Network work group
Internet Draft
Intended status: Standards Track

Fatai Zhang
Young Lee
Jianrui Han
Huawei
G. Bernstein
Grotto Networking
Yunbin Xu
CATR

Expires: September 5, 2015

March 6, 2015

OSPF-TE Extensions for General Network Element Constraints

draft-ietf-ccamp-gmpls-general-constraints-ospf-te-10.txt

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., SONET/SDH, 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 Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

This Internet-Draft will expire on September 5, 2015.

Copyright Notice

Copyright (c) 2015 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 (http://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.

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].

Table of Contents

1. Introduction....3
2. Node Information....4
   2.1. Connectivity Matrix....4
3. Link Information....4
   3.1. Port Label Restrictions....5
4. Routing Procedures....5
5. Scalability and Timeliness....6
   5.1. Different Sub-TLVs into Multiple LSAs....6
   5.2. Decomposing a Connectivity Matrix into Multiple Matrices..7
6. Security Considerations....7

7. Manageability.....8
8. IANA Considerations.....8
8.1. Node Information.....8
8.2. Link Information.....9
9. References.....9
9.1. Normative References.....9
9.2. Informative References.....10
10. Authors' Addresses 10
Acknowledgment.....12

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 Network (WSON) [RFC6163]. Such constraints can lead to the requirement for link by link label availability in path computation and label assignment.

[GEN-Encode] provides efficient encodings of information needed by the routing and label assignment process in technologies such as WSON and are potentially applicable to a wider range of technologies. The encoding provided in [GEN-Encode] 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 [GEN-Encode] to enhance the Traffic Engineering (TE) properties of GMPLS TE which are defined in [RFC3630], [RFC4202], and [RFC4203]. The enhancements to the TE properties of GMPLS TE links can be advertised in OSPF TE 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 (1) Router Address [RFC3630], (2) Link [RFC3630], (3) 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 are not defined in this document. [GEN-Encode] provides encoding details.

2. Node Information

According to [GEN-Encode], the additional node information representing node switching asymmetry constraints includes Node ID and connectivity matrix. Except for the Node ID, which should comply with Routing Address described in [RFC3630], the other pieces of information are defined in this document.

Per [GEN-Encode], this document defines the Connectivity Matrix Sub-TLV of the Node Attribute TLV defined in [RFC5786]. The new Sub-TLV has Type TBD1 (to be assigned by IANA).

In some specific technologies, e.g., WSON networks, the Connectivity Matrix sub-TLV may be optional, which depends on the control plane implementations. Usually, for example, in WSON networks, 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 is 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 is used to identify these restrictions, which can represent either the potential connectivity matrix for asymmetric switches (e.g., ROADMs and such) or fixed connectivity for an asymmetric device such as a multiplexer as defined in [RFC7446].

The Connectivity Matrix is a sub-TLV of the Node Attribute 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 2.1 of [GEN-Encode]. 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], [RFC4203]. For example, Link ID, Administrative Group, Interface Switching Capability Descriptor (ISCD), Link Protection Type, Shared Risk Link Group Information

(SRLG), and Traffic Engineering Metric are among the typical link sub-TLVs.

Per [GEN-Encode], this document defines the Port Label Restrictions Sub-TLV of the Link TLV defined in [RFC3630]. The new Sub-TLV has Type TBD2 (to be assigned by IANA).

Generally all the sub-TLVs above are optional, which depends on the control plane implementations. 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 describes the wavelength restrictions that the link and various optical devices such as OXCs, ROADMs, and waveband multiplexers may impose on a port in WSON. These restrictions represent what wavelength may or may not be used on a link and are relatively static. The detailed information about port label restrictions is described 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 [GEN-Encode]. 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 the 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 ([RFC3630 and RFC5876]), 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 information 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 [RFC3630] Section 3 for related details.

5. Scalability and Timeliness

This document has defined two sub-TLVs for describing generic routing contraints. The examples given in [GEN-Encode] show that very large systems, in terms of label count or ports, can be very efficiently encoded. However there has been concern expressed that some possible systems may produce LSAs that exceed the IP Maximum Transmission Unit (MTU) and that methods 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 (Node Attribute TLV)
2. Port Label Restrictions (Link TLV)

The Connectivity Matrix can be carried in the Node Attribute TLV as defined in [RFC5786] while the Port Label Restrictions can be carried in an 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 [GEN-Encode] a connectivity matrix just consists of pairs of input and output ports that can reach each other and hence such this decomposition would be straightforward. Each of these sub-matrices would get a unique matrix identifier per [GEN-Encode].

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 then 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 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 Generalized Multiprotocol Label Switching (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 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, Management, and Administration (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 is requested to allocate 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". IANA is requested to assign a new code point as follows:

Type	Sub-TLV	Reference
TBD1	Connectivity Matrix	[This.I-D]

8.2. Link Information

IANA maintains the "Open Shortest Path First (OSPF) Traffic Engineering TLVs" registry with a sub-registry called "Types for sub-LVs of TE Link TLV". IANA is requested to assign a new code point as follows:

Type	Sub-TLV	Reference
TBD2	Port Label Restrictions	[This.I-D]

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

[RFC3630] Katz, D., Kompella, K., and Yeung, D., "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

[RFC4202] Kompella, K., Ed., and Y. Rekhter, Ed., "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005

[RFC4203] Kompella, K., Ed., and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.

[RFC5250] L. Berger, I. Bryskin, A. Zinin, R. Coltun "The OSPF Opaque LSA Option", RFC 5250, July 2008.

[RFC5786] R. Aggarwal and K. Kompella,"Advertising a Router's Local Addresses in OSPF Traffic Engineering (TE) Extensions", RFC 5786, March 2010.

[GEN-Encode] G. Bernstein, Y. Lee, D. Li, W. Imajuku, " General Network Element Constraint Encoding for GMPLS Controlled Networks", work in progress: draft-ietf-ccamp-general-constraint-encode.

9.2. Informative References

[RFC6020] M. Bjorklund, Ed., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, October 2010.

[RFC6163] Y. Lee, G. Bernstein, W. Imajuku, "Framework for GMPLS and PCE Control of Wavelength Switched Optical Networks (WSON)", RFC 6163, February 2011.

[RFC6241] R. Enns, Ed., M. Bjorklund, Ed., Schoenwaelder, Ed., A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, June 2011.

[RFC6825] M. Miyazawa, T. Otani, K. Kumaki, T. Nadeau, "Traffic Engineering Database Management Information Base in Support of MPLS-TE/GMPLS", RFC 6825, January 2013.

[RFC7446] Y. Lee, G. Bernstein, D. Li, W. Imajuku, "Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks", RFC 7446, February 2015.

[RFC5920] L. Fang, Ed., "Security Framework for MPLS and GMPLS Networks", RFC 5920, July 2010.

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

10. Contributors

Guoying Zhang
China Academy of Telecommunication Research of MII
11 Yue Tan Nan Jie Beijing, P.R.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 P.R.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 P.R.China

Phone: +86-755-28972912
Email: zhangfatai@huawei.com

Young Lee
Huawei Technologies
5360 Legacy Drive, Building 3
Plano, TX 75023
USA

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 P.R.China

Phone: +86-755-28977943
Email: hanjianrui@huawei.com

Greg Bernstein
Grotto Networking
Fremont CA, USA

Phone: (510) 573-2237
Email: gregb@grotto-networking.com

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

Acknowledgment

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