

LSR Working Group
Internet-Draft
Intended status: Standards Track
Expires: July 8, 2019

A. Wang
China Telecom
A. Lindem
Cisco Systems
J. Dong
Huawei Technologies
K. Talaulikar
P. Psenak
Cisco Systems
January 4, 2019

OSPF Extension for Prefix Originator
draft-wang-lsr-ospf-prefix-originator-ext-01

Abstract

This document describes OSPFv2 and OSPFv3 encodings to advertise the router-id of the originator of inter-area prefixes for OSPFv2 and OSPFv3 LSAs, which are needed in several use cases in several multi-area OSPF use cases.

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 July 8, 2019.

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

- [1.](#) Introduction [2](#)
- [2.](#) Conventions used in this document [3](#)
- [3.](#) Scenario Description [3](#)
- [4.](#) Prefix Source Router-ID sub-TLV [4](#)
- [5.](#) Extended LSA Elements of Procedure [5](#)
- [6.](#) Security Considerations [5](#)
- [7.](#) IANA Considerations [5](#)
- [8.](#) Acknowledgement [6](#)
- [9.](#) References [6](#)
 - [9.1.](#) Normative References [6](#)
 - [9.2.](#) Informative References [7](#)
- [Appendix A.](#) Inter-Area Topology Retrieval Process [7](#)
- [Appendix B.](#) Special Considerations on Inter-Area Topology Retrieval [8](#)
- Authors' Addresses [8](#)

[1.](#) Introduction

[I-D.ietf-ospf-mpls-elc] defines mechanisms to signal Entropy Label Capability (ELC) and Entropy Readable Label Depth (ERLD) for ingress LSR to discover each LSR's capability of reading the maximum label stack depth and performing EL-based load-balancing in MPLS networks. The ingress LSR can use this information to push the appropriate label stack for specific FEC trafffic, 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 ELC and 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 [[I-D.ietf-ospf-segment-routing-msd](#)] 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, BGP-LS [[RFC7752](#)] describes mechanisms using the BGP protocol to advertise Link-State information. This can enable an 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 Area Border Router (ABR) that connects multiple areas will hide the detailed topology information for these non-backbone areas, and the router in backbone area that runs the BGP-LS protocol can only 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 IS-IS "IPv4/IPv6 Source Router IDs" 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 Link state Advertisements (LSAs) must be extended.

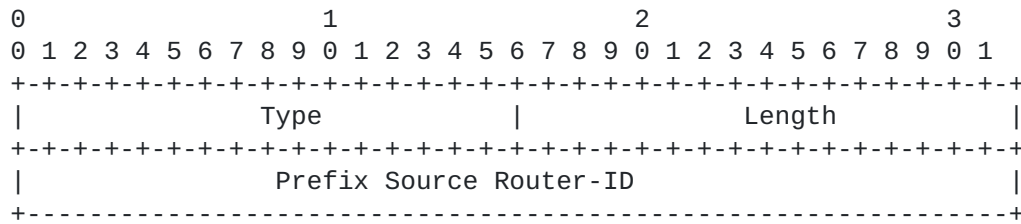
This draft provides such solution for the OSPFv2 and OSPFv3 protocols.

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. Scenario Description

Fig.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.



- o Source Router-ID Sub-TLV Type: TBD1[RFC7684] or TBD2 [[RFC8362](#)]
- o Length: 4
- o Value: Router-ID of OSPFv2/OSPFv3 source router

This sub-TLV can be included in the "OSPFv2 Extended Prefix Opaque LSA" [[RFC7684](#)] or the "E-Inter-Area-Prefix-LSA" [[RFC8362](#)].

5. 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 ELC, 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](#).

6. Security Considerations

TBD.

7. IANA Considerations

TBD.

8. Acknowledgement

Many thanks to Les Ginsberg for his valuable suggestions on this draft. And also thanks Jeff Tantsura and Rob Shakir for their valuable comments on this draft.

9. References

9.1. Normative 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-11](#)
(work in progress), October 2018.
- [I-D.ietf-ospf-mpls-elc]
Xu, X., Kini, S., Sivabalan, S., Filsfils, C., and S.
Litkowski, "Signaling Entropy Label Capability and Entropy
Readable Label-stack Depth Using OSPF", [draft-ietf-ospf-
mpls-elc-07](#) (work in progress), September 2018.
- [I-D.ietf-ospf-segment-routing-msd]
Tantsura, J., Chunduri, U., Aldrin, S., and P. Psenak,
"Signaling MSD (Maximum SID Depth) using OSPF", [draft-
ietf-ospf-segment-routing-msd-25](#) (work in progress),
October 2018.
- [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>>.
- [RFC2328] Moy, J., "OSPF Version 2", STD 54, [RFC 2328](#),
DOI 10.17487/RFC2328, April 1998,
<<https://www.rfc-editor.org/info/rfc2328>>.
- [RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF
for IPv6", [RFC 5340](#), DOI 10.17487/RFC5340, July 2008,
<<https://www.rfc-editor.org/info/rfc5340>>.
- [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>>.

9.2. Informative References

- [I-D.wang-idr-bgpls-inter-as-topology-ext]
Wang, A. and H. Chen, "BGP-LS Extension for Inter-AS Topology Retrieval", [draft-wang-idr-bgpls-inter-as-topology-ext-02](#) (work in progress), August 2018.

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: wangaj.bri@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