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 is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc8814.

Copyright Notice

Copyright (c) 2020 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
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.

[RFC8664] defines how to signal MSD in the Path Computation Element Protocol (PCEP). The OSPF and IS-IS extensions for the signaling of MSD are defined in [RFC8476] and [RFC8491], respectively.
However, if PCEP is not supported/configured on the head-end of an SR tunnel or a Binding-SID anchor node, and the 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 expose 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 MSDs are known to be useful. For example, [OSPF-ELC] and [ISIS-ELC] define Entropy Readable Label Depth (ERLD), which 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

MSD: Maximum SID Depth - the number of SIDs supported by a node or a link on a node

PCE: Path Computation Element

PCEP: Path Computation Element Protocol

SID: Segment Identifier as defined in [RFC8402]

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
- 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 [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 based on their MSD capabilities. When a BGP-LS speaker is originating the topology learnt via link-state routing protocols such as 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 extensions introduced in this document allow for advertisement of different MSD-Types, which are defined elsewhere and were introduced 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

The Node MSD ([RFC8476] [RFC8491]) is encoded in a new Node Attribute TLV [RFC7752] to carry the provisioned SID depth of the router identified by the corresponding Router-ID. Node MSD is the smallest MSD supported by the node on the set of interfaces configured for use. MSD values may be learned via a hardware API or may be provisioned. The following format is used:

```
+---------------+---------------+---------------+---------------+---------------+---------------+---------------+---------------+
|                |                |                |                |                |                |                |
|  Type          |  Length       |  MSD-Type     |  MSD-Value    |  MSD-Type...  |  MSD-Value... |
| +---------------+---------------+---------------+---------------+---------------+---------------+---------------+---------------+
```

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 "IGP MSD-Types" registry defined in [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

The Link MSD ([RFC8476] [RFC8491]) is defined to carry the MSD of the interface associated with the link. It is encoded in a new Link Attribute TLV [RFC7752] using the following format:

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | 1 |
| +---------------------------------------------+---------------------------------------------+
| | Type                                         | Length                                      |
| +---------------------------------------------+---------------------------------------------+
| MSD-Type | MSD-Value | MSD-Type... | MSD-Value... |
| +---------------------------------------------+---------------------------------------------+

Figure 2: Link MSD TLV 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 "IGP MSD-Types" registry defined in [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

IANA has assigned code points from the registry "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs" based on the table below.

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

Table 1: BGP-LS MSD TLV Code Points
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 Section 6 (Manageability Considerations) of [RFC7752]. Specifically, the malformed attribute tests for syntactic checks in Section 6.2.2 (Fault Management) 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 Network Layer Reachability Information (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 to Sections 1 and 2 of [RFC7752]).

This document only introduces new Attribute TLVs, and any syntactic error in them would result in the BGP-LS Attribute being discarded [RFC7752]. The MSD information introduced in BGP-LS by this specification, may be used by BGP-LS consumer applications like an SR 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 head-end 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 head-end 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 is an ongoing work based on the [BGP-MODEL].

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 procedures and protocol extensions defined in this document do not affect the BGP security model. See the “Security Considerations” Section of [RFC4271] for a discussion of BGP security. Also, refer to [RFC4272] and [RFC6952] for analyses of security issues for BGP. Security considerations for acquiring and distributing BGP-LS information are discussed in [RFC7752]. The TLVs introduced in this document are used to propagate the MSD IGP extensions defined in [RFC8476] and [RFC8491]. It is assumed that the IGP instances originating these TLVs will support
all the required security (as described in [RFC8476] and [RFC8491]) in order to prevent any
security issues when propagating the TLVs into BGP-LS. The advertisement of the node and link
attribute information defined in this document presents no significant additional risk beyond
that associated with the existing node and link attribute information already supported in
[RFC7752].

8. References

8.1. Normative References

[RFC2119] Bradner, S., “Key words for use in RFCs to Indicate Requirement Levels”, BCP 14,

Distribution of Link-State and Traffic Engineering (TE) Information Using BGP",

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP

Depth (MSD) Using OSPF", RFC 8476, DOI 10.17487/RFC8476, December 2018,

Depth (MSD) Using IS-IS", RFC 8491, DOI 10.17487/RFC8491, November 2018,

8.2. Informative References

Provider Networks", Work in Progress, Internet-Draft, draft-ietf-idr-bgp-model-09, 28 June 2020,

Entropy Label Capability and Entropy Readable Label Depth Using IS-IS", Work
in Progress, Internet-Draft, draft-ietf-isis-mpls-elc-13, 28 May 2020,

Entropy Label Capability and Entropy Readable Label Depth Using OSPF", Work
in Progress, Internet-Draft, draft-ietf-ospf-mpls-elc-15, 1 June 2020,
Acknowledgements

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

Contributors

Siva Sivabalan
Cisco Systems Inc.
Canada
Email: msiva@cisco.com

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
Amazon Web Services
Email: nikost@amazon.com