Network Working Group
Internet Draft
Category: Standards Track
Updates: 3630
Expires: April 2004

K. Kompella, Editor
Y. Rekhter, Editor
Juniper Networks
October 2003

OSPF Extensions in Support of Generalized Multi-Protocol Label Switching

draft-ietf-ccamp-ospf-gmpls-extensions-12.txt

Status of this Memo

Copyright Notice

Abstract

This document specifies encoding of extensions to the OSPF routing protocol in support of Generalized Multi-Protocol Label Switching.

Specification of Requirements

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].

1. Introduction

This document specifies extensions to the OSPF routing protocol in support of carrying link state information for Generalized Multi-Protocol Label Switching (GMPLS). The set of required enhancements to OSPF are outlined in [GMPLS-ROUTING].

In this section we define the enhancements to the TE properties of GMPLS TE links that can be announced in OSPF TE LSAs. The Traffic Engineering (TE) LSA, which is an opaque LSA with area flooding scope [OSPF-TE], 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 two values (1) Router Address or (2) Link. In this document, we enhance the sub-TLVs for the Link TLV in support of GMPLS. Specifically, we add the following sub-TLVs to the Link TLV:

Sub-TLV Type	Length	Name
11	8	Link Local/Remote Identifiers
14	4	Link Protection Type
15	variable	Interface Switching Capability Descriptor
16	variable	Shared Risk Link Group

1.1. Link Local/Remote Identifiers

A Link Local/Remote Identifiers is a sub-TLV of the Link TLV. The type of this sub-TLV is 11, and length is eight octets. The value field of this sub-TLV contains four octets of Link Local Identifier followed by four octets of Link Remote Idenfier (see Section "Support for unnumbered links" of [GMPLS-ROUTING]). If the Link Remote Identifier is unknown, it is set to 0.

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                  Link Local Identifier                        |
```



```
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                  Link Remote Identifier                       |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

A node can communicate its Link Local Identifier to its neighbor using a link local Opaque LSA, as described in Section "Exchanging Link Local TE Information".

1.2. Link Protection Type

The Link Protection Type is a sub-TLV of the Link TLV. The type of this sub-TLV is 14, and length is four octets.

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |Protection Cap |                    Reserved                   |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The first octet is a bit vector describing the protection capabilities of the link (see Section "Link Protection Type" of [GMPLS-ROUTING]). They are:

0x01   Extra Traffic

0x02   Unprotected

0x04   Shared

0x08   Dedicated 1:1

0x10   Dedicated 1+1

0x20   Enhanced

0x40   Reserved

0x80 Reserved

The remaining three octets SHOULD be set to zero by the sender, and SHOULD be ignored by the receiver.

The Link Protection Type sub-TLV may occur at most once within the Link TLV.

1.3. Shared Risk Link Group (SRLG)

The SRLG is a sub-TLV (of type 16) of the Link TLV. The length is the length of the list in octets. The value is an unordered list of 32 bit numbers that are the SRLGs that the link belongs to. The format of the value field is as shown below:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Shared Risk Link Group Value                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        ............                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Shared Risk Link Group Value                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

This sub-TLV carries the Shared Risk Link Group information (see Section "Shared Risk Link Group Information" of [GMPLS-ROUTING]).

The SRLG sub-TLV may occur at most once within the Link TLV.

1.4. Interface Switching Capability Descriptor

The Interface Switching Capability Descriptor is a sub-TLV (of type 15) of the Link TLV. The length is the length of value field in octets. The format of the value field is as shown below:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Switching Cap |   Encoding    |           Reserved            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 0               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 1               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 2               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 3               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 4               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 5               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 6               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 7               |
```



```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Switching Capability-specific information              |
|                  (variable)                                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The Switching Capability (Switching Cap) field contains one of the following values:

```
1      Packet-Switch Capable-1 (PSC-1)
2      Packet-Switch Capable-2 (PSC-2)
3      Packet-Switch Capable-3 (PSC-3)
4      Packet-Switch Capable-4 (PSC-4)
51     Layer-2 Switch Capable  (L2SC)
100    Time-Division-Multiplex Capable (TDM)
150    Lambda-Switch Capable   (LSC)
200    Fiber-Switch Capable    (FSC)
```

The Encoding field contains one of the values specified in Section 3.1.1 of [GMPLS-SIG].

Maximum LSP Bandwidth is encoded as a list of eight 4 octet fields in the IEEE floating point format [IEEE], with priority 0 first and priority 7 last. The units are bytes (not bits!) per second.

The content of the Switching Capability specific information field depends on the value of the Switching Capability field.

When the Switching Capability field is PSC-1, PSC-2, PSC-3, or PSC-4, the Switching Capability specific information field includes Minimum LSP Bandwidth, Interface MTU, and padding.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Minimum LSP Bandwidth                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Interface MTU       |            Padding            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The Minimum LSP Bandwidth is is encoded in a 4 octets field in the IEEE floating point format. The units are bytes (not bits!) per second. The Interface MTU is encoded as a 2 octets integer. The padding is 2 octets, and is used to make the Interface Switching Capability Descriptor sub-TLV 32-bits aligned. It SHOULD be set to zero by the sender and SHOULD be ignored by the receiver.

When the Switching Capability field is L2SC, there is no Switching Capability specific information field present.

When the Switching Capability field is TDM, the Switching Capability specific information field includes Minimum LSP Bandwidth, an indication whether the interface supports Standard or Arbitrary SONET/SDH, and padding.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Minimum LSP Bandwidth                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Indication  |                    Padding                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The Minimum LSP Bandwidth is encoded in a 4 octets field in the IEEE floating point format. The units are bytes (not bits!) per second. The indication whether the interface supports Standard or Arbitrary SONET/SDH is encoded as 1 octet. The value of this octet is 0 if the interface supports Standard SONET/SDH, and 1 if the interface supports Arbitrary SONET/SDH. The padding is 3 octets, and is used to make the Interface Switching Capability Descriptor sub-TLV 32-bits aligned. It SHOULD be set to zero by the sender and SHOULD be ignored by the receiver.

When the Switching Capability field is LSC, there is no Switching Capability specific information field present.

To support interfaces that have more than one Interface Switching Capability Descriptor (see Section "Interface Switching Capability Descriptor" of [GMPLS-ROUTING]) the Interface Switching Capability Descriptor sub-TLV may occur more than once within the Link TLV.

2. Implications on Graceful Restart

The restarting node should follow the OSPF restart procedures [OSPF-RESTART], and the RSVP-TE restart procedures [GMPLS-RSVP].

When a restarting node is going to originate its TE LSAs, the TE LSAs containing Link TLV should be originated with 0 unreserved bandwidth, Traffic Engineering metric set to 0xffffffff, and if the Link has LSC or FSC as its Switching Capability then also with 0 as Max LSP Bandwidth, until the node is able to determine the amount of unreserved resources taking into account the resources reserved by the already established LSPs that have been preserved across the restart. Once the restarting node determines the amount of

unreserved resources, taking into account the resources reserved by the already established LSPs that have been preserved across the restart, the node should advertise these resources in its TE LSAs.

In addition in the case of a planned restart prior to restarting, the restarting node SHOULD originate the TE LSAs containing Link TLV with 0 as unreserved bandwidth, and if the Link has LSC or FSC as its Switching Capability then also with 0 as Max LSP Bandwidth. This

would discourage new LSP establishment through the restarting router.

Neighbors of the restarting node should continue advertise the actual unreserved bandwidth on the TE links from the neighbors to that node.

Regular graceful restart should not be aborted if a TE LSA or TE topology changes. TE graceful restart need not be aborted if a TE LSA or TE topology changes.

3. Exchanging Link Local TE Information

It is often useful for a node to communicate some Traffic Engineering information for a given interface to its neighbors on that interface. One example of this is a Link Local Identifier. If nodes X and Y are connected by an unnumbered point-to-point interface I, then X's Link Local Identifier for I is Y's Link Remote Identifier for I. X can communicate its Link Local Identifer for I by exchanging with Y a TE link local opaque LSA described below. Note that this information need only be exchanged over interface I, hence the use of a link local Opaque LSA.

A TE Link Local LSA is an opaque LSA of type 9 (link-local flooding scope) with Opaque Type [TBD] and Opaque ID of 0.

```
  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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |            LS age             |    Options    |       9       |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |  Opaque Type  |               Opaque ID                       |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                     Advertising Router                        |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                     LS sequence number                        |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |         LS checksum           |             length            |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                                                               |
 +-                            TLVs                             -+
```



```
 |                             ...                               |
```

The format of the TLVs that make up the body of the TE Link Local LSA is the same as that of the TE TLVs: a 2-octet Type field followed by a 2-octet Length field which indicates the length of the Value field in octets. The Value field is zero-padded at the end to a four octet boundary.

The only TLV defined here is the Link Local Identifier TLV, with Type 1, Length 4 and Value the 32 bit Link Local Identifier for the link over which the TE Link Local LSA is exchanged.

4. Contributors

Ayan Banerjee
Calient Networks
5853 Rue Ferrari
San Jose, CA 95138
Phone: +1.408.972.3645
Email: abanerjee@calient.net

John Drake
Calient Networks
5853 Rue Ferrari
San Jose, CA 95138
Phone: +1.408.972.3720
Email: jdrake@calient.net

Greg Bernstein
Ciena Corporation
10480 Ridgeview Court
Cupertino, CA 94014
Phone: +1.408.366.4713
Email: greg@ciena.com

Don Fedyk
Nortel Networks Corp.
600 Technology Park Drive
Billerica, MA 01821
Phone: +1.978.288.4506
Email: dwfedyk@nortelnetworks.com

Eric Mannie
Independent Consultant
E-mail: eric_mannie@hotmail.com

Debanjan Saha
Tellium Optical Systems
2 Crescent Place
P.O. Box 901
Ocean Port, NJ 07757
Phone: +1.732.923.4264
Email: dsaha@tellium.com

Vishal Sharma
Metanoia, Inc.
335 Elan Village Lane, Unit 203
San Jose, CA 95134-2539
Phone: +1.408.943.1794
Email: v.sharma@ieee.org

5. Acknowledgements

The authors would like to thank Suresh Katukam, Jonathan Lang, Quaizar Vohra, and Alex Zinin for their comments on the draft.

6. Security Considerations

This document specifies the contents of Opaque LSAs in OSPFv2. As Opaque LSAs are not used for SPF computation or normal routing, the extensions specified here have no direct effect on IP routing. Tampering with GMPLS TE LSAs may have an effect on the underlying transport (optical and/or SONET-SDH) network. [OSPF-TE] suggests mechanisms such as [OSPF-SIG] to protect the transmission of this information, and those or other mechanisms should be used to secure and/or authenticate the information carried in the Opaque LSAs.

IANA Considerations

The memo introduces 4 new sub-TLVs of the TE Link TLV in the TE Opaque LSA for OSPF v2; [OSPF-TE] says that the sub-TLVs of the TE Link TLV in the range 10-32767 must be assigned by Expert Review, and must be registered with IANA.

The memo has four suggested values for the four sub-TLVs of the TE Link TLV; it is strongly recommended that the suggested values be granted, as there are interoperable implementations using these values.

Normative References

[GMPLS-ROUTING] Kompella, K., and Rekhter, Y. (Editors), "Routing Extensions in Support of Generalized Multi-Protocol Label Switching", (work in progress) [draft-ietf-ccamp-gmpls-routing-08.txt]

[GMPLS-RSVP] Berger, L., (Editor), "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003

[GMPLS-SIG] Berger, L. (Editor), "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003

[IEEE] IEEE, "IEEE Standard for Binary Floating-Point Arithmetic", Standard 754-1985, 1985 (ISBN 1-5593-7653-8).

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

[OSPF-RESTART] Moy, J., Pillay-Esnault, P., Lindem, A., "Graceful OSPF Restart", (work in progress) [draft-ietf-ospf-hitless-restart-08.txt]

[OSPF-SIG] Murphy, S., Badger, M., and B. Wellington, "OSPF with Digital Signatures", RFC 2154, June 1997.

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

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

Authors' Information

Kireeti Kompella
Juniper Networks, Inc.
1194 N. Mathilda Ave
Sunnyvale, CA 94089
Email: kireeti@juniper.net

Yakov Rekhter
Juniper Networks, Inc.
1194 N. Mathilda Ave
Sunnyvale, CA 94089
Email: yakov@juniper.net

Intellectual Property Rights Notices

copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this standard. Please address the information to the IETF Executive Director.

Full Copyright Statement