Network work group   Fatai Zhang
Internet Draft   Young Lee
Intended status: Standards Track   Jianrui Han
                                                           Huawei
                                                     G. Bernstein
                                                 Grotto Networking
                                                        Yunbin Xu
                                                             CATR

Expires: December 19, 2012   June 19, 2012

           OSPF-TE Extensions for General Network Element Constraints

             draft-ietf-ccamp-gmpls-general-constraints-ospf-te-03.txt

Status of this Memo

Abstract

Internet-Draft   Generic Constraint OSPF-TE   June 2012

   Generalized Multiprotocol Label Switching can be used to control a
   wide variety of technologies including packet switching (e.g., MPLS),
   time-division (e.g., SONET/SDH, 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 OSPF routing protocol
   extensions to support these kinds of constraints under the control
   of Generalized MPLS (GMPLS).

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

. Introduction

   Some data plane technologies that wish to make use of a GMPLS
   control plane contain 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
   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.

   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 Traffic Engineering (TE) properties of GMPLS
   TE links can be announced 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) Generic
   Node Attribute defined in Section 2. In this document, we enhance
   the sub-TLVs for the Link TLV and define a new top-level TLV
   (Generic Node Attribute 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 detail.

2. Node Information

   According to [GEN-Encode], the additional node information
   representing node switching asymmetry constraints includes Node ID,
   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.

This document defines a new top TLV named the Generic Node Attribute
TLV which carries attributes related to a general network element.
This Generic Node Attribute TLV contains one or more sub-TLVs

Per [GEN-Encode], we have identified the following new Sub-TLVs to
the Generic Node Attribute TLV. Detail description for each newly
defined Sub-TLV is provided in subsequent sections:

       Sub-TLV Type   Length   Name

          TBD   variable   Connectivity Matrix

   In some specific technologies, e.g., WSON networks, 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 appear in the LSAs because WSON
   switches are asymmetric at present. It is assumed that the switches
   are symmetric switching, if there is no Connectivity Matrix sub-TLV
   in the LSAs.

2.1. Connectivity Matrix

   It is necessary to identify which ingress ports and labels can be
   switched to some specific labels on a specific egress port, if the
   switching devices in some technology are highly asymmetric.

   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 [WSON-
   Info].

   The Connectivity Matrix is a sub-TLV (the type is TBD by IANA) of
   the Generic Node Attribute TLV. The length is the length of value
   field in octets. The meaning and format of this sub-TLV are defined
   in Section 5.3 of [GEN-Encode]. One sub-TLV contains one matrix. The
   Connectivity Matrix sub-TLV may occur more than once to contain
   multi-matrices within the Generic Node Attribute TLV. In addition a
   large connectivity matrix can be decomposed into smaller separate
   matrices for transmission in multiple LSAs as described in Section 5.

3. Link Information

The most common link sub-TLVs nested to link top-level 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], we add the following additional link sub-TLVs to
the link-TLV in this document.

Sub-TLV Type	Length	Name
TBD	variable	Port Label Restrictions

TBD	variable	Available Labels
TBD	variable	Shared Backup Labels

Generally all the sub-TLVs above are optional, which depends on the
control plane implementations. If it is default no restrictions on
labels, Port Label Restrictions sub-TLV may not appear in the LSAs.
In order to be able to compute label assignment, Available Labels
sub-TLV may appear in the LSAs. For example, in WSON networks,
without available wavelength information, path computation need
guess what lambdas may be available (high blocking probability or
distributed wavelength assignment may be used).

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. More dynamic information
is contained in the information on available labels. Port label
restrictions are specified relative to the port in general or to a
specific connectivity matrix for increased modeling flexibility.

For example, 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 [WSON-Info].

The Port Label Restrictions is a sub-TLV (the type is TBD by IANA)
of the Link TLV. The length is the length of value field in octets.
The meaning and format of this sub-TLV are defined in Section 5.4 of
[GEN-Encode]. The Port Label Restrictions sub-TLV may occur more
than once to specify a complex port constraint within the link TLV.

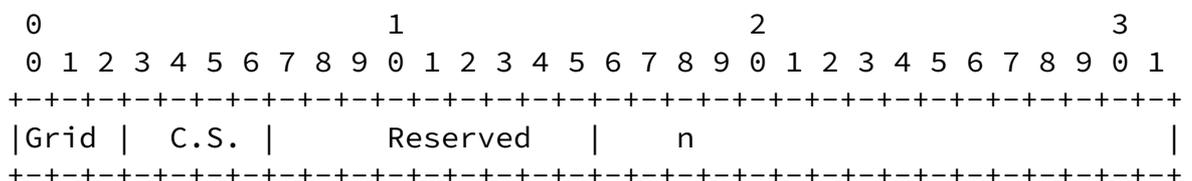## 3.2. Available Labels

Available Labels indicates the labels available for use on a link as
described in [GEN-Encode]. The Available Labels is a sub-TLV (the
type is TBD by IANA) of the Link TLV. The length is the length of
value field in octets. The meaning and format of this sub-TLV are
defined in Section 5.1 of [GEN-Encode]. The Available Labels sub-TLV
may occur at most once within the link TLV.

Note that there are five approaches for Label Set which is used to
represent the Available Labels described in [GEN-Encode]. Usually,

it depends on the implementation to one of the approaches. In WSON
networks, considering that the continuity of the available or
unavailable wavelength set can be scattered for the dynamic
wavelength availability, so it may burden the routing to reorganize
the wavelength set information when the Inclusive (/Exclusive) List
(/Range) approaches are used to represent Available Wavelengths
information. Therefore, it is RECOMMENDED that only the Bitmap Set
be used for representation Available Wavelengths information.

The "Base Label" and "Last Label" in label set defined in [GEN-
Encode] corresponds to base wavelength label and last wavelength
label in WSON, the format of which is described as follows:

```
     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
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |Grid | C.S. |    Reserved   |  n                              |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The detailed information related to wavelength label can be referred

to [RFC6205].

3.3. Shared Backup Labels

Shared Backup Labels indicates the labels available for shared
backup use on a link as described in [GEN-Encode].

The Shared Backup Labels is a sub-TLV (the type is TBD by IANA) of
the Link TLV. The length is the length of value field in octets. The
meaning and format of this sub-TLV are defined in Section 5.2 of
[GEN-Encode]. The Shared Backup Labels sub-TLV may occur at most
once within the link TLV.

4. Routing Procedures

All the sub-TLVs are nested to top-level TLV(s) and contained in
Opaque LSAs. The flooding of Opaque LSAs must follow the rules
specified in [RFC2328], [RFC5250], [RFC3630], [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 standard-compliant approach is to separate the dynamic
information sub-TLVs from the static information sub-TLVs, each

nested to 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 Generic 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 and Available Labels information sub-TLV can be
nested to the top level link TLVs and advertised in the separate
LSAs.

Note that as with other TE information, an implementation SHOULD
take measures to avoid rapid and frequent updates of routing

information that could cause the routing network to become swamped.
A threshold mechanism MAY be applied such that updates are only
flooded when a number of changes have been made to the label
availability information (e.g., wavelength availability) within a
specific time. Such mechanisms MUST be configurable if they are
implemented.

5. Scalability and Timeliness

   This document has defined four 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 LSA 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

   Four sub-TLVs are defined in this document:

      1. Connectivity Matrix (Generic Node Attribute TLV)
      2. Port Label Restrictions (Link TLV)
      3. Available Labels (Link TLV)
      4. Shared Backup Labels (Link TLV)

   Except for the Connectivity Matrix all these are carried in an Link
   TLV of which there can be at most one in an LSA [RFC3630]. Of these
   sub-TLVs the Port Label Restrictions are relatively static, i.e.,

---

   only would change with hardware changes or significant system
   reconfiguration. While the Available Labels and Shared Backup Labels
   are dynamic, meaning that they may change with LSP setup or teardown
   through the system. The most important technique for scalability and
   OSPF bandwidth reduction is to separate the dynamic information sub-
   TLVs from the static information sub-TLVs and advertise them in
   separate OSPF TE LSAs[RFC3630 and RFC5250].

   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.

6. Security Considerations

   This document does not introduce any further security issues other
   than those discussed in [RFC 3630], [RFC 4203].

7. IANA Considerations

   [RFC3630] says that the top level Types in a TE LSA and Types for
   sub-TLVs for each top level Types must be assigned by Expert Review,
   and must be registered with IANA.

   IANA is requested to allocate new Types for the TLV or sub-TLVs as
   defined in Sections 2 and 3 as follows:

7.1. Node Information

   This document introduces a new Top Level Node TLV (Generic Node
   Attribute TLV) under the OSPF TE LSA defined in [RFC3630].

      Value   TLV Type

```
        TBA       Generic Node Attribute

   This document also introduces the following sub-TLVs of Generic Node
   Attribute TLV:

        Type      sub-TLV

        TBD       Connectivity Matrix
```

[7.2](). Link Information

```
   This document introduces the following sub-TLVs of TE Link TLV
   (Value 2):

        Type      sub-TLV

        TBD       Port Label Restrictions

        TBD       Available Labels

        TBD       Shared Backup Labels
```

[8](). References

[8.1](). 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.

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

   [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.

   [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, May 2011.

   [RFC6205] T. Otani, H. Guo, K. Miyazaki, D. Caviglia, " Generalized
             Labels for Lambda-Switching Capable Label Switching
             Routers", RFC 6205, January 2011.

8.2. Informative References

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

   [WSON-Info] Y. Lee, G. Bernstein, D. Li, W. Imajuku, "Routing and
             Wavelength Assignment Information Model for Wavelength
             Switched Optical Networks", work in progress: draft-ietf-
             ccamp-rwa-info, September 2011.

9. 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
1700 Alma Drive, Suite 100
Plano, TX 75075
USA

Phone: (972) 509-5599 (x2240)
Email: ylee@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

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

Acknowledgment

Intellectual Property

Full Copyright Statement