribute (23)
      Tunnel Type: SR Policy (15)
          Binding SID
          SRv6 Binding SID
          Preference
          Priority
          Policy Name
          Policy Candidate Path 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 Encapsulation Attribute (23)
      Tunnel Type: SR Policy (15)
          Binding SID
          SRv6 Binding SID
          Preference
          Priority
          Policy Name
          Policy Candidate Path 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.

Once the IFIT attributes are signalled, if a packet arrives at the headend and, based on the types of steering described in [RFC9256], it may get steered into an SR Policy where IFIT methods are applied. Therefore it will be managed consequently with the corresponding IOAM or Alternate Marking information according to the enabled IFIT methods.

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:

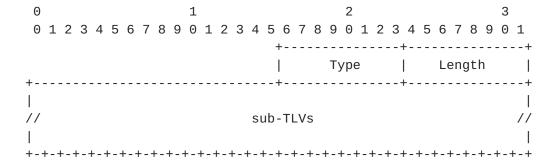


Figure 1: Fig. 1 IFIT Attributes Sub-TLV

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.

In case of empty IFIT Attributes Sub-TLV, i.e. no further IFIT sub-TLV and Length=0, IFIT methods will not be activated. If two conflicting IOAM sub-TLVs are present (e.g. Pre-allocated Trace Option and Incremental Trace Option) it means that they are not usable and none of the two methods will be activated. The same applies if there is more than one instance of the sub-TLV of the same type. Anyway the validation of the individual fields of the IFIT Attributes sub-TLVs are handled by the SRPM (SR Policy Module).

The process of stopping IFIT methods can be done by setting empty IFIT Attributes Sub-TLV, while, for modifying IFIT methods parameters, the IFIT Attributes Sub-TLVs can be updated accordingly. Additionally the backward compatibility is guaranteed, since an implementation that does not understand IFIT Attributes Sub-TLV can simply ignore it.

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:

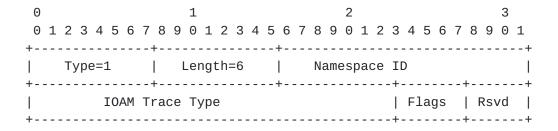


Figure 2: Fig. 2 IOAM Pre-allocated Trace Option Sub-TLV

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 [RFC9197].

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 <u>section 4.4 of [RFC9197]</u>.

Flags: A 4-bit field. The definition is the same as described in [RFC9322] and section 4.4 of [RFC9197].

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

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:

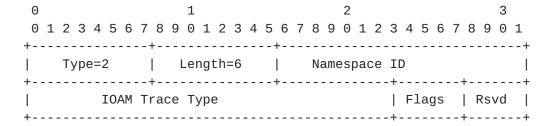


Figure 3: Fig. 3 IOAM Incremental Trace Option Sub-TLV

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 definition is the same as the pre-allocated trace option sub-TLV in $\underline{\text{section 4.1}}$.

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:

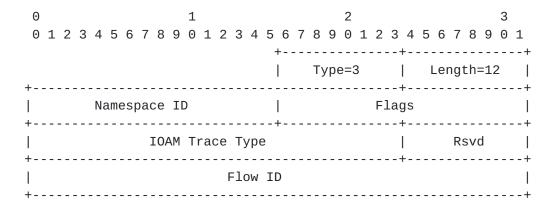


Figure 4: Fig. 4 IOAM Directly Export Option Sub-TLV

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 [RFC9197].

Flags: A 16-bit field. The definition is the same as described in section 3.2 of [RFC9326].

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 <u>section 4.4 of [RFC9197]</u>.

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

Flow ID: A 32-bit flow identifier. The definition is the same as described in <u>section 3.2 of [RFC9326]</u>.

<u>5.4</u>. 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:

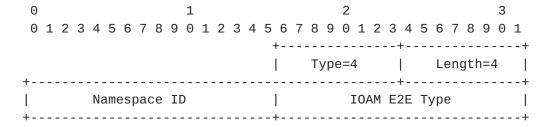


Figure 5: Fig. 5 IOAM Edge-to-Edge Option Sub-TLV

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 <u>section 4.6 of [RFC9197]</u>.

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 <u>section 4.6 of [RFC9197]</u>.

5.5. Enhanced Alternate Marking (EAM) sub-TLV

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

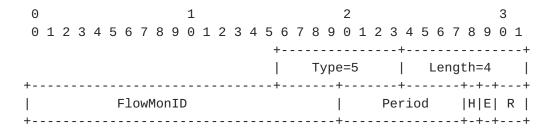


Figure 6: Fig. 6 Enhanced Alternate Marking Sub-TLV

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 <u>section 5.3 of [RFC9343]</u>.

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

H: A flag indicating that the measurement is Hop-By-Hop.

E: A flag indicating that the measurement is end to end.

R: A 2-bit field reserved for further usage. It MUST be zero and ignored on receipt.

6. SR Policy Operations with IFIT Attributes

The details of SR Policy installation and use are specified in [RFC9256]. 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], indeed A BGP Speaker MUST perform the syntactic validation of the SR Policy NLRI to determine if it is malformed, including the TLVs/sub-TLVs. In case of any error detected, either at the attribute or its TLV/sub-TLV level, the "treat-as-withdraw" strategy MUST be applied.

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].

The 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 RFC 9197 [RFC9197] and Alternate Marking [RFC9341]. 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.

IFIT data MUST be propagated in a limited domain in order to avoid malicious attacks and solutions to ensure this requirement are respectively discussed in [RFC9197] and [RFC9343].

IFIT methods (IOAM and Alternate Marking) are applied within a controlled domain where the network nodes are locally administered. A limited administrative domain provides the network administrator with the means to select, monitor and control the access to the network, making it a trusted domain also for the BGP extensions defined in this document.

9. Contributors

The following people provided relevant contributions to this document:

Yali Wang Huawei

Email: wangyali11@huawei.com

10. Acknowledgements

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

11. References

11.1. Normative References

> <https://datatracker.ietf.org/doc/html/draft-ietf-idrsegment-routing-te-policy-25>.

- [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)", <u>RFC 7799</u>, 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.

- [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>.
- [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>.
- [RFC8799] Carpenter, B. and B. Liu, "Limited Domains and Internet Protocols", <u>RFC 8799</u>, DOI 10.17487/RFC8799, July 2020, https://www.rfc-editor.org/info/rfc8799>.
- [RFC9012] Patel, K., Van de Velde, G., Sangli, S., and J. Scudder,
 "The BGP Tunnel Encapsulation Attribute", RFC 9012,
 DOI 10.17487/RFC9012, April 2021,
 https://www.rfc-editor.org/info/rfc9012>.

- [RFC9322] Mizrahi, T., Brockners, F., Bhandari, S., Gafni, B., and
 M. Spiegel, "In Situ Operations, Administration, and
 Maintenance (IOAM) Loopback and Active Flags", RFC 9322,
 DOI 10.17487/RFC9322, November 2022,
 <https://www.rfc-editor.org/info/rfc9322>.
- [RFC9326] Song, H., Gafni, B., Brockners, F., Bhandari, S., and T.
 Mizrahi, "In Situ Operations, Administration, and
 Maintenance (IOAM) Direct Exporting", RFC 9326,
 DOI 10.17487/RFC9326, November 2022,
 <https://www.rfc-editor.org/info/rfc9326>.

[RFC9486] Bhandari, S., Ed. and F. Brockners, Ed., "IPv6 Options for In Situ Operations, Administration, and Maintenance (IOAM)", RFC 9486, DOI 10.17487/RFC9486, September 2023, https://www.rfc-editor.org/info/rfc9486.

11.2. Informative References

[I-D.ietf-pce-pcep-ifit]

Yuan, H., , Yang, P., Li, W., and G. Fioccola, "Path Computation Element Communication Protocol (PCEP) Extensions to Enable IFIT", Work in Progress, Internet-Draft, draft-ietf-pce-pcep-ifit-03, 7 July 2023, https://datatracker.ietf.org/doc/html/draft-ietf-pce-pcep-ifit-03>.

[I-D.song-opsawg-ifit-framework]

Song, H., Qin, F., Chen, H., Jin, J., and J. Shin, "Framework for In-situ Flow Information Telemetry", Work in Progress, Internet-Draft, draft-song-opsawg-ifit-framework-20, draft-song-opsawg-ifit-framework-20.

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 Shunxing Yang China Telecom Guangzhou China

Email: yangsx@chinatelecom.cn

Tianran Zhou Huawei 156 Beiqing Rd., Haidian District Beijing China Email: zhoutianran@huawei.com

Giuseppe Fioccola Huawei Palazzo Verrocchio, Centro Direzionale Milano 2 20054 Segrate (Milan) Italy

Email: giuseppe.fioccola@huawei.com