Network Working Group N. Wu
Internet-Draft Z. Zhuang
Intended status: Standards Track Huawei
Expires: April 21, 2016 October 19, 2015

BGP Extensions for Segment Allocation
draft-wu-idr-bgp-segment-allocation-ext-00

Abstract

This document defines extensions to the BGP-LS to distribute/push the segment information to its administrative SR domain and describes some use cases.

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 http://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 21, 2016.

Table of Contents

1. Introduction

In those networks with a central controller, it may be beneficial to allocate and manage SIDs for the network since the controller has the whole link-state database in mind. This document proposes BGP extensions to allocate SIDs in a centralized manner instead of distribution way.

2. Terminology

- MPP: MPLS Path Programming
- RR: Route Reflector
- SID: Segment Identifier
- SR-Path: Segment Routing Path

3. Motivation

3.1. Allocating Segment in BGP Networks

It is possible that BGP may be the only routing protocol in some networks, such as the one described in [I-D.ietf-rtgwg-bgp-routing-large-dc]. If Segment Routing [I-D.ietf-spring-segment-routing] is going to be used for in the dataplane, it will be better to allocate SIDs in a centralized manner since no IGP flooding mechanism to advertise now.

In order to allocating SIDs, the centralized allocator SHOULD collect BGP network topology database ahead, which at least consists of BGP speakers, prefixes and adjacencies among them. No concrete technique for collecting this database has been specified in this document.

3.2. Allocating Segment in IGP Networks

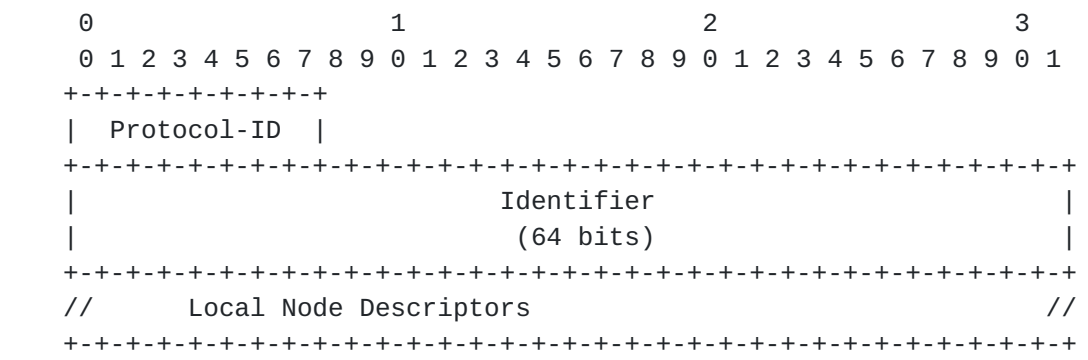
In the scenario SR & LDP interoperation described in [I-D.ietf-spring-segment-routing-ldp-interop], if mapping entries are allocated in a centralized manner, e.g. a controller, it is possible that Binding SIDs will be populated to a designated SRMS through a protocol instead of IGP, no matter whether the SRMS is a dedicated server or function module.

4. Protocol Extensions

This section defines a new Protocol-ID called as BGP-Segment-Allocation (TBA) in the BGP-LS specification. The use of a new Protocol-ID allows separation and differentiation between the NLRIs carrying Segment Allocation information from the NLRIs carrying IGP link-state information as defined in [I-D.ietf-idr-ls-distribution].

4.1. Node NLRI for Segment Allocation

This section describes the Node NLRI used for allocating the Node-SID. The Node NLRI Type uses descriptors and attributes already defined in [I-D.ietf-idr-ls-distribution]. The format of the Node NLRI Type is as follows:



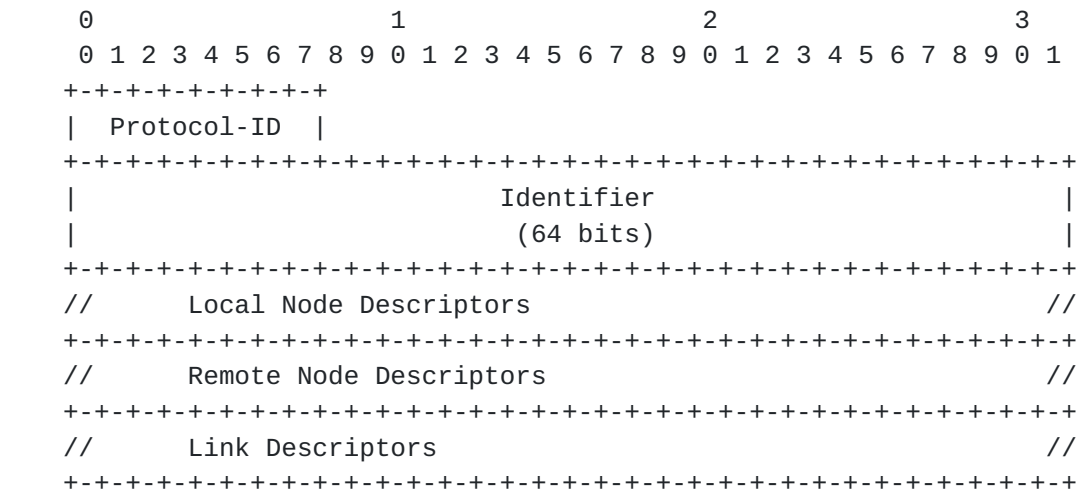
Where:

- Protocol-ID set to the new Protocol-ID: BGP-Segment-Allocation
- Node Descriptors defined in [I-D.ietf-idr-ls-distribution] can be reused

This NLRI MAY contain BGP-LS-SR TLV 1033 (SID/Label Binding) as its attribute.

4.2. Link NLRI for Segment Allocation

This section describes the Link NLRI used for allocating the Adj-SID. The format of the Link NLRI Type is as follows:



Where:

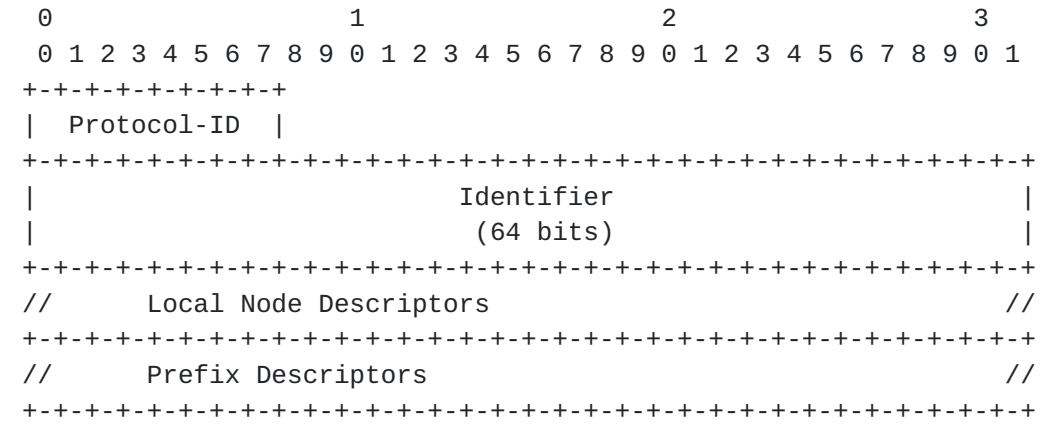
- Protocol-ID set to the new Protocol-ID: BGP-Segment-Allocation
- Node Descriptors and Link Descriptors defined in [I-D.ietf-idr-ls-distribution] can be reused.

Following TLV will be used in Link Attribute:

- BGP-LS-SR TLV 1034: SR Capabilities
- BGP-LS-SR TLV 1035: SR Algorithm
- BGP-LS-SR TLV 1099: Adj-SID
- BGP-LS-SR TLV 1036: Peer-SID
- BGP-LS-SR TLV 1037: Peer-Set-SID

4.3. Prefix NLRI for Segment Allocation

This section describes the Prefix NLRI used for Allocating the Prefix-SID. The format of the Link NLRI Type is as follows:



Where:

- Protocol-ID set to the new Protocol-ID: BGP-Segment-Allocation
- Node Descriptors and Prefix Descriptors defined in [I-D.ietf-idr-ls-distribution] can be reused.

Following TLV will be used in Prefix Attribute:

- BGP-LS-SR TLV 1034: SR Capabilities
- BGP-LS-SR TLV 1035: SR Algorithm
- BGP-LS-SR TLV 1158: Prefix SID

5. Applications

5.1. Allocating Segments for BGP Networks

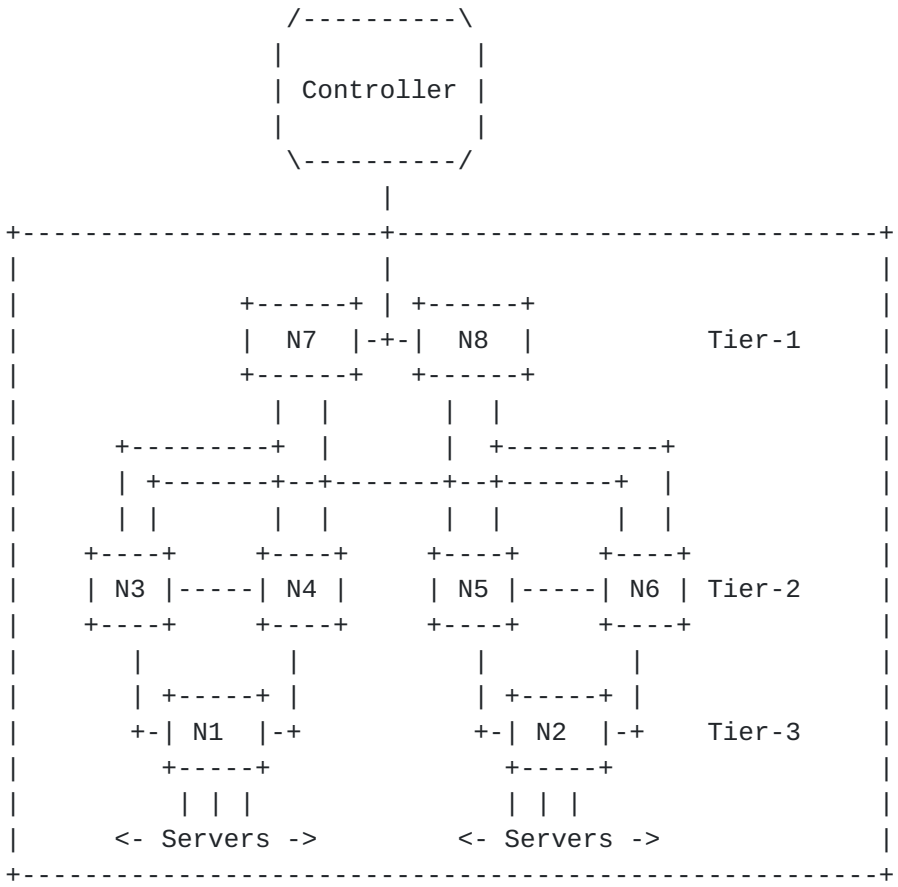
As shown below, we assume:

Each node is its own AS (Node X has AS X). The loopback of Node X is 1.1.1.x/32.

Each node peers with its neighbors via BGP session.

Each node peers with Controller via BGP session.

Local BGP-LS Identifier in Node X is set to X0000.



When the controller has collected the topology information of this BGP network, it can start segment allocation to the network.

5.1.1. Node-SID Distribution via a Prefix NLRI

A Node-SID represents a Node and has a global significance, something like a loopback of a router. Like an operator assigns a loopback's to their routers, it's expected that the Node-SID value will be assigned to every node. The assigned value can be an absolute or Index value and must be globally unique. In order to push a Node-SID for a router(e.g., N7), Controller advertise a Prefix NLRI to all the routers of the BGP-SR Network, where:

- Protocol-ID set to the new Protocol-ID: BGP-Segment-Allocation
- Local Node Descriptors contains
  - BGP Router-ID: 7.7.7.7
  - Local ASN: AS7
  - BGP-LS Identifier: 70000
- Prefix Descriptors
  - 7.7.7.7/32
- Prefix Attribute contains
  - BGP-LS-SR TLV 1034: SR Capabilities
  - BGP-LS-SR TLV 1035: SR Algorithm
  - BGP-LS-SR TLV 1158: Prefix SID, With the N-flag (node-SID flag) set.
  - Other Prefix Attributes.

5.1.2. Adj-SID Distribution via a Link NLRI

In order to push a Adj-SID for a router(e.g., N7 connects to N8), Controller advertise a Link NLRI to all the routers of the BGP-SR Network, where:

- Protocol-ID set to the new Protocol-ID: BGP-Segment-Allocation
- Local Node Descriptors contains
  - BGP Router-ID: 7.7.7.7
  - Local ASN: AS7

    * BGP-LS Identifier: 70000

  o Remote Node Descriptors contains

    * BGP Router-ID: 8.8.8.8

    * Local ASN: AS8

    * BGP-LS Identifier: 80000

  o Link Descriptors

    * BGP session IPv4 local address: 7.7.7.7

    * BGP session IPv4 peer address: 8.8.8.8

  o Link Attribute contains

    * BGP-LS-SR TLV 1034: SR Capabilities

    * BGP-LS-SR TLV 1035: SR Algorithm

    * BGP-LS-EPE TLV 1036: Peer-Node-SID

    * Other Prefix Attributes.

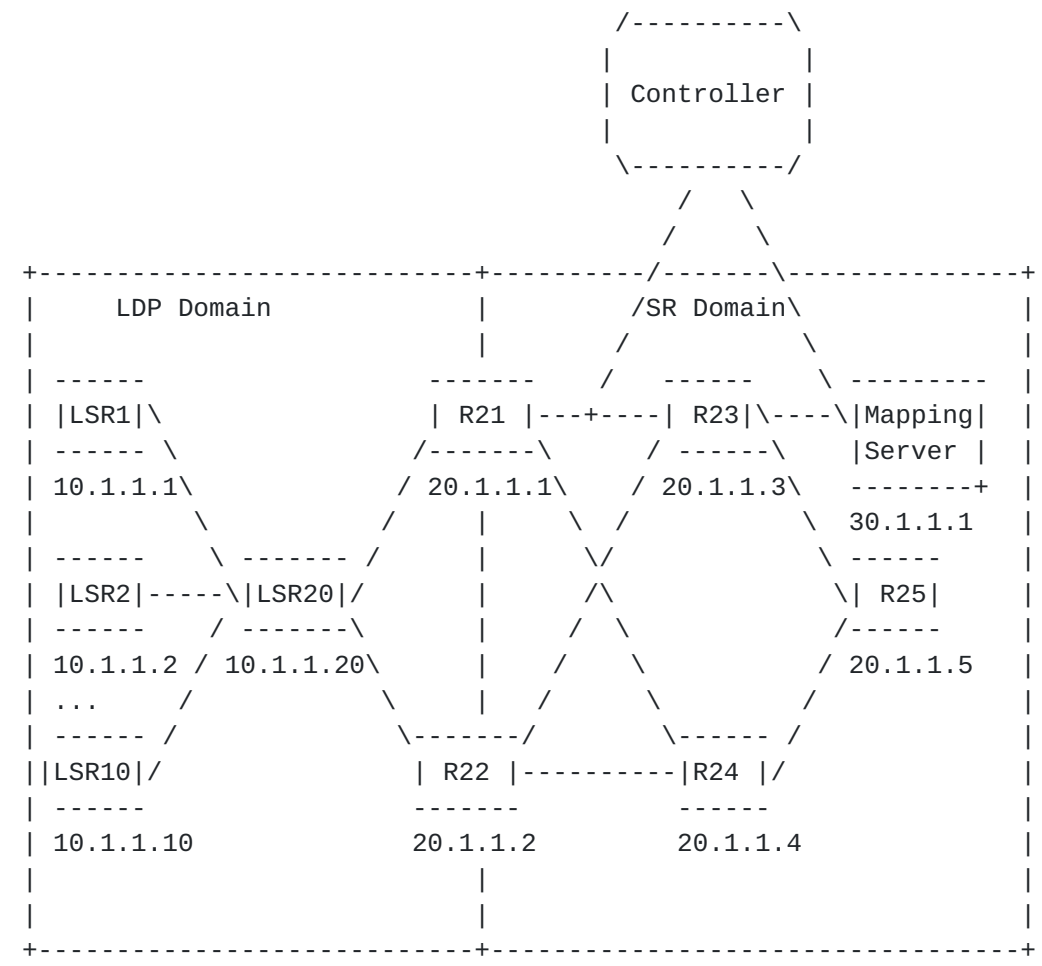
  In the similar way, the controller can distribute Peer-Adj-SID and Peer-Set-SID.

5.2. Allocating Segments for IGP Networks

  In IGP networks deployed with SR, the method defined in [I-D.ietf-idr-ls-distribution] to populate the topology database and the SRGB to the controller.

  A controller may use the extensions defined in this document to populate mapping entries to the SRMS. Then the SRMS will advertise this mapping to all the SR Nodes via IGP.

  In the following figure, LSR1-10 and LSR20 are only running LDP and R21-to-R25 Routers are SR capable Routers. R21 and R22 will be running both SR and LDP as they are on the border between SR and LDP. The whole network is running single IGP let's say IS-IS.



The Node-SIDs and their corresponding label value mapping could be like this:

Prefix	Index Value	Range
10.1.1.1/32	1001	10
10.1.1.20/32	1020	1
20.1.1.1/32	2001	5

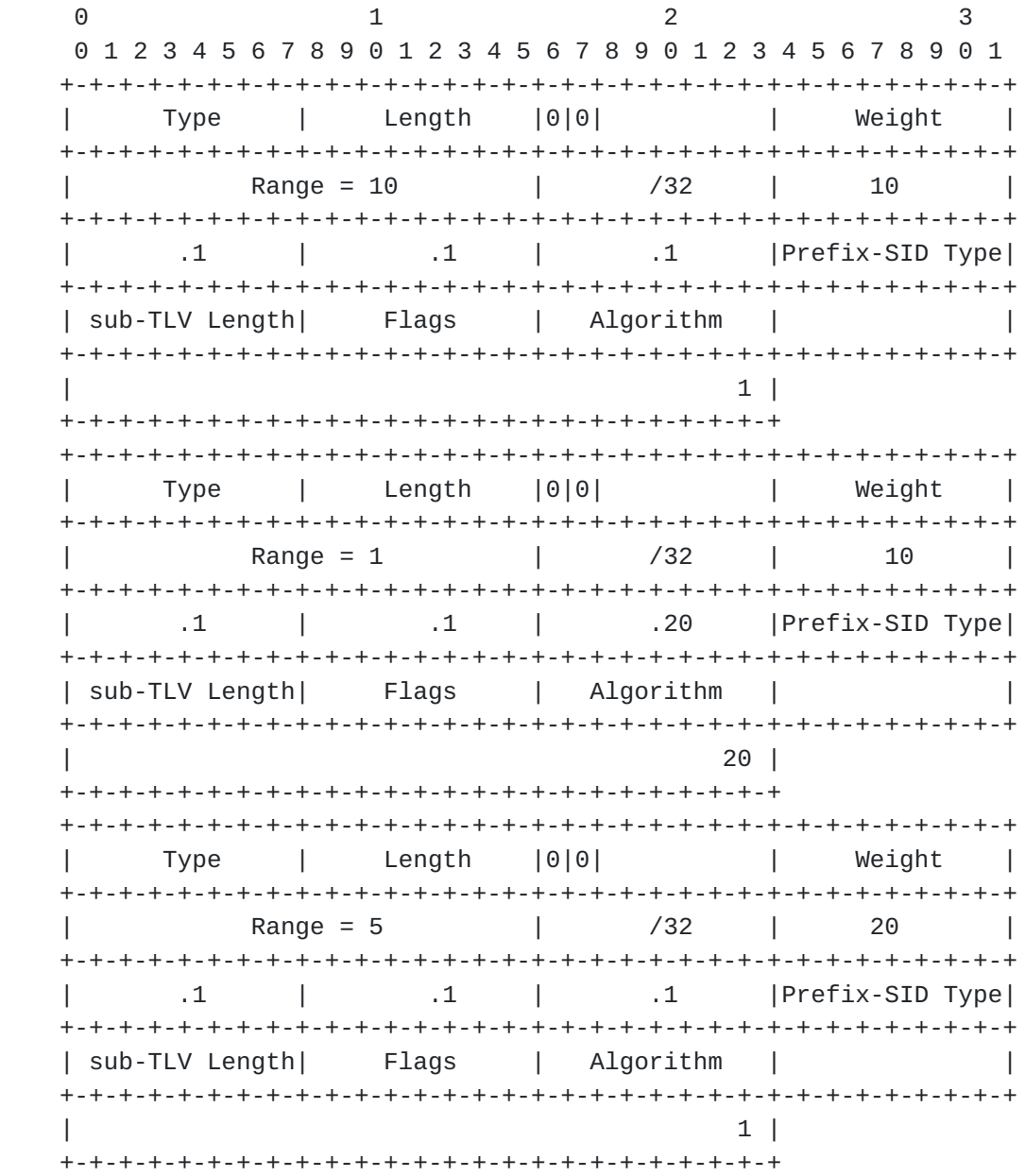
The controller will advertise a node NLRI to Mapping Server, where:

- Protocol-ID set to the new Protocol-ID: BGP-Segment-Allocation
- Local Node Descriptors contains
  - Mapping Server's node descriptor
- Node Attribute contains

* BGP-LS-SR TLV-1033: SID/Label Binding TLV

* Other Prefix Attributes

Mapping Server will convert BGP-LS-SR TLV-1033 to IS-IS TLV-149, and advertise this mapping to all the SR Nodes via IS-IS.



A node receiving a MS entry for a prefix MUST check the existence of such prefix in its link-state database prior to consider and use the associated SID. This has been defined in [I-D.ietf-isis-segment-routing-extensions].

6. IANA Considerations

TBD.

7. Security Considerations

TBD.

8. Acknowledgements

TBD.

9. References

9.1. Normative References

[I-D.ietf-idr-ls-distribution]
Gredler, H., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and TE Information using BGP", draft-ietf-idr-ls-distribution-13 (work in progress), October 2015.

[I-D.ietf-isis-segment-routing-extensions]
Previdi, S., 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-05 (work in progress), June 2015.

[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-07 (work in progress), August 2015.

[I-D.ietf-spring-segment-routing]
Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and r. rjs@rob.sh, "Segment Routing Architecture", draft-ietf-spring-segment-routing-06 (work in progress), October 2015.

[I-D.ietf-spring-segment-routing-ldp-interop]
Filsfils, C., Previdi, S., Bashandy, A., Decraene, B., and S. Litkowski, "Segment Routing interoperability with LDP", draft-ietf-spring-segment-routing-ldp-interop-00 (work in progress), October 2015.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

9.2. 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., Filsfils, C., Ray, S., Patel, K., Dong, J., and M. Chen, "Segment Routing Egress Peer Engineering BGP-LS Extensions", draft-ietf-idr-bgpls-segment-routing-epe-00 (work in progress), June 2015.

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