Internet Engineering Task Force (IETF)
Request for Comments: 7580
Category: Standards Track
ISSN: 2070-1721

F. Zhang
Y. Lee
J. Han
Huawei
G. Bernstein
Grotto Networking
Y. Xu
CATR

June 2015

OSPF-TE Extensions for General Network Element Constraints

Abstract

Generalized Multiprotocol Label Switching (GMPLS) can be used to control a wide variety of technologies including packet switching (e.g., MPLS), time division (e.g., Synchronous Optical Network / Synchronous Digital Hierarchy (SONET/SDH) and Optical Transport Network (OTN)), wavelength (lambdas), and spatial switching (e.g., incoming port or fiber to outgoing port or fiber). In some of these technologies, network elements and links may impose additional routing constraints such as asymmetric switch connectivity, non-local label assignment, and label range limitations on links. This document describes Open Shortest Path First (OSPF) routing protocol extensions to support these kinds of constraints under the control of GMPLS.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7580.

Table of Contents

1. Introduction

Some data-plane technologies that require the use of a GMPLS control plane impose additional constraints on switching capability and label assignment. In addition, some of these technologies should be capable of performing non-local label assignment based on the nature of the technology, e.g., wavelength continuity constraint in Wavelength Switched Optical Networks (WSONs) [RFC6163]. Such constraints can lead to the requirement for link-by-link label availability in path computation and label assignment.

[RFC7579] provides efficient encodings of information needed by the routing and label assignment process in technologies such as WSON. These encodings are potentially applicable to a wider range of technologies as well. The encoding provided in [RFC7579] is protocol-neutral and can be used in routing, signaling, and/or Path Computation Element communication protocol extensions.

This document defines extensions to the OSPF routing protocol based on [RFC7579] to enhance the Traffic Engineering (TE) properties of GMPLS TE that are defined in [RFC3630], [RFC4202], and [RFC4203]. The enhancements to the TE properties of GMPLS TE links can be advertised in OSPF-TE Link State Advertisements (LSAs). The TE LSA, which is an opaque LSA with area flooding scope [RFC3630], has only one top-level Type-Length-Value (TLV) triplet and has one or more nested sub-TLVs for extensibility. The top-level TLV can take one of three values: Router Address [RFC3630], Link [RFC3630], or Node Attribute [RFC5786]. In this document, we enhance the sub-TLVs for the Link TLV in support of the general network element constraints under the control of GMPLS.

The detailed encoding of OSPF extensions is not defined in this document. [RFC7579] provides encoding details.

1.1. 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 RFC 2119 [RFC2119].

2. Node Information

According to [RFC7579], the additional node information representing node switching asymmetry constraints includes device type and connectivity matrix. Except for the device type, which is defined in [RFC7579], the other pieces of information are defined in this document.

Per [RFC7579], this document defines the Connectivity Matrix sub-TLV of the Node Attribute TLV defined in [RFC5786]. The new sub-TLV has Type 14.

Depending on the control-plane implementation being used, the Connectivity Matrix sub-TLV may be optional in some specific technologies, e.g., WSON networks. Usually, for example, in WSON networks, the Connectivity Matrix sub-TLV may be advertised in the LSAs since WSON switches are currently asymmetric. If no Connectivity Matrix sub-TLV is included, it is assumed that the switches support symmetric switching.

2.1. Connectivity Matrix

If the switching devices supporting certain data-plane technology are asymmetric, it is necessary to identify which input ports and labels can be switched to some specific labels on a specific output port.

The connectivity matrix, which can represent either the potential connectivity matrix for asymmetric switches (e.g., Reconfigurable Optical Add/Drop Multiplexers (ROADMs) and such) or fixed connectivity for an asymmetric device such as a multiplexer as defined in [RFC7446], is used to identify these restrictions.

The Connectivity Matrix is a sub-TLV of the Node Attribute TLV. The length is the length of the value field in octets. The meaning and format of this sub-TLV value field are defined in Section 2.1 of [RFC7579]. One sub-TLV contains one matrix. The Connectivity Matrix sub-TLV may occur more than once to contain multiple matrices within the Node Attribute TLV. In addition, a large connectivity matrix can be decomposed into smaller sub-matrices for transmission in multiple LSAs as described in Section 5.

3. Link Information

The most common link sub-TLVs nested in the top-level Link TLV are already defined in [RFC3630] and [RFC4203]. For example, Link ID, Administrative Group, Interface Switching Capability Descriptor (ISCD), Link Protection Type, Shared Risk Link Group (SRLG), and Traffic Engineering Metric are among the typical link sub-TLVs.

Per [RFC7579], this document defines the Port Label Restrictions sub-TLV of the Link TLV defined in [RFC3630]. The new sub-TLV has Type 34.

Generally, all the sub-TLVs above are optional, depending on control-plane implementations being used. The Port Label Restrictions sub-TLV will not be advertised when there are no restrictions on label assignment.

3.1. Port Label Restrictions

Port label restrictions describe the label restrictions that the network element (node) and link may impose on a port. These restrictions represent what labels may or may not be used on a link and are intended to be relatively static. For increased modeling flexibility, port label restrictions may be specified relative to the port in general or to a specific connectivity matrix.

For example, the port label restrictions describe the wavelength restrictions that the link and various optical devices such as Optical Cross-Connects (OXCs), ROADMs, and waveband multiplexers may impose on a port in WSON. These restrictions represent which wavelengths may or may not be used on a link and are relatively static. Detailed information about port label restrictions is provided in [RFC7446].

The Port Label Restrictions sub-TLV is a sub-TLV of the Link TLV. The length is the length of value field in octets. The meaning and format of this sub-TLV value field are defined in Section 2.2 of [RFC7579]. The Port Label Restrictions sub-TLV may occur more than once to specify a complex port constraint within the Link TLV.

4. Routing Procedures

All sub-TLVs are nested in top-level TLV(s) and contained in Opaque LSAs. The flooding rules of Opaque LSAs are specified in [RFC2328], [RFC5250], [RFC3630], and [RFC4203].

Considering the routing scalability issues in some cases, the routing protocol should be capable of supporting the separation of dynamic information from relatively static information to avoid unnecessary updates of static information when dynamic information is changed. A standards-compliant approach is to separate the dynamic information sub-TLVs from the static information sub-TLVs, each nested in a separate top-level TLV (see [RFC3630] and [RFC5786]), and advertise them in the separate OSPF-TE LSAs.

For node information, since the connectivity matrix information is static, the LSA containing the Node Attribute TLV can be updated with a lower frequency to avoid unnecessary updates.

For link information, a mechanism MAY be applied such that static information and dynamic information of one TE link are contained in separate Opaque LSAs. For example, the Port Label Restrictions sub-TLV could be nested in separate top-level Link TLVs and advertised in the separate LSAs.

As with other TE information, an implementation typically takes measures to avoid rapid and frequent updates of routing information that could cause the routing network to become swamped. See Section 3 of [RFC3630] for related details.

5. Scalability and Timeliness

This document defines two sub-TLVs for describing generic routing constraints. The examples given in [RFC7579] show that very large systems, in terms of label count or ports, can be very efficiently encoded. However, because there has been concern expressed that some possible systems may produce LSAs that exceed the IP Maximum Transmission Unit (MTU), methods should be given to allow for the splitting of general constraint LSAs into smaller LSAs that are under the MTU limit. This section presents a set of techniques that can be used for this purpose.

5.1. Different Sub-TLVs into Multiple LSAs

Two sub-TLVs are defined in this document:

1. Connectivity Matrix (carried in the Node Attribute TLV)

2. Port Label Restrictions (carried in the Link TLV)

The Connectivity Matrix sub-TLV can be carried in the Node Attribute TLV (as defined in [RFC5786]), whereas the Port Label Restrictions sub-TLV can be carried in a Link TLV, of which there can be at most one in an LSA (as defined in [RFC3630]). Note that the port label restrictions are relatively static, i.e., only would change with hardware changes or significant system reconfiguration.

5.2. Decomposing a Connectivity Matrix into Multiple Matrices

In the highly unlikely event that a Connectivity Matrix sub-TLV by itself would result in an LSA exceeding the MTU, a single large matrix can be decomposed into sub-matrices. Per [RFC7579], a connectivity matrix just consists of pairs of input and output ports that can reach each other; hence, this decomposition would be straightforward. Each of these sub-matrices would get a unique matrix identifier per [RFC7579].

From the point of view of a path computation process, prior to receiving an LSA with a Connectivity Matrix sub-TLV, no connectivity restrictions are assumed, i.e., the standard GMPLS assumption of any port to any port reachability holds. Once a Connectivity Matrix sub-TLV is received, path computation would know that connectivity is restricted and use the information from all Connectivity Matrix sub-TLVs received to understand the complete connectivity potential of the system. Prior to receiving any Connectivity Matrix sub-TLVs, path computation may compute a path through the system when, in fact, no path exists. In between the reception of an additional Connectivity Matrix sub-TLV, path computation may not be able to find a path through the system when one actually exists. Both cases are currently encountered and handled with existing GMPLS mechanisms. Due to the reliability mechanisms in OSPF, the phenomena of late or missing Connectivity Matrix sub-TLVs would be relatively rare.

In the case where the new sub-TLVs or their attendant encodings are malformed, the proper action would be to log the problem and ignore just the sub-TLVs in GMPLS path computations rather than ignoring the entire LSA.

6. Security Considerations

This document does not introduce any further security issues other than those discussed in [RFC3630], [RFC4203], and [RFC5250].

For general security aspects relevant to GMPLS-controlled networks, please refer to [RFC5920].

7. Manageability

No existing management tools handle the additional TE parameters as defined in this document and distributed in OSPF-TE. The existing MIB module contained in [RFC6825] allows the TE information distributed by OSPF-TE to be read from a network node; this MIB module could be augmented (possibly by a sparse augmentation) to report this new information.

The current environment in the IETF favors the Network Configuration Protocol (NETCONF) [RFC6241] and YANG [RFC6020] over SNMP and MIB modules. Work is in progress in the TEAS working group to develop a YANG module to represent the generic TE information that may be present in a Traffic Engineering Database (TED). This model may be extended to handle the additional information described in this document to allow that information to be read from network devices or exchanged between consumers of the TED. Furthermore, links state

export using BGP [BGP-LS] enables the export of TE information from a network using BGP. Work could realistically be done to extend BGP-LS to also carry the information defined in this document.

It is not envisaged that the extensions defined in this document will place substantial additional requirements on Operations, Administration, and Maintenance (OAM) mechanisms currently used to diagnose and debug OSPF systems. However, tools that examine the contents of opaque LSAs will need to be enhanced to handle these new sub-TLVs.

8. IANA Considerations

IANA has allocated new sub-TLVs as defined in Sections 2 and 3 as follows:

8.1. Node Information

IANA maintains the "Open Shortest Path First (OSPF) Traffic Engineering TLVs" registry with a sub-registry called "Types for sub-TLVs of TE Node Attribute TLV (Value 5)". IANA has assigned a new code point as follows:

Type	Sub-TLV	Reference
14	Connectivity Matrix	[RFC7580]

8.2. Link Information

IANA maintains the "Open Shortest Path First (OSPF) Traffic Engineering TLVs" registry with a sub-registry called "Types for sub-TLVs of TE Link TLV (Value 2)". IANA has assigned a new code point as follows:

Type	Sub-TLV	Reference
34	Port Label Restrictions	[RFC7580]

9. References

9.1. 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, <http://www.rfc-editor.org/info/rfc2119>.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, DOI 10.17487/RFC2328, April 1998, <http://www.rfc-editor.org/info/rfc2328>.

[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>.

[RFC4202] Kompella, K., Ed., and Y. Rekhter, Ed., "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, DOI 10.17487/RFC4202, October 2005, <http://www.rfc-editor.org/info/rfc4202>.

[RFC4203] Kompella, K., Ed., and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, DOI 10.17487/RFC4203, October 2005, <http://www.rfc-editor.org/info/rfc4203>.

[RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", RFC 5250, DOI 10.17487/RFC5250, July 2008, <http://www.rfc-editor.org/info/rfc5250>.

[RFC5786] Aggarwal, R. and K. Kompella, "Advertising a Router's Local Addresses in OSPF Traffic Engineering (TE) Extensions", RFC 5786, DOI 10.17487/RFC5786, March 2010, <http://www.rfc-editor.org/info/rfc5786>.

[RFC7579] Bernstein, G., Ed., Lee, Y., Ed., Li, D., Imajuku, W., and J. Han, "General Network Element Constraint Encoding for GMPLS-Controlled Networks", RFC 7579, DOI 10.17487/RFC7579, June 2015, <http://www.rfc-editor.org/info/rfc7579>.

9.2. Informative References

[BGP-LS] Gredler, H., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and TE Information using BGP", Work in Progress, draft-ietf-idr-ls-distribution-11, June 2015.

[RFC6020] Bjorklund, M., Ed., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, DOI 10.17487/RFC6020, October 2010, <http://www.rfc-editor.org/info/rfc6020>.

[RFC6163] Lee, Y., Ed., Bernstein, G., Ed., and W. Imajuku, "Framework for GMPLS and Path Computation Element (PCE) Control of Wavelength Switched Optical Networks (WSONs)", RFC 6163, DOI 10.17487/RFC6163, April 2011, <http://www.rfc-editor.org/info/rfc6163>.

[RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <http://www.rfc-editor.org/info/rfc6241>.

[RFC6825] Miyazawa, M., Otani, T., Kumaki, K., and T. Nadeau, "Traffic Engineering Database Management Information Base in Support of MPLS-TE/GMPLS", RFC 6825, DOI 10.17487/RFC6825, January 2013, <http://www.rfc-editor.org/info/rfc6825>.

[RFC7446] Lee, Y., Ed., Bernstein, G., Ed., Li, D., and W. Imajuku, "Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks", RFC 7446, DOI 10.17487/RFC7446, February 2015, <http://www.rfc-editor.org/info/rfc7446>.

[RFC5920] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", RFC 5920, DOI 10.17487/RFC5920, July 2010, <http://www.rfc-editor.org/info/rfc5920>.

Acknowledgments

We thank Ming Chen and Yabin Ye from DICONNET Project who provided valuable information for this document.

Contributors

Guoying Zhang
China Academy of Telecommunication Research of MII
11 Yue Tan Nan Jie
Beijing
China
Phone: +86-10-68094272
EMail: zhangguoying@mail.ritt.com.cn

Dan Li
Huawei Technologies Co., Ltd.
F3-5-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129
China
Phone: +86-755-28973237
EMail: danli@huawei.com

Ming Chen
European Research Center
Huawei Technologies
Riesstr. 25, 80992
Munchen
Germany
Phone: 0049-89158834072
EMail: minc@huawei.com

Yabin Ye
European Research Center
Huawei Technologies
Riesstr. 25, 80992
Munchen
Germany
Phone: 0049-89158834074
EMail: yabin.ye@huawei.com

Authors' Addresses

Fatai Zhang
Huawei Technologies
F3-5-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129
China
Phone: +86-755-28972912
EMail: zhangfatai@huawei.com

Young Lee
Huawei Technologies
5360 Legacy Drive, Building 3
Plano, TX 75023
United States
Phone: (469)277-5838
EMail: leeyoung@huawei.com

Jianrui Han
Huawei Technologies Co., Ltd.
F3-5-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129
China
Phone: +86-755-28977943
EMail: hanjianrui@huawei.com

Greg Bernstein
Grotto Networking
Fremont, CA
United States
Phone: (510) 573-2237
EMail: gregb@grotto-networking.com

Yunbin Xu
China Academy of Telecommunication Research of MII
11 Yue Tan Nan Jie
Beijing
China
Phone: +86-10-68094134
EMail: xuyunbin@mail.ritt.com.cn