Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: March 6, 2020

X. Xu
Alibaba Inc
S. Kini

P. Psenak
C. Filsfils
Cisco
S. Litkowski
Orange

September 3, 2019

Signaling Entropy Label Capability and Entropy Readable Label Depth Using IS-IS

draft-ietf-isis-mpls-elc-08

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 Entropy Label Capability (ELC), on that tunnel. 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 IS-IS. These mechanisms are particularly useful, where label advertisements are done via protocols like IS-IS.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

[RFC6790] describes a method to load-balance Multiprotocol Label Switching (MPLS) traffic flows using Entropy Labels (EL). "The Use of Entropy Labels in MPLS Forwarding" [RFC6790] introduces the concept of Entropy Label Capability (ELC) and defines the signalings of this capability via MPLS signaling protocols. Recently, mechanisms have been defined to signal labels via link-state Interior Gateway Protocols (IGP) such as IS-IS [I-D.ietf-isis-segment-routing-extensions]. In such scenarios, the defined signaling mechanisms are inadequate. This draft defines a mechanism to signal the ELC using IS-IS. This mechanism is useful when the label advertisement is also done via IS-IS.

In addition, in the cases where LSPs are used for whatever reasons (e.g., SR-MPLS [I-D.ietf-spring-segment-routing-mpls]), 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 [I-D.ietf-mpls-spring-entropy-label] may

be used by ingress LSRs to determine the position of the EL label in the stack, and whether it's necessary to insert multiple ELs at different positions in the label stack.

2. Terminology

This memo makes use of the terms defined in [RFC6790] and [RFC4971].

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.

3. Advertising ERLD Using IS-IS

A new MSD-type of the Node MSD sub-TLV [RFC8491], called ERLD is defined to advertise the ERLD of a given router. As shown in Figure 2, it is formatted as described in [RFC8491] with a new MSD-Type code to be assigned by IANA (the type code of 2 is desired) and the Value field is set to the ERLD in the range between 0 to 255. The scope of the advertisement depends on the application. If a router has multiple line-cards with different capabilities of reading the maximum label stack depth, the router MUST advertise the smallest one.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| MSD-Type=TBD2 |      ERLD     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: ERLD MSD-Type Format

4. Advertising ELC Using IS-IS

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 a multi-area network, 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.

One bit of the "Bit Values for Prefix Attribute Flags Sub-TLV" registry defined in [RFC7794] (Bit 3 is desired) is to be assigned by the IANA for the ELC. If a router has multiple line cards, the router MUST NOT announce the ELC for any prefixes that are locally attached unless all of its line-cards are capable of processing ELs.

If a router supports ELs on all of its line-cards, it SHOULD set the ELC for every local host prefix it advertises in IS-IS.

When a router leaks a prefix between two levels (upwards or downwards), it MUST preserve the ELC signaling for this prefix.

```
 0 1 2 3 4 5 6 7...
+-+-+-+-+-+-+-+-+...
|X|R|N|E|        ...
+-+-+-+-+-+-+-+-+...
```

When redistributing a prefix between two IS-IS protocol instances or redistributing from another protocol to an IS-IS protocol instance, a router SHOULD preserve the ELC signaling for that prefix. The exact mechanism used to exchange ELC between protocol instances running on an ASBR is outside of the scope of this document and is implementation specific.

5. Acknowledgements

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

6. BGP-LS Extension

The IS-IS extensions defined in this document can be advertised via BGP-LS [RFC7752] using existing BGP-LS TLVs.

The ELC Flag included in the Prefix Attribute Flags sub-TLV, as defined in Section 4, is advertised using the Prefix Attribute Flags TLV (TLV 1170) of the BGP-LS IPv4/IPv6 Prefix NLRI Attribute as defined in section 2.3.2 of [I-D.ietf-idr-bgp-ls-segment-routing-ext].

The ERLD MSD-type introduced for IS-IS in Section 3 is advertised using the Node MSD TLV (TLV 266) of the BGP-LS Node NLRI Attribute as defined in section 3 of [I-D.ietf-idr-bgp-ls-segment-routing-ext].

7. IANA Considerations

IANA is requested to allocate the E-bit (bit position 3 is desired) from the "Bit Values for Prefix Attribute Flags Sub-TLV" registry.

IANA is requested to allocate a MSD type (the type code of 2 is desired) from the "IGP MSD Types" registry for ERLD.

8. Security Considerations

The security considerations as described in [RFC4971] nd [I-D.ietf-mpls-spring-entropy-label] are applicable to this document.

Incorrectly setting the E flag (ELC capable) (during origination, leaking or redistribution) may lead to black-holing of the traffic on the egress node.

9. 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-isis-segment-routing-extensions]
Previdi, S., Ginsberg, L., Filsfils, C., Bashandy, A., Gredler, H., and B. Decraene, "IS-IS Extensions for Segment Routing", draft-ietf-isis-segment-routing-extensions-25 (work in progress), May 2019.

[I-D.ietf-mpls-spring-entropy-label]
Kini, S., Kompella, K., Sivabalan, S., Litkowski, S., Shakir, R., and J. Tantsura, "Entropy label for SPRING tunnels", draft-ietf-mpls-spring-entropy-label-12 (work in progress), July 2018.

[I-D.ietf-spring-segment-routing-mpls]
Bashandy, A., Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with MPLS data plane", draft-ietf-spring-segment-routing-mpls-22 (work in progress), May 2019.

[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>.

[RFC4971] Vasseur, JP., Ed., Shen, N., Ed., and R. Aggarwal, Ed., "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information", RFC 4971, DOI 10.17487/RFC4971, July 2007, <https://www.rfc-editor.org/info/rfc4971>.

[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>.

[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>.

[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>.

[RFC8491] Tantsura, J., Chunduri, U., Aldrin, S., and L. Ginsberg, "Signaling Maximum SID Depth (MSD) Using IS-IS", RFC 8491, DOI 10.17487/RFC8491, November 2018, <https://www.rfc-editor.org/info/rfc8491>.

Authors' Addresses

Xiaohu Xu
Alibaba Inc

Email: xiaohu.xxh@alibaba-inc.com

Sriganesh Kini

Email: sriganeshkini@gmail.com

Peter Psenak
Cisco

Email: ppsenak@cisco.com

Clarence Filsfils
Cisco

Email: cfilsfil@cisco.com

Stephane Litkowski
Orange

Email: stephane.litkowski@orange.com