

CCAMP Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 18, 2011

D. Ceccarelli
D. Caviglia
Ericsson
S. Belotti
P. Grandi
Alcatel-Lucent
F. Zhang
D. Li
Huawei Technologies
J. Drake
Juniper
October 15, 2010

Technology Agnostic OSPF Traffic Engineering Extensions for Generalized
MPLS (GMPLS)
[draft-bccdg-ccamp-gmpls-ospf-agnostic-00](#)

Abstract

This document defines a new approach to Generalized Multiprotocol Label Switching (GMPLS) bandwidth advertisement aiming at providing the Network Elements (NEs) and Path Computation Elements (PCEs) with all the data required for crank-backs minimization and scalability optimization.

A new Open Shortest Path First - Traffic Engineering (OSPF-TE) routing protocol sub-tlv is defined for bandwidth advertisement per service type.

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 April 18, 2011.

Copyright Notice

Internet-Draft

Technology Agnostic OSPF-TE

October 2010

Copyright (c) 2010 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.

Table of Contents

1.	Introduction	3
1.1.	Conventions Used in This Document	3
2.	OSPF Extensions	3
2.1.	Bandwidth Accounting sub-TLV	4
3.	LSA composition	8
4.	Examples	8
5.	Applicability	12
6.	Compatibility Considerations	16
7.	Security Considerations	16
8.	IANA Considerations	16
9.	Contributors	16
10.	Acknowledgements	17
11.	References	17
11.1.	Normative References	17
11.2.	Informative References	17
	Authors' Addresses	17

1. Introduction

An Opaque OSPF (Open Shortest Path First) LSA (Link State Advertisements) carrying application-specific information can be generated and advertised to other nodes following the flooding procedures defined in [\[RFC5250\]](#). Three types of opaque LSA are defined, i.e. type 9 - link-local flooding scope, type 10 - area-local flooding scope, type 11 - AS flooding scope.

Traffic Engineering (TE) LSA using type 10 opaque LSA is defined in [\[RFC3630\]](#) for TE purposes. This type of LSA is composed of a standard LSA header and a payload including one top-level TLV (Type/Length/Value triplet) and possible several nested sub-TLVs. [\[RFC3630\]](#) defines two top-level TLVs: Router Address TLV and Link TLV; and nine possible sub-TLVs for the Link TLV, used to carry link related TE information.

The Link type sub-TLVs are enhanced by [\[RFC4203\]](#) in order to support GMPLS networks and related specific link information.

In GMPLS networks each node generates TE LSAs to advertise its TE information and capabilities (link-specific or node-specific), through the network. The TE information carried in the LSAs are collected by the other nodes of the network and stored into their local Traffic Engineering Databases (TED).

In GMPLS networks, routing serves as the foundation for automatically establishing Label Switched Paths (LSPs) through GMPLS RSVP-TE signaling.

This document describes technology agnostic OSPF LSA extensions to support connection oriented transport networks under the control of GMPLS (e.g. OTN, SDH, MPLS-TP). In particular a new OSPF-TE LSP is defined for bandwidth advertisement per service type taking into account priorities and technology specific capabilities.

[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 [[RFC2119](#)].

[2.](#) OSPF Extensions

Each TE LSA can carry a top-level link TLV with several nested sub-TLVs to describe different attributes of a TE link. Two top-level TLVs are defined in [[RFC 3630](#)]. (1) The Router Address TLV (referred

to as the Node TLV) and (2) the TE link TLV. One or more sub-TLVs can be nested into the two top-level TLVs. The sub-TLV set for the two top-level TLVs are also defined in [[RFC 3630](#)] and [[RFC 4203](#)].

This document defines a new link sub-TLV, called Bandwidth Accounting (BA) sub-TLV (Sub-tlv value TBA by IANA, suggested 26).

One or more component links can be bundled as a TE link. In case of link bundling a single BA sub-TLV will be used to describe several component links.

[2.1.](#) Bandwidth Accounting sub-TLV

The BA sub-TLV has a so generic format that it can be used for the advertisement of any type of transport technology, from SDH/SONET to OTN, from L2SC to PSC etc. The main difference from the ISCD defined in [[RFC4202](#)] is the fact that unreserved bandwidth is advertised per service type per priority. The format of the BA sub-TLV is based on "8 bytes" data blocks repeated for each service type/priority/technology specific capability combination as is illustrated in Figure 1.

```

      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           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Service Type  | M | T.S. Flags |           Reserved           |Prior|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Bandwidth                  |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Service Type  | M | T.S. Flags |           Reserved           |Prior|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Bandwidth                  |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Service Type  | M | T.S. Flags |           Reserved           |Prior|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Bandwidth                  |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
~                                     ...                                     ~
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Service Type  | M | T.S. Flags |           Reserved           |Prior|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     Bandwidth                  |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```


interface), a data block for each value MUST be advertised. For example, when advertising an ODUFlex service type in an OTN network, both Unreserved bandwidth and MAX LSDP bandwidth are advertised as illustrated in Figure 2 (assuming supported priorities: P1 and P5).

[EDITOR NOTE]: Under Discussion - M=2 - Available bandwidth at priority P_i , Where Available bandwidth is defined as the unused link bandwidth available for additional non-traffic engineered IP/LDP forwarding and can be used as input to a node equal cost multipath load balancing function

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																				
ODUFlex										M=0										T.S. Flags										Reserved										P1
										Unreserved Bandwidth @ P1																														

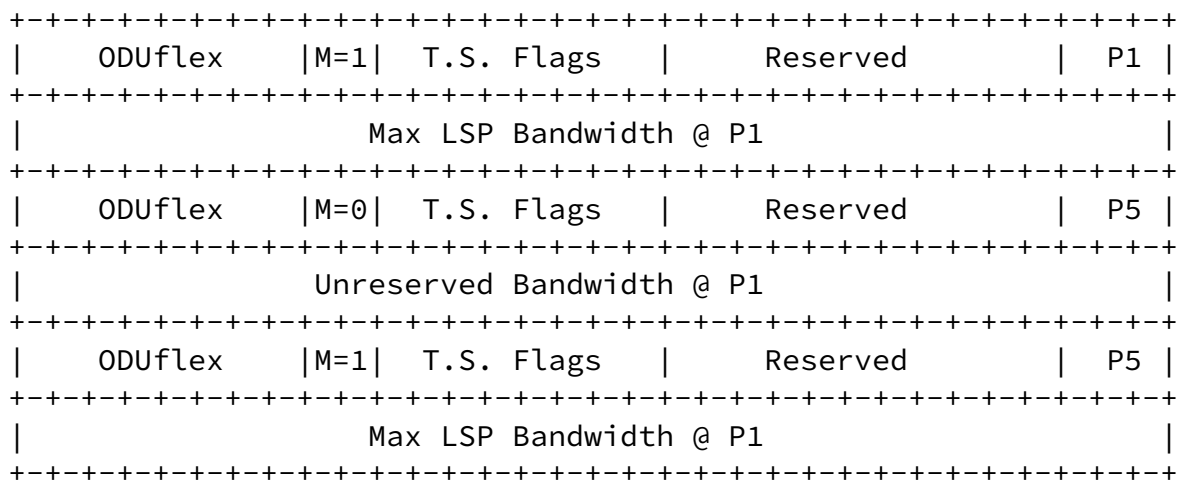


Figure 2: M field utilization example

- o Technology Specific Flags (8 bits): These bits are used for the advertisement of technology specific interface capabilities and are defined in companion technology specific IDs. Depending on the technology it could be possible to have different data block advertised for different cabability flags.
- o Reserved (11 bits): Reserved bits MUST be set to zero.
- o Priority (3 bits): Indicates the priority related to the advertised service type. Only supported priorities MUST be advertised.
- o Bandwidth (32 bits): Independently on the type of bandwidth being advertised (see M field), this field is expressed in Bytes/sec in IEEE floating point format unless differently stated in technology specific documents.

The maximum bandwidth that an LSP can occupy in a TE link is determined by the component link with the maximum unreserved bandwidth in such TE link. For example, if two OTN OTU3 component links are bundled in a TE link, the unreserved bandwidth of the first component link is 20*1.25 Gbps, and the unreserved bandwidth of the second component link is 24*1.25Gbps, then the unreserved bandwidth of this TE link is 44*1.25Gbps, but the maximum bandwidth an LSP can

occupy in this TE link is 24*1,25Gbps, not 44*1,25Gbps.

All the reserved fields MUST be set to zero and SHOULD be ignored when received.

3. LSA composition

Each NE generates an LSA to describe the attributes of each TE link. If we suppose to have unnumbered link IDs, the LSA should carry a link TLV with the following nested minimal sub-TLVs:

`< Link > ::= < Link Type > < Link ID > < Link
Local/Remote Identifiers > < Generalized-ISCD >`

- o Link Type sub-TLV: Defined in [\[RFC 3630\]](#).
- o Link ID sub-TLV: Defined in [\[RFC 3630\]](#), for point-to-point link, indicates the remote router ID.
- o Link Local/Remote Identifiers sub-TLV: Defined in [\[RFC 4203\]](#), indicates the local link ID and the remote link ID.
- o Bandwidth Accounting sub-TLV: Defined in this document, carries the Bandwidth related information of the advertised TE-link.

4. Examples

The examples in the following pages are not normative and are not intended to infer or mandate any specific implementation. Moreover they aim at giving a general idea of the utilization of the BA sub-TLV in a technology agnostic scenario.

Figure 3 shows the case of a TE-link composed of two component links.

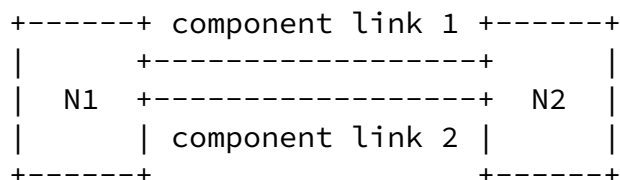


Figure 3: Example

The nominal bandwidth of the two component links is 10Gbps and 40Gbps respectively. The former has the capability of carrying service types A and B, while the latter, service types B and C, where A and C are fixed bandwidth service types (just unreserved bandwidth is advertised) and B variable bandwidth service types (unreserved bandwidth and Max LSP bandwidth advertised). The supported priorities are:0 and 3.

In this example the two component links are bundled as a TE link but it could also be possible to consider each of them as separate TE links.

If the two component links are bundled together, N1 and N2 should assign a link local ID to the TE link and then N1 can get the link remote ID automatically or manually.

Just after the creation of the TE Link comprising the two component links, the BA sub-TLV would be advertised as follows:

Internet-Draft

Technology Agnostic OSPF-TE

October 2010

```

      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      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| S.Type(A)     |M=0| T.S. Flags     |      Reserved      | P0 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Unreserved Bandwidth = 10 Gbps      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| S.Type(A)     |M=0| T.S. Flags     |      Reserved      | P3 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Unreserved Bandwidth = 10 Gbps      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| S.Type(B)     |M=0| T.S. Flags     |      Reserved      | P0 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Unreserved Bandwidth = 50 Gbps      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| S.Type(B)     |M=1| T.S. Flags     |      Reserved      | P0 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Max LSP Bandwidth = 40 Gbps      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| S.Type(B)     |M=0| T.S. Flags     |      Reserved      | P3 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Unreserved Bandwidth = 50 Gbps      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| S.Type(B)     |M=1| T.S. Flags     |      Reserved      | P3 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Max LSP Bandwidth = 40 Gbps      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| S.Type(C)     |M=0| T.S. Flags     |      Reserved      | P0 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Unreserved Bandwidth = 40 Gbps      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| S.Type(C)     |M=0| T.S. Flags     |      Reserved      | P3 |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Unreserved Bandwidth = 40 Gbps      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Figure 4: Example - BA sub-TLV(to)

Suppose that at time t1 an service type B LSP is created allocating 35 Gbps at priority 3. The BA sub-TLV will be modified 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								
Switching Cap										Encoding										Reserved																			
S.Type(A) M=0										T.S. Flags										Reserved P0																			
Unreserved Bandwidth = 10 Gbps																																							
S.Type(A) M=0										T.S. Flags										Reserved P3																			
Unreserved Bandwidth = 10 Gbps																																							
S.Type(B) M=0										T.S. Flags										Reserved P0																			
Unreserved Bandwidth = 50 Gbps																																							
S.Type(B) M=1										T.S. Flags										Reserved P0																			
Max LSP Bandwidth = 40 Gbps																																							
S.Type(B) M=0										T.S. Flags										Reserved P3																			
Unreserved Bandwidth = 15 Gbps																																							
S.Type(B) M=1										T.S. Flags										Reserved P3																			
Max LSP Bandwidth = 10 Gbps																																							
S.Type(C) M=0										T.S. Flags										Reserved P0																			
Unreserved Bandwidth = 40 Gbps																																							

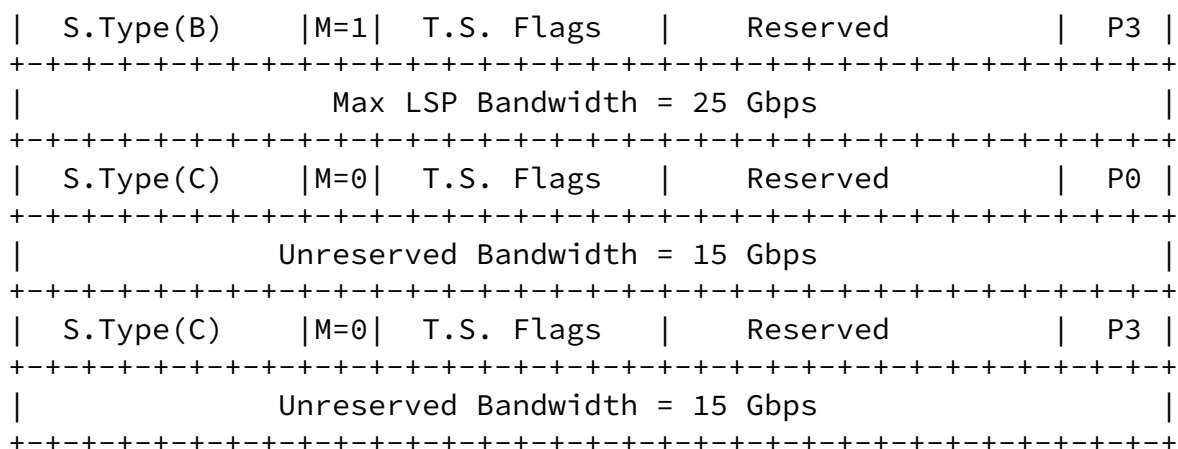


Figure 6: Example - BA sub-TLV (t2)

5. Applicability

The goal of this section is providing a comparison in term of bandwidth utilization between the BA sub-TLV based advertisement and the [\[RFC4203\]](#) based one. In order to provide a meaningful comparison between the two solutions (i.e. with same type and quantity of

information carried) it is necessary to assume [\[RFC4203\]](#) tools properly extended.

In other words it is assumed that both unreserved bandwidth and max LSP bandwidth are advertised per signal type. The unreserved bandwidth per signal type could be advertised by means of an unreserved bandwidth sub-tlv per signal type (1 header word + 8 body words) or using the technology specific part of the ISCD (8 words). In this example the utilization of the technology specific part of the ISCD is considered in order to take into account the most optimized option.

The following example is based on the advertisement of a simple link supporting six different types of fixed bandwidth service types (A,B,C,D,E,F) and a variable length service type (G).

+-----+

+-----+

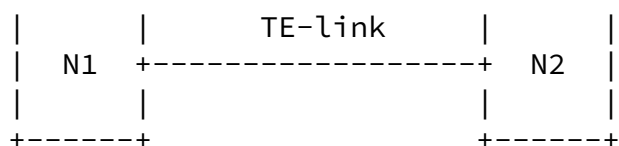


Figure 7: Example

Three different cases are analyzed:

- 8 priorities supported
- 5 priorities supported
- 1 priorities supported

In the first case, [\[RFC4203\]](#) approach would use 1 ISCD per signal type. The ISCD would need to be extended as follows:

[illegible]

	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		
+	+	+	+
	Technology Specific Part		
+	+	+	S
	Technology Specific Part		W
+	+	+	I
	Unreserved Bandwidth at priority 0		T
+	+	+	C
	Unreserved Bandwidth at priority 1		H.
+	+	+	
	Unreserved Bandwidth at priority 2		C
+	+	+	A
	Unreserved Bandwidth at priority 3		P.
+	+	+	
	Unreserved Bandwidth at priority 4		S
+	+	+	P
	Unreserved Bandwidth at priority 5		E
+	+	+	C.
	Unreserved Bandwidth at priority 6		
+	+	+	I
	Unreserved Bandwidth at priority 7		N
+	+	+	F.

Figure 8: Example

The amount of words used per ISCD is 20 for a total amount of 140

words. On the other side, using the BA sub-TLV these words would be used:

- 1 word for type/length declaration

- 1 word for sub-tlv header
- 2 words per (fixed) service type per priority = $2 \times 6 \times 8 = 96$
- 4 words per (variable) service type per priority = $4 \times 1 \times 8 = 32$

Total words used with 8 priorities: 140 ([RFC4203](#)) vs 130 (BA sub-TLV).

Performing the same computation in a scenario where 5 priorities are supported, the number of words used in the [[RFC4203](#)] approach would be the same (140), while in the BA sub-TLV would be:

- 1 word for type/length declaration
- 1 word for sub-tlv header
- 2 words per (fixed) service type per priority = $2 \times 6 \times 5 = 60$
- 4 words per (variable) service type per priority = $4 \times 1 \times 5 = 20$

Total words used with 5 priorities: 140 ([RFC4203](#)) vs 82 (BA sub-TLV).

The difference is significantly higher as the number of supported priorities decreases. Considering the case of single priority, the number of words used by the BA sub-TLV approach would be:

- 1 word for type/length declaration
- 1 word for sub-tlv header
- 2 words per (fixed) service type per priority = $2 \times 6 \times 1 = 12$
- 4 words per (variable) service type per priority = $4 \times 1 \times 1 = 4$

Total words used with 1 priority: 140 ([RFC4203](#)) vs 18 (BA sub-TLV).

It is worth considering that using the Unreserved bandwidth sub-TLV for unreserved bandwidth advertisement would increase the difference between the two solutions due to the fact that a higher number of headers is needed and at least a new word per sub-TLV would be required for the identification of the service type.

6. Compatibility Considerations

Backward compatibility issues are addressed in technology specific documents.

7. 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. [[RFC3630](#)] suggests mechanisms such as [[RFC2154](#)] 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.

8. IANA Considerations

TBD

9. Contributors

Francesco Fondelli, Ericsson

Email: francesco.fondelli@ericsson.com

Eve Varma, Alcatel-Lucent

Email: eve.varma@alcatel-lucent.com

Jonathan Sadler, Tellabs

Email: Jonathan.Sadler@tellabs.com

Lyndon Ong, Ciena

Email: Lyong@Ciena.com

Internet-Draft

Technology Agnostic OSPF-TE

October 2010

[10.](#) Acknowledgements

TBD

[11.](#) References

[11.1.](#) Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC2154] Murphy, S., Badger, M., and B. Wellington, "OSPF with Digital Signatures", [RFC 2154](#), June 1997.
- [RFC2370] Coltun, R., "The OSPF Opaque LSA Option", [RFC 2370](#), July 1998.
- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", [RFC 3630](#), September 2003.
- [RFC4201] Kompella, K., Rekhter, Y., and L. Berger, "Link Bundling in MPLS Traffic Engineering (TE)", [RFC 4201](#), October 2005.
- [RFC4202] Kompella, K. and Y. Rekhter, "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", [RFC 4202](#), October 2005.
- [RFC4203] Kompella, K. and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", [RFC 4203](#), October 2005.
- [RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", [RFC 5250](#), July 2008.
- [RFC5339] Le Roux, J.L. and D. Papadimitriou, "Evaluation of Existing GMPLS Protocols against Multi-Layer and Multi-Region Networks (MLN/MRN)", [RFC 5339](#), September 2008.

11.2. Informative References

Ceccarelli, et al. Expires April 18, 2011 [Page 17]

Internet-Draft Technology Agnostic OSPF-TE October 2010

Authors' Addresses

Daniele Ceccarelli
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy

Email: daniele.ceccarelli@ericsson.com

Diego Caviglia
Ericsson
Via A. Negrone 1/A
Genova - Sestri Ponente
Italy

Email: diego.caviglia@ericsson.com

Sergio Belotti
Alcatel-Lucent
Via Trento, 30
Vimercate
Italy

Email: sergio.belotti@alcatel-lucent.com

Pietro Vittorio Grandi
Alcatel-Lucent
Via Trento, 30
Vimercate

Italy

Email: pietro_vittorio.grandi@alcatel-lucent.com

Fatai Zhang

Huawei Technologies

F3-5-B R&D Center, Huawei Base

Shenzhen 518129 P.R.China Bantian, Longgang District

Phone: +86-755-28972912

Email: zhangfatai@huawei.com

Ceccarelli, et al.

Expires April 18, 2011

[Page 18]

Internet-Draft

Technology Agnostic OSPF-TE

October 2010

Dan Li

Huawei Technologies

F3-5-B R&D Center, Huawei Base

Shenzhen 518129 P.R.China Bantian, Longgang District

Phone: +86-755-28973237

Email: danli@huawei.com

John E Drake

Juniper

Email: jdrake@juniper.net

