```
IDR Working Group                                               N. Wu
Internet-Draft                                               Z. Zhuang
Intended status: Standards Track                              H. Chen
Expires: April 25, 2019                                        Huawei
                                                     October 22, 2018
```

BGP Extensions for IDs Allocation
draft-wu-idr-bgp-segment-allocation-ext-01

Abstract

   This document describes extensions to the BGP-LS for IDs allocation.
   The IDs are SIDs for segment routing and indirection IDs for flow
   redirect. They are distributed to their domains if needed.

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 April 25, 2019.

Table of Contents

1. Introduction

In a network with a central controller, the controller has the link state information of the network, including traffic engineering information. In addition, the controller allocates and manages the resources of the network in general. It is natural and beneficial for the controller to allocate and manage IDs as a kind of network resources.

This document proposes extensions to the BGP-LS as a controller for allocating Segment Identifers (SIDs) for segment routing (SR) and indirection identifiers (IIDs) for flow redirect. If needed, some IDs will be distributed into their network domains.

2. Terminology

The following terminology is used in this document.

SR: Segment Routing.

SID: Segment Identifier.

IID: Indirection Identifier.

SR-Path: Segment Routing Path.

SR-Tunnel: Segment Routing Tunnel.

RR: Route Reflector.

MPP: MPLS Path Programming.

NAI: Node or Adjacency Identifier.

GTID: Global Tunnel Identifier. It is used to identify a tunnel in
   a network.

TED: Traffic Engineering Database.

3. Protocol Extensions

This section defines a new Protocol-ID, called IDs-Allocation (IDA),
in the Procotol-ID field of Link State NLRI. The use of the new
Protocol-ID allows separation and differentiation between the NLRIs
carrying IDs Allocation information from the NLRIs carrying IGP link-
state information defined in [RFC7752].

3.1. Node NLRI for IDs Allocation

The Node NLRI with the new Protocol-ID is used for allocating the IDs
associated with a node. It has the same format (refer to the
Figure below) as that defined in [RFC7752] and may contain the
descriptor and attributes defined in [RFC7752].
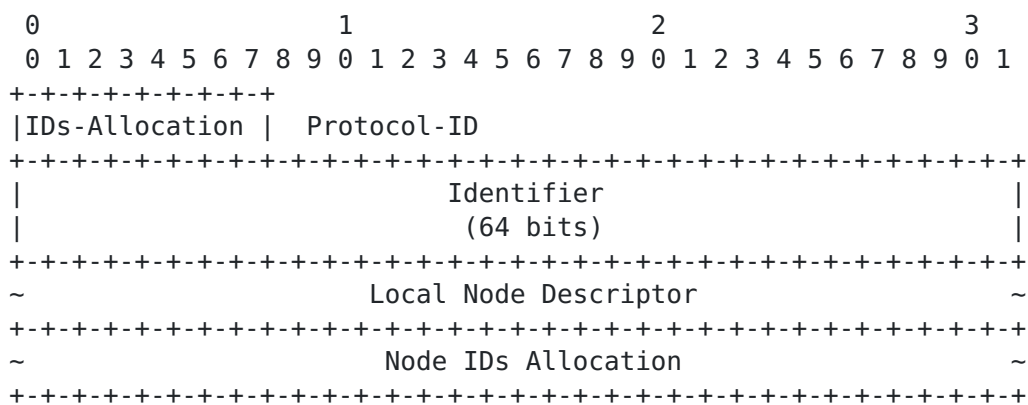
```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+
|IDs-Allocation |  Protocol-ID
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Identifier                            |
|                          (64 bits)                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                     Local Node Descriptor                     ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                     Node IDs Allocation                       ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where:

Protocol-ID: set to IDs-Allocation, a new Protocol-ID.

Local Node Descriptor: defined in [RFC7752], can be reused.

Node IDs Allocation: TLVs or sub-TLVs for IDs associated with the
  node given by the Local Node Descriptor.

Node IDs Allocation field may contain some of the followings:

SR-Capabilities TLV (1034): It contains the Segment Routing Global
  Base (SRGB) range(s) allocated for the node.

SR Local Block TLV (1036): The SR Local Block (SRLB) TLV contains
  the range(s) of SIDs/labels allocated to the node for local SIDs.

IID Flow Redirect TLV (TBD): A new TLV that contains an
  Indireection ID (IID) and a Flow Specification for redirecting
  traffic flow given by the Flow Specification to the tunnel
  indicated by the IID. Its format is shown 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: Flow Redirect (TBD)   |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Reserved           | Flags(1 octet)|   ID-Type     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Indirection ID                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                       Flow Specification                      ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IID Flow Redirect TLV

Where Flags, ID-Type and Indirection ID are the same as those defined
in [I-D.ietf-idr-flowspec-path-redirect]. The contents of Flow
Specification is the same as that defined in [RFC5575].

3.2. Link NLRI for IDs Allocation

The Link NLRI with the new Protocol-ID is used for allocating the IDs
associated with a link. It has the same format (refer to the
Figure below) as that defined in in [RFC7752] and may contain the
descriptors and attributes defined in [RFC7752].

```
      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
     +-+-+-+-+-+-+-+-+
     |IDs-Allocation |  Protocol-ID
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                         Identifier                           |
     |                         (64 bits)                            |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     ~                     Local Node Descriptor                    ~
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     ~                     Remote Node Descriptor                   ~
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     ~                        Link Descriptor                       ~
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     ~                      Link IDs Allocation                     ~
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
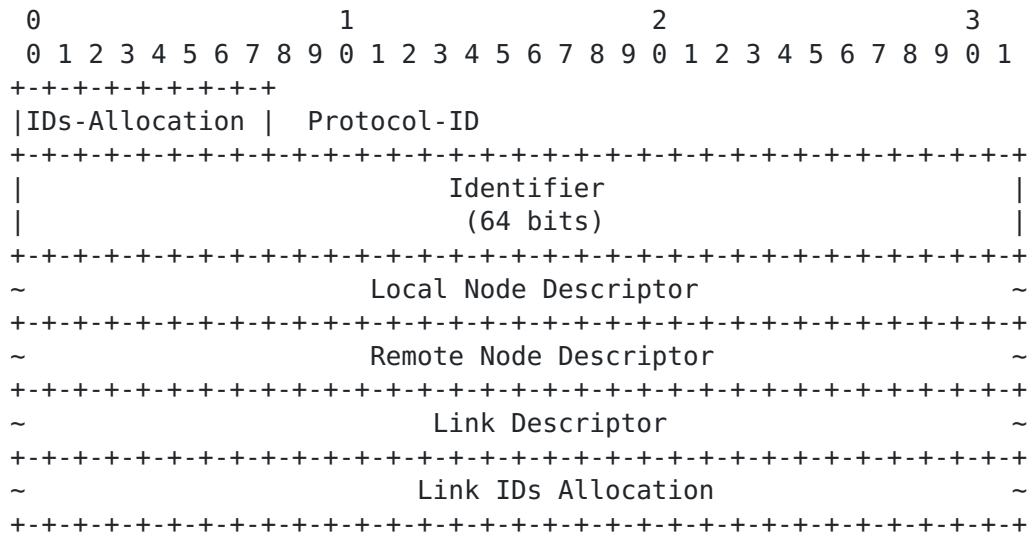
   Where:

      Protocol-ID: set to IDs-Allocation, a new Protocol-ID.

      Node Descriptors: defined in [RFC7752], can be reused.

      Link Descriptor: defined in [RFC7752], can be reused.

      Link IDs Allocation: TLVs or sub-TLVs for IDs associated with the
       link given by the Link Descriptor.

   Link IDs Allocation field may contain some of the followings:

      Adj-SID TLV (1099): It contains the Segment Identifier (SID)
       allocated for the link/adjacency.

      LAN Adj-SID TLV (1100): It contains the Segment Identifier (SID)
       allocated for the adjacency/link to a non-DR router on a
       broadcast, NBMA, or hybrid link.

3.3. Prefix NLRI for IDs Allocation

   The Prefix NLRI with the new Protocol-ID is used for allocating the
   IDs associated with a prefix. It has the same format (refer to the
   Figure below) as that defined in in [RFC7752] and may contain the
   descriptors and attributes defined in [RFC7752].

```
 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
+-+-+-+-+-+-+-+-+
|IDs-Allocation |  Protocol-ID
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Identifier                            |
|                         (64 bits)                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                     Local Node Descriptor                     ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                       Prefix Descriptor                       ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                      Prefix IDs Allocation                    ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where:

   Protocol-ID: set to IDs-Allocation, a new Protocol-ID.

   Local Node Descriptor: defined in [RFC7752], can be reused.

   Prefix Descriptor: defined in [RFC7752], can be reused.

   Prefix IDs Allocation: TLVs or sub-TLVs for IDs associated with
    the prefix given by the Prefix Descriptor.

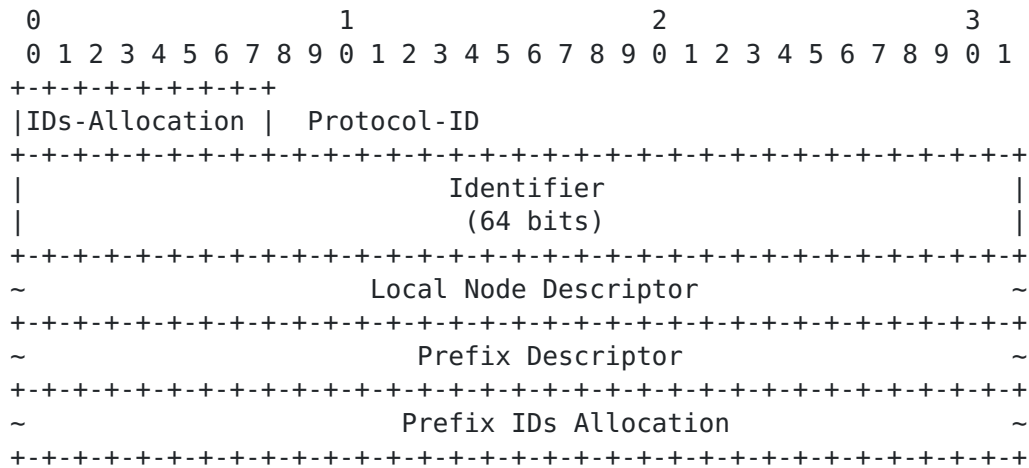Prefix IDs Allocation field may contain some of the followings:

   Prefix-SID TLV (1158): It contains the Segment Identifier (SID)
    allocated for the prefix.

   Prefix Range TLV (1159): It contains a range of prefixes and the
    Segment Identifier (SID)s allocated for the prefixes.

4. Application Examples

4.1. Allocating IDs for DC BGP Network
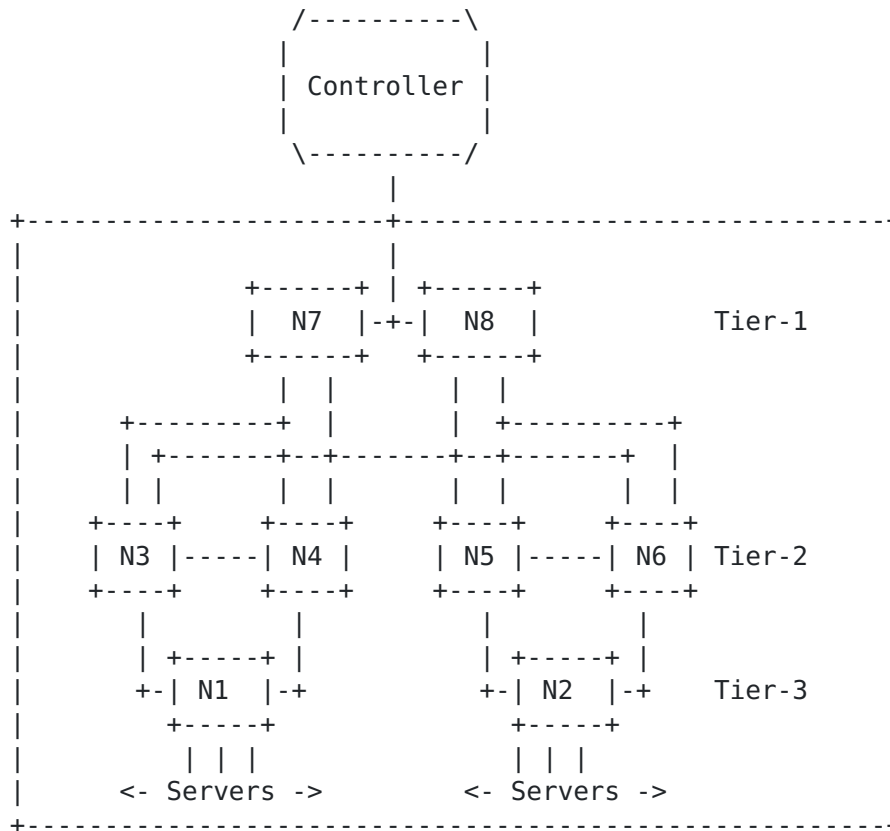
   BGP may be the only routing protocol in some networks, such as the DC
   network described in [I-D.ietf-rtgwg-bgp-routing-large-dc]. If SR is
   used in this case, it is better for a controller to allocate SIDs and
   distribute them in the network.

   In the DC network shown below, it is assumed:

   Node NX has a loopback 1.1.1.X/32. For example, N7 has loopback
   1.1.1.7/32.

Each node peers with its neighbors and the Controller via BGP
session.

The BGP Router ID and AS of Node NX are 1.1.1.X and X respectively.

```
                         /----------\
                         |          |
                         | Controller |
                         |          |
                          \----------/
                               |
    +----------------------+-----------------------------+
    |                      |                             |
    |        +------+ | +------+                         |
    |        | N7   |-+-| N8   |            Tier-1        |
    |        +------+   +------+                          |
    |          | |        | |                            |
    |    +---------+ |      | +----------+               |
    |    | +-------+--+-------+--+-------+   |            |
    |    | | |       | |      | |      | |   |           |
    |  +----+     +----+    +----+    +----+               |
    |  | N3 |-----| N4 |    | N5 |-----| N6 | Tier-2      |
    |  +----+     +----+    +----+    +----+               |
    |    |          |         |         |                |
    |    | +-----+ |          | +-----+ |                |
    |    +-| N1   |-+          +-| N2   |-+    Tier-3      |
    |      +-----+                +-----+                 |
    |      | | |                  | | |                  |
    |     <- Servers ->         <- Servers ->             |
    +--------------------------------------------------------+
```

After the controller gets the topology information of this BGP
network, it can start to allocate SIDs and distribute them in the
network.

4.1.1. Node-SID Distribution via Prefix NLRI

In order to distribute a Node-SID for a router (e.g., N7), Controller
allocates a SID for its loopback and advertises a Prefix NLRI
containing the SID for the loopback as a prefix to all the routers in
the BGP Network. The following is an example of Prefix NLRI for Node
N7's SID.

o Protocol-ID set to the new Protocol-ID: IDs-Allocation

o Local Node Descriptor contains

   * BGP Router-ID: 1.1.1.7

* Local ASN: 7

o Prefix Descriptor

    * 1.1.1.7/32

o Prefix Attribute contains

    * Prefix-SID TLV (1158): SID for 1.1.1.7/32 with N-Flag=1 (Node-
       Flag).

    * Other Prefix Attributes.

4.1.2. Adj-SID Distribution via Link NLRI

In order to distribute a Adj-SID for a router, Controller allocates a
SID for the adjacency/link and advertise a Link NLRI containing the
SID for the adjacency to all the routers of the BGP Network. The
following is an example of Link NLRI for the SID of the adjacency/
link from N7 to N8.

o Protocol-ID set to the new Protocol-ID: IDs-Allocation

o Local Node Descriptors contains

    * BGP Router-ID: 1.1.1.7

    * Local ASN: 7

o Remote Node Descriptor contains

    * BGP Router-ID: 1.1.1.8

    * Local ASN: 8

o Link Descriptor

    * BGP session IPv4 local address: 1.1.1.7

    * BGP session IPv4 peer address: 1.1.1.8

o Link Attribute contains

    * Adj-SID TLV (1099): SID for Adjacency from N7 to N8

    * Other Link Attributes.

4.1.3. IID Distribution via Node NLRI

In order to distribute an IID Flow Redirect, Controller allocates an
IID for a Flow Specification and advertises a Node NLRI containing
the IID and the Flow Specification to all the routers in the BGP
Network. The following is an example of Node NLRI for the IID Flow
Redirect, where Local Node Descriptor describes the Controller (its
BGP Router-ID is 7.7.7.7 and ASN is 7777).

o Protocol-ID set to the new Protocol-ID: IDs-Allocation

o Local Node Descriptor contains

   * BGP Router-ID: 7.7.7.7

   * Local ASN: 7777

o Node Attribute contains

   * IID Flow Redirect TLV (TBD): IID and Flow Specification

   * Other Node Attributes.

4.2. Allocating IDs for IGP Network

In an IGP network with SR capability, a controller may allocate SIDs
for SR and send the SIDs to their responding nodes in the network.
Each of the nodes distributes the SIDs it receives to the other nodes
in the network through IGP.

For example, for a node A, the controller may allocate a SID for a
loopback of the node A and send the SID to the node A through using a
Prefix NLRI. The node A distributes the SID to the other nodes in
the network via IGP. The Prefix NLRI contains:

o Protocol-ID set to the new Protocol-ID: IDs-Allocation

o Local Node Descriptor indicating the node A

o Prefix Descriptor describing the loopback as prefix

o Prefix Attribute containing

   * Prefix-SID TLV (1158): SID for the loopback with N-Flag=1
      (Node-Flag).

   * Other Prefix Attributes.

5. IANA Considerations

   This document requests assigning a code-point from the registry "BGP-
   LS Protocol-IDs" as follows:

```
   +-------------+----------------------------------+-------------+
   | Protocol-ID | Description                      | Reference   |
   +-------------+----------------------------------+-------------+
   |     TBD     | IDs Allocation                   | Section 3   |
   +-------------+----------------------------------+-------------+
```

   This document requests assigning a code-point from the registry "BGP-
   LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute
   TLVs" as follows:

```
   +----------------+----------------------------------+-------------+
   | TLV Code Point | Description                      | Reference   |
   +----------------+----------------------------------+-------------+
   |      TBD       | IID Flow Redirect                | Section 3.1 |
   +----------------+----------------------------------+-------------+
```

6. Security Considerations

   Protocol extensions defined in this document do not affect the BGP
   security other than those as discussed in the Security Considerations
   section of [RFC7752].

7. Acknowledgements

   The authors would like to thank Robin Li, and others for their
   valuable suggestions and comments on this draft.

8. References

8.1. Normative References

   [I-D.ietf-idr-flowspec-path-redirect]
              Velde, G., Patel, K., and Z. Li, "Flowspec Indirection-id
              Redirect", draft-ietf-idr-flowspec-path-redirect-06 (work
              in progress), June 2018.

   [I-D.ietf-isis-segment-routing-extensions]
              Previdi, S., Ginsberg, L., Filsfils, C., Bashandy, A.,
              Gredler, H., Litkowski, S., Decraene, B., and J. Tantsura,
              "IS-IS Extensions for Segment Routing", draft-ietf-isis-
              segment-routing-extensions-19 (work in progress), July
              2018.

[I-D.ietf-rtgwg-bgp-routing-large-dc]
          Lapukhov, P., Premji, A., and J. Mitchell, "Use of BGP for
          routing in large-scale data centers", draft-ietf-rtgwg-
          bgp-routing-large-dc-11 (work in progress), June 2016.

[I-D.ietf-spring-segment-routing]
          Filsfils, C., Previdi, S., Ginsberg, L., Decraene, B.,
          Litkowski, S., and R. Shakir, "Segment Routing
          Architecture", draft-ietf-spring-segment-routing-15 (work
          in progress), January 2018.

[I-D.ietf-spring-segment-routing-ldp-interop]
          Bashandy, A., Filsfils, C., Previdi, S., Decraene, B., and
          S. Litkowski, "Segment Routing interworking with LDP",
          draft-ietf-spring-segment-routing-ldp-interop-15 (work in
          progress), September 2018.

[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>.

[RFC5575] Marques, P., Sheth, N., Raszuk, R., Greene, B., Mauch, J.,
          and D. McPherson, "Dissemination of Flow Specification
          Rules", RFC 5575, DOI 10.17487/RFC5575, August 2009,
          <https://www.rfc-editor.org/info/rfc5575>.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and
          S. Ray, "North-Bound Distribution of Link-State and
          Traffic Engineering (TE) Information Using BGP", RFC 7752,
          DOI 10.17487/RFC7752, March 2016,
          <https://www.rfc-editor.org/info/rfc7752>.

8.2. Informative References

[I-D.gredler-idr-bgp-ls-segment-routing-extension]
          Gredler, H., Ray, S., Previdi, S., Filsfils, C., Chen, M.,
          and J. Tantsura, "BGP Link-State extensions for Segment
          Routing", draft-gredler-idr-bgp-ls-segment-routing-
          extension-02 (work in progress), October 2014.

[I-D.ietf-idr-bgpls-segment-routing-epe]
          Previdi, S., Talaulikar, K., Filsfils, C., Patel, K., Ray,
          S., and J. Dong, "BGP-LS extensions for Segment Routing
          BGP Egress Peer Engineering", draft-ietf-idr-bgpls-
          segment-routing-epe-17 (work in progress), October 2018.

Authors' Addresses

    Nan Wu
    Huawei
    Huawei Bld., No.156 Beiqing Rd.
    Beijing 100095
    China

    Email: eric.wu@huawei.com

    Shunwan Zhuang
    Huawei
    Huawei Bld., No.156 Beiqing Rd.
    Beijing 100095
    China

    Email: zhuangshunwan@huawei.com

    Huaimo Chen
    Huawei
    Boston, MA
    USA

    Email: Huaimo.chen@huawei.com