

Workgroup: Network Working Group  
Internet-Draft:  
draft-li-ldr-bgp-request-cp-sr-te-policy-04  
Published: 12 October 2022  
Intended Status: Standards Track  
Expires: 15 April 2023  
Authors: Z. Li      Q. Gao      H. Chen      Y. Fan  
         Huawei   Huawei   Futurewei   Casa Systems  
         X. Liu                   L. Liu  
         Volta Networks   Fujitsu  
**BGP Request for Candidate Paths of SR TE Policies**

## Abstract

An SR Policy is a set of candidate paths. The headend of an SR Policy may learn multiple candidate paths for an SR Policy via a number of different mechanisms, e.g., CLI, NetConf, PCEP, or BGP. This document defines extensions to BGP for the headend to request BGP speaker (controller) for advertising the candidate paths.

## 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](#)].

## 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 15 April 2023.

## Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.

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

- 1. [Introduction](#)
- 2. [Terminology](#)
- 3. [Overview of BGP Request for SR-TE Paths](#)
- 4. [BGP Request UPDATE Message](#)
  - 4.1. [Extention of SR Policy NLRI](#)
  - 4.2. [New SR Policy Sub-TLVs](#)
    - 4.2.1. [SR Path Attributes Sub-TLV](#)
    - 4.2.2. [Synchronization Sub-TLV](#)
    - 4.2.3. [Metric Sub-TLV](#)
    - 4.2.4. [Include Route Sub-TLV](#)
    - 4.2.5. [Load Balance Sub-TLV](#)
    - 4.2.6. [Request Parameter Sub-TLV](#)
- 5. [IANA](#)
- 6. [Contributors](#)
- 7. [Acknowledgments](#)
- 8. [References](#)
  - 8.1. [Normative References](#)
  - 8.2. [Informative References](#)
- [Authors' Addresses](#)

## 1. Introduction

An SR Policy defined in [[I-D.ietf-spring-segment-routing-policy](#)] is a set of candidate paths. The headend of an SR Policy may be informed by various means including: Configuration, PCEP [[RFC8281](#)] or BGP [[I-D.ietf-idr-segment-routing-te-policy](#)]. All these mechanisms are Controller initiated, but in some situations the headend may want to pull a set of candidate paths from Controller rather than receive all information passively. Actually PCEP can use request and reply messages defined in [[RFC5440](#)] to match this requirement, but the mechanism is not clear when controller advertises candidate paths via BGP.

This document defines a way to request controller (BGP speaker) to advertise candidate paths via BGP update messages. This makes BGP have the mechanism with request and reply similar to PCEP.

## 2. Terminology

RP: Request Parameters

LSPA: LSP Attributes

SVEC: Synchronization VEctor

IRO: Include Route Object

ERO: Explicit Route Object

MSD: Base MPLS Imposition Maximum SID Depth, as defined in [RFC8491]

NAI: Node or Adjacency Identifier

PCC: Path Computation Client

PCE: Path Computation Element

PCEP: Path Computation Element Communication Protocol

SID: Segment Identifier

SR: Segment Routing

SR-TE: Segment Routing Traffic Engineering

## 3. Overview of BGP Request for SR-TE Paths

[[I-D.ietf-idr-segment-routing-te-policy](#)] defines the extensions to BGP for a headend to receive candidate paths in a BGP UPDATE message from a controller (BGP speaker). In some situations a headend just wants to get these candidate paths when some special event occurs (for example, when it receives a customer route (VPN route) with a special color or special BGP attribute). This document defines the mechanism in which the headend requests the controller to advertise the expected SR policy with the candidate paths when this special situation occurs.

At first, the headend decides to get a new candidate path from the controller based on some trigger event. This trigger mechanism is out of scope of this document.

Then, the headend creates a new BGP request UPDATE message (defined below in this document) and sends it to the controller. The message contains the constraints/attributes of SR-TE paths such as affinity, metric, SRLG, and so on. This special request UPDATE message is called request message or request for short. It SHOULD NOT be used for BGP best path selection.

After receiving the request message, the controller will calculate one or a set of paths (i.e., segment lists) according to the request from the headend and advertise the SR Policy with the paths computed to the headend using [[I-D.ietf-idr-segment-routing-te-policy](#)]. How to calculate the paths is out of scope of this document.

#### 4. BGP Request UPDATE Message

A BGP request UPDATE message is based on the update message defined in [[I-D.ietf-idr-segment-routing-te-policy](#)] with some extensions described below.

##### 4.1. Extention of SR Policy NLRI

The SR Policy NLRI defined in [[I-D.ietf-idr-segment-routing-te-policy](#)] has the following format:

```
+-----+
|  NLRI Length   | 1 octet
+-----+
|  Distinguisher  | 4 octets
+-----+
|  Policy Color   | 4 octets
+-----+
|  Endpoint       | 4 or 16 octets
+-----+
```

where:

\*NLRI Length: 1 octet of length expressed in bits as defined in [[RFC4760](#)].

\*Distinguisher: 4-octet value uniquely identifying the policy in the context of <color, endpoint> tuple. The distinguisher has no semantic value and is solely used by the SR Policy originator to make unique (from an NLRI perspective) multiple occurrences of the same SR Policy.

\*Policy Color: 4-octet value identifying (with the endpoint) the policy. The color is used to match the color of the destination prefixes to steer traffic into the SR Policy [[I-D.ietf-spring-segment-routing-policy](#)]

\*Endpoint: identifies the endpoint of a policy. The Endpoint may represent a single node or a set of nodes (e.g., an anycast address). The Endpoint is an IPv4 (4-octet) address or an IPv6 (16-octet) address according to the AFI of the NLRI.

NLRI Length, Policy Color, Endpoint field remains unchanged, while the Distinguisher field will be set to FF:FF:FF:FF to indicate that

the UPDATE message with this NLRI is a request message to the controller.

#### 4.2. New SR Policy Sub-TLVs

The content of the SR Policy is encoded in the Tunnel Encapsulation Attribute TLV of type 23 defined in [[I-D.ietf-idr-tunnel-encaps](#)] containing a new Tunnel Type TLV of type 15. The SR Policy Encoding structure is as follows:

SR Policy SAFI NLRI: <Distinguisher, Policy-Color, Endpoint>

Attributes:

Tunnel Encaps Attribute (23)  
 Tunnel Type (15): SR Policy  
 <Sub-TLVs>

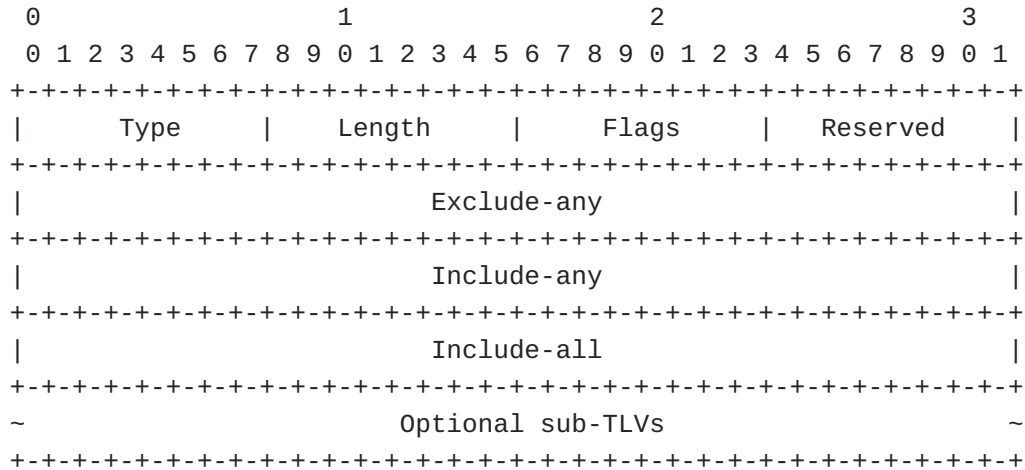
Preference, Binding SID, Priority, Policy Name, ENLP, Segment List, Weight and Segment Sub-TLVs are all defined in [[I-D.ietf-idr-segment-routing-te-policy](#)] for a SR Policy to be advertised to a headend.

Additional 6 new Sub-TLVs are defined below for the request mechanism. They are SR Path Attributes, Synchronization, Metric, Include Route, Load Balance, and Request Parameters Sub-TLVs.

##### 4.2.1. SR Path Attributes Sub-TLV

A SR Path Attributes Sub-TLV contains the attributes of the SR paths requested, which are similar to an LSP Attributes Object defined in [[RFC5440](#)] and [[RFC3209](#)].

It is optional and specifies various attributes or constraints of the paths requested. Its format is shown below.



where:

\*Type: TBD1

\*Length: specifies the length of the value field not including Type and Length fields.

\*Flags (8 bits): No flag is currently defined. Undefined flags MUST be set to zero on transmission and be ignored on receipt.

\*Reserved (8 bits): This field MUST be set to zero on transmission and be ignored on receipt.

\*Exclude-any: A 32-bit vector representing a set of attribute filters associated with a path any of which renders a link unacceptable.

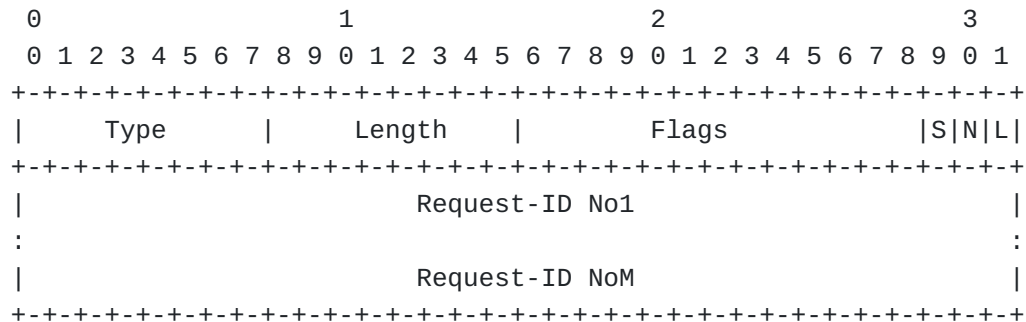
\*Include-any: A 32-bit vector representing a set of attribute filters associated with a path any of which renders a link acceptable (with respect to this test). A null set (all bits set to zero) automatically passes.

\*Include-all: A 32-bit vector representing a set of attribute filters associated with a path all of which must be present for a link to be acceptable (with respect to this test). A null set (all bits set to zero) automatically passes.

\*Optional sub-TLVs: No optional sub-TLV is currently defined.

#### 4.2.2. Synchronization Sub-TLV

A Synchronization Sub-TLV allows a headend to request the synchronization of a set of M dependent or independent SR path requests. This TLV is similar to the SVEC Object defined in [RFC5440]. It is optional and has the following format.



where:

\*Type: TBD2

\*Length: specifies the length of the value field not including Type and Length fields.

\*Flags (16 bits): Defines the potential dependency among a set of SR paths (i.e., segment lists). Three flags are defined as follows:

-L (Link diverse) bit: when set, it indicates that the computed SR paths (i.e., segment lists) MUST NOT have any link in common.

-N (Node diverse) bit: when set, it indicates that the computed SR paths (i.e., segment lists) MUST NOT have any node in common.

-S (SRLG diverse) bit: when set, it indicates that the computed SR paths (i.e., segment lists) MUST NOT share any SRLG (Shared Risk Link Group).

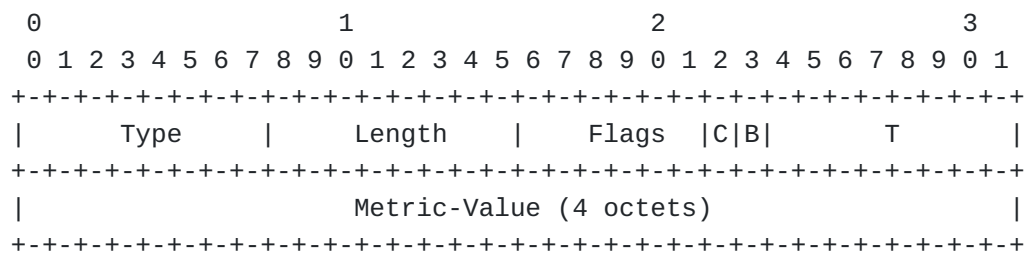
\*Request-ID No1, ..., NoM: each of which uniquely identifies one of M SR path requests.

In case of M synchronized independent path requests, the bits L, N, and S are set to zero.

Unassigned flags MUST be set to zero on transmission and MUST be ignored on receipt.

#### 4.2.3. Metric Sub-TLV

A Metric Sub-TLV carries the same content as a Metric Object defined in [RFC5440] and [I-D.ietf-pce-segment-routing]. It has following format:



\*Type: TBD3.

\*Length: specifies the length of the value field not including Type and Length fields.

\*Flags (8 bits): Two flags are currently defined:

- B (Bound - 1 bit): When set, the metric-value indicates a bound (a maximum) for the path metric that must not be exceeded for the headend to consider the computed path as acceptable. The path metric must be less than or equal to the value specified in the metric-value field. When the B flag is cleared, the metric-value field is not used to reflect a bound constraint.
- C (Computed Metric - 1 bit): When set, it indicates that the controller MUST provide the computed path metric value (should a path satisfying the constraints be found) in the advertisement message for the corresponding metric.
- Unassigned flags MUST be set to zero on transmission and MUST be ignored on receipt.

\*T (Type - 8 bits): Specifies the metric type. Four metric types are currently defined:

- T=1: IGP metric
- T=2: TE metric
- T=3: Hop Counts
- T=11: Maximum SID Depth of the requested path

\*Metric-Value (32 bits): It is a metric value encoded in 32 bits IEEE floating point format (see [IEEE.754.1985]).

#### **4.2.4. Include Route Sub-TLV**

The Include Route Sub-TLV is optional and can be used to specify that the computed candidate path MUST traverse a set of specified network elements. The Include Route Sub-TLV carries the same content as IRO Object defined in [\[RFC5440\]](#), [\[RFC3209\]](#) and SR-ERO defined in [\[I-D.ietf-pce-segment-routing\]](#)

The Include Route Sub-TLV has following format:



```

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   |   NT   |      Flags   |F|S|C|M|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
~                               SID (optional)                               ~
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
~                               NAI (variable, optional)                               ~
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Where:

\*Type: TBD4.

\*Length: It specifies the length of the value field not including Type and Length fields.

\*NAI Type (NT): It indicates the type and format of the NAI contained, if any is present. If the F bit is set to zero, then the NT field has no meaning and MUST be ignored by the receiver. This document describes the following NT values:

-NT=0: The NAI is absent.

-NT=1: The NAI is an IPv4 node ID.

-NT=2: The NAI is an IPv6 node ID.

-NT=3: The NAI is an IPv4 adjacency.

-NT=4: The NAI is an IPv6 adjacency with global IPv6 addresses.

-NT=5: The NAI is an unnumbered adjacency with IPv4 node IDs.

-NT=6: The NAI is an IPv6 adjacency with link-local IPv6 addresses.

\*SID and NAI are the same as SR-ERO defined in [[I-D.ietf-pce-segment-routing](#)]

#### 4.2.5. Load Balance Sub-TLV

A Load Balance Sub-TLV specifies how many SR paths (i.e., segment lists) should be computed for a path request. It has following format:

```

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   |   Flag   |   Max-Slist   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
~                               Optional sub-TLVs                               ~
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Where:

\*Type: TBD5.

\*Length: It specifies the length of the value field not including Type and Length fields.

\*Flags (8 bits): No flag is currently defined. The Flags field MUST be set to zero on transmission and MUST be ignored on receipt.

\*Max-Slist (8 bits): It indicates the maximum number of SR paths (i.e., segment lists) to be computed for the request. The load is distributed among these SR paths.

\*Optional sub-TLVs: No Optional sub-TLV is currently defined.

#### 4.2.6. Request Parameter Sub-TLV

A Request Parameter (RP) Sub-TLV specifies the request identifier and other parameters for a path request. It has the format below.

```

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   |   Flags   |0|B|R|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Request-ID                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
~                               Optional sub-TLVs                               ~
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Where:

\*Type: TBD6.

\*Length: It specifies the length of the value field not including Type and Length fields.

\*Flags (16 bits): Three flag bits are currently defined as follows:

- R (Reoptimization - 1 bit): when set, it indicates that the SR path request message is for the reoptimization of an existing SR path, which is represented by a segment list Sub-TLV in the message.
- B (Bi-directional - 1 bit): when set, it indicates that the SR path request relates to bi-directional paths that has the same traffic engineering requirements including fate sharing, TE links, and other requirements (such as latency and jitter) in each direction.
- O (strict/loose - 1 bit): when set, it indicates that a loose path is acceptable. Otherwise (i.e., when cleared), it indicates that a path exclusively made of strict hops is required.

## 5. IANA

Under Existing Registry Name: "BGP Tunnel Encapsulation Attribute Sub-TLVs", IANA is requested to assign new Sub-TLV values for SR Path Request as follows:

Type	Value	Sub-TLV Name	Reference
TBD1		SR Path Attributes Sub-TLV	This document
TBD2		Synchronization Sub-TLV	This document
TBD3		Metric Sub-TLV	This document
TBD4		Include Route Sub-TLV	This document
TBD5		Load Balance Sub-TLV	This document
TBD6		Request Parameters Sub-TLV	This document

## 6. Contributors

TBD

## 7. Acknowledgments

TBD

## 8. References

### 8.1. Normative References

**[I-D.ietf-idr-segment-routing-te-policy]**

Previdi, S., Filsfils, C., Talaulikar, K., Mattes, P., Jain, D., and S. Lin, "Advertising Segment Routing Policies in BGP", Work in Progress, Internet-Draft, draft-ietf-idr-segment-routing-te-policy-20, 27 July 2022, <<https://www.ietf.org/archive/id/draft-ietf-idr-segment-routing-te-policy-20.txt>>.

**[I-D.ietf-spring-segment-routing-policy]**

Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", Work in Progress, Internet-Draft, draft-ietf-spring-segment-routing-policy-22, 22 March 2022, <<https://www.ietf.org/archive/id/draft-ietf-spring-segment-routing-policy-22.txt>>.

**[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>>.

**[RFC5440]** Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <<https://www.rfc-editor.org/info/rfc5440>>.

**[RFC8281]** Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for PCE-Initiated LSP Setup in a Stateful PCE Model", RFC 8281, DOI 10.17487/RFC8281, December 2017, <<https://www.rfc-editor.org/info/rfc8281>>.

### 8.2. Informative References

**[I-D.ietf-idr-tunnel-encaps]** Patel, K., Velde, G. V. D., Sangli, S. R., and J. Scudder, "The BGP Tunnel Encapsulation Attribute", Work in Progress, Internet-Draft, draft-ietf-idr-tunnel-encaps-22, 7 January 2021, <<https://www.ietf.org/archive/id/draft-ietf-idr-tunnel-encaps-22.txt>>.

**[I-D.ietf-pce-segment-routing]**

Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Extensions for Segment Routing", Work in Progress, Internet-Draft, draft-ietf-

pce-segment-routing-16, 4 March 2019, <<https://www.ietf.org/archive/id/draft-ietf-pce-segment-routing-16.txt>>.

**[RFC3209]** Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <<https://www.rfc-editor.org/info/rfc3209>>.

**[RFC4090]** Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, DOI 10.17487/RFC4090, May 2005, <<https://www.rfc-editor.org/info/rfc4090>>.

**[RFC4760]** Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <<https://www.rfc-editor.org/info/rfc4760>>.

**[RFC5420]** Farrel, A., Ed., Papadimitriou, D., Vasseur, JP., and A. Ayyangar, "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", RFC 5420, DOI 10.17487/RFC5420, February 2009, <<https://www.rfc-editor.org/info/rfc5420>>.

#### Authors' Addresses

Zhenbin Li  
Huawei  
156 Beiqing Road  
Beijing, 100095  
P.R. China

Email: [lizhenbin@huawei.com](mailto:lizhenbin@huawei.com)

Qiangzhou Gao  
Huawei  
156 Beiqing Road  
Beijing, 100095  
P.R. China

Email: [gaoqiangzhou@huawei.com](mailto:gaoqiangzhou@huawei.com)

Huaimo Chen  
Futurewei  
Boston, MA,  
United States of America

Email: [Huaimo.chen@futurewei.com](mailto:Huaimo.chen@futurewei.com)

Yanhe Fan  
Casa Systems  
United States of America

Email: [yfan@casa-systems.com](mailto:yfan@casa-systems.com)

Xufeng Liu  
Volta Networks  
McLean, VA  
United States of America

Email: [xufeng.liu.ietf@gmail.com](mailto:xufeng.liu.ietf@gmail.com)

Lei Liu  
Fujitsu  
United States of America

Email: [liulei.kddi@gmail.com](mailto:liulei.kddi@gmail.com)