LSR Working Group   A. Wang
Internet-Draft   China Telecom
Intended status: Standards Track   A. Lindem
Expires: February 26, 2020   Cisco Systems
                                                             J. Dong
                                                  Huawei Technologies
                                                        K. Talaulikar
                                                            P. Psenak
                                                        Cisco Systems
                                                      August 25, 2019

OSPF Extension for Prefix Originator
draft-ietf-lsr-ospf-prefix-originator-02

Abstract

   This document describes Open Shortest Path First (OSPF) v2 and OSPFv3
   encodings to advertise the router-id of the originator of inter-area
   prefixes for OSPFv2 and OSPFv3 Link-State Advertisement (LSA), which
   are needed in several use cases in multi-area OSPF use cases.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

[I-D.ietf-ospf-mpls-elc] defines mechanisms to Entropy Readable Label
Depth (ERLD) for ingress Label Switching Router (LSR) to discover
each LSR's capability of performing Entropy Label (EL) -based load-
balancing in Multi Protocol Label Switch (MPLS) networks. The
ingress LSR can use this information to push the appropriate label
stack for specific traffic, especially in segment routing
environments and other stacked LSPs scenarios.

However, in inter-area scenarios, the Area Border Router (ABR) does
not advertise the originating OSPF router-id for inter-area prefixes.
An OSPF router in one area doesn't know where the prefixes really
came from and can't determine the router that originated inter-area
prefixes and then can't judge the ERLD capabilities of the
destination. It is necessary to transfer the originator information
of these inter-area prefixes to ensure the ingress LSR constructs the
right Label stack.

More generally, draft [RFC8476] defines a mechanism to advertise
multiple types of supported Maximum SID Depths (MSD) at node and/or
link granularity. This information will be referred when the head-

end router starts to send traffic to destination prefixes. In inter-
area scenario, it is also necessary for the sender to learn the
capabilities of the receivers associated with the inter-area
prefixes.

There is also another scenario where knowing the originator of inter-
area prefixes is useful. For example, Border Gateway Protocol Link-
State (BGP-LS) [RFC7752] describes mechanisms using the BGP protocol
to advertise Link-State information. This can enable an Soft
Definition Network (SDN) controller to collect the underlay network
topology automatically.

But if the underlay network is divided into multiple areas and
running the OSPF protocol, it is not easy for the SDN controller to
rebuild the multi-area topology, because normally an ABR that
connects multiple areas will hide the detailed topology information
for these non-backbone areas. If only the router in backbone area
runs the BGP-LS protocol, it just learn and report the summary
network information from the non-backbone areas. If the SDN
controller can learn the originator of the inter-area prefixes, it is
possible for them to rebuild the inter-area topology automatically.

[RFC7794] introduces the Intermediate System to Intermediate System
(IS-IS) "IPv4/IPv6 Source Router IDs" Type-Length-Value (TLV) to
advertise the source of the prefixes redistributed from a different
IS-IS level. This TLV can be used in the above scenarios. Such
solution can also be applied in networks that run the OSPF protocol,
but the related LSAs TLV must be extended.

This draft provides such solution for the OSPFv2 and OSPFv3
protocols.

2. Conventions used in this document

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 [RFC2119] .

3. Terminology

The following terms are used in this document:

o   ABR: Area Border Router

o   ERLD: Entropy Readable Label Depth

o   EL: Entropy Label

   o IS-IS: Intermediate System to Intermediate System

   o LSA: Link-State Advertisement

   o MSD: Maximum SID Depths

   o OSPF: Open Shortest Path First

   o SID: Segment IDentifier

4. Conventions used in this document

   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 [RFC2119] .

5. Scenario Description

   Figure 1 illustrates the topology scenario when OSPF is running in
   multi-area. R0-R4 are routers in backbone area, S1-S4,T1-T4 are
   internal routers in area 1 and area 2 respectively. R1 and R3 are
   area border routers between area 0 and area 1. R2 and R4 are area
   border routers between area 0 and area 2. N1 is the network between
   router S1 and S2 and N2 is the network between router T1 and T2. Ls1
   is the loopback address of Node S1 and Lt1 is the loopback address of
   Node T1.

```
                        +-----------------+
                        |IP SDN Controller|
                        +--------+--------+
                                 |
                                 | BGP-LS
                                 |
      +--------------------+------+--------+-----+-------------+
      | +--+       +--+   ++-+   ++-+    +-++   + -+        +--+|
      | |S1+--------+S2+---+R1+---|R0+----+R2+---+T1+--------+T2||
      | +-++   N1   +-++   ++-+   +--+    +-++   ++++   N2   +-++|
      |  |          |      |             |     ||          | |
      |  |          |      |             |     ||          | |
      | +-++       +-++   ++-+           +-++   ++++        +-++|
      | |S4+--------+S3+---+R3+-----------+R4+---+T3+--------+T4||
      | +--+       +--+   ++-+           +-++   ++-+        +--+|
      |                    |             |                     |
      |                    |             |                     |
      |      Area 1        |    Area 0   |     Area 2          |
      +--------------------+-------------+---------------------+
```
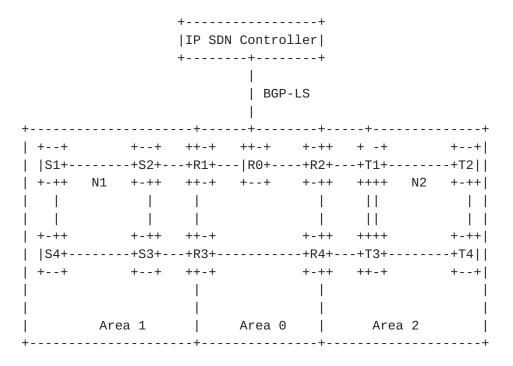
              Figure 1: OSPF Inter-Area Prefix Originator Scenario

If S1 wants to send traffic to prefix Lt1 that is connected T1 in
another area, it should know the ERLD, and MSD values that are
associated with the node T1, and then construct the right label stack
at the ingress node for the target traffic.

In another scenario, If R0 has some method to learn the originator of
network N1 and reports such information to IP SDN controller, then it
is possible for the controller to retrieval the topology in non-
backbone area. The topology retrieval process and its usage
limitation are described in the Appendix A and Appendix B.

From the above scenarios, we can conclude it is useful to introduce
and define the prefix originator sub TLV within OSPF.

6. Prefix Source Router-ID sub-TLV

[RFC7684] and [RFC8362] define the TLV extensions for OSPFv2 and
OSPFv3 respectively. These documents facilitate addition of new
attributes for prefixes. Based on these formats, we can define new
sub-TLV to advertise the "Prefix Source Router ID", as that defined
in [RFC7794].

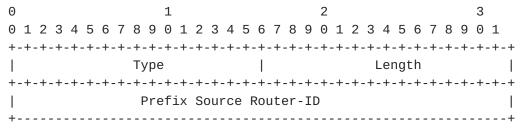The "Prefix Source Router-ID" sub-TLV has the following format:

```
 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             |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |              Prefix Source Router-ID                          |
 +--------------------------------------------------------------+
        Figure 2: Prefix Source Router-ID sub-TLV Format
```

o Source Router-ID Sub-TLV Type: TBD1[RFC7684] or TBD2 [RFC8362]

o Length: 4

o Value: Router-ID of OSPFv2/OSPFv3 source router

For OSPFv2, this sub-TLV is a sub-TLV of OSPFv2 Extended Prefix TLV,
which is included in the "OSPFv2 Extended Prefix Opaque LSA"
[RFC7684].

For OSPFv3, this sub-TLV is a sub-TLV of "Inter-Area-Prefix TLV",
which is included in the "E-Inter-Area-Prefix-LSA".

7. Extended LSA Elements of Procedure

   When an ABR, for example R2 in Fig.1, receives the Router-LSA
   announcement in area 2, it should originate the corresponding "OSPFv2
   Extended Prefix Opaque LSA" for OSPFv2 or "E-Inter-Area-Prefix-LSA"
   for OSPFv3 that includes the Source Router-ID sub-TLV for the network
   prefixes, e.g., for prefix Lt1, N2. etc., which identifies the source
   router that advertised the prefix.

   When S1 in another area receives such LSA, it then can learn that
   prefix Lt1 is associated with node T1, check the ERLD, or MSD value
   according to its necessity, and construct the right label stack at
   the ingress node S1 for the traffic destined to Lt1.

   When R0 receives such LSA, it learns the Prefix Source Router-id and
   includes it in the prefix information advertised to an SDN controller
   as described in[I-D.ietf-idr-bgp-ls-segment-routing-ext]. The SDN
   controller can then use such information to build the inter-area
   topology according to the process described in the Appendix A. The
   topology retrieval process may not suitable for some environments as
   stated in Appendix B.

8. Security Considerations

   Security concerns for OSPF are addressed in [RFC5709]

   Advertisement of the additional information defined in this document
   introduces no new security concerns

9. IANA Considerations

   This specification defines one Prefix Source Router-ID sub-TLV as
   described in Section 6. This value should be added to the existing
   OSPFv2 Extended Prefix TLV Sub-TLVs registry and OSPFv3 Extended-LSA
   Sub-TLVs registry respectively.

   Th following new sub-TLV is added to the registry of "OSPFv2 Extended
   Prefix TLV Sub-TLVs". The allocation policy is IETF Review that
   defined in [RFC7684]

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-++++++
|   Code Point  |   Description      |        Status          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-++++++
|     TBD     | Prefix Source Sub-TLV |  Allocation from IANA     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-++++++
```
Figure 3: Prefix Source sub-TLV CodePoint from OSPFv2 Extended Prefix TLV Sub-
TLVs

The following sub-TLV is added to the registry of "OSPFv3 Extended-
LSA Sub-TLVs". The allocation is IETF Review that defined in
[RFC8362]

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+++++++
|   Code Point  |  Description         |         Status            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+++++++
|     TBD       | Prefix Source Sub-TLV |   Allocation from IANA    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+++++++
```
Figure 4: Prefix Source sub-TLV CodePoint from OSPFv3 Extended-LSA Sub-TLVs

10. Acknowledgement

Many thanks to Les Ginsberg for his valuable suggestions on this
draft. And also thanks Jeff Tantsura,Rob Shakir, Van De Velde
Gunter, Goethals Dirk, Shaofu Peng, John E Drake for their valuable
comments on this draft.

11. References

11.1. 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>.

[RFC5709]   Bhatia, M., Manral, V., Fanto, M., White, R., Barnes, M.,
            Li, T., and R. Atkinson, "OSPFv2 HMAC-SHA Cryptographic
            Authentication", RFC 5709, DOI 10.17487/RFC5709, October
            2009, <https://www.rfc-editor.org/info/rfc5709>.

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

[RFC7794]   Ginsberg, L., Ed., Decraene, B., Previdi, S., Xu, X., and
            U. Chunduri, "IS-IS Prefix Attributes for Extended IPv4
            and IPv6 Reachability", RFC 7794, DOI 10.17487/RFC7794,
            March 2016, <https://www.rfc-editor.org/info/rfc7794>.

   [RFC8362] Lindem, A., Roy, A., Goethals, D., Reddy Vallem, V., and
              F. Baker, "OSPFv3 Link State Advertisement (LSA)
              Extensibility", RFC 8362, DOI 10.17487/RFC8362, April
              2018, <https://www.rfc-editor.org/info/rfc8362>.

   [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>.

11.2. Informative References

   [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-16
              (work in progress), June 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.

Appendix A. Inter-Area Topology Retrieval Process

   When an IP SDN Controller receives this information, it should
   compare the prefix NLRI that included in the BGP-LS packet. When it
   encounters the same prefix but with different source router ID, it
   should extract the corresponding area-ID, rebuild the link between
   these two different source routers in non-backbone area. Belows is
   one example that based on the Fig.1:

   Assuming we want to rebuild the connection between router S1 and
   router S2 that locates in area 1:

   a. Normally, router S1 will advertise prefix N1 within its router-
       LSA.

   b. When this router-LSA reaches the ABR router R1, it will convert
       it into summary-LSA, add the Prefix Source Router-ID sub-TLV,
       which is router id of S1 in this example.

   c. R1 then floods this extension summary-LSA to R0, which is running
       BGP-LS protocol with IP SDN Controller. The controller then
       knows the prefixes of N1 is from S1.

d. Router S2 will do the similar process, and the controller will
    also learn that prefixes N1 is also from S2.

e. Then it can reconstruct the link between S1 and S2, using the
    prefix N1. The topology within Area 1 can then be reconstructed
    accordingly.

Iterating the above process continuously, an IP SDN controller can
retrieve a detailed topology that spans multiple areas.

Appendix B. Special Considerations on Inter-Area Topology Retrieval

The above topology retrieval process can be applied in the case where
each link between routers is assigned a unique prefix. However,
there are some situations where this heuristic cannot be applied.
Specifically, the cases where the link is unnumbered or the prefix
corresponding to the link is an anycast prefix and is not unique.

The Appendix A heuristic to rebuild the topology can normally be used
if all links are numbered and the anycast prefixes correspond to
loopbacks and have a host prefix length, i.e., 32 for IPv4 prefixes
and 128 for IPv6 prefixes.

Authors' Addresses

Aijun Wang
China Telecom
Beiqijia Town, Changping District
Beijing 102209
China

Email: wangaj3@chinatelecom.cn

Acee Lindem
Cisco Systems
301 Midenhall Way
Cary, NC 27513
USA

Email: acee@cisco.com

Jie Dong
Huawei Technologies
Beijing
China

Email: jie.dong@huawei.com

Ketan Talaulikar
Cisco Systems
S.No. 154/6, Phase I, Hinjawadi
Pune   411 057
India

Email: ketant@cisco.com

Peter Psenak
Cisco Systems
Pribinova Street 10
Bratislava, Eurovea Centre, Central 3   81109
Slovakia

Email: ppsenak@cisco.com