

IDR Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: March 8, 2020

J. Tantsura  
Apstra, Inc.  
U. Chunduri  
Futurewei Technologies  
K. Talaulikar  
Cisco Systems  
G. Mirsky  
ZTE Corp.  
N. Triantafyllis  
Apstra, Inc.  
September 5, 2019

**Signaling MSD (Maximum SID Depth) using Border Gateway Protocol Link-  
State  
draft-ietf-idr-bgp-ls-segment-routing-msd-06**

**Abstract**

This document defines a way for a Border Gateway Protocol Link-State (BGP-LS) speaker to advertise multiple types of supported Maximum SID Depths (MSDs) at node and/or link granularity.

Such advertisements allow entities (e.g., centralized controllers) to determine whether a particular Segment Identifier (SID) stack can be supported in a given network.

**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 March 8, 2020.

## Copyright Notice

Copyright (c) 2019 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 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

<a href="#">1.</a>	Introduction . . . . .	<a href="#">2</a>
<a href="#">1.1.</a>	Conventions used in this document . . . . .	<a href="#">3</a>
<a href="#">1.1.1.</a>	Terminology . . . . .	<a href="#">3</a>
<a href="#">1.1.2.</a>	Requirements Language . . . . .	<a href="#">4</a>
<a href="#">2.</a>	Advertisement of MSD via BGP-LS . . . . .	<a href="#">4</a>
<a href="#">3.</a>	Node MSD TLV . . . . .	<a href="#">4</a>
<a href="#">4.</a>	Link MSD TLV . . . . .	<a href="#">5</a>
<a href="#">5.</a>	IANA Considerations . . . . .	<a href="#">6</a>
<a href="#">6.</a>	Manageability Considerations . . . . .	<a href="#">6</a>
<a href="#">7.</a>	Security Considerations . . . . .	<a href="#">7</a>
<a href="#">8.</a>	Contributors . . . . .	<a href="#">7</a>
<a href="#">9.</a>	Acknowledgements . . . . .	<a href="#">8</a>
<a href="#">10.</a>	References . . . . .	<a href="#">8</a>
<a href="#">10.1.</a>	Normative References . . . . .	<a href="#">8</a>
<a href="#">10.2.</a>	Informative References . . . . .	<a href="#">8</a>
	Authors' Addresses . . . . .	<a href="#">9</a>

## [1.](#) Introduction

When Segment Routing (SR) [[RFC8402](#)] paths are computed by a centralized controller, it is critical that the controller learns the Maximum SID Depth (MSD) that can be imposed at each node/link on a given SR path. This ensures that the Segment Identifier (SID) stack depth of a computed path doesn't exceed the number of SIDs the node is capable of imposing.

[I-D.ietf-pce-segment-routing] defines how to signal MSD in the Path Computation Element Protocol (PCEP). The OSPF and IS-IS extensions for signaling of MSD are defined in [[RFC8476](#)] and [[RFC8491](#)] respectively.



However, if PCEP is not supported/configured on the head-end of a SR tunnel or a Binding-SID anchor node, and controller does not participate in IGP routing, it has no way of learning the MSD of nodes and links. BGP-LS [[RFC7752](#)] defines a way to advertise topology and associated attributes and capabilities of the nodes in that topology to a centralized controller. This document defines extensions to BGP-LS to advertise one or more types of MSDs at node and/or link granularity.

Other types of MSD are known to be useful. For example, [[I-D.ietf-ospf-mpls-elc](#)] and [[I-D.ietf-isis-mpls-elc](#)] define Readable Label Depth Capability (RLDC) that is used by a head-end to insert an Entropy Label (EL) at a depth that can be read by transit nodes.

In the future, it is expected that new MSD-Types will be defined to signal additional capabilities, e.g., ELs, SIDs that can be imposed through recirculation, or SIDs associated with another data plane such as IPv6. MSD advertisements MAY be useful even if SR itself is not enabled. For example, in a non-SR MPLS network, MSD defines the maximum label depth.

## **1.1. Conventions used in this document**

### **1.1.1. Terminology**

BGP-LS: Distribution of Link-State and TE Information using Border Gateway Protocol

MSD: Maximum SID Depth

PCC: Path Computation Client

PCE: Path Computation Element

PCEP: Path Computation Element Protocol

SID: Segment Identifier

SR: Segment routing

Label Imposition: Imposition is the act of modifying and/or adding labels to the outgoing label stack associated with a packet. This includes:

- o replacing the label at the top of the label stack with a new label.



- o pushing one or more new labels onto the label stack The number of labels imposed is then the sum of the number of labels that are replaced and the number of labels that are pushed. See [[RFC3031](#)] for further details.

### **1.1.2. Requirements Language**

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "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 .

## **2. Advertisement of MSD via BGP-LS**

This document describes extensions that enable BGP-LS speakers to signal the MSD capabilities of nodes and their links in a network to a BGP-LS consumer of network topology such as a centralized controller. The centralized controller can leverage this information in computation of SR paths and their instantiation on network nodes based on their MSD capabilities. When a BGP-LS speaker is originating the topology learnt via link-state routing protocols like OSPF or IS-IS, the MSD information for the nodes and their links is sourced from the underlying extensions as defined in [[RFC8476](#)] and [[RFC8491](#)] respectively. The BGP-LS speaker may also advertise the MSD information for the local node and its links when not running any link-state IGP protocol e.g. when running BGP as the only routing protocol.

The extensions introduced in this document allow for advertisement of different MSD-Types. This document does not define these MSD-Types but leverages the definition, guidelines and the code-point registry specified in [[RFC8491](#)]. This enables sharing of MSD-Types that may be defined in the future by the IGPs in BGP-LS.

## **3. Node MSD TLV**

Node MSD is encoded in a new Node Attribute TLV [[RFC7752](#)] using the following format:









- o Length: variable (multiple of 2); represents the total length of the value field in octets.
- o Value : consists of one or more pairs of a 1-octet MSD-Type and 1-octet MSD-Value.
  - \* MSD-Type : one of the values defined in the IANA registry titled "IGP MSD-Types" under the "Interior Gateway Protocol (IGP) Parameters" registry created by [[RFC8491](#)].
  - \* MSD-Value : a number in the range of 0-255. For all MSD-Types, 0 represents the lack of ability to impose an MSD stack of any depth; any other value represents that of the link when used as an outgoing interface.

## 5. IANA Considerations

This document requests assigning code-points from the registry "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs" based on table below. Early allocation for these code-points have been done by IANA.

+-----+-----+-----+-----+		+-----+-----+-----+-----+	
Code Point	Description	IS-IS TLV/Sub-TLV	
+-----+-----+-----+-----+		+-----+-----+-----+-----+	
266	Node MSD	242/23	
267	Link MSD	(22,23,25,141,222,223)/15	
+-----+-----+-----+-----+		+-----+-----+-----+-----+	

## 6. Manageability Considerations

The new protocol extensions introduced in this document augment the existing IGP topology information that is distributed via [[RFC7752](#)]. Procedures and protocol extensions defined in this document do not affect the BGP protocol operations and management other than as discussed in the Manageability Considerations section of [[RFC7752](#)]. Specifically, the malformed attribute tests for syntactic checks in the Fault Management section of [[RFC7752](#)] now encompass the new BGP-LS Attribute 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 BGP-LS Attribute 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 retrieves this information over a BGP-LS session (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.

This document only introduces new Attribute TLVs and any syntactic error in them would result in the BGP-LS Attribute being discarded with an error log. The MSD information introduced in BGP-LS by this specification, may be used by BGP-LS consumer applications like a SR path computation engine (PCE) to learn the SR SID-stack handling capabilities of the nodes in the topology. This can enable the SR PCE to perform path computations taking into consideration the size of SID Stack that the specific headend node may be able to impose. Errors in the encoding or decoding of the MSD information may result in the unavailability of such information to the SR PCE or incorrect information being made available to it. This may result in the headend node not being able to instantiate the desired SR path in its forwarding and provide the SR based optimization functionality. The handling of such errors by applications like SR PCE may be implementation specific and out of scope of this document.

The extensions, specified in this document, do not introduce any new configuration or monitoring aspects in BGP or BGP-LS other than as discussed in [[RFC7752](#)]. The manageability aspects of the MSD features are covered by [[I-D.ietf-spring-sr-yang](#)].

## **7. Security Considerations**

The advertisement of an incorrect MSD value may have negative consequences. If the value is smaller than supported, path computation may fail to compute a viable path. If the value is larger than supported, an attempt to instantiate a path that can't be supported by the head-end (the node performing the SID imposition) may occur. The presence of this information may also inform an attacker of how to induce any of the aforementioned conditions.

This document does not introduce security issues beyond those discussed in [[RFC7752](#)], [[RFC8476](#)] and [[RFC8491](#)]

## **8. Contributors**

Siva Sivabalan  
Cisco Systems Inc.  
Canada

Email: [msiva@cisco.com](mailto:msiva@cisco.com)



## 9. Acknowledgements

We like to thank Acee Lindem, Stephane Litkowski and Bruno Decraene for their reviews and valuable comments.

## 10. References

### 10.1. 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>>.
- [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>>.
- [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>>.
- [RFC8476] Tantsura, J., Chunduri, U., Aldrin, S., and P. Psenak, "Signaling Maximum SID Depth (MSD) Using OSPF", [RFC 8476](#), DOI 10.17487/RFC8476, December 2018, <<https://www.rfc-editor.org/info/rfc8476>>.
- [RFC8491] Tantsura, J., Chunduri, U., Aldrin, S., and L. Ginsberg, "Signaling Maximum SID Depth (MSD) Using IS-IS", [RFC 8491](#), DOI 10.17487/RFC8491, November 2018, <<https://www.rfc-editor.org/info/rfc8491>>.

### 10.2. Informative References

- [I-D.ietf-isis-mpls-elc]  
Xu, X., Kini, S., Psenak, P., Filsfils, C., and S. Litkowski, "Signaling Entropy Label Capability and Entropy Readable Label Depth Using IS-IS", [draft-ietf-isis-mpls-elc-08](#) (work in progress), September 2019.
- [I-D.ietf-ospf-mpls-elc]  
Xu, X., Kini, S., Psenak, P., Filsfils, C., and S. Litkowski, "Signaling Entropy Label Capability and Entropy Readable Label-stack Depth Using OSPF", [draft-ietf-ospf-mpls-elc-08](#) (work in progress), May 2019.



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

Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "PCEP Extensions for Segment Routing", [draft-ietf-pce-segment-routing-16](#) (work in progress), March 2019.

[I-D.ietf-spring-sr-yang]

Litkowski, S., Qu, Y., Lindem, A., Sarkar, P., and J. Tantsura, "YANG Data Model for Segment Routing", [draft-ietf-spring-sr-yang-13](#) (work in progress), July 2019.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", [RFC 3031](#), DOI 10.17487/RFC3031, January 2001, <<https://www.rfc-editor.org/info/rfc3031>>.

[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>>.

Authors' Addresses

Jeff Tantsura  
Apstra, Inc.

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

Uma Chunduri  
Futurewei Technologies

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

Ketan Talaulikar  
Cisco Systems

Email: [ketant@cisco.com](mailto:ketant@cisco.com)

Greg Mirsky  
ZTE Corp.

Email: [gregimirsky@gmail.com](mailto:gregimirsky@gmail.com)





Nikos Triantafyllis  
Apstra, Inc.

Email: [nikos@apstra.com](mailto:nikos@apstra.com)