Signaling MSD (Maximum SID Depth) using Border Gateway Protocol Link-State
draft-ietf-idr-bgp-ls-segment-routing-msd-08

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 19, 2020.
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:

- replacing the label at the top of the label stack with a new label.
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 (described in [RFC8491]) 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:
Figure 1: Node MSD TLV Format

Where:

- **Type**: 266
- **Length**: variable (multiple of 2); represents the total length of the value field in octets.
- **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 node. This value MUST represent the lowest value supported by any link configured for use by the advertising protocol instance.

4. Link MSD TLV

Link MSD is encoded in a new Link Attribute TLV [RFC7752] using the following format:

Where:

- **Type**: 267
Length: variable (multiple of 2); represents the total length of the value field in octets.

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.

<table>
<thead>
<tr>
<th>Code Point</th>
<th>Description</th>
<th>IS-IS TLV/Sub-TLV</th>
</tr>
</thead>
<tbody>
<tr>
<td>266</td>
<td>Node MSD</td>
<td>242/23</td>
</tr>
<tr>
<td>267</td>
<td>Link MSD</td>
<td>(22,23,25,141,222,223)/15</td>
</tr>
</tbody>
</table>

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 specify any new configuration or monitoring aspects in BGP or BGP-LS. The specification of BGP models BGP and BGP-LS models is an ongoing work based on the [I-D.ietf-idr-bgp-model]. The management of the MSD features within an ietf segment-routing stack is also an ongoing work based on the [I-D.ietf-spring-sr-yang]. Management of the segment routing in IGP is ongoing work for ISIS [I-D.ietf-isis-sr-yang], and OSPF [I-D.ietf-ospf-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.

The document does not introduce additional security issues beyond discussed in [RFC7752], [RFC8476] and [RFC8491]. However, [RFC7752] is being revised in [I-D.ietf-idr-rfc7752bis] to provide additional clarification in several portions of the specification after receiving feedback from implementers. One of the places that is being clarified is the error handling and security. It is expected that after [I-D.ietf-idr-rfc7752bis] is released that implementers will update all BGP-LS base implementations improving the error handling for protocol work (including this document) that depend on this function.
8. Contributors

Siva Sivabalan
Cisco Systems Inc.
Canada

Email: 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


10.2. Informative References

[I-D.ietf-idr-rfc7752bis]

[I-D.ietf-isis-mpls-elc]

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

[I-D.ietf-ospf-mpls-elc]

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

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

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

Decraene, B., Litkowski, S., and R. Shakir, "Segment
Routing Architecture", RFC 8402, DOI 10.17487/RFC8402,

Authors’ Addresses

Jeff Tantsura
Apstra, Inc.
Email: jefftant.ietf@gmail.com

Uma Chunduri
Futurewei Technologies
Email: umac.ietf@gmail.com

Ketan Talaulikar
Cisco Systems
Email: ketant@cisco.com

Greg Mirsky
ZTE Corp.
Email: gregimirsky@gmail.com

Nikos Triantafillis
Apstra, Inc.
Email: nikos@apstra.com