OSPF Working Group   J. Tantsura
Internet-Draft   Individual
Intended status: Standards Track   U. Chunduri
Expires: September 9, 2017   Huawei Technologies
                                                           S. Aldrin
                                                         Google, Inc
                                                           P. Psenak
                                                       Cisco Systems
                                                       March 8, 2017

             Signaling MSD (Maximum SID Depth) using OSPF
                 draft-ietf-ospf-segment-routing-msd-01

Abstract

   This document proposes a way to signal Maximum SID Depth (MSD)
   supported by a node at node and/or link granularity by an OSPF
   Router. In a Segment Routing (SR) enabled network a centralized
   controller that programs SR tunnels needs to know the MSD supported
   by the head-end at node and/or link granularity to push the SID stack
   of an appropriate depth. MSD is relevant to the head-end of a SR
   tunnel or Binding-SID anchor node where Binding-SID expansions might
   result in creation of a new SID stack. Here the term OSPF means both
   OSPFv2 and OSPFv3.

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 September 9, 2017.

Table of Contents

1. Introduction

When Segment Routing tunnels are computed by a centralized
controller, it is critical that the controller learns the MSD
"Maximum SID Depth" of the node or link SR tunnel exits over, so the
SID stack depth of a path computed doesn't exceed the number of SIDs
the node is capable of imposing. This document describes how to use
OSPF to signal the MSD of a node or link to a centralized controller.

PCEP SR extensions draft [I-D.ietf-pce-segment-routing] signals MSD
in SR PCE Capability TLV and METRIC Object. 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 to learn the MSD of nodes and links which has been
configured. BGP-LS [RFC7752] defines a way to expose topology and
associated attributes and capabilities of the nodes in that topology
to a centralized controller. MSD signaling by BGP-LS has been
defined in [I-D.tantsura-idr-bgp-ls-segment-routing-msd]. Typically,
BGP-LS is configured on a small number of nodes, that do not

necessarily act as head-ends. In order, for BGP-LS to signal MSD for
the all nodes and links in the network MSD is relevant, MSD
capabilites SHOULD be distributed to every OSPF router in the
network.

[I-D.ietf-ospf-mpls-elc] defines Readable Label Depth Capability
(RLDC) that is used by a head-end to insert Entropy Label (EL) at
appropriate depth, so it could be read by transit nodes. MSD in
contrary signals ability to push SID's stack of a particular depth.

MSD of type 1 (IANA Registry) is used to signal the number of SIDs a
node is capable of imposing, to be used by a path computation
element/controller and is only relevant to the part of the stack
created as the result of the computation. In case, there are
additional labels (e.g. service) that are to be pushed to the stack -
MSD SHOULD be adjusted to reflect that. In the future, new MSD types
could be defined to signal additional capabilities: entropy labels,
labels that can be pushed thru recirculation, or another dataplane
e.g IPv6.

[I-D.ietf-ospf-mpls-elc] defines, RLDC which indicates how many
labels a node can read to take a decision to insert an Entropy Label
(EL) and is different than how many labels a node can push as defined
by MSD in this draft.

1.1. Conventions used in this document

1.1.1. Terminology

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

OSPF: Open Shortest Path First

MSD: Maximum SID Depth

PCC: Path Computation Client

PCE: Path Computation Element

PCEP: Path Computation Element Protocol

SID: Segment Identifier

SR: Segment routing

1.2. 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].

2. Terminology

This memo makes use of the terms defined in [RFC4970].

3. Node MSD TLV

A new TLV within the body of the OSPF RI Opaque LSA, called Node MSD
TLV is defined to carry the provisioned SID depth of the router
originating the RI LSA. Node MSD is the lowest MSD supported by the
node.

```
 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             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Sub-Type and Value ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ ...
```
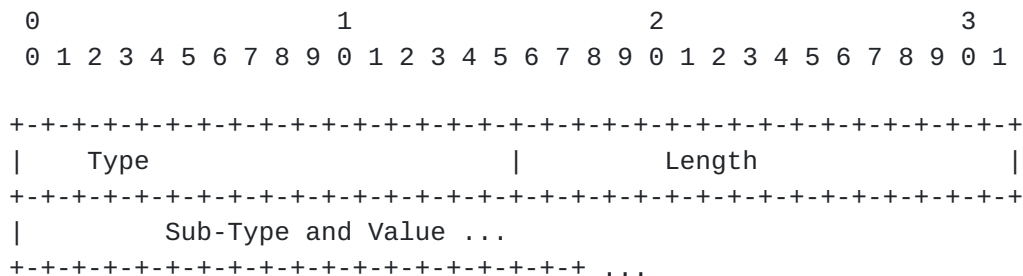
Figure 1: Node MSD TLV

The Type (2 bytes) of this TLV is 12 (Suggested value - to be
assigned by IANA).

Length is variable (minimum of 2, multiple of 2 octets) and
represents the total length of value field.

Value field consists of a 1 octet sub-type (IANA Registry) and 1
octet value.

Sub-Type 1 (IANA Section), MSD and the Value field contains maximum
MSD of the router originating the RI LSA. Node Maximum MSD is a
number in the range of 0-254. 0 represents lack of the ability to
push MSD of any depth; any other value represents that of the node.
This value SHOULD represent the lowest value supported by node.

Other Sub-types other than defined above are reserved for future
extensions.

This TLV is applicable to OSPFv2 and to OSPFv3 [RFC5838] and is
optional. The scope of the advertisement is specific to the
deployment.

4. LINK MSD sub-TLV

A new sub-TLV called Link MSD sub-TLV is defined to carry the
provisioned SID depth of the interface associated with the link.

```
 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             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Sub-Type and Value ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ ...
```
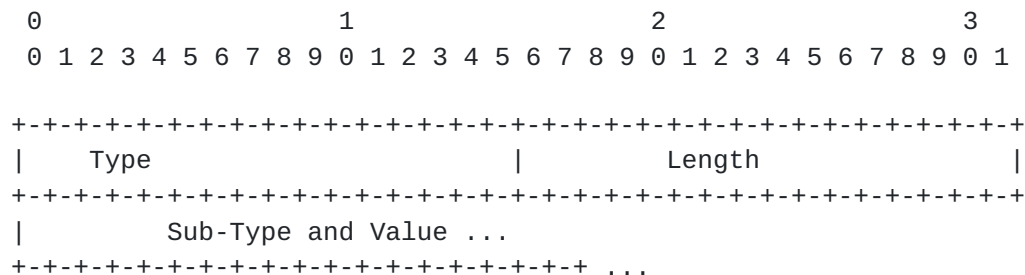
Figure 2: Link MSD Sub-TLV

The Type (2 bytes) of this TLV:

For OSPFv2, the Link level MSD value is advertised as an optional
Sub-TLV of OSPFv2 Extended Link TLV as defined in [RFC7684], and the
value is 5 (Suggested value - to be assigned by IANA)

For OSPFv3, the Link level MSD value is advertised as an optional
Sub-TLV of the Router-Link TLV as defined in
[I-D.ietf-ospf-ospfv3-lsa-extend], and the value is 3 (Suggested
value - to be assigned by IANA).

Length is variable and similar to what is defined in Section 3.

Value field consists of a 1 octet sub-type (IANA Registry) and 1
octet value.

Sub-Type 1 (IANA Section), MSD and the Value field contains Link MSD
of the router originating the corresponding LSA as specified for
OSPFv2 and OSPFv3. Link MSD is a number in the range of 0-254. 0
represents lack of the ability to push MSD of any depth; any other
value represents that of the particular link MSD value.

Other Sub-types other than defined above are reserved for future
extensions.

5. Node MSD vs Link MSD conflict resolution

   When both Node MSD and Link MSD are present, the value in the Link
   MSD MUST be used.

6. IANA Considerations

   This document includes a request to IANA to allocate TLV type codes
   for the new TLV proposed in Section 3 of this document from OSPF
   Router Information (RI) TLVs Registry as defined by [RFC4970]. Also
   for link MSD, we request IANA to allocate new sub-TLV codes as
   proposed in Section 4 from OSPFv2 Extended Link Opaque LSAs Extended
   Link TLV registry and from Router-Link TLV defined in OSPFv3 Extend-
   LSA Sub-TLV registry.

   This document also request IANA to create a new Sub-type registry as
   proposed in Section 3, Section 4.

      Value   Name   Reference
      -----   --------------------   -------------
      0   Reserved   This document
      1   MSD   This document
      2-250   Unassigned   This document
      251-254   Experimental   This document
      255   Reserved   This document

              Figure 3: MSD Sub-type Codepoints Registry

7. Security Considerations

   This document describes a mechanism to signal Segment Routing MSD
   supported at node and/or link granularity through OSPF LSA's and does
   not introduce any new security issues.

8. Contributors

   The following people contributed to this document:

   Les Ginsberg

   Email: ginsberg@cisco.com

9. Acknowledgements

   The authors would like to thank 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,
              <http://www.rfc-editor.org/info/rfc2119>.

   [RFC4970] Lindem, A., Ed., Shen, N., Vasseur, JP., Aggarwal, R., and
              S. Shaffer, "Extensions to OSPF for Advertising Optional
              Router Capabilities", RFC 4970, DOI 10.17487/RFC4970, July
              2007, <http://www.rfc-editor.org/info/rfc4970>.

10.2. Informative References

   [I-D.ietf-ospf-mpls-elc]
              Xu, X., Kini, S., Sivabalan, S., Filsfils, C., and S.
              Litkowski, "Signaling Entropy Label Capability Using
              OSPF", draft-ietf-ospf-mpls-elc-04 (work in progress),
              November 2016.

   [I-D.ietf-ospf-ospfv3-lsa-extend]
              Lindem, A., Mirtorabi, S., Roy, A., and F. Baker, "OSPFv3
              LSA Extendibility", draft-ietf-ospf-ospfv3-lsa-extend-13
              (work in progress), October 2016.

   [I-D.ietf-pce-segment-routing]
              Sivabalan, S., Medved, J., Filsfils, C., Crabbe, E.,
              Raszuk, R., Lopez, V., Tantsura, J., Henderickx, W., and
              J. Hardwick, "PCEP Extensions for Segment Routing", draft-
              ietf-pce-segment-routing-08 (work in progress), October
              2016.

   [I-D.tantsura-idr-bgp-ls-segment-routing-msd]
              Tantsura, J., Chunduri, U., Mirsky, G., and S. Sivabalan,
              "Signaling Maximum SID Depth using Border Gateway Protocol
              Link-State", draft-tantsura-idr-bgp-ls-segment-routing-
              msd-02 (work in progress), January 2017.

   [RFC5838] Lindem, A., Ed., Mirtorabi, S., Roy, A., Barnes, M., and
              R. Aggarwal, "Support of Address Families in OSPFv3",
              RFC 5838, DOI 10.17487/RFC5838, April 2010,
              <http://www.rfc-editor.org/info/rfc5838>.

   [RFC7684] Psenak, P., Gredler, H., Shakir, R., Henderickx, W.,
              Tantsura, J., and A. Lindem, "OSPFv2 Prefix/Link Attribute
              Advertisement", RFC 7684, DOI 10.17487/RFC7684, November
              2015, <http://www.rfc-editor.org/info/rfc7684>.

   [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,
              <http://www.rfc-editor.org/info/rfc7752>.

Authors' Addresses

   Jeff Tantsura
   Individual

   Email: jefftant.ietf@gmail.com

   Uma Chunduri
   Huawei Technologies

   Email: uma.chunduri@huawei.com

   Sam Aldrin
   Google, Inc

   Email: aldrin.ietf@gmail.com

   Peter Psenak
   Cisco Systems

   Email: ppsenak@cisco.com