

TEWG
Internet Draft

Francois Le Faucheur
Editor
Cisco Systems,
Inc.
December 2004

Document: [draft-ietf-tewg-diff-te-proto-08.txt](#)

Expires: June 2005

**Protocol extensions for support of
Differentiated-Service-aware MPLS Traffic Engineering**

Status of this Memo

This document is an Internet-Draft and is subject to all provisions of [section 3 of RFC 3667](#). By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she become aware will be disclosed, in accordance with [RFC 3668](#).

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 June 13, 2005.

Copyright Notice

Copyright (C) The Internet Society (2004). All Rights Reserved.

Abstract

This document specifies the protocol extensions for support of Differentiated-Service-aware MPLS Traffic Engineering (DS-TE). This includes generalization of the semantic of a number of IGP extensions

already defined for existing MPLS Traffic Engineering in [RFC3630](#) and RFC-TBD as well as additional IGP extensions beyond those. This also includes extensions to RSVP-TE signaling beyond those already specified in [RFC3209](#) for existing MPLS Traffic Engineering. These extensions address the Requirements for DS-TE spelt out in [RFC3564](#).

<RFC-Editor-note> To be removed by the RFC editor at the time of publication:
Please replace "TBD" above by the actual RFC number when
an RFC number is allocated to [[ISIS-TE](#)]
</RFC-Editor-note>

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 [[RFC2119](#)].

Table of Contents

1.	Introduction.....	3
2.	Contributing Authors.....	4
3.	Definitions.....	5
4.	Configurable Parameters.....	5
4.1.	Link Parameters.....	5
4.1.1.	Bandwidth Constraints (BCs) 5	
4.1.2.	Overbooking6	
4.2.	LSR Parameters.....	6
4.2.1.	TE-Class Mapping 7	
4.3.	LSP Parameters.....	8
4.3.1.	Class-Type 8	
4.3.2.	Setup and Holding Preemption Priorities 8	
4.3.3.	Class-Type/Preemption Relationship 8	
4.4.	Examples of Parameters Configuration.....	8
4.4.1.	Example 1 8	
4.4.2.	Example 2 9	
4.4.3.	Example 3 10	
4.4.4.	Example 4 10	
4.4.5.	Example 5 11	
5.	IGP Extensions for DS-TE.....	11
5.1.	Bandwidth Constraints.....	11
5.2.	Unreserved Bandwidth.....	14
6.	RSVP-TE Extensions for DS-TE.....	15
6.1.	DS-TE related RSVP Messages Format.....	15
6.1.1.	Path Message Format 15	
6.2.	CLASSTYPE Object.....	16
6.2.1.	CLASSTYPE object 16	
6.3.	Handling CLASSTYPE Object.....	16

6.4. Non-support of the CLASSTYPE Object.....	19
---	--------------------

6.5. Error Codes For Diff-Serv-aware TE.....	19
7. DS-TE support with MPLS extensions.....	20
7.1. DS-TE support and references to preemption priority.....	21
7.2. DS-TE support and references to Maximum Reservable Bandwidth	21
8. Constraint Based Routing.....	21
9. Diff-Serv scheduling.....	22
10. Existing TE as a Particular Case of DS-TE.....	22
11. Computing "Unreserved TE-Class [i]" and Admission Control Rules	22
11.1. Computing "Unreserved TE-Class [i]".....	23
11.2. Admission Control Rules.....	23
12. Security Considerations.....	24
13. Acknowledgments.....	24
14. IANA Considerations.....	24
14.1. A new name space for Bandwidth Constraints Model Identifiers	24
14.2. A new name space for Error Values under the "Diff-Serv-aware TE Error".....	24
14.3. Assignments made in this Document.....	25
14.3.1. Bandwidth Constraints sub-TLV for OSPF version 2	25
14.3.2. Bandwidth Constraints sub-TLV for ISIS	25
14.3.3. CLASSTYPE object for RSVP	26
14.3.4. "Diff-Serv-aware TE Error" Error Code	26
14.3.5. Error Values for "Diff-Serv-aware TE Error"	26
15. Intellectual Property Considerations.....	27
16. Normative References.....	28
17. Informative References.....	28
18. Editor's Address:.....	29
19. Full Copyright Statement.....	29
Appendix A - Prediction for Multiple Path Computation.....	30
Appendix B - Solution Evaluation.....	31
Appendix C - Interoperability with non DS-TE capable LSRs.....	32
Disclaimer of Validity.....	35
Copyright Statement.....	35
Acknowledgment.....	35

1.Introduction

[DSTE-REQ] presents the Service Providers requirements for support of Differentiated-Service (Diff-Serv)-aware MPLS Traffic Engineering (DS-TE). This includes the fundamental requirement to be able to enforce different bandwidth constraints for different classes of traffic.

This document specifies the IGP and RSVP-TE signaling extensions (beyond those already specified for existing MPLS Traffic Engineering [[OSPF-TE](#)][[ISIS-TE](#)][[RSVP-TE](#)]) for support of the DS-TE requirements spelled out in [[DSTE-REQ](#)] including environments relying on

distributed Constraint Based Routing (e.g. path computation involving Head-end LSRs).

[DSTE-REQ] provides a definition and examples of Bandwidth Constraints Models. The present document does not specify nor assume a particular Bandwidth Constraints model. Specific Bandwidth Constraints model are outside the scope of this document. While the extensions for DS-TE specified in this document may not be sufficient to support all the conceivable Bandwidth Constraints models, they do support the "Russian Dolls" Model specified in [[DSTE-RDM](#)], the "Maximum Allocation" Model specified in [[DSTE-MAM](#)] and the "Maximum Allocation with Reservation" Model specified in [[DSTE-MAR](#)].

There may be differences between the quality of service expressed and obtained with Diffserv without DS-TE and with DS-TE. Because DS-TE uses Constraint Based Routing, and because of the type of admission control capabilities it adds to Diffserv, DS-TE has capabilities for traffic that Diffserv does not: Diffserv does not indicate preemption, by intent, whereas DS-TE describes multiple levels of preemption for its Class Types. Also, Diffserv does not support any means of explicitly controlling overbooking, while DS-TE allows this. When considering a complete quality of service environment, with Diffserv routers and DS-TE, it is important to consider these differences carefully.

2. Contributing Authors

This document was the collective work of several. The text and content of this document was contributed by the editor and the co-authors listed below. (The contact information for the editor appears in [Section 18](#), and is not repeated below.)

Jim Boyle
Protocol Driven Networks, Inc.
1381 Kildaire Farm Road #288
Cary, NC 27511, USA
Phone: (919) 852-5160
Email: jboyle@pdnets.com

Kireeti Kompella
Juniper Networks, Inc.
1194 N. Mathilda Ave.
Sunnyvale, CA 94099
Email: kireeti@juniper.net

William Townsend
Tenor Networks
100 Nagog Park
Acton, MA 01720
Phone: +1-978-264-4900
Email:
btownsend@tenornetworks.com

Thomas D. Nadeau
Cisco Systems, Inc.
250 Apollo Drive
Chelmsford, MA 01824
Phone: +1-978-244-3051
Email: tnadeau@cisco.com

Darek Skalecki

Nortel Networks

Le Faucheur, et al.

[Page 4]

3500 Carling Ave,
Nepean K2H 8E9
Phone: +1-613-765-2252
Email: dareks@nortelnetworks.com

3. Definitions

For readability a number of definitions from [[DSTE-REQ](#)] are repeated here:

Traffic Trunk: an aggregation of traffic flows of the same class [i.e. which are to be treated equivalently from the DS-TE perspective] which are placed inside a Label Switched Path.

Class-Type (CT): the set of Traffic Trunks crossing a link that is governed by a specific set of Bandwidth constraints. CT is used for the purposes of link bandwidth allocation, constraint based routing and admission control. A given Traffic Trunk belongs to the same CT on all links.

TE-Class: A pair of:

- i. a Class-Type
- ii. a preemption priority allowed for that Class-Type. This means that an LSP transporting a Traffic Trunk from that Class-Type can use that preemption priority as the set-up priority, as the holding priority or both.

Definitions for a number of MPLS terms are not repeated here. Those can be found in [[MPLS-ARCH](#)].

4. Configurable Parameters

This section only discusses the differences with the configurable parameters supported for MPLS Traffic Engineering as per [[TE-REQ](#)], [[ISIS-TE](#)], [[OSPF-TE](#)], and [[RSVP-TE](#)]. All other parameters are unchanged.

4.1. Link Parameters

4.1.1. Bandwidth Constraints (BCs)

[[DSTE-REQ](#)] states that "Regardless of the Bandwidth Constraints Model, the DS-TE solution MUST allow support for up to 8 BCs."

For DS-TE, the existing "Maximum Reservable link bandwidth" parameter is retained but its semantic is generalized and interpreted as the

aggregate bandwidth constraints across all Class-Types, so that, independently of the Bandwidth Constraints Model in use:

$\text{SUM}(\text{Reserved}(\text{CTc})) \leq \text{Max Reservable Bandwidth},$

where the SUM is across all values of "c" in the range $0 \leq c \leq 7$.

Additionally, on every link, a DS-TE implementation MUST provide for configuration of up to 8 additional link parameters which are the eight potential Bandwidth Constraints i.e. BC0, BC1, ..., BC7. The LSR MUST interpret these Bandwidth Constraints in accordance with the supported Bandwidth Constraints Model (i.e. what bandwidth constraint applies to what Class-Type and how).

Where the Bandwidth Constraints Model imposes some relationship among the values to be configured for these Bandwidth Constraints, the LSR MUST enforce those at configuration time. For example, when the "Russian Doll" Bandwidth Constraints Model ([\[DSTE-RDM\]](#)) is used, the LSR MUST ensure that BC_i is configured smaller or equal to BC_j, where i is greater than j, and ensure that BC0 is equal to the Maximum Reservable Bandwidth. As another example, when the Maximum Allocation Model ([\[DSTE-MAM\]](#)) is used, the LSR MUST ensure that all BC_i are configured smaller or equal to the Maximum Reservable Bandwidth.

4.1.2.Overbooking

DS-TE enables a network administrator to apply different overbooking (or underbooking) ratios for different CTs.

The principal methods to achieve this are the same as historically used in existing TE deployment, which are :

- (i) To take into account the overbooking/underbooking ratio appropriate for the OA/CT associated with the considered LSP at the time of establishing the bandwidth size of a given LSP. We refer to this method as the "LSP Size Overbooking method". AND/OR
- (ii) To take into account the overbooking/underbooking ratio at the time of configuring the Maximum Reservable Bandwidth/Bandwidth Constraints and use values which are larger(overbooking) or smaller(underbooking) than actually supported by the link. We refer to this method as the "Link Size Overbooking method".

The "LSP Size Overbooking" method and the "Link Size Overbooking" method are expected to be sufficient in many DS-TE environments and require no additional configurable parameters. Other overbooking methods may involve such additional configurable parameters but are beyond the scope of this document.

4.2.LSR Parameters

4.2.1. TE-Class Mapping

In line with [[DSTE-REQ](#)], the preemption attributes defined in [TE-REQ] are retained with DS-TE and applicable within, and across, all Class Types. The preemption attributes of setup priority and holding priority retain existing semantics, and in particular these semantics are not affected by the LSP Class Type. This means that if LSP1 contends with LSP2 for resources, LSP1 may preempt LSP2 if LSP1 has a higher set-up preemption priority (i.e. lower numerical priority value) than LSP2 holding preemption priority regardless of LSP1 CT and LSP2 CT.

DS-TE LSRs MUST allow configuration of a TE-Class mapping whereby the Class-Type and preemption level are configured for each of (up to) 8 TE-Classes.

This mapping is referred to as :

$$\text{TE-Class}[i] \quad <--> \quad < \text{CT}_c, \text{preemption } p >$$

Where $0 \leq i \leq 7$, $0 \leq c \leq 7$, $0 \leq p \leq 7$

Two TE-Classes MUST NOT be identical (i.e. have both the same Class-Type and the same preemption priority).

There are no other restrictions on how any of the 8 Class-Types can be paired up with any of the 8 preemption priorities to form a TE-class. In particular, one given preemption priority can be paired up with two (or more) different Class-Types to form two (or more) TE-classes. Similarly, one Class-Type can be paired up with two (or more) different preemption priorities to form two (or more) TE-Classes. Also, there is no mandatory ordering relationship between the TE-Class index (i.e. "i" above) and the Class-Type (i.e. "c" above) or the preemption priority (i.e. "p" above) of the TE-Class.

Where the network administrator uses less than 8 TE-Classes, the DS-TE LSR MUST allow remaining ones to be configured as "Unused". Note that "Configuring all the 8 TE-Classes as "Unused" effectively results in disabling TE/DS-TE since no TE/DS-TE LSP can be established (nor even configured, since as described in [section 4.3.3](#) below, the CT and preemption priorities configured for an LSP MUST form one of the configured TE-Classes)".

To ensure coherent DS-TE operation, the network administrator MUST configure exactly the same TE-Class Mapping on all LSRs of the DS-TE domain.

When the TE-class mapping needs to be modified in the DS-TE domain, care ought to be exercised during the transient period of reconfiguration during which some DS-TE LSRs may be configured with

the new TE-class mapping while others are still configured with the old TE-class mapping. It is recommended that active tunnels do not use any of the TE-classes which are being modified during such a transient reconfiguration period.

4.3.LSP Parameters

4.3.1.Class-Type

With DS-TE, LSRs MUST support, for every LSP, an additional configurable parameter which indicates the Class-Type of the Traffic Trunk transported by the LSP.

There is one and only one Class-Type configured per LSP.

The configured Class-Type indicates, in accordance with the supported Bandwidth Constraints Model, the Bandwidth Constraints that MUST be enforced for that LSP.

4.3.2.Setup and Holding Preemption Priorities

As per existing TE, DS-TE LSRs MUST allow every DS-TE LSP to be configured with a setup and holding priority, each with a value between 0 and 7.

4.3.3.Class-Type/Preemption Relationship

With DS-TE, the preemption priority configured for the setup priority of a given LSP and the Class-Type configured for that LSP MUST be such that, together, they form one of the (up to) 8 TE-Classes configured in the TE-Class Mapping specified in [section 4.2.1](#) above.

The preemption priority configured for the holding priority of a given LSP and the Class-Type configured for that LSP MUST also be such that, together, they form one of the (up to) 8 TE-Classes configured in the TE-Class Mapping specified in [section 4.2.1](#) above.

The LSR MUST enforce these two rules at configuration time.

4.4.Examples of Parameters Configuration

For illustrative purposes, we now present a few examples of how these configurable parameters may be used. All these examples assume that different bandwidth constraints need to be enforced for different sets of Traffic Trunks (e.g. for Voice and for Data) so that two, or more, Class-Types need to be used.

4.4.1.Example 1

The Network Administrator of a first network using two Class Types (CT1 for Voice and CT0 for Data), may elect to configure the following TE-Class Mapping to ensure that Voice LSPs are never driven away from their shortest path because of Data LSPs:

```
TE-Class[0]  <-->  < CT1 , preemption 0 >
TE-Class[1]  <-->  < CT0 , preemption 1 >
TE-Class[i]  <-->  unused,    for 2 <= i <= 7
```

Voice LSPs would then be configured with:

- CT=CT1, set-up priority =0, holding priority=0

Data LSPs would then be configured with:

- CT=CT0, set-up priority =1, holding priority=1

A new Voice LSP would then be able to preempt an existing Data LSP in case they contend for resources. A Data LSP would never preempt a Voice LSP. A Voice LSP would never preempt another Voice LSP. A Data LSP would never preempt another Data LSP.

4.4.2.Example 2

The Network Administrator of another network may elect to configure the following TE-Class Mapping in order to optimize global network resource utilization by favoring placement of large LSPs closer to their shortest path:

```
TE-Class[0]  <-->  < CT1 , preemption 0 >
TE-Class[1]  <-->  < CT0 , preemption 1 >
TE-Class[2]  <-->  < CT1 , preemption 2 >
TE-Class[3]  <-->  < CT0 , preemption 3 >
TE-Class[i]  <-->  unused,    for 4 <= i <= 7
```

Large size Voice LSPs could be configured with:

- CT=CT1, set-up priority =0, holding priority=0

Large size Data LSPs could be configured with:

- CT=CT0, set-up priority = 1, holding priority=1

Small size Voice LSPs could be configured with:

- CT=CT1, set-up priority = 2, holding priority=2

Small size Data LSPs could be configured with:

- CT=CT0, set-up priority = 3, holding priority=3.

A new large size Voice LSP would then be able to preempt a small size Voice LSP or any Data LSP in case they contend for resources.

A new large size Data LSP would then be able to preempt a small size Data LSP or a small size Voice LSP in case they contend for

resources, but it would not be able to preempt a large size Voice LSP.

4.4.3.Example 3

The Network Administrator of another network may elect to configure the following TE-Class Mapping in order to ensure that Voice LSPs are never driven away from their shortest path because of Data LSPs while also achieving some optimization of global network resource utilization by favoring placement of large LSPs closer to their shortest path:

```
TE-Class[0]  <-->  < CT1 , preemption 0 >
TE-Class[1]  <-->  < CT1 , preemption 1 >
TE-Class[2]  <-->  < CT0 , preemption 2 >
TE-Class[3]  <-->  < CT0 , preemption 3 >
TE-Class[i]  <-->  unused,    for 4 <= i <= 7
```

Large size Voice LSPs could be configured with:

- CT=CT1, set-up priority = 0, holding priority=0.

Small size Voice LSPs could be configured with:

- CT=CT1, set-up priority = 1, holding priority=1.

Large size Data LSPs could be configured with:

- CT=CT0, set-up priority = 2, holding priority=2.

Small size Data LSPs could be configured with:

- CT=CT0, set-up priority = 3, holding priority=3.

A Voice LSP could preempt a Data LSP if they contend for resources. A Data LSP would never preempt a Voice LSP. A Large size Voice LSP could preempt a small size Voice LSP if they contend for resources. A Large size Data LSP could preempt a small size Data LSP if they contend for resources.

4.4.4.Example 4

The Network Administrator of another network may elect to configure the following TE-Class Mapping in order to ensure that no preemption occurs in the DS-TE domain:

```
TE-Class[0]  <-->  < CT1 , preemption 0 >
TE-Class[1]  <-->  < CT0 , preemption 0 >
TE-Class[i]  <-->  unused,    for 2 <= i <= 7
```

Voice LSPs would then be configured with:

- CT=CT1, set-up priority =0, holding priority=0

Data LSPs would then be configured with:

- CT=CT0, set-up priority =0, holding priority=0

No LSP would then be able to preempt any other LSP.

4.4.5.Example 5

The Network Administrator of another network may elect to configure the following TE-Class Mapping in view of increased network stability through a more limited use of preemption:

```
TE-Class[0]  <-->  < CT1 , preemption 0 >
TE-Class[1]  <-->  < CT1 , preemption 1 >
TE-Class[2]  <-->  < CT0 , preemption 1 >
TE-Class[3]  <-->  < CT0 , preemption 2 >
TE-Class[i]  <-->  unused,    for 4 <= i <= 7
```

Large size Voice LSPs could be configured with:

- CT=CT1, set-up priority = 0, holding priority=0.

Small size Voice LSPs could be configured with:

- CT=CT1, set-up priority = 1, holding priority=0.

Large size Data LSPs could be configured with:

- CT=CT0, set-up priority = 2, holding priority=1.

Small size Data LSPs could be configured with:

- CT=CT0, set-up priority = 2, holding priority=2.

A new large size Voice LSP would be able to preempt a Data LSP in case they contend for resources, but it would not be able to preempt any Voice LSP even a small size Voice LSP.

A new small size Voice LSP would be able to preempt a small size Data LSP in case they contend for resources, but it would not be able to preempt a large size Data LSP or any Voice LSP.

A Data LSP would not be able to preempt any other LSP.

5.IGP Extensions for DS-TE

This section only discusses the differences with the IGP advertisement supported for (aggregate) MPLS Traffic Engineering as per [[OSPF-TE](#)] and [[ISIS-TE](#)]. The rest of the IGP advertisement is unchanged.

5.1.Bandwidth Constraints

As detailed above in [section 4.1.1](#), up to 8 Bandwidth Constraints (BCb, $0 \leq b \leq 7$) are configurable on any given link.

With DS-TE, the existing "Maximum Reservable Bandwidth" sub-TLV ([[OSPF-TE](#)], [[ISIS-TE](#)]) is retained with a generalized semantic so that it MUST now be interpreted as the aggregate bandwidth constraint across all Class-Types [i.e.

SUM (Reserved (CTc)) <= Max Reservable Bandwidth], independently of the Bandwidth Constraints Model.

This document also defines the following new optional sub-TLV to advertise the eight potential Bandwidth Constraints (BC0 to BC7):

"Bandwidth Constraints" sub-TLV:

- Bandwidth Constraints Model Id (1 octet)
- Reserved (3 octets)
- Bandwidth Constraints (Nx4 octets)

Where:

- With OSPF, the sub-TLV is a sub-TLV of the "Link TLV" and its sub-TLV type is TBD
- With ISIS, the sub-TLV is a sub-TLV of the "extended IS reachability TLV" and its sub-TLV type is TBD ().

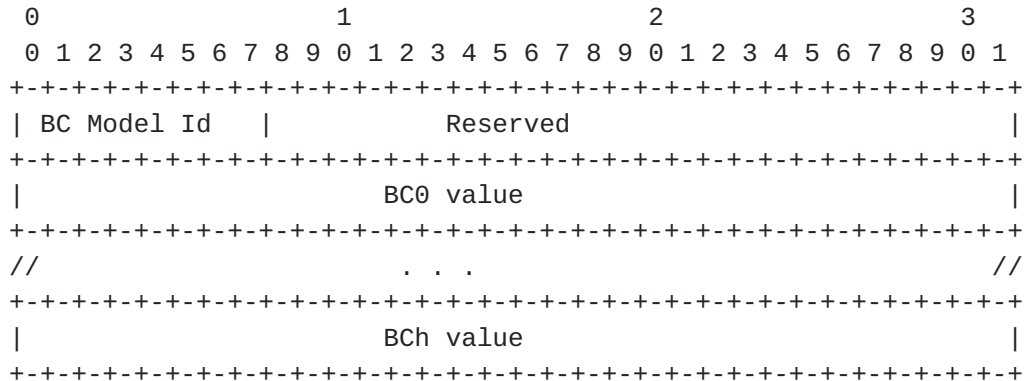
<IANA-note> To be removed by the RFC editor at the time of publication: When the sub-TLV numbers are allocated by IANA for OSPF and ISIS, replace "TBD" in the two bullet points above by the respective assigned value. See IANA Considerations section for allocation request.

</IANA-note>

- Bandwidth Constraints Model Id: 1 octet identifier for the Bandwidth Constraints Model currently in use by the LSR initiating the IGP advertisement. See the IANA Considerations section below for assignment of values in this name space.
- Reserved: a 3-octet field. This field should be set to zero by the LSR generating the sub-TLV and should be ignored by the LSR receiving the sub-TLV.
- Bandwidth Constraints: contains BC0, BC1, ... BC(N-1). Each Bandwidth Constraint is encoded on 32 bits in IEEE floating point format. The units are bytes (not bits!) per second. Where the configured TE-class mapping and the Bandwidth Constraints model in use are such that BCh+1, BCh+2, ...and BC7 are not relevant to any of the Class-Types associated with a configured TE-class, it is RECOMMENDED that only the Bandwidth Constraints from BC0 to BCh be advertised, in order to minimize the impact on IGP scalability.

All relevant generic TLV encoding rules (including TLV format, padding and alignment, as well as IEEE floating point format encoding) defined in [[OSPF-TE](#)] and [[ISIS-TE](#)] are applicable to this new sub-TLV.

The "Bandwidth Constraints" sub-TLV format is illustrated below:



A DS-TE LSR MAY optionally advertise Bandwidth Constraints.

A DS-TE LSR which does advertise Bandwidth Constraints MUST use the new "Bandwidth Constraints" sub-TLV (in addition to the existing Maximum Reservable Bandwidth sub-TLV) to do so. For example, considering the case where a Service Provider deploys DS-TE with TE-classes associated with CT0 and CT1 only, and where the Bandwidth Constraints Model is such that only BC0 and BC1 are relevant to CT0 and CT1: a DS-TE LSR which does advertise Bandwidth Constraints would include in the IGP advertisement the Maximum Reservable Bandwidth sub-TLV as well as the "Bandwidth Constraints" sub-TLV, where the former should contain the aggregate bandwidth constraint across all CTs and the latter would contain BC0 and BC1.

A DS-TE LSR receiving the "Bandwidth Constraints" sub-TLV with a Bandwidth Constraints Model Id which does not match the Bandwidth Constraints Model it currently uses, SHOULD generate a warning to the operator/management-system reporting the inconsistency between Bandwidth Constraints Models used on different links. Also, in that case, if the DS-TE LSR does not support the Bandwidth Constraints Model designated by the Bandwidth Constraints Model Id, or if the DS-TE LSR does not support operations with multiple simultaneous Bandwidth Constraints Models, the DS-TE LSR MAY discard the corresponding TLV. If the DS-TE LSR does support the Bandwidth Constraints Model designated by the Bandwidth Constraints Model Id and if the DS-TE LSR does support operations with multiple simultaneous Bandwidth Constraints Models, the DS-TE LSR MAY accept the corresponding TLV and allow operations with different Bandwidth Constraints Models used in different parts of the DS-TE domain.

5.2.Unreserved Bandwidth

With DS-TE, the existing "Unreserved Bandwidth" sub-TLV is retained as the only vehicle to advertise dynamic bandwidth information necessary for Constraint Based Routing on Head-ends, except that it is used with a generalized semantic. The Unreserved Bandwidth sub-TLV still carries eight bandwidth values but they now correspond to the unreserved bandwidth for each of the TE-Class (instead of for each preemption priority as per existing TE).

More precisely, a DS-TE LSR MUST support the Unreserved Bandwidth sub-TLV with a definition which is generalized into the following:

The Unreserved Bandwidth sub-TLV specifies the amount of bandwidth not yet reserved for each of the eight TE-classes, in IEEE floating point format arranged in increasing order of TE-Class index, with unreserved bandwidth for TE-Class [0] occurring at the start of the sub-TLV, and unreserved bandwidth for TE-Class [7] at the end of the sub-TLV. The unreserved bandwidth value for TE-Class [i] ($0 \leq i \leq 7$) is referred to as "Unreserved TE-Class [i]". It indicates the bandwidth that is available, for reservation, to an LSP which :

- transports a Traffic Trunk from the Class-Type of TE-Class[i], and
- has a setup priority corresponding to the preemption priority of TE-Class[i].

The units are bytes per second.

Since the bandwidth values are now ordered by TE-class index and thus can relate to different CTs with different bandwidth constraints and can relate to any arbitrary preemption priority, a DS-TE LSR MUST NOT assume any ordered relationship among these bandwidth values.

With existing TE, since all preemption priorities reflect the same (and only) bandwidth constraints and since bandwidth values are advertised in preemption priority order, the following relationship is always true, and is often assumed by TE implementations:

If $i < j$, then "Unreserved Bw [i]" \geq "Unreserved Bw [j]"

With DS-TE, no relationship is to be assumed so that:

If $i < j$, then any of the following relationship may be true
 "Unreserved TE-Class [i]" = "Unreserved TE-Class [j]"
 OR
 "Unreserved TE-Class [i]" > "Unreserved TE-Class [j]"
 OR
 "Unreserved TE-Class [i]" < "Unreserved TE-Class [j]".

Rules for computing "Unreserved TE-Class [i]" are specified in [section 11](#).

If TE-Class[i] is unused, the value advertised by the IGP in "Unreserved TE-Class [i]" MUST be set to zero by the LSR generating the IGP advertisement, and MUST be ignored by the LSR receiving the IGP advertisement.

6. RSVP-TE Extensions for DS-TE

In this section we describe extensions to RSVP-TE for support of Diff-Serv-aware MPLS Traffic Engineering. These extensions are in addition to the extensions to RSVP defined in [\[RSVP-TE\]](#) for support of (aggregate) MPLS Traffic Engineering and to the extensions to RSVP defined in [\[DIFF-MPLS\]](#) for support of Diff-Serv over MPLS.

6.1. DS-TE related RSVP Messages Format

One new RSVP Object is defined in this document: the CLASSTYPE Object. Detailed description of this Object is provided below. This new Object is applicable to Path messages. This specification only defines the use of the CLASSTYPE Object in Path messages used to establish LSP Tunnels in accordance with [\[RSVP-TE\]](#) and thus containing a Session Object with a C-Type equal to LSP_TUNNEL_IPv4 and containing a LABEL_REQUEST object.

Restrictions defined in [\[RSVP-TE\]](#) for support of establishment of LSP Tunnels via RSVP-TE are also applicable to the establishment of LSP Tunnels supporting DS-TE. For instance, only unicast LSPs are supported and Multicast LSPs are for further study.

This new CLASSTYPE object is optional with respect to RSVP so that general RSVP implementations not concerned with MPLS LSP set up do not have to support this object.

An LSR supporting DS-TE MUST support the CLASSTYPE Object.

6.1.1. Path Message Format

The format of the Path message is as follows:

```
<Path Message> ::=
    <Common Header> [ <INTEGRITY> ]
    <SESSION> <RSVP_HOP>
    <TIME_VALUES>
    [ <EXPLICIT_ROUTE> ]
    <LABEL_REQUEST>
    [ <SESSION_ATTRIBUTE> ]
    [ <DIFFSERV> ]
    [ <CLASSTYPE> ]
    [ <POLICY_DATA> ... ]
    [ <sender descriptor> ]
```


6.2.CLASSTYPE Object

```
<IANA-note> To be removed by the RFC editor at the time of
publication:      When the RSVP Class-Num is assigned by IANA
replace "TBD"
      above by the assigned value. See IANA Considerations section
      for allocation request.
</IANA-note>
```

```
<IANA-note> To be removed by the RFC editor at the time of
publication:      When the RSVP Class Number is assigned by IANA
replace "TBD"
      above by the assigned value. See IANA Considerations section
      for allocation request.
</IANA-note>
```

Indicates the Class-Type. Values currently allowed are 1, 2, ... , 7. Value of 0 is Reserved.

To establish an LSP tunnel with RSVP, the sender LSR creates a Path message with a session type of LSP_Tunnel_IPv4 and with a

LABEL_REQUEST object as per [[RSVP-TE](#)]. The sender LSR may also include the DIFFSERV object as per [[DIFF-MPLS](#)].

If the LSP is associated with Class-Type 0, the sender LSR MUST NOT include the CLASSTYPE object in the Path message. This allows backward compatibility with non-DSTE-configured or non-DSTE-capable LSRs as discussed below in [section 10](#) and [Appendix C](#).

If the LSP is associated with Class-Type N ($1 \leq N \leq 7$), the sender LSR MUST include the CLASSTYPE object in the Path message with the Class-Type (CT) field set to N.

If a path message contains multiple CLASSTYPE objects, only the first one is meaningful; subsequent CLASSTYPE object(s) MUST be ignored and MUST NOT be forwarded.

Each LSR along the path MUST record the CLASSTYPE object, when present, in its path state block.

If the CLASSTYPE object is not present in the Path message, the LSR MUST associate the Class-Type 0 to the LSP.

The destination LSR responding to the Path message by sending a Resv message MUST NOT include a CLASSTYPE object in the Resv message (whether the Path message contained a CLASSTYPE object or not).

During establishment of an LSP corresponding to the Class-Type N, the LSR MUST perform admission control over the bandwidth available for that particular Class-Type.

An LSR that recognizes the CLASSTYPE object and that receives a path message which:

- contains the CLASSTYPE object, but
- which does not contain a LABEL_REQUEST object or which does not have a session type of LSP_Tunnel_IPv4,

MUST send a PathErr towards the sender with the error code "Diff-Serv-aware TE Error" and an error value of "Unexpected CLASSTYPE object". Those are defined below in [section 6.5](#).

An LSR receiving a Path message with the CLASSTYPE object, which:

- recognizes the CLASSTYPE object, but
- does not support the particular Class-Type,

MUST send a PathErr towards the sender with the error code "Diff-Serv-aware TE Error" and an error value of "Unsupported Class-Type". Those are defined below in [section 6.5](#).

An LSR receiving a Path message with the CLASSTYPE object, which:

- recognizes the CLASSTYPE object, but
- determines that the Class-Type value is not valid (i.e. Class-Type value 0),

MUST send a PathErr towards the sender with the error code "Diff-Serv-aware TE Error" and an error value of "Invalid Class-Type value". Those are defined below in [section 6.5](#).

An LSR receiving a Path message with the CLASSTYPE object, which:

- recognizes the CLASSTYPE object,
- supports the particular Class-Type, but
- determines that the tuple formed by (i) this Class-Type and (ii) the set-up priority signaled in the same Path message, is not one of the eight TE-classes configured in the TE-class mapping,

MUST send a PathErr towards the sender with the error code "Diff-Serv-aware TE Error" and an error value of "CT and setup priority do not form a configured TE-Class". Those are defined below in [section 6.5](#).

An LSR receiving a Path message with the CLASSTYPE object, which:

- recognizes the CLASSTYPE object,
- supports the particular Class-Type, but
- determines that the tuple formed by (i) this Class-Type and (ii) the holding priority signaled in the same Path message, is not one of the eight TE-classes configured in the TE-class mapping,

MUST send a PathErr towards the sender with the error code "Diff-Serv-aware TE Error" and an error value of "CT and holding priority do not form a configured TE-Class". Those are defined below in [section 6.5](#).

An LSR receiving a Path message with the CLASSTYPE object, which:

- recognizes the CLASSTYPE object,
- supports the particular Class-Type, but
- determines that the tuple formed by (i) this Class-Type and (ii) the setup priority signaled in the same Path message, is not one of the eight TE-classes configured in the TE-class mapping, AND
- determines that the tuple formed by (i) this Class-Type and (ii) the holding priority signaled in the same Path message, is not one of the eight TE-classes configured in the TE-class mapping

MUST send a PathErr towards the sender with the error code "Diff-Serv-aware TE Error" and an error value of "CT and setup priority do not form a configured TE-Class AND CT and holding priority do not form a configured TE-Class". Those are defined below in [section 6.5](#).

An LSR receiving a Path message with the CLASSTYPE object and with the DIFFSERV object for an L-LSP, which:

- recognizes the CLASSTYPE object,
- has local knowledge of the relationship between Class-Types and PSC (e.g. via configuration)

- based on this local knowledge, determines that the PSC signaled in the DIFFSERV object is inconsistent with the Class-Type signaled in the CLASSTYPE object,

MUST send a PathErr towards the sender with the error code "Diff-Serv-aware TE Error" and an error value of "Inconsistency between signaled PSC and signaled CT". Those are defined below in [section 6.5](#).

An LSR receiving a Path message with the CLASSTYPE object and with the DIFFSERV object for an E-LSP, which:

- recognizes the CLASSTYPE object,
- has local knowledge of the relationship between Class-Types and PHBs (e.g. via configuration)
- based on this local knowledge, determines that the PHBs signaled in the MAP entries of the DIFFSERV object are inconsistent with the Class-Type signaled in the CLASSTYPE object,

MUST send a PathErr towards the sender with the error code "Diff-Serv-aware TE Error" and an error value of "Inconsistency between signaled PHBs and signaled CT". Those are defined below in [section 6.5](#).

An LSR MUST handle the situations where the LSP can not be accepted for other reasons than those already discussed in this section, in accordance with [[RSVP-TE](#)] and [[DIFF-MPLS](#)] (e.g. a reservation is rejected by admission control, a label can not be associated).

6.4.Non-support of the CLASSTYPE Object

An LSR that does not recognize the CLASSTYPE object Class-Num MUST behave in accordance with the procedures specified in [[RSVP](#)] for an unknown Class-Num whose format is 0bbbbbbb (i.e. it MUST send a PathErr with the error code "Unknown object class" toward the sender).

An LSR that recognizes the CLASSTYPE object Class-Num but does not recognize the CLASSTYPE object C-Type, MUST behave in accordance with the procedures specified in [[RSVP](#)] for an unknown C-type (i.e. it MUST send a PathErr with the error code "Unknown object C-Type" toward the sender).

In both situations, this causes the path set-up to fail. The sender SHOULD notify the operator/management-system that an LSP cannot be established and possibly might take action to retry reservation establishment without the CLASSTYPE object.

6.5.Error Codes For Diff-Serv-aware TE

In the procedures described above, certain errors are reported as a "Diff-Serv-aware TE Error". The value of the "Diff-Serv-aware TE Error" error code is (TBD).

<IANA-note> To be removed by the RFC editor at the time of publication: When the "Diff-Serv-aware TE Error" Error Code is assigned by

IANA, replace "TBD" above by the assigned value.

See IANA Considerations section for allocation request.

</IANA-note>

The following defines error values for the Diff-Serv-aware TE Error:

Value	Error
1	Unexpected CLASSTYPE object
2	Unsupported Class-Type
3	Invalid Class-Type value
4	Class-Type and setup priority do not form a configured TE-Class
5	Class-Type and holding priority do not form a configured TE-Class
6	Class-Type and setup priority do not form a configured TE-Class AND Class-Type and holding priority do not form a configured TE-Class
7	Inconsistency between signaled PSC and signaled Class-Type
8	Inconsistency between signaled PHBs and signaled Class-Type

See the IANA Considerations section for allocation of additional Values.

7.DS-TE support with MPLS extensions.

There are a number of extensions to the initial base specification for signaling [[RSVP-TE](#)] and IGP support for TE [[OSPF-TE](#)][[ISIS-TE](#)]. Those include enhancements for generalization ([[GMPLS-SIG](#)] and [[GMPLS-ROUTE](#)]), as well as for additional functionality such as LSP hierarchy [[HIERARCHY](#)], link bundling [[BUNDLE](#)] and fast restoration [[REROUTE](#)]. These specifications may reference how to encode information associated with certain preemption priorities, how to treat LSPs at different preemption priorities, or otherwise specify encodings or behavior that have a different meaning for a DS-TE router.

In order for an implementation to support both this specification for Diff-Serv-aware TE and a given MPLS enhancement such as those listed above (but not limited to those), it MUST treat references to

"preemption priority" and to "Maximum Reservable Bandwidth" in a generalized manner, such as it is used in this specification.

Additionally, current and future MPLS enhancements may include more precise specification for how they interact with Diff-Serv-aware TE.

7.1.DS-TE support and references to preemption priority

When a router supports both Diff-Serv-aware TE and one of the MPLS protocol extensions such as those mentioned above, encoding of values of preemption priority in signaling or encoding of information associated with preemption priorities in IGP defined for the MPLS extension, MUST be considered to be an encoding of the same information for the corresponding TE-Class. For instance, if an MPLS enhancement specifies advertisement in IGP of a parameter for routing information at preemption priority N, in a DS-TE environment it MUST actually be interpreted as specifying advertisement of the same routing information but for TE-Class [N]. On receipt, DS-TE routers MUST interpret it as such as well.

When there is discussion on how to comparatively treat LSPs of different preemption priority, a DS-TE LSR MUST treat the preemption priorities in this context as the preemption priorities associated with the TE-Classes of the LSPs in question.

7.2.DS-TE support and references to Maximum Reservable Bandwidth

When a router supports both Diff-Serv-aware TE and MPLS protocol extensions such as those mentioned above, advertisements of Maximum Reservable Bandwidth MUST be done with the generalized interpretation defined above in [section 4.1.1](#) as the aggregate bandwidth constraint across all Class-Types and MAY also allow the optional advertisement of all Bandwidth Constraints.

8.Constraint Based Routing

Let us consider the case where a path needs to be computed for an LSP whose Class-Type is configured to CTc and whose set-up preemption priority is configured to p.

Then the pair of CTc and p will map to one of the TE-Classes defined in the TE-Class mapping. Let us refer to this TE-Class as TE-Class[i].

The Constraint Based Routing algorithm of a DS-TE LSR is still only required to perform path computation satisfying a single bandwidth constraint which is to fit in "Unreserved TE-Class [i]" as advertised by the IGP for every link. Thus, no changes are required to the existing TE Constraint Based Routing algorithm itself.

The Constraint Based Routing algorithm MAY also take into account, when used, the optional additional information advertised in IGP such as the Bandwidth Constraints and the Maximum Reservable Bandwidth. As an example, the Bandwidth Constraints MIGHT be used as a tie-breaker criteria in situations where multiple paths, otherwise equally attractive, are possible.

9. Diff-Serv scheduling

The Class-Type signaled at LSP establishment MAY optionally be used by DS-TE LSRs to dynamically adjust the resources allocated to the Class-Type by the Diff-Serv scheduler. In addition, the Diff-Serv information (i.e. the PSC) signaled by the TE-LSP signaling protocols as specified in [[DIFF-MPLS](#)], if used, MAY optionally be used by DS-TE LSRs to dynamically adjust the resources allocated to a PSC/OA within a Class Type by the Diff-Serv scheduler.

10. Existing TE as a Particular Case of DS-TE

We observe that existing TE can be viewed as a particular case of DS-TE where:

- (i) a single Class-Type is used,
- (ii) all 8 preemption priorities are allowed for that Class-Type, and
- (iii) the following TE-Class Mapping is used:
TE-Class[i] <--> < CT0 , preemption i >
Where $0 \leq i \leq 7$.

In that case, DS-TE behaves as existing TE.

As with existing TE, the IGP advertises:

- Unreserved Bandwidth for each of the 8 preemption priorities

As with existing TE, the IGP may advertise:

- Maximum Reservable Bandwidth containing a bandwidth constraint applying across all LSPs

Since all LSPs transport traffic from CT0, RSVP-TE signaling is done without explicit signaling of the Class-Type (which is only used for other Class-Types than CT0 as explained in [section 6](#)) as with existing TE.

11. Computing "Unreserved TE-Class [i]" and Admission Control Rules

11.1.Computing "Unreserved TE-Class [i]"

We first observe that, for existing TE, details on admission control algorithms for TE LSPs, and consequently details on formulas for computing the unreserved bandwidth, are outside the scope of the current IETF work. This is left for vendor differentiation. Note that this does not compromise interoperability across various implementations since the TE schemes rely on LSRs to advertise their local view of the world in terms of Unreserved Bw to other LSRs. This way, regardless of the actual local admission control algorithm used on one given LSR, Constraint Based Routing on other LSRs can rely on advertised information to determine whether an additional LSP will be accepted or rejected by the given LSR. The only requirement is that an LSR advertises unreserved bandwidth values which are consistent with its specific local admission control algorithm and take into account the holding preemption priority of established LSPs.

In the context of DS-TE, again, details on admission control algorithms are left for vendor differentiation and formulas for computing the unreserved bandwidth for TE-Class[i] are outside the scope of this specification. However, DS-TE places the additional requirement on the LSR that the unreserved bandwidth values advertised MUST reflect all of the Bandwidth Constraints relevant to the CT associated with TE-Class[i] in accordance with the Bandwidth Constraints Model. Thus, formulas for computing "Unreserved TE-Class [i]" depend on the Bandwidth Constraints Model in use and MUST reflect how bandwidth constraints apply to CTs. Example formulas for computing "Unreserved TE-Class [i]" Model are provided for the Russian Dolls Model and Maximum Allocation Model respectively in [\[DSTE-RDM\]](#) and [\[DSTE-MAM\]](#).

As with existing TE, DS-TE LSRs MUST consider the holding preemption priority of established LSPs (as opposed to their set-up preemption priority) for the purpose of computing the unreserved bandwidth for TE-Class [i].

11.2.Admission Control Rules

A DS-TE LSR MUST support the following admission control rule:

Regardless of how the admission control algorithm actually computes the unreserved bandwidth for TE-Class[i] for one of its local link, an LSP of bandwidth B, of set-up preemption priority p and of Class-Type CT_c is admissible on that link if, and only if,:

$$B \leq \text{Unreserved Bandwidth for TE-Class}[i]$$

Where

- TE-Class [i] maps to < CTc , p > in the TE-Class mapping configured on the LSR.

12. Security Considerations

This document does not introduce additional security threats beyond those described for Diff-Serv ([\[DIFF-ARCH\]](#)) and MPLS Traffic Engineering ([\[TE-REQ\]](#), [\[RSVP-TE\]](#), [\[OSPF-TE\]](#), [\[ISIS-TE\]](#)) and the same security measures and procedures described in these documents apply here. For example, the approach for defense against theft- and denial-of-service attacks discussed in [\[DIFF-ARCH\]](#), which consists of the combination of traffic conditioning at DS boundary nodes along with security and integrity of the network infrastructure within a Diff-Serv domain, may be followed when DS-TE is in use. Also, as stated in [\[TE-REQ\]](#), it is specifically important that manipulation of administratively configurable parameters (such as those related to DS-TE LSPs) be executed in a secure manner by authorized entities.

13. Acknowledgments

We thank Martin Tatham, Angela Chiu and Pete Hicks for their earlier contribution in this work. We also thank Sanjaya Choudhury for his thorough review and suggestions.

14. IANA Considerations

This document creates two new name spaces which are to be managed by IANA. Also, a number of assignments from existing name spaces have been made by IANA in this document. Those are discussed below.

14.1. A new name space for Bandwidth Constraints Model Identifiers

This document defines in [section 5.1](#) a "Bandwidth Constraints Model Id" field (name space) within the "Bandwidth Constraints" sub-TLV, both for OSPF and ISIS. IANA is requested to create and maintain this new name space. The field for this namespace is 1 octet, and IANA guidelines for assignments for this field are as follows:

- o values in the range 0-239 are to be assigned according to the "Specification Required" policy defined in [\[IANA-CONS\]](#).

- o values in the range 240-255 are reserved for "Private Use" as defined in [\[IANA-CONS\]](#).

14.2. A new name space for Error Values under the "Diff-Serv-aware TE Error"

An Error Code is an 8-bit quantity defined in [[RSVP](#)] that appears in an ERROR_SPEC object to broadly define an error condition. With each Error Code there may be a 16-bit Error Value (which depends on the Error Code) that further specifies the cause of the error.

This document defines in [section 6.5](#) a new RSVP error code, the "Diff-Serv-aware TE Error" (see [section 14.3.4](#)). The Error Values for the "Diff-Serv-aware TE Error" constitute a new name space to be managed by IANA.

This document defines, in [section 6.5](#), values 1 through 7 in that name space (see [section 14.3.5](#)).

Future allocations of values in this name space are to be assigned by IANA using the "Specification Required" policy defined in [IANA-CONS].

14.3.Assignments made in this Document

14.3.1.Bandwidth Constraints sub-TLV for OSPF version 2

[OSPF-TE] creates a name space for the sub-TLV types within the "Link TLV" of the Traffic Engineering LSA and rules for management of this name space by IANA.

This document defines in [section 5.1](#) a new sub-TLV, the "Bandwidth Constraints" sub-TLV, for the OSPF "Link" TLV. In accordance with the IANA considerations provided in [[OSPF-TE](#)], a sub-TLV type in the range 10 to 32767 was requested and the value TBD has been assigned by IANA for the "Bandwidth Constraints" sub-TLV.

<IANA-note> To be removed by the RFC editor at the time of publication:

When the sub-TLV Type is assigned by IANA replace "TBD" above by the assigned value.

</IANA-note>

14.3.2.Bandwidth Constraints sub-TLV for ISIS

[ISIS-TE] creates a name space for the sub-TLV types within the ISIS "Extended IS Reachability" TLV and rules for management of this name space by IANA.

This document defines in [section 5.1](#) a new sub-TLV, the "Bandwidth Constraints" sub-TLV, for the ISIS "Extended IS Reachability" TLV. In accordance with the IANA considerations provided in [[ISIS-TE](#)], a sub-TLV type was requested and the value TBD has been assigned by IANA for the "Bandwidth Constraints" sub-TLV.

<IANA-note> To be removed by the RFC editor at the time of publication:
When the sub-TLV Type is assigned by IANA replace "TBD" above by the assigned value.
</IANA-note>

14.3.3. CLASSTYPE object for RSVP

[RSVP] defines the Class Number name space for RSVP object which is managed by IANA. Currently allocated Class Numbers are listed at "<http://www.iana.org/assignments/rsvp-parameters>"

This document defines in [section 6.2.1](#) a new RSVP object, the CLASSTYPE object. IANA was requested to assign a Class Number for this RSVP object from the range defined in section 3.10 of [\[RSVP\]](#) for those objects which, if not understood, cause the entire RSVP message to be rejected with an error code of "Unknown Object Class". Such objects are identified by a zero in the most significant bit of the class number (i.e. Class-Num = 0bbbbbbb).
IANA assigned Class-Number TBD to the CLASSTYPE object. C_Type 1 is defined in this document for the CLASSTYPE object.

<IANA-note> To be removed by the RFC editor at the time of publication:
When the RSVP Class-Num is assigned by IANA replace "TBD" above by the assigned value.
</IANA-note>

14.3.4. "Diff-Serv-aware TE Error" Error Code

[RSVP] defines the Error Code name space and rules for management of this name space by IANA. Currently allocated Error Codes are listed at "<http://www.iana.org/assignments/rsvp-parameters>"

This document defines in [section 6.5](#) a new RSVP Error Code, the "Diff-Serv-aware TE Error". In accordance with the IANA considerations provided in [\[RSVP\]](#), Error Code TBD was assigned by IANA to the "Diff-Serv-aware TE Error".

<IANA-note> To be removed by the RFC editor at the time of publication:
When the RSVP Class-Num is assigned by IANA replace "TBD" above by the assigned value.
</IANA-note>

14.3.5. Error Values for "Diff-Serv-aware TE Error"

An Error Code is an 8-bit quantity defined in [\[RSVP\]](#) that appears in an ERROR_SPEC object to broadly define an error condition. With each

Error Code there may be a 16-bit Error Value (which depends on the Error Code) that further specifies the cause of the error.

This document defines in [section 6.5](#) a new RSVP error code, the "Diff-Serv-aware TE Error" (see [section 14.3.4](#)). The Error Values for the "Diff-Serv-aware TE Error" constitute a new name space to be managed by IANA.

This document defines, in [section 6.5](#), the following Error Values for the "Diff-Serv-aware TE Error":

Value	Error
-------	-------

- | | |
|---|---|
| 1 | Unexpected CLASSTYPE object |
| 2 | Unsupported Class-Type |
| 3 | Invalid Class-Type value |
| 4 | Class-Type and setup priority do not form a configured TE-Class |
| 5 | Class-Type and holding priority do not form a configured TE-Class |
| 6 | Class-Type and setup priority do not form a configured TE-Class AND Class-Type and holding priority do not form a configured TE-Class |
| 7 | Inconsistency between signaled PSC and signaled Class-Type |
| 8 | Inconsistency between signaled PHBs and signaled Class-Type |

See [section 14.2](#) for allocation of other values in that name space.

15. Intellectual Property Considerations

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

16. Normative References

[DSTE-REQ] Le Faucheur et al, Requirements for support of Diff-Serv-aware MPLS Traffic Engineering, [RFC3564](#).

[MPLS-ARCH] Rosen et al., "Multiprotocol Label Switching Architecture", [RFC3031](#).

[DIFF-ARCH] Blake et al., "An Architecture for Differentiated Services", [RFC2475](#).

[TE-REQ] Awduche et al., "Requirements for Traffic Engineering Over MPLS", [RFC2702](#).

[OSPF-TE] Katz et al., "Traffic Engineering (TE) Extensions to OSPF Version 2", [RFC3630](#).

[ISIS-TE] Smit, Li, "Intermediate System to Intermediate System (IS-IS) extensions for Traffic Engineering (TE)", [RFC 3784](#).

[RSVP-TE] Awduche et al, "RSVP-TE: Extensions to RSVP for LSP Tunnels", [RFC 3209](#).

[RSVP] Braden et al, "Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification", [RFC 2205](#).

[DIFF-MPLS] Le Faucheur et al, "MPLS Support of Diff-Serv", [RFC3270](#).

[RFC2119] S. Bradner, Key words for use in RFCs to Indicate Requirement Levels, [RFC2119](#).

[IANA-CONS], T. Narten et al, "Guidelines for Writing an IANA Considerations Section in RFCs", [RFC2434](#).

17. Informative References

[DSTE-RDM] Le Faucheur et al., "Russian Dolls Bandwidth Constraints Model for DS-TE", [draft-ietf-tewg-diff-te-russian-07.txt](#), work in progress.

[DSTE-MAM] Le Faucheur, Lai, "Maximum Allocation Bandwidth Constraints Model for DS-TE", [draft-ietf-tewg-diff-te-mam-04.txt](#), work in progress .

[DSTE-MAR] Ash, "Max Allocation with Reservation Bandwidth Constraints Model for MPLS/DiffServ TE & Performance Comparisons", [draft-ietf-tewg-diff-te-mar-03.txt](#), work in progress .

[GMPLS-SIG] Berger et. al., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", [RFC3471](#)

[GMPLS-ROUTE] Kompella et. al., "Routing Extensions in Support of Generalized MPLS", [draft-ietf-ccamp-gmpls-routing-09.txt](#), work in progress.

[BUNDLE] Kompella, Rekhter, Berger, "Link Bundling in MPLS Traffic Engineering", [draft-ietf-mpls-bundle-04.txt](#), work in progress.

[HIERARCHY] Kompella, Rekhter, "LSP Hierarchy with Generalized MPLS TE", [draft-ietf-mpls-lsp-hierarchy-08.txt](#), work in progress.

[REROUTE] Pan et. al., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", [draft-ietf-mpls-rsvp-lsp-fastreroute-07.txt](#), work in progress.

18.Editor's Address:

Francois Le Faucheur
Cisco Systems, Inc.
Village d'Entreprise Green Side - Batiment T3
400, Avenue de Roumanille
06410 Biot-Sophia Antipolis
France
Phone: +33 4 97 23 26 19
Email: flefauch@cisco.com

19.Full Copyright Statement

Copyright (C) The Internet Society (2004). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other

Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Appendix A - Prediction for Multiple Path Computation

There are situations where a Head-End needs to compute paths for multiple LSPs over a short period of time. There are potential advantages for the Head-end in trying to predict the impact of the n -th LSP on the unreserved bandwidth when computing the path for the $(n+1)$ -th LSP, before receiving updated IGP information. One example would be to perform better load-distribution of the multiple LSPs across multiple paths. Another example would be to avoid CAC rejection when the $(n+1)$ -th LSP would no longer fit on a link after establishment of the n -th LSP. While there are also a number of conceivable scenarios where doing such predictions might result in a worse situation, it is more likely to improve the situation. As a matter of fact, a number of network administrators have elected to use such predictions when deploying existing TE.

Such predictions are local matters, are optional and are outside the scope of this specification.

Where such predictions are not used, the optional Bandwidth Constraint sub-TLV and the optional Maximum Reservable Bandwidth sub-TLV need not be advertised in IGP for the purpose of path computation since the information contained in the Unreserved Bw sub-TLV is all that is required by Head-Ends to perform Constraint Based Routing.

Where such predictions are used on Head-Ends, the optional Bandwidth Constraints sub-TLV and the optional Maximum Reservable Bandwidth sub-TLV MAY be advertised in IGP. This is in order for the Head-ends to predict as accurately as possible how an LSP affects unreserved bandwidth values for subsequent LSPs.

Remembering that actual admission control algorithms are left for vendor differentiation, we observe that predictions can only be

performed effectively when the Head-end LSR predictions are based on the same (or a very close) admission control algorithm as used by other LSRs.

Appendix B - Solution Evaluation

1. Satisfying Detailed Requirements

This DS-TE Solution addresses all the scenarios presented in [DSTE-REQ].

It also satisfies all the detailed requirements presented in [DSTE-REQ].

The objective set out in the last paragraph of section "4.7 overbooking" of [DSTE-REQ] is only partially addressed by this DS-TE solution. Through support of the "LSP Size Overbooking" and "Link Size Overbooking" methods, this DS-TE solution effectively allows CTs to have different overbooking ratios and simultaneously allows overbooking to be tweaked differently (collectively across all CTs) on different links. But, in a general sense, it does not allow the effective overbooking ratio of every CT to be tweaked differently in different parts of the network independently of other CTs, while maintaining accurate bandwidth accounting of how different CTs mutually affect each other through shared Bandwidth Constraints (such as the Maximum Reservable Bandwidth).

2. Flexibility

This DS-TE solution supports 8 CTs. It is entirely flexible as to how Traffic Trunks are grouped together into a CT.

3. Extendibility

A maximum of 8 CTs is considered by the authors of this document as more than comfortable. A maximum of 8 TE-classes is considered by the authors of this document as sufficient. However, this solution could be extended to support more CTs or more TE-classes if deemed necessary in the future; This would necessitate additional IGP extensions beyond those specified in this document.

Although the prime objective of this solution is support of Diff-Serv-aware Traffic Engineering, its mechanisms are not tightly coupled with Diff-Serv. This makes the solution amenable, or more easily extendable, for support of potential other future Traffic Engineering applications.

4. Scalability

This DS-TE solution is expected to have a very small scalability impact compared to existing TE.

From an IGP viewpoint, the amount of mandatory information to be advertised is identical to existing TE. One additional sub-TLV has been specified, but its use is optional and it only contains a limited amount of static information (at most 8 Bandwidth Constraints).

We expect no noticeable impact on LSP Path computation since, as with existing TE, this solution only requires CSPF to consider a single unreserved bandwidth value for any given LSP.

From a signaling viewpoint we expect no significant impact due to this solution since it only requires processing of one additional information (the Class-Type) and does not significantly increase the likelihood of CAC rejection. Note that DS-TE has some inherent impact on LSP signaling in the sense that it assumes that different classes of traffic are split over different LSPs so that more LSPs need to be signaled; but this is due to the DS-TE concept itself and not to the actual DS-TE solution discussed here.

5. Backward Compatibility/Migration

This solution is expected to allow smooth migration from existing TE to DS-TE. This is because existing TE can be supported as a particular configuration of DS-TE. This means that an "upgraded" LSR with a DS-TE implementation can directly interwork with an "old" LSR supporting existing TE only.

This solution is expected to allow smooth migration when increasing the number of CTs actually deployed since it only requires configuration changes. However, these changes need to be performed in a coordinated manner across the DS-TE domain.

Appendix C - Interoperability with non DS-TE capable LSRs

This DSTE solution allows operations in a hybrid network where some LSRs are DS-TE capable while some LSRs are not DS-TE capable, which may occur during migration phases. This Appendix discusses the constraints and operations in such hybrid networks.

We refer to the set of DS-TE capable LSRs as the DS-TE domain. We refer to the set of non DS-TE capable (but TE capable) LSRs as the TE-domain.

Hybrid operations requires that the TE-class mapping in the DS-TE domain is configured so that:

- a TE-class exist for CT0 for every preemption priority actually used in the TE domain
- the index in the TE-class mapping for each of these TE-classes is equal to the preemption priority.

For example, imagine the TE domain uses preemption 2 and 3. Then, DS-TE can be deployed in the same network by including the following TE-classes in the TE-class mapping:

i	<--->	CT	preemption
=====			
2		CT0	2
3		CT0	3

Another way to look at this is to say that, the whole TE-class mapping does not have to be consistent with the TE domain, but the subset of this TE-Class mapping applicable to CT0 has to effectively be consistent with the TE domain.

Hybrid operations also requires that:

- non DS-TE capable LSRs be configured to advertise the Maximum Reservable Bandwidth
- DS-TE capable LSRs be configured to advertise Bandwidth Constraints (using the Max Reservable Bandwidth sub-TLV as well as the Bandwidth Constraints sub-TLV, as specified in [section 5.1](#) above).

This allows DS-TE capable LSRs to unambiguously identify non DS-TE capable LSRs.

Finally hybrid operations require that non DS-TE capable LSRs be able to accept Unreserved Bw sub-TLVs containing non decreasing bandwidth values (ie with Unreserved [p] < Unreserved [q] with p < q).

In such hybrid networks :

- CT0 LSPs can be established by both DS-TE capable LSRs and non-DSTE capable LSRs
- CT0 LSPs can transit via (or terminate at) both DS-TE capable LSRs and non-DSTE capable LSRs
- LSPs from other CTs can only be established by DS-TE capable LSRs
- LSPs from other CTs can only transit via (or terminate at) DS-TE capable LSRs

Let us consider, the following example to illustrate operations:

LSR0-----LSR1-----LSR2
 Link01 Link12

Where:

LSR0 is a non-DS-TE capable LSR

LSR1 and LSR2 are DS-TE capable LSRs

Let's assume again that preemption 2 and 3 are used in the TE-domain and that the following TE-class mapping is configured on LSR1 and LSR2:

i	<--->	CT	preemption
=====			
0		CT1	0
1		CT1	1
2		CT0	2
3		CT0	3
rest		unused	

LSR0 is configured with a Max Reservable bandwidth=m01 for Link01.
 LSR1 is configured with a BC0=x0 a BC1=x1(possibly=0), and a Max Reservable Bandwidth=m10(possibly=m01) for Link01.

LSR0 will advertise in IGP for Link01:

- Max Reservable Bw sub-TLV = <m01>
- Unreserved Bw sub-TLV =
 <CT0/0,CT0/1,CT0/2,CT0/3,CT0/4,CT0/5,CT0/6,CT0/7>

On receipt of such advertisement, LSR1 will:

- understand that LSR0 is not DS-TE capable because it advertised a Max Reservable Bw sub-TLV and no Bandwidth Constraints sub-TLV
- conclude that only CT0 LSPs can transit via LSR0 and that only the values CT0/2 and CT0/3 are meaningful in the Unreserved Bw sub-TLV. LSR1 may effectively behave as if the six other values contained in the Unreserved Bw sub-TLV were set to zero.

LSR1 will advertise in IGP for Link01:

- Max Reservable Bw sub-TLV = <m10>
- Bandwidth Constraints sub-TLV = <BC Model ID, x0,x1>
- Unreserved Bw sub-TLV = <CT1/0,CT1/1,CT0/2,CT0/3,0,0,0,0>

On receipt of such advertisement, LSR0 will:

- Ignore the Bandwidth Constraints sub-TLV (unrecognized)
- Correctly process CT0/2 and CT0/3 in the Unreserved Bw sub-TLV and use these values for CT0 LSP establishment
- Incorrectly believe that the other values contained in the Unreserved Bw sub-TLV relates to other preemption priorities for CT0, but will actually never use those since we assume that only preemption 2 and 3 are used in the TE domain.

Disclaimer of Validity

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Copyright Statement

Copyright (C) The Internet Society (2004). This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.

