

Inter-Domain Routing
Internet-Draft
Intended status: Standards Track
Expires: April 22, 2019

S. Previdi, Ed.
Individual
K. Talaulikar
C. Filsfils
Cisco Systems, Inc.
K. Patel
Arrcus, Inc.
S. Ray
Individual Contributor
J. Dong
Huawei Technologies
October 19, 2018

BGP-LS extensions for Segment Routing BGP Egress Peer Engineering
draft-ietf-idr-bgpls-segment-routing-epe-17

Abstract

Segment Routing (SR) leverages source routing. A node steers a packet through a controlled set of instructions, called segments, by prepending the packet with an SR header. A segment can represent any instruction, topological or service-based. SR segments allow steering a flow through any topological path and service chain while maintaining per-flow state only at the ingress node of the SR domain.

This document describes an extension to BGP Link State (BGP-LS) for advertisement of BGP Peering Segments along with their BGP peering node information so that efficient BGP Egress Peer Engineering (EPE) policies and strategies can be computed based on Segment Routing.

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-Draft Segment Routing EPE BGP-LS Extensions October 2018

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 22, 2019.

Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](https://trustee.ietf.org/license-info) 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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	3
2.	Segment Routing Documents	4
3.	BGP Peering Segments	4
4.	BGP-LS NLRI for BGP	5
4.1.	BGP Router ID and Member ASN	6
4.2.	Mandatory BGP Node Descriptors	7
4.3.	Optional BGP Node Descriptors	7
5.	BGP-LS Attributes for BGP Peering Segments	8
5.1.	Peer-Node-SID	10
5.2.	Peer-Adj-SID	11
5.3.	Peer-Set-SID	12
6.	Illustration	12
6.1.	Reference Diagram	13
6.2.	Peer-Node-SID for Node D	15
6.3.	Peer-Node-SID for Node F	15
6.4.	Peer-Node-SID for Node E	15
6.5.	Peer-Adj-SID for Node E, Link 1	16
6.6.	Peer-Adj-SID for Node E, Link 2	16
7.	Implementation Status	16

8.	IANA Considerations	17
8.1.	New BGP-LS Protocol-ID	18
8.2.	Node Descriptors and Link Attribute TLVs	18
9.	Manageability Considerations	19
10.	Security Considerations	20

11.	Contributors	20
12.	Acknowledgements	21
13.	References	21
13.1.	Normative References	21
13.2.	Informative References	22
	Authors' Addresses	22

[1.](#) Introduction

Segment Routing (SR) leverages source routing. A node steers a packet through a controlled set of instructions, called segments, by prepending the packet with an SR header with segment identifiers (SID). A SID can represent any instruction, topological or service-based. SR segments allows to enforce a flow through any topological path or service function while maintaining per-flow state only at the ingress node of the SR domain.

The SR architecture [[RFC8402](#)] defines three types of BGP Peering Segments that may be instantiated at a BGP node:

- o Peer Node Segment (Peer-Node-SID) : instruction to steer to a specific peer node
- o Peer Adjacency Segment (Peer-Adj-SID) : instruction to steer over a specific local interface towards a specific peer node
- o Peer Set Segment (Peer-Set-SID) : instruction to load-balance to a set of specific peer nodes

SR can be directly applied to either an MPLS dataplane (SR/MPLS) with no change on the forwarding plane or to a modified IPv6 forwarding plane (SRv6).

This document describes extensions to the Link State NLRI and the BGP-LS Attribute defined for BGP-LS [[RFC7752](#)] for advertising BGP peering segments from a BGP node along with its peering topology

information (i.e. its peers, interfaces, and peering ASs) to enable computation of efficient BGP Egress Peer Engineering (BGP-EPE) policies and strategies using the SR/MPLS dataplane. The corresponding extensions for SRv6 are specified in [\[I-D.dawra-idr-bgpls-srv6-ext\]](#).

One use-case for these BGP Peering Segments is to enable computation of SR paths that enable Central BGP-EPE as described in [\[I-D.ietf-spring-segment-routing-central-epe\]](#). This use-case comprises of a centralized controller that learns the BGP Peering SIDs via BGP-LS and then uses this information to program a SR policy [\[I-D.ietf-spring-segment-routing-policy\]](#) at any node in the domain to

perform traffic steering via a specific BGP egress node to a specific EBGP peer(s) optionally also over a specific interface.

This document introduces a new BGP protocol type for BGP-LS NLRI and defines new BGP-LS Node and Link description TLVs to facilitate advertising BGP-LS Link NLRI that represent the BGP peering topology. Further, it specifies the BGP-LS Attribute TLVs for advertisement of the BGP Peering Segments (i.e. Peer Node SID, Peer Adjacency SID, and Peer Set SID) to be advertised in the same BGP-LS Link NLRI.

[2.](#) Segment Routing Documents

The main reference is the SR architecture defined in [\[RFC8402\]](#).

The SR BGP-EPE architecture and use-case is described in [\[I-D.ietf-spring-segment-routing-central-epe\]](#).

[3.](#) BGP Peering Segments

As described in [\[I-D.ietf-spring-segment-routing-central-epe\]](#), a BGP-EPE enabled Egress PE node MAY advertise SIDs corresponding to its attached peers. These SIDs are called BGP peering segments or BGP Peering SIDs. In case of EBGP, they enable the expression of source-routed inter-domain paths.

An ingress border router of an AS may compose a list of SIDs to steer a flow along a selected path within the AS, towards a selected egress border router C of the AS, and to a specific EBGP peer. At minimum, a BGP-EPE policy applied at an ingress PE involves two SIDs: the Node

SID of the chosen egress PE and then the BGP Peering SID for the chosen egress PE peer or peering interface.

Each BGP session MUST be described by a Peer Node SID. The description of the BGP session MAY be augmented by additional Peer Adjacency SIDs. Finally, multiple Peer Node SIDs or Peer Adjacency SIDs MAY be part of the same group/set in order to group EPE resources under a common Peer-Set SID.

When the extensions defined in this document are applied to the EPE use-case defined in [[I-D.ietf-spring-segment-routing-central-epe](#)], then the following BGP Peering SIDs need to be instantiated on a BGP router for each of its BGP peer sessions that are enabled for EPE:

- o One Peer-Node-SID MUST be instantiated to describe the BGP peer session.
- o One or more Peer-Adj-SID MAY be instantiated corresponding to the underlying link(s) to the directly connected BGP peer session.

- o A Peer-Set-SID MAY be instantiated and additionally associated and shared between one or more Peer-Node-SIDs or Peer-Adj-SIDs.

While an egress point in a topology usually refers to EBGp sessions between external peers, there's nothing in the extensions defined in this document that would prevent the use of these extensions in the context of IBGP sessions. However, unlike EBGp sessions which are generally between directly connected BGP routers which are also along the traffic forwarding path, IBGP peer sessions may be setup to BGP routers which are not in the forwarding path. As such, when the IBGP design includes sessions with route-reflectors, a BGP router SHOULD NOT instantiate a BGP Peering SID for those sessions to peer nodes which are not in the forwarding path since the purpose of BGP Peering SID is to steer traffic to that specific peers. Thus, the applicability for IBGP peering may be limited to only those deployments where the IBGP peer is also along with forwarding data path. Further details and the use-cases of BGP Peering SIDs and their BGP-LS extensions to IBGP deployments are beyond the scope of this document.

The BGP Peering SIDs instantiated as described above are then advertised via BGP-LS Link NLRI as described in the sections below.

4. BGP-LS NLRI for BGP

This section describes the BGP-LS NLRI encodings that describe the BGP peering and link connectivity between BGP routers.

This document specifies the advertisement of BGP peering topology information via BGP-LS NLRI which requires use of a new BGP protocol identifier.

Protocol-ID : BGP (codepoint 7 Early Allocation by IANA [Section 8](#) from the registry "BGP-LS Protocol-IDs")

The use of a new Protocol-ID allows separation and differentiation between the BGP-LS NLRI carrying BGP information from the NLRI carrying IGP link-state information as defined in [\[RFC7752\]](#).

The BGP Peering information along with their Peering Segments are advertised using BGP-LS Link NLRI with the protocol ID set to BGP. The BGP-LS Link NLRI uses the descriptor TLVs and BGP-LS Attribute TLVs as defined in [\[RFC7752\]](#). In order to correctly describe BGP nodes, new TLVs are defined in this section.

[RFC7752] defines Link NLRI Type is 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
+-----+-----+-----+-----+
| Protocol-ID |
+-----+-----+-----+-----+
|                               Identifier                               |
|                               (64 bits)                               |
+-----+-----+-----+-----+
//      Local Node Descriptors      //
+-----+-----+-----+-----+
//      Remote Node Descriptors      //
+-----+-----+-----+-----+
//      Link Descriptors              //
+-----+-----+-----+-----+
```

Figure 1: BGP-LS Link NLRI

Node Descriptors and Link Descriptors are defined in [[RFC7752](#)].

[4.1.](#) BGP Router ID and Member ASN

Two new Node Descriptors TLVs are defined in this document:

- o BGP Router Identifier (BGP Router-ID):

Type: 516 (Early Allocation by IANA [Section 8](#) from the registry "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs").

Length: 4 octets

Value: 4 octet unsigned non-zero integer representing the BGP Identifier as defined in [[RFC4271](#)] and [[RFC6286](#)].

- o Confederation Member ASN (Member-ASN)

Type: 517 (Early Allocation by IANA [Section 8](#) from the registry "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs").

Length: 4 octets

Value: 4 octet unsigned non-zero integer representing the Member ASN inside the Confederation [[RFC5065](#)].

[4.2.](#) Mandatory BGP Node Descriptors

The following Node Descriptors TLVs MUST be included in BGP-LS NLRI as Local Node Descriptors when distributing BGP information:

- o BGP Router-ID, which contains a valid BGP Identifier of the local BGP node.

- o Autonomous System Number, which contains the ASN or confederation identifier (ASN), if confederations are used, of the local BGP node.

Note that [[RFC6286](#)] ([section 2.1](#)) requires the BGP identifier (router-id) to be unique within an Autonomous System and non-zero. Therefore, the <ASN, BGP Router-ID> tuple is globally unique.

The following Node Descriptors TLVs MUST be included in BGP-LS Link NLRI as Remote Node Descriptors when distributing BGP information:

- o BGP Router-ID, which contains the valid BGP Identifier of the peer BGP node.
- o Autonomous System Number, which contains the ASN or the confederation identifier (ASN), if confederations are used, of the peer BGP node.

[4.3](#). Optional BGP Node Descriptors

The following Node Descriptors TLVs MAY be included in BGP-LS NLRI as Local Node Descriptors when distributing BGP information:

- o Member-ASN, which contains the ASN of the confederation member, if BGP confederations are used, of the local BGP node.
- o Node Descriptors as defined in [[RFC7752](#)].

The following Node Descriptors TLVs MAY be included in BGP-LS Link NLRI as Remote Node Descriptors when distributing BGP information:

- o Member-ASN, which contains the ASN of the confederation member, if BGP confederations are used, of the peer BGP node.
- o Node Descriptors as defined in defined in [[RFC7752](#)].

[5](#). BGP-LS Attributes for BGP Peering Segments

This section defines the BGP-LS Attributes corresponding to the following BGP Peer Segment SIDs:

Peer Node Segment Identifier (Peer-Node-SID)

Peer Adjacency Segment Identifier (Peer-Adj-SID)

Peer Set Segment Identifier (Peer-Set-SID)

The following new BGP-LS Link attributes TLVs are defined for use with BGP-LS Link NLRI for advertising BGP Peering SIDs:

TLV Code Point	Description	Length
1101	Peer Node Segment Identifier (Peer-Node-SID)	variable
1102	Peer Adjacency Segment Identifier (Peer-Adj-SID)	variable
1103	Peer Set Segment Identifier (Peer-Set-SID)	variable

Figure 2: BGP-LS TLV code points for BGP-EPE

Peer-Node-SID, Peer-Adj-SID, and Peer-Set-SID have all the same format defined here 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										Weight										Reserved																			
SID/Label/Index (variable)																																							

Figure 3: BGP-LS Peering SIDs TLV Format

- o Type: 1101, 1102 or 1103 (Early Allocation by IANA ([Section 8](#)) from the registry "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs").

- o Length: variable.
- o Flags: one octet of flags with the following definition:

```

 0 1 2 3 4 5 6 7
+---+---+---+---+
|V|L|B|P| Rsvd  |
+---+---+---+---+

```

Figure 4: Peering SID TLV Flags Format

- * V-Flag: Value flag. If set, then the SID carries a label value. By default the flag is SET.
 - * L-Flag: Local Flag. If set, then the value/index carried by the SID has local significance. By default the flag is SET.
 - * B-Flag: Backup Flag. If set, the SID refers to a path that is eligible for protection.
 - * P-Flag: Persistent Flag: If set, the SID is persistently allocated, i.e., the SID value remains consistent across router restart and session/interface flap.
 - * Rsvd bits: Reserved for future use and MUST be zero when originated and ignored when received.
- o Weight: 1 octet. The value represents the weight of the SID for the purpose of load balancing. An example use of the weight is described in [[RFC8402](#)].
 - o SID/Index/Label. According to the TLV length and to the V and L flags settings, it contains either:
 - * A 3 octet local label where the 20 rightmost bits are used for encoding the label value. In this case, the V and L flags MUST be SET.
 - * A 4 octet index defining the offset in the SRGB (Segment Routing Global Block as defined in [[RFC8402](#)] advertised by this router. In this case, the SRGB MUST be advertised using the extensions defined in [[I-D.ietf-idr-bgp-ls-segment-routing-ext](#)].

The values of the Peer-Node-SID, Peer-Adj-SID, and Peer-Set-SID Sub-

TLVs SHOULD be persistent across router restart.

The Peer-Node-SID TLV MUST be included in the BGP-LS Attribute for the BGP-LS Link NLRI when advertising BGP peering information for the use case described in [[I-D.ietf-spring-segment-routing-central-epe](#)] and MAY be omitted for other use cases.

The Peer-Adj-SID and Peer-Set-SID TLVs MAY be included in the BGP-LS Attribute for the BGP-LS Link NLRI when advertising BGP peering information for the use case described in [[I-D.ietf-spring-segment-routing-central-epe](#)] and MAY be omitted for other use cases.

Additional BGP-LS Link Attribute TLVs, as defined in [[RFC7752](#)] MAY be included with the BGP-LS Link NLRI in order to advertise the characteristics of the peering link.

[5.1.](#) Peer-Node-SID

The Peer-Node-SID TLV includes a SID associated with the BGP peer node that is described by a BGP-LS Link NLRI as specified in [Section 4](#).

The Peer-Node-SID, at the BGP node advertising it, has the following semantics:

- o SR header operation: NEXT (as defined in [[RFC8402](#)]).
- o Next-Hop: the connected peering node to which the segment is associated.

The Peer-Node-SID is advertised with a BGP-LS Link NLRI, where:

- o Local Node Descriptors include:
 - * Local BGP Router-ID (TLV 516) of the BGP-EPE enabled egress PE.
 - * Local ASN (TLV 512).
- o Remote Node Descriptors include:

- * Peer BGP Router-ID (TLV 516) (i.e.: the peer BGP ID used in the BGP session)
- * Peer ASN (TLV 512).
- o Link Descriptors include the addresses used by the BGP session encoded using TLVs as defined in [[RFC7752](#)]:

- * IPv4 Interface Address (TLV 259) contains the BGP session IPv4 local address.
- * IPv4 Neighbor Address (TLV 260) contains the BGP session IPv4 peer address.
- * IPv6 Interface Address (TLV 261) contains the BGP session IPv6 local address.
- * IPv6 Neighbor Address (TLV 262) contains the BGP session IPv6 peer address.
- o Link Attribute TLVs include the Peer-Node-SID TLV as defined in Figure 3.

[5.2.](#) Peer-Adj-SID

The Peer-Adj-SID TLV includes a SID associated with the underlying link to the BGP peer node that is described by a BGP-LS Link NLRI as specified in [Section 4](#).

The Peer-Adj-SID, at the BGP node advertising it, has the following semantics:

- o SR header operation: NEXT (as defined in [[RFC8402](#)]).
- o Next-Hop: the interface peer address.

The Peer-Adj-SID is advertised with a BGP-LS Link NLRI, where:

- o Local Node Descriptors include:

- * Local BGP Router-ID (TLV 516) of the BGP-EPE enabled egress PE.
- * Local ASN (TLV 512).
- o Remote Node Descriptors include:
 - * Peer BGP Router-ID (TLV 516) (i.e. the peer BGP ID used in the BGP session).
 - * Peer ASN (TLV 512).
- o Link Descriptors MUST include the following TLV, as defined in [\[RFC7752\]](#):
 - * Link Local/Remote Identifiers (TLV 258) contains the 4-octet Link Local Identifier followed by the 4-octet Link Remote

Identifier [\[RFC5307\]](#). The value 0 is used by default when the link remote identifier is unknown.

- o Additional Link Descriptors TLVs, as defined in [\[RFC7752\]](#), MAY also be included to describe the addresses corresponding to the link between the BGP routers:
 - * IPv4 Interface Address (Sub-TLV 259) contains the address of the local interface through which the BGP session is established.
 - * IPv6 Interface Address (Sub-TLV 261) contains the address of the local interface through which the BGP session is established.
 - * IPv4 Neighbor Address (Sub-TLV 260) contains the IPv4 address of the peer interface used by the BGP session.
 - * IPv6 Neighbor Address (Sub-TLV 262) contains the IPv6 address of the peer interface used by the BGP session.
- o Link Attribute TLVs include the Peer-Adj-SID TLV as defined in Figure 3.

[5.3.](#) Peer-Set-SID

The Peer-Set-SID TLV includes a SID that is shared amongst BGP peer nodes or the underlying links that are described by BGP-LS Link NLRI as specified in [Section 4](#).

The Peer-Set-SID, at the BGP node advertising it, has the following semantics:

- o SR header operation: NEXT (as defined in [[RFC8402](#)]).
- o Next-Hop: load balance across any connected interface to any peer in the associated peer set.

The Peer-Set-SID TLV containing the same SID value (encoded as defined in Figure 3) is included in the BGP-LS Attribute for all of the BGP-LS Link NLRI corresponding to the Peer Node or Peer Adjacency segments associated with the peer set.

[6](#). Illustration

[6.1](#). Reference Diagram

The following reference diagram is used throughout this section. The solution is illustrated for IPv6 with MPLS-based SIDs and the BGP-EPE topology is based on EBGP sessions between external peers.

This illustration is non-normative text provided as an example for implementers and describes the BGP-LS advertisements for the Central EPE use-case.

As stated in [Section 3](#), the solution illustrated hereafter is equally applicable to an IBGP session topology. In other words, the solution also applies to the case where C, D, F, and E are in the same AS and run IBGP sessions between each other.

```
      +-----+
      |         |
+---D |         | H
      |         |
      +-----+
```

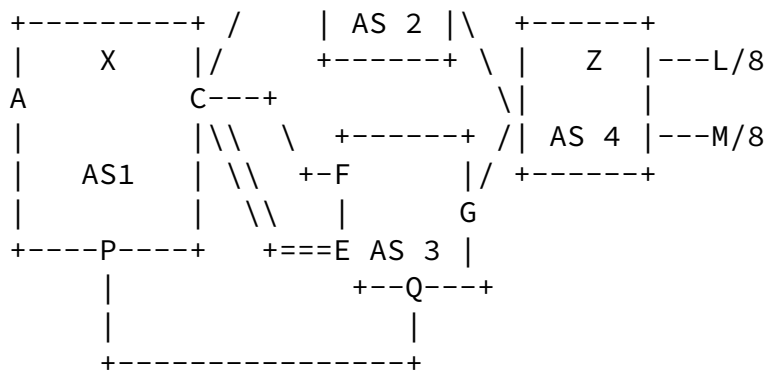


Figure 5: Reference Diagram

IP addressing:

- o C's IP address of interface to D: 2001:db8:cd::c/64, D's interface: 2001:db8:cd::d/64
- o C's IP address of interface to F: 2001:db8:cf::c/64, F's interface: 2001:db8:cf::f/64
- o C's IP address of upper interface to E: 2001:db8:ce1::c/64, E's interface: 2001:db8:ce1::e
- o C's local identifier of upper interface to E: 0.0.0.1.0.0.0.0
- o C's IP address of lower interface to E: 2001:db8:ce2::c, E's interface: 2001:db8:ce2::e
- o C's local identifier of lower interface to E: 0.0.0.2.0.0.0.0

- o Loopback of E used for EBGp multi-hop peering to C: 2001:db8:e::e/128
- o C's loopback is 2001:db8:c::c/128 with SID 64

BGP Router-IDs are C, D, F and E.

- o C's BGP Router-ID: 192.0.2.3
- o D's BGP Router-ID: 192.0.2.4

- o E's BGP Router-ID: 192.0.2.5
- o F's BGP Router-ID: 192.0.2.6

C's BGP peering:

- o Single-hop EBGp peering with neighbor 2001:db8:cd::d (D)
- o Single-hop EBGp peering with neighbor 2001:db8:cf::f (F)
- o Multi-hop EBGp peering with E on ip address 2001:db8:e::e (E)

C's resolution of the multi-hop EBGp session to E:

- o Static route 2001:db8:e::e/128 via 2001:db8:ce1::e
- o Static route 2001:db8:e::e/128 via 2001:db8:ce2::e

Node C configuration is such that:

- o A Peer-Node-SID is allocated to each peer (D, F and E).
- o An Peer-Adj-SID is defined for each recursing interface to a multi-hop peer (CE upper and lower interfaces).
- o A Peer-Set-SID is defined to include all peers in AS3 (peers F and E).

A BGP-LS Link NLRI is used in order to encode C's connectivity. The Link NLRI uses the Protocol-ID for BGP (value 7 as per Early Allocation by IANA).

Once the BGP-LS update is originated by C, it may be advertised to internal (IBGP) as well as external (EBGP) neighbors supporting the BGP-LS EPE extensions defined in this document. Note that the BGP-LS sessions may be completely separate and different from the normal BGP routing sessions described above - e.g. to a central EPE controller.

[6.2.](#) Peer-Node-SID for Node D

Descriptor TLVs used in the BGP-LS Link NLRI:

- o Local Node Descriptors (BGP Router-ID, local ASN): 192.0.2.3, AS1
- o Remote Node Descriptors (BGP Router-ID, peer ASN): 192.0.2.4, AS2
- o Link Descriptors (BGP session IPv6 local address, BGP session IPv6 neighbor address): 2001:db8:cd::c, 2001:db8:cd::d

Link Attribute TLVs used in the BGP-LS Attribute associated with the BGP-LS Link NLRI above:

- o Peer-Node-SID: 1012
- o Other Link Attributes: see [section 3.3.2 of \[RFC7752\]](#)

[6.3.](#) Peer-Node-SID for Node F

Descriptor TLVs used in the BGP-LS Link NLRI:

- o Local Node Descriptors (BGP Router-ID, ASN): 192.0.2.3, AS1
- o Remote Node Descriptors (BGP Router-ID, ASN): 192.0.2.6, AS3
- o Link Descriptors (BGP session IPv6 local address, BGP session IPv6 peer address): 2001:db8:cf::c, 2001:db8:cf::f

Link Attribute TLVs used in the BGP-LS Attribute associated with the BGP-LS Link NLRI above:

- o Peer-Node-SID: 1022
- o Peer-Set-SID: 1060
- o Other Link Attributes: see [section 3.3.2 of \[RFC7752\]](#)

[6.4.](#) Peer-Node-SID for Node E

Descriptor TLVs used in the BGP-LS Link NLRI:

- o Local Node Descriptors (BGP Router-ID, ASN): 192.0.2.3, AS1
- o Remote Node Descriptors (BGP Router-ID, ASN): 192.0.2.5, AS3
- o Link Descriptors (BGP session IPv6 local address, BGP session IPv6 peer address): 2001:db8:c::c, 2001:db8:e::e

Link Attribute TLVs used in the BGP-LS Attribute associated with the BGP-LS Link NLRI above:

- o Peer-Node-SID: 1052
- o Peer-Set-SID: 1060

[6.5.](#) Peer-Adj-SID for Node E, Link 1

Descriptor TLVs used in the BGP-LS Link NLRI:

- o Local Node Descriptors (BGP Router-ID, ASN): 192.0.2.3, AS1
- o Remote Node Descriptors (BGP Router-ID, ASN): 192.0.2.5, AS3
- o Link Descriptors (local interface identifier, IPv6 peer interface address): 0.0.0.1.0.0.0.0 , 2001:db8:ce1::e

Link Attribute TLVs used in the BGP-LS Attribute associated with the BGP-LS Link NLRI above:

- o Peer-Adj-SID: 1032
- o Other Link Attributes: see [section 3.3.2 of \[RFC7752\]](#)

[6.6.](#) Peer-Adj-SID for Node E, Link 2

Descriptor TLVs used in the BGP-LS Link NLRI:

- o Local Node Descriptors (BGP Router-ID, ASN): 192.0.2.3, AS1
- o Remote Node Descriptors (BGP Router-ID, ASN): 192.0.2.5, AS3
- o Link Descriptors (local interface identifier, IPv6 peer interface address): 0.0.0.2.0.0.0.0 , 2001:db8:ce2::e

Link Attribute TLVs used in the BGP-LS Attribute associated with the BGP-LS Link NLRI above:

- o Peer-Adj-SID: 1042
- o Other Link Attributes: see [section 3.3.2 of \[RFC7752\]](#)

[7.](#) Implementation Status

Note to RFC Editor: Please remove this section prior to publication, as well as the reference to [RFC 7942](#).

This section records the status of known implementations of the protocol defined by this specification at the time of posting of this Internet-Draft, and is based on a proposal described in [\[RFC7942\]](#). The description of implementations in this section is intended to assist the IETF in its decision processes in progressing drafts to RFCs. Please note that the listing of any individual implementation here does not imply endorsement by the IETF. Furthermore, no effort has been spent to verify the information presented here that was supplied by IETF contributors. This is not intended as, and must not be construed to be, a catalog of available implementations or their features. Readers are advised to note that other implementations may exist.

According to [\[RFC7942\]](#), "this will allow reviewers and working groups to assign due consideration to documents that have the benefit of running code, which may serve as evidence of valuable experimentation and feedback that have made the implemented protocols more mature. It is up to the individual working groups to use this information as they see fit".

Several early implementations exist and will be reported in detail in a forthcoming version of this document. For purposes of early interoperability testing, when no FCFS code point was available, implementations have made use of the following values:

+-----+-----+-----+-----+-----+			
Codepoint Description			
+-----+-----+-----+-----+-----+			
	7	Protocol-ID BGP	
	516	BGP Router-ID	
	517	BGP Confederation Member	
	1101	Peer-Node-SID	
	1102	Peer-Adj-SID	
	1103	Peer-Set-SID	
+-----+-----+-----+-----+-----+			

Figure 6: BGP-LS New Codepoints

IANA has now confirmed the Early Allocation of the above codepoints. See [Section 8](#).

[8.](#) IANA Considerations

This document defines:

A new Protocol-ID: BGP. The codepoint is from the "BGP-LS Protocol-IDs" registry.

Two new TLVs: BGP-Router-ID and BGP Confederation Member. The codepoints are in the "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs" registry.

Three new BGP-LS Attribute TLVs: Peer-Node-SID, Peer-Adj-SID and Peer-Set-SID. The codepoints are in the "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs" registry.

[8.1.](#) New BGP-LS Protocol-ID

This document defines a new value in the registry "BGP-LS Protocol-IDs":

+-----+			
Codepoint	Description	Status	
+-----+			
7	BGP	Early Allocation by IANA	
+-----+			

Figure 7: BGP Protocol Codepoint

[8.2.](#) Node Descriptors and Link Attribute TLVs

This document defines 5 new TLVs in the registry "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs":

- o Two new node descriptor TLVs
- o Three new link attribute TLVs

All the new 5 codepoints are in the same registry: "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs".

The following new Node Descriptors TLVs are defined:

Codepoint	Description	Status
516	BGP Router-ID	Early Allocation by IANA
517	BGP Confederation Member	Early Allocation by IANA

Figure 8: BGP-LS Descriptor TLVs Codepoints

The following new Link Attribute TLVs are defined:

Codepoint	Description	Status
1101	Peer-Node-SID	Early Allocation by IANA
1102	Peer-Adj-SID	Early Allocation by IANA
1103	Peer-Set-SID	Early Allocation by IANA

Figure 9: BGP-LS Attribute TLVs Codepoints

9. Manageability Considerations

The new protocol extensions introduced in this document augment the existing IGP topology information BGP-LS distribution [RFC7752] by adding support for distribution of BGP peering topology information. As such, the Manageability Considerations section of [RFC7752] applies to these new extensions as well.

Specifically, the malformed NLRI attribute tests for syntactic checks in the Fault Management section of [RFC7752] now apply to the TLVs for the BGP-LS NLRI TLVs defined in this document. The semantic or content checking for the TLVs specified in this document and their association with the BGP-LS NLRI types or their associated BGP-LS Attributes is left to the consumer of the BGP-LS information (e.g. an application or a controller) and not the BGP protocol.

A consumer of the BGP-LS information is retrieving this information from a BGP protocol component, that is doing the signaling over a BGP-LS session, via some APIs or a data model (refer [Section 1](#) and 2 of [\[RFC7752\]](#)). The handling of semantic or content errors by the consumer would be dictated by the nature of its application usage and hence is beyond the scope of this document. It may be expected that an error detected in the NLRI descriptor TLVs would result in that specific NLRI update being unusable and hence its update to be discarded along with an error log. While an error in Attribute TLVs would result in only that specific attribute being discarded with an error log.

The operator MUST be provided with the options of configuring, enabling, and disabling the advertisement of each of the Peer-Node-SID, Peer-Adj-SID, and Peer-Set-SID as well as control of which information is advertised to which internal or external peer. This is not different from what is required by a BGP speaker in terms of information origination and advertisement.

BGP Peering Segments are associated with the normal BGP routing peering sessions. However, the BGP peering information along with these Peering Segments themselves are advertised via a distinct BGP-

LS peering session. It is expected that this isolation as described in [\[RFC7752\]](#) is followed when advertising BGP peering topology information via BGP-LS.

BGP-EPE functionality enables the capability for instantiation of an SR path for traffic engineering a flow via an egress BGP router to a specific peer, bypassing the normal BGP best path routing for that flow and any routing policies implemented in BGP on that egress BGP router. As with any traffic engineering solution, the controller or application implementing the policy needs to ensure that there is no looping or mis-routing of traffic. Traffic counters corresponding to the MPLS label of the BGP Peering SID on the router would indicate the traffic being forwarded based on the specific EPE path. Monitoring these counters and the flows hitting the corresponding MPLS forwarding entry would help identify issues, if any, with traffic engineering over the EPE paths.

[10](#). Security Considerations

[RFC7752] defines BGP-LS NLRI to which the extensions defined in this document apply. The Security Considerations section of [[RFC7752](#)] also applies to these extensions.

BGP-EPE enables engineering of traffic when leaving the administrative domain via an egress BGP router. Therefore precaution is necessary to ensure that the BGP peering information collected via BGP-LS is limited to specific controllers or applications in a secure manner. By default, Segment Routing operates within a trusted domain (refer Security Considerations section in [[RFC8402](#)] for more detail) and its security considerations also apply to BGP Peering Segments. The BGP-EPE policies are expected to be used entirely within this trusted SR domain (e.g. between multiple AS/domains within a single provider network).

The isolation of BGP-LS peering sessions is also required to ensure that BGP-LS topology information (including the newly added BGP peering topology) is not advertised to an external BGP peering session outside an administrative domain.

[11.](#) Contributors

Mach (Guoyi) Chen
Huawei Technologies
China

Email: mach.chen@huawei.com

Previdi, et al.

Expires April 22, 2019

[Page 20]

Internet-Draft

Segment Routing EPE BGP-LS Extensions

October 2018

Acee Lindem
Cisco Systems Inc.
US

Email: acee@cisco.com

[12.](#) Acknowledgements

The authors would like to thank Jakob Heitz, Howard Yang, Hannes Gredler, Peter Psenak, Arjun Sreekantiah and Bruno Decraene for their feedback and comments. The authors would also like to thank Susan Hares for her substantial contributions in improving the clarity of

the document during her shepherd's review.

13. References

13.1. 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>>.
- [RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", [RFC 4271](#), DOI 10.17487/RFC4271, January 2006, <<https://www.rfc-editor.org/info/rfc4271>>.
- [RFC5065] Traina, P., McPherson, D., and J. Scudder, "Autonomous System Confederations for BGP", [RFC 5065](#), DOI 10.17487/RFC5065, August 2007, <<https://www.rfc-editor.org/info/rfc5065>>.
- [RFC5307] Kompella, K., Ed. and Y. Rekhter, Ed., "IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", [RFC 5307](#), DOI 10.17487/RFC5307, October 2008, <<https://www.rfc-editor.org/info/rfc5307>>.
- [RFC6286] Chen, E. and J. Yuan, "Autonomous-System-Wide Unique BGP Identifier for BGP-4", [RFC 6286](#), DOI 10.17487/RFC6286, June 2011, <<https://www.rfc-editor.org/info/rfc6286>>.
- [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>>.

Previdi, et al.

Expires April 22, 2019

[Page 21]

Internet-Draft

Segment Routing EPE BGP-LS Extensions

October 2018

13.2. Informative References

- [I-D.dawra-idr-bgpls-srv6-ext]
Dawra, G., Filsfils, C., Talaulikar, K., Chen, M.,
daniel.bernier@bell.ca, d., Uttaro, J., Decraene, B., and

H. Elmalky, "BGP Link State extensions for IPv6 Segment Routing(SRv6)", [draft-dawra-idr-bgpls-srv6-ext-04](#) (work in progress), September 2018.

[I-D.ietf-idr-bgp-ls-segment-routing-ext]

Previdi, S., Talaulikar, K., Filsfils, C., Gredler, H., and M. Chen, "BGP Link-State extensions for Segment Routing", [draft-ietf-idr-bgp-ls-segment-routing-ext-09](#) (work in progress), October 2018.

[I-D.ietf-spring-segment-routing-central-epe]

Filsfils, C., Previdi, S., Dawra, G., Aries, E., and D. Afanasiev, "Segment Routing Centralized BGP Egress Peer Engineering", [draft-ietf-spring-segment-routing-central-epe-10](#) (work in progress), December 2017.

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

Filsfils, C., Sivabalan, S., daniel.voyer@bell.ca, d., bogdanov@google.com, b., and P. Mattes, "Segment Routing Policy Architecture", [draft-ietf-spring-segment-routing-policy-01](#) (work in progress), June 2018.

[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>>.

[RFC7942] Sheffer, Y. and A. Farrel, "Improving Awareness of Running Code: The Implementation Status Section", [BCP 205](#), [RFC 7942](#), DOI 10.17487/RFC7942, July 2016, <<https://www.rfc-editor.org/info/rfc7942>>.

Authors' Addresses

Stefano Previdi (editor)
Individual

Email: stefano@previdi.net

Ketan Talaulikar
Cisco Systems, Inc.

Email: ketant@cisco.com

Clarence Filsfils
Cisco Systems, Inc.
Brussels
Belgium

Email: cfilsfil@cisco.com

Keyur Patel
Arrcus, Inc.

Email: Keyur@arrcus.com

Saikat Ray
Individual Contributor

Email: raysaikat@gmail.com

Jie Dong
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

Email: jie.dong@huawei.com

