OSPF WG
Internet-Draft
Intended status: Standards Track
Expires: January 17, 2018

S. Hegde
C. Bowers
Juniper Networks

July 16, 2017

Advertising TE protocols in OSPF
draft-hegde-ospf-advertising-te-protocols-01

Abstract

This document defines a mechanism to indicate which traffic engineering protocols are enabled on a link in OSPF. It does so by introducing a new Traffic-Engineering Protocol sub-TLV for the Link TLV in the OSPFv2 TE Opaque LSA. This document also describes mechanisms to address backward compatibility issues for routers that have not yet been upgraded to software that understands this new sub-TLV.

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].

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 January 17, 2018.

Table of Contents

1. Introduction

OSPF extensions for traffic engineering are specified in [RFC3630]. [RFC3630] defines several link attributes such as administrative group, maximum link bandwidth, and shared risk link groups (SRLGs) which can be used by traffic engineering applications. Additional link attributes for traffic engineering have subsequently been defined in other documents as well. Most recently [RFC7471] defined link attributes for delay, loss, and measured bandwidth utilization. All of the TE link attributes specified in [RFC3630] and [RFC7471] are carried in sub-TLVs in the Link TLV of the TE Opaque LSA.

The primary consumers of these traffic engineering link attributes have been RSVP-based applications that use the advertised link attributes to compute paths which will subsequently be signalled using RSVP-TE. However, these traffic engineering link attributes

have also been used by other applications, such as IP/LDP fast-reroute using loop-free alternates as described in [RFC7916]. In the future, it is likely that traffic engineering applications based on Segment Routing [I-D.ietf-spring-segment-routing] will also use these link attributes.

Existing OSPF standards do not provide a mechanism to explicitly indicate whether or not RSVP has been enabled on a link. In general, implementations have used the presence of the Link TLV in the TE Opaque LSA to infer that RSVP is enabled on a link.

This document defines a standard way to indicate whether or not RSVP, segment routing, or another future protocol is enabled on a link. In this way, implementations will not have to infer whether or not RSVP is enabled based on the presence of different sub-TLVs, but can use the explicit indication. When network operators want to use a non-RSVP traffic engineering application (such as IP/LDP FRR or segment routing), they will be able to advertise traffic engineer sub-TLVs and explicitly indicate what traffic engineering protocols are enabled on a link.

2. Motivation

The motivation of this document is to provide a mechanism to advertise TE attributes for current and future applications without ambiguity. The following objectives help to accomplish this in a range of deployment scenarios.

1. Advertise TE attributes for the link for variety of applications.

2. Allow the solution to be backward compatible so that nodes that do not understand the new advertisement do not cause issues for existing RSVP deployment.

3. Allow the solution to be extensible for any new applications that need to look at TE attributes.

4. Allow the TE protocol enabled on a link to be communicated unambiguously.

5. The solution should try to limit any increases to the quantity and size of link state advertisements.

2.1. Explicit and unambiguous indication of TE protocol

Communicating unambiguously which TE protocol is enabled on a link is important to be able to share this information with other consumers through other protocols, aside from just the IGP. For example, for a

network running both RSVP-TE and SR, it will be useful to communicate which TE protocols are enabled on which links via BGP-LS [RFC7752] to a central controller. Typically, BGP-LS relies on the IGP to distribute IGP topology and traffic engineering information so that only a few BGP-LS sessions with the central controller are needed. In order for a router running a BGP-LS session to a central controller to correctly communicate what TE protocols are enabled on the links in the IGP domain, that information first needs to be communicated unambiguously within the IGP itself.

2.2. Limit increases in link state advertisements

Over the years, the size of the networks running OSPF has grown both in terms of the total number of nodes as well as the number of links interconnecting those nodes. OSPF has proven to be quite scalable. With the advent of cloud scale computing, we expect the demands placed on OSPF by network operators to continue to grow as networks become larger and more richly interconnected. If we expect OSPF to continue to scale to meet this challenge, then as we evolve OSPF, we should be careful to limit the increases in both the quantity and size of link state advertisements to the amount necessary to solve the problem at hand. The solution described in this draft attempts to do that.

3. Solution

3.1. Traffic-engineering protocol sub-TLV

A new Traffic-Engineering Protocol sub-TLV is added to the Link TLV in the OSPFv2 TE Opaque LSA. The Traffic-Engineering Protocol sub-TLV indicates the protocols enabled on the link. The sub-TLV has flags in the value field to indicate the protocol enabled on the link. The length field is variable to allow the flags field to grow for future requirements.

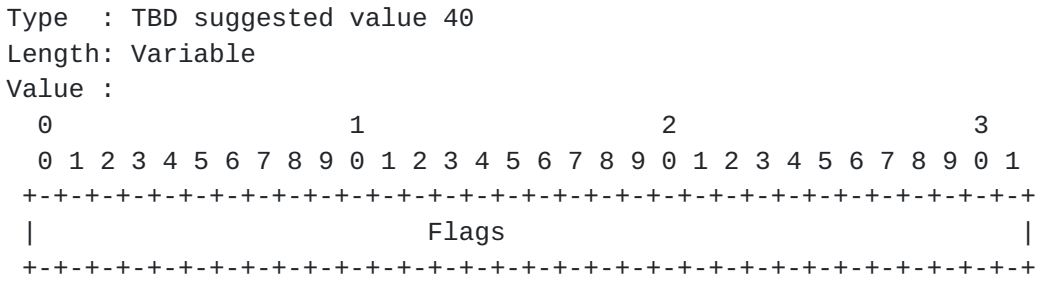

Figure 1: Traffic-Engineering Protocol sub-TLV

Type : TBA (suggested value 40)

Length: variable (in bytes)

Value: The value field consists of bits indicating the protocols enabled on the link. This document defines the two protocol values below.

Value	Protocol Name
0x01	RSVP
0x02	Segment Routing

Figure 2: Flags for the protocols

The RSVP flag is set to one to indicate that RSVP-TE is enabled on a link. The RSVP flag is set to zero to indicate that RSVP-TE is not enabled on a link.

The Segment Routing flag is set to one to indicate that Segment Routing is enabled on a link. The Segment Routing flag is set to zero to indicate that Segment Routing is not enabled on a link.

All undefined flags MUST be set to zero on transmit and ignored on receipt.

An implementation that supports the TE Protocol sub-TLV and sends the Link TLV MUST advertise the TE protocol sub-TLV in the Link TLV, even when both the RSVP and SR flags are set to zero. In other words, whenever the TE protocol sub-TLV is supported, it MUST be sent, even if no TE protocols are enabled on the link. This allows a receiving router to determine whether or not the sending router is capable of sending the TE Protocol sub-TLV.

A router supporting the TE protocol sub-TLV which receives an advertisement for a link containing the Link TLV with the TE protocol sub-TLV present SHOULD respect the values of the flags in the TE protocol sub-TLV. The receiving router SHOULD only consider links with a given TE protocol enabled for inclusion in a path using that TE protocol. Conversely, links for which the TE protocol sub-TLV is present, but for which the TE protocol flag is not set to one, SHOULD NOT be included in any TE CSPF computations on the receiving router for the protocol in question.

However, if the SR protocol flag is set to zero on a link but the adjacency-sids are advertised for that link, applications MAY use the adjacency-sid for other purposes, for example OAM.

The ability for a receiving router to determine whether or not the sending router is capable of sending the TE protocol sub-TLV is also used for backward compatibility as described in Section 4.

An implementation that supports the TE protocol sub-TLV SHOULD be able to advertise TE attribute sub-TLVs without enabling RSVP-TE signalling on the link.

4. Backward compatibility

Routers running older software that do not understand the new Traffic-Engineering protocol sub-TLV will continue to interpret the presence of the Link TLV in the TE Opaque LSA to mean that RSVP is enabled a link. A network operator may not want to or be able to upgrade all routers in the domain at the same time. There are two backward compatibility scenarios to consider depending on whether the router that doesn't understand the new TE protocol sub-TLV is an RSVP-TE ingress router or an RSVP-TE transit router.

4.1. Scenario with upgraded RSVP-TE transit router but RSVP-TE ingress router not upgraded

An upgraded RSVP-TE transit router is able to explicitly indicate that RSVP is not enabled on a link by advertising the TE protocol sub-TLV with the RSVP flag set to zero. However, an RSVP-TE ingress router that has not been upgraded to understand the new TE protocol sub-TLV will not understand that RSVP-TE is not enabled on the link, and may include the link on a path computed for RSVP-TE. When the network tries to signal an explicit path LSP using RSVP-TE through that link, it will fail. In order to avoid this scenario, an operator can use the mechanism described below.

For this scenario, the basic idea is to use the existing administrative group link attribute as a means of preventing existing RSVP implementations from using a link. The network operator defines an administrative group to mean that RSVP is not enabled on a link. We refer to this admin group the RSVP-not-enabled admin group. If the operator needs to advertise a TE sub-TLV (maximum link bandwidth, for example) on a link, but doesn't want to enable RSVP on that link, then the operator also advertises the RSVP-not-enabled admin group on that link. The operator can then use existing mechanisms to exclude links advertising the RSVP-not-enabled admin group from the constrained shortest path first (CSPF) computation used by RSVP. This will prevent RSVP implementations from attempting to signal

RSVP-TE LSPs across links that do not have RSVP enabled. Once the entire network domain is upgraded to understand the TE protocol sub-TLV in this draft, the configuration involving the RSVP-not-enabled admin group is no longer needed for this network.

To be clear, the RSVP-not-enabled admin group is an arbitrary admin group chosen by a network operator for this purpose. It is not a value that would need to be standardized.

4.2. Scenario with upgraded RSVP-TE ingress router but RSVP-TE transit router not upgraded

The other scenario to consider is when the RSVP-TE ingress router has been upgraded to understand the TE protocol sub-TLV, but the RSVP-TE transit router has not. In this case, the transit router has not been upgraded, so it is not yet capable of sending the TE protocol sub-TLV. If the transit router has RSVP-TE enabled on a link, we would like for the RSVP-TE ingress router to still be able to use the link for RSVP-TE paths. While it is possible to describe a solution for this scenario that makes use of administrative groups, we describe a simpler solution below.

The solution for this scenario relies on the following observation. If the RSVP-TE ingress router can understand that the transit router is not capable of sending the TE protocol sub-TLV, then it can continue inferring whether or not RSVP-TE is enabled on the transit router links based on the presence of the Link TLV in the TE Opaque LSA, just as it does today.

To accomplish this, we require an upgraded router to send the TE protocol sub-TLV if it sends the OSPF TE Link TLV, even when both the RSVP and SR flags are set to zero. In other words, whenever the TE protocol sub-TLV is supported, it MUST be sent, even if no TE protocols are enabled on the link. see Section 3. This allows the receiving router to interpret the absence of the TE-protocol sub-TLV in the OSPF TE Link TLV to mean that the sending router has not been upgraded. This allows the upgraded RSVP-TE ingress router to distinguish between transit routers that have been upgraded and those that haven't. When the transit router has been upgraded, then the RSVP-TE ingress router uses the information in the TE protocol sub-TLV. When the transit router has not been upgraded, then RSVP-TE ingress router contines to infer whether or not RSVP-TE is enabled on the transit router links based on the presence of TE sub-TLVs, just as it does today. The solution for this scenario requires no configuration on the part of network operators.

4.3. Need for a long term solution

The use of the adminstrative group link attribute to prevent an RSVP-TE ingress router from computing a path using a given link is an effective short term workaround to allow networks to incrementally upgrade the routers to software that understands the new TE-protocol sub-TLV. One might also consider a long term solution based solely on the use of operator-defined adminstrative groups to communicate the TE protocol enabled on a link. However, we do not consider this workaround to be an effective long term solution because it relies on operator configuration that would have to be maintained in the long term. As discussed in Section 2, continuing to have to infer which TE protocol is enabled on a link would also limit our ability to communicate this information unambiguously in an interoperable manner for use by other applications such as central controllers.

4.4. Interaction with the Extended Link Opaque LSA

The Extended Link TLV and the Extended Link Opaque LSA were introduced in [RFC7684] with the initial purpose of associating Adjacency SIDs with links for segment routing. A pure segment routing deployment that does not make use of any of the traffic engineering attributes carried in the Link TLV in the TE Opaque LSA does not need to advertise the Link TLV in the TE Opaque LSA. It only needs to advertise Extended Link TLV in the Extended Link Opaque LSA for the link. If the operator wants to make use of any traffic engineering attributes defined for the Link TLV in the TE Opaque LSA, then the routers in the network need to advertise the Link TLV in the TE Opaque LSA to carry the TE attributes as well the Extended Link TLV in the Extended Link Opaque LSA to carry the Adjacency SIDs.

5. Security Considerations

This document does not introduce any further security issues other than those discussed in [RFC3630].

6. IANA Considerations

This specification updates one OSPF registry:

The Types for sub-TLVs of the TE Link TLV Registry

i) Traffic-engineering Protocol sub-tlv = Suggested value 35

7. References

7.1. Normative References

[I-D.ietf-spring-segment-routing]
Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-09 (work in progress), July 2016.

[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>.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, DOI 10.17487/RFC3630, September 2003, <http://www.rfc-editor.org/info/rfc3630>.

[RFC7471] Giacalone, S., Ward, D., Drake, J., Atlas, A., and S. Previdi, "OSPF Traffic Engineering (TE) Metric Extensions", RFC 7471, DOI 10.17487/RFC7471, March 2015, <http://www.rfc-editor.org/info/rfc7471>.

7.2. Informative References

[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>.

[RFC7916] Litkowski, S., Ed., Decraene, B., Filsfils, C., Raza, K., Horneffer, M., and P. Sarkar, "Operational Management of Loop-Free Alternates", RFC 7916, DOI 10.17487/RFC7916, July 2016, <http://www.rfc-editor.org/info/rfc7916>.

Authors' Addresses

Shraddha Hegde
Juniper Networks
Embassy Business Park
Bangalore, KA 560093
India

Email: shraddha@juniper.net

Chris Bowers
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
US

Email: cbowers@juniper.net