LSR Working Group X. Xu
Internet-Draft Alibaba Inc
Intended status: Standards Track S. Kini
Expires: December 3, 2020
P. Psenak
C. Filsfils
S. Litkowski
Cisco Systems, Inc.
M. Bocci
Nokia
June 1, 2020

Signaling Entropy Label Capability and Entropy Readable Label Depth Using OSPF
draft-ietf-ospf-mpls-elc-15

Abstract

Multiprotocol Label Switching (MPLS) has defined a mechanism to load-balance traffic flows using Entropy Labels (EL). An ingress Label Switching Router (LSR) cannot insert ELs for packets going into a given Label Switched Path (LSP) unless an egress LSR has indicated via signaling that it has the capability to process ELs, referred to as the Entropy Label Capability (ELC), on that LSP. In addition, it would be useful for ingress LSRs to know each LSR's capability for reading the maximum label stack depth and performing EL-based load-balancing, referred to as Entropy Readable Label Depth (ERLD). This document defines a mechanism to signal these two capabilities using OSPFv2 and OSPFv3 and BGP-LS.

Table of Contents

1. Introduction

[RFC6790] describes a method to load-balance Multiprotocol Label Switching (MPLS) traffic flows using Entropy Labels (EL). It also introduces the concept of Entropy Label Capability (ELC) and defines the signaling of this capability via MPLS signaling protocols. Recently, mechanisms have been defined to signal labels via link-state Interior Gateway Protocols (IGP) such as OSPFv2 [RFC8665] and OSPFv3 [RFC8666]. This draft defines a mechanism to signal the ELC using OSPFv2 and OSPFv3.

In cases where Segment Routing (SR) is used with the MPLS Data Plane (e.g., SR-MPLS [RFC8660]), it would be useful for ingress LSRs to know each intermediate LSR's capability of reading the maximum label stack depth and performing EL-based load-balancing. This capability,

referred to as Entropy Readable Label Depth (ERLD) as defined in [RFC8662], may be used by ingress LSRs to determine the position of the EL label in the stack, and whether it is necessary to insert multiple ELs at different positions in the label stack. This document defines a mechanism to signal the ERLD using OSPFv2 and OSPFv3.

2. Terminology

This memo makes use of the terms defined in [RFC6790], and [RFC8662].

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.

The key word OSPF is used throughout the document to refer to both OSPFv2 and OSPFv3.

3. Advertising ELC Using OSPF

Even though ELC is a property of the node, in some cases it is advantageous to associate and advertise the ELC with a prefix. In multi-area networks, routers may not know the identity of the prefix originator in a remote area, or may not know the capabilities of such originator. Similarly, in a multi domain network, the identity of the prefix originator and its capabilities may not be known to the ingress LSR.

If a router has multiple interfaces, the router MUST NOT announce ELC unless all of its interfaces are capable of processing ELs.

If the router supports ELs on all of its interfaces, it SHOULD advertise the ELC with every local host prefix it advertises in OSPF.

3.1. Advertising ELC Using OSPFv2

[RFC7684] defines the OSPFv2 Extended Prefix TLV to advertise additional attributes associated with a prefix. The OSPFv2 Extended Prefix TLV includes a one-octet Flags field. A new flag in the Flags field is used to signal the ELC for the prefix:

0x20 - E-Flag (ELC Flag): Set by the advertising router to indicate that the prefix originator is capable of processing ELs.

The ELC signaling MUST be preserved when an OSPF Area Border Router (ABR) distributes information between areas. To do so, an ABR MUST

originate an OSPFv2 Extended Prefix Opaque LSA [RFC7684] including the received ELC setting.

When an OSPF Autonomous System Boundary Router (ASBR) redistributes a prefix from another instance of OSPF or from some other protocol, it SHOULD preserve the ELC signaling for the prefix if it exists. To do so, an ASBR SHOULD originate an Extended Prefix Opaque LSA [RFC7684] including the ELC setting of the redistributed prefix. The flooding scope of the Extended Prefix Opaque LSA MUST match the flooding scope of the LSA that an ASBR originates as a result of the redistribution. The exact mechanism used to exchange ELC between protocol instances on an ASBR is outside of the scope of this document.

3.2. Advertising ELC Using OSPFv3

[RFC5340] defines the OSPFv3 PrefixOptions field to indicate capabilities associated with a prefix. A new bit in the OSPFv3 PrefixOptions is used to signal the ELC for the prefix:

0x40 - E-Flag (ELC Flag): Set by the advertising router to indicate that the prefix originator is capable of processing ELs.

The ELC signaling MUST be preserved when an OSPFv3 Area Border Router (ABR) distributes information between areas. The setting of the ELC Flag in the Inter-Area-Prefix-LSA [RFC5340] or in the Inter-Area-Prefix TLV [RFC8362], generated by an ABR, MUST be the same as the value the ELC Flag associated with the prefix in the source area.

When an OSPFv3 Autonomous System Boundary Router (ASBR) redistributes a prefix from another instance of OSPFv3 or from some other protocol, it SHOULD preserve the ELC signaling for the prefix if it exists. The setting of the ELC Flag in the AS-External-LSA, NSSA-LSA [RFC5340] or in the External-Prefix TLV [RFC8362], generated by an ASBR, MUST be the same as the value of the ELC Flag associated with the prefix in the source domain. The exact mechanism used to exchange ELC between protocol instances on the ASBR is outside of the scope of this document.

4. Advertising ERLD Using OSPF

The ERLD is advertised in a Node MSD TLV [RFC8476] using the ERLD-MSD type defined in [I-D.ietf-isis-mpls-elc].

If a router has multiple interfaces with different capabilities of reading the maximum label stack depth, the router MUST advertise the smallest value found across all of its interfaces.

The absence of ERLD-MSD advertisements indicates only that the advertising node does not support advertisement of this capability.

When the ERLD-MSD type is received in the OSPFv2 or OSPFv3 Link MSD Sub-TLV [RFC8476], it MUST be ignored.

The considerations for advertising the ERLD are specified in [RFC8662].

5. Signaling ELC and ERLD in BGP-LS

The OSPF extensions defined in this document can be advertised via BGP-LS (Distribution of Link-State and TE Information Using BGP) [RFC7752] using existing BGP-LS TLVs.

The ELC is advertised using the Prefix Attribute Flags TLV as defined in [I-D.ietf-idr-bgp-ls-segment-routing-ext].

The ERLD-MSD is advertised using the Node MSD TLV as defined in [I-D.ietf-idr-bgp-ls-segment-routing-msd].

6. IANA Considerations

Early allocation has been done by IANA for this document as follows:

- Flag 0x20 in the OSPFv2 Extended Prefix TLV Flags registry has been allocated by IANA to the E-Flag (ELC Flag).

- Bit 0x40 in the "OSPFv3 Prefix Options (8 bits)" registry has been allocated by IANA to the E-Flag (ELC Flag).

7. Security Considerations

This document specifies the ability to advertise additional node capabilities using OSPF and BGP-LS. As such, the security considerations as described in [RFC5340], [RFC7770], [RFC7752], [RFC7684], [RFC8476], [RFC8662], [I-D.ietf-idr-bgp-ls-segment-routing-ext] and [I-D.ietf-idr-bgp-ls-segment-routing-msd] are applicable to this document.

Incorrectly setting the E flag during origination, propagation or redistribution may lead to poor or no load-balancing of the MPLS traffic or black-holing of the MPLS traffic on the egress node.

Incorrectly setting of the ERLD value may lead to poor or no load-balancing of the MPLS traffic.

8. Contributors

The following people contributed to the content of this document and should be considered as co-authors:

Gunter Van de Velde (editor)
Nokia
Antwerp
BE

Email: gunter.van_de_velde@nokia.com

Wim Henderickx
Nokia
Belgium

Email: wim.henderickx@nokia.com

Keyur Patel
Arrcus
USA

Email: keyur@arrcus.com

9. Acknowledgements

The authors would like to thank Yimin Shen, George Swallow, Acee Lindem, Les Ginsberg, Ketan Talaulikar, Jeff Tantsura , Bruno Decraene and Carlos Pignataro for their valuable comments.

10. References

10.1. 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-16 (work in progress), June 2019.

[I-D.ietf-idr-bgp-ls-segment-routing-msd]
Tantsura, J., Chunduri, U., Talaulikar, K., Mirsky, G., and N. Triantafillis, "Signaling MSD (Maximum SID Depth) using Border Gateway Protocol - Link State", draft-ietf-idr-bgp-ls-segment-routing-msd-18 (work in progress), May 2020.

[I-D.ietf-isis-mpls-elc]
Xu, X., Kini, S., Psenak, P., Filsfils, C., Litkowski, S., and M. Bocci, "Signaling Entropy Label Capability and Entropy Readable Label Depth Using IS-IS", draft-ietf-isis-mpls-elc-13 (work in progress), May 2020.

[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>.

[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>.

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, DOI 10.17487/RFC6790, November 2012, <https://www.rfc-editor.org/info/rfc6790>.

[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>.

[RFC7770] Lindem, A., Ed., Shen, N., Vasseur, JP., Aggarwal, R., and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities", RFC 7770, DOI 10.17487/RFC7770, February 2016, <https://www.rfc-editor.org/info/rfc7770>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[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>.

[RFC8662] Kini, S., Kompella, K., Sivabalan, S., Litkowski, S., Shakir, R., and J. Tantsura, "Entropy Label for Source Packet Routing in Networking (SPRING) Tunnels", RFC 8662, DOI 10.17487/RFC8662, December 2019, <https://www.rfc-editor.org/info/rfc8662>.

10.2. Informative References

[RFC8660] Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with the MPLS Data Plane", RFC 8660, DOI 10.17487/RFC8660, December 2019, <https://www.rfc-editor.org/info/rfc8660>.

[RFC8665] Psenak, P., Ed., Previdi, S., Ed., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", RFC 8665, DOI 10.17487/RFC8665, December 2019, <https://www.rfc-editor.org/info/rfc8665>.

[RFC8666] Psenak, P., Ed. and S. Previdi, Ed., "OSPFv3 Extensions for Segment Routing", RFC 8666, DOI 10.17487/RFC8666, December 2019, <https://www.rfc-editor.org/info/rfc8666>.

Authors' Addresses

Xiaohu Xu
Alibaba Inc

Email: xiaohu.xxh@alibaba-inc.com

Sriganesh Kini

Email: sriganeshkini@gmail.com

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

Email: ppsenak@cisco.com

Clarence Filsfils
Cisco Systems, Inc.
Brussels
Belgium

Email: cfilsfil@cisco.com

Stephane Litkowski
Cisco Systems, Inc.
La Rigourdiere
Cesson Sevigne
France

Email: slitkows@cisco.com

Matthew Bocci
Nokia
Shoppenhangers Road
Maidenhead, Berks
UK

Email: matthew.bocci@nokia.com