Francois Le Faucheur
Cisco Systems, Inc.

IETF Internet Draft
Expires: April, 2002
Document: draft-lefaucheur-diff-te-mam-00.txt February, 2003

Maximum Allocation Bandwidth Constraints Model for Diff-Serv-aware MPLS Traffic Engineering

Status of this Memo

Abstract

This document provides specification for one Bandwidth Constraints model for Diff-Serv-aware MPLS Traffic Engineering, which is referred to as the Maximum Allocation Model.

Summary for Sub-IP related Internet Drafts

RELATED DOCUMENTS:
draft-ietf-tewg-diff-te-reqts-07.txt
draft-ietf-tewg-diff-te-proto-03.txt

WHERE DOES IT FIT IN THE PICTURE OF THE SUB-IP WORK
This ID is a Working Group document of the TE Working Group.

WHY IS IT TARGETED AT THIS WG(s)
TEWG is responsible for specifying protocol extensions for support of Diff-Serv-aware MPLS Traffic Engineering.

JUSTIFICATION
The TEWG charter states that "This will entail verification and review of the Diffserv requirements in the WG Framework document and initial specification of how these requirements can be met through use and potentially expansion of existing protocols."
In line with this, the TEWG is specifying bandwidth constraints model for Diff-Serv-aware MPLS Traffic Engineering. This document specifies one particular bandwidth constraints model.

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].

1. Introduction

[DSTE-REQ] presents the Service Providers requirements for support of 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.

[DSTE-REQ] also defines the concept of Bandwidth Constraint Models for DS-TE and states that "The DS-TE technical solution MUST specify at least one bandwidth constraint model and MAY specify multiple Bandwidth Constraints models."

This document provides a detailed description of one particular Bandwidth Constraint model for DS-TE which is introduced in [DSTE-REQ] and called the Maximum Allocation Model (MAM).

[DSTE-PROTO] specifies the IGP and RSVP-TE signaling extensions for support of DS-TE. These extensions support MAM.

2. Definitions

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

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.

"Reserved (CTc)": For a given Class-Type CTc (0 <= c <= MaxCT) ,let us define "Reserved(CTc)" as the sum of the bandwidth reserved by all established LSPs which belong to CTc.

The following definition from [DSTE-PROTO] is also repeated here:

Normalised(CTc) : let us define "Normalised(CTc)" as "Reserved(CTc)/LOM(c)", where LOM (c) is the Local Overbooking Multiplier for CTc defined in [DSTE-PROTO].

We also introduce the following definitions:

Reserved(CTb,q) : let us define "Reserved(CTb,q)" as the sum of the bandwidth reserved by all established LSPs which belong to CTb and have a holding priority of q. Note that if q and CTb do not form one of the 8 possible configured TE-Classes, then there can not be any established LSP which belong to CTb and have a holding priority of q, so in that case Reserved(CTb,q)=0.

Normalised(CTc,q) let us define "ormalised(CTc,q)" as "Reserved(CTc/q) / LOM(c)", where LOM (c) is the Local Overbooking Multiplier for CTc defined in [DSTE-PROTO].

3. Maximum Allocation Model Definition

MAM is defined in the following manner (assuming for now that the optional per-CT Local Overbooking Multipliers defined in [DSTE-PROTO] are not used - i.e. LOM[c]=1 , 0<=c<=7):

- o Maximum Number of Bandwidth Constraints (MaxBC)= Maximum Number of Class-Types (MaxCT) = 8
- o for each value of b in the range 0 <= b <= (MaxCT - 1):
  Reserved (CTb) <= BCb,

A DS-TE LSR implementing MAM MUST support enforcement of bandwidth constraints in compliance with this definition.

Where 8 Class-Types are active, the MAM bandwidth constraints can also be expressed in the following way:

- All LSPs from CT7 use no more than BC7
- All LSPs from CT6 use no more than BC6
- All LSPs from CT5 use no more than BC5
- etc.
- All LSPs from CT0 use no more than BC0

Purely for illustration purposes, the diagram below represents MAM in a pictorial manner when 3 Class-Types are active:


```
I-------------II------------II------------I
I    CT2      II   CT1      II     CT0    I
I-------------II------------II------------I

<-----BC2------><-----BC1-----><-----BC0----->
```

While more flexible/sophisticated Bandwidth Constraints models can be defined (and are indeed defined - see [RDM]) , the Maximum Allocation Model is attractive in some DS-TE environments for the following reasons:

- Network administrators generally find MAM extremely simple and intuitive
- MAM matches simple bandwidth control policies that Network Administrators may want to enforce (i.e. set aside a fixed chunk of bandwidth for a given type of traffic (aka. Class-Type).
- MAM can be used in a way which ensures very strict isolation

across Class-Types, so that each Class-Type is guaranteed its share of bandwidth no matter the level of contention by other classes, whether preemption is used or not.
- MAM can simultaneously achieve isolation, bandwidth efficiency and protection against QoS degradation of the premium CT.
- MAM only requires limited protocol extensions such as the ones defined in [DSTE-PROTO].

MAM may not be attractive in some DS-TE environments because:
- MAM cannot simultaneously achieve isolation, bandwidth efficiency and protection against QoS degradation of CTs other than the Premium CT.

Additional considerations on the properties of MAM can be found in [BC-CONS] and [BC-MODEL].

As a very simple example usage of the MAM Model, a network administrator using one CT for Voice (CT1) and one CT for Data (CT0) might configure on a given 2.5 Gb/s link:
- BC0 = 1.5 Gb/s (i.e. Data is limited to 1.5 Gb/s)
- BC1 = 1 Gb/s   (i.e. Voice is limited to 1 Gb/s).

4. Example Formulas for Computing "Unreserved TE-Class [i]" with Maximum Allocation Model

As specified in [DSTE-PROTO], formulas for computing "Unreserved TE-Class [i]" MUST reflect all of the Bandwidth Constraints relevant to the CT associated with TE-Class[i], and thus, depend on the Bandwidth Constraints Model. Thus, a DS-TE LSR implementing MAM MUST reflect the MAM bandwidth constraints defined in section 3 above when computing "Unreserved TE-Class [i]".

Keeping in mind, as explained in [DSTE-PROTO], that details of admission control algorithms as well as formulas for computing "Unreserved TE-Class [i]" are outside the scope of the IETF work, we provide in this section, for illustration purposes, an example of how values for the unreserved bandwidth for TE-Class[i] might be computed with MAM, assuming:

- the basic admission control algorithm which simply deducts the exact bandwidth of any established LSP from all of the Bandwidth Constraints relevant to the CT associated with that LSP.
- the optional per-CT Local Overbooking Multipliers are not used (.i.e. LOM[c]=1, 0<= c <=7).

Then:

```
 "Unreserved TE-Class [i]" =
[ BCc - SUM ( Reserved(CTc,q) ) ] for q <= p
```

where:

```
     TE-Class [i] <--> < CTc , preemption p>
```

in the configured TE-Class mapping.

5. Support of Optional Local Overbooking Method

We remind the reader that, as discussed in [DSTE-PROTO], the "LSP/link size overbooking" method (which does not use the Local Overbooking Multipliers - LOMs-) is expected to be sufficient in many DS-TE environments. It is expected that the optional Local Overbooking method (and LOMs) would only be used in specific environments, in particular where different overbooking ratios need to be enforced on different links of the DS-TE domain and cross-effect of overbooking across CTs needs to be accounted for very accurately.

This section discusses the impact of the optional local overbooking method on MAM and associated rules and formula. This is only applicable in the cases where the optional local overbooking method is indeed supported by the DS-TE LSRs and actually deployed.

5.1. Maximum Allocation Model Definition With Local Overbooking

As specified in [DSTE-PROTO], when the optional Local Overbooking method is supported, the bandwidth constraints MUST be applied to "Normalised(CTc)" rather than to "Reserved(CTc)". Thus, when the optional Local Overbooking method is supported, the MAM definition is extended in the following manner:

- Maximum Number of Bandwidth Constraints (MaxBC)= Maximum Number of Class-Types (MaxCT) = 8
- for each value of b in the range 0 <= b <= (MaxCT - 1):
  Normalised(CTb) <= BCb,

A DS-TE LSR implementing MAM and implementing the optional Local Overbooking method MUST support enforcement of bandwidth constraints in compliance with this extended definition.

Purely for illustration purposes, the diagram below represents the Russian Doll Bandwidth Constraints model in a pictorial manner when 3 Class-Types are active and the local overbooking method is used:

```
I-------------------II------------------II--------------------I
I  Normalised(CT2)  II  Normalised(CT1)  II   Normalised(CT0)  I
I-------------------II------------------II--------------------I

<--------BC2--------><--------BC1--------><--------BC0--------->
```

5.2. Example Formulas for Computing "Unreserved TE-Class [i]" With Local Overbooking

A DS-TE LSR implementing MAM and implementing the optional Local Overbooking method MUST reflect the MAM bandwidth constraints defined in section 5.1 above when computing "Unreserved TE-Class [i]".

Again, keeping in mind that details of admission control algorithms as well as formulas for computing "Unreserved TE-Class [i]" are outside the scope of the IETF work, we provide in this section, for illustration purposes, an example of how values for the unreserved bandwidth for TE-Class[i] might be computed with MAM, assuming:

- the basic admission control algorithm which simply deducts the exact bandwidth of any established LSP from all of the Bandwidth Constraints relevant to the CT associated with that LSP.
- the optional per-CT Local Overbooking Multipliers are used.

When the optional local overbooking method is supported, the example generalized formula of section 4 becomes:

```
 "Unreserved TE-Class [i]" =
LOM(c) x [ BCc - SUM ( Normalised(CTc,q) ) ] for q <= p ,
```

Or, equivalently:

```
 "Unreserved TE-Class [i]" =
[ LOM(c) x  BCc ] - SUM ( Reserved  (CTc,q) ) for q <= p ,
```

where:

```
     TE-Class [i] <--> < CTc , preemption p>
```

in the configured TE-Class mapping.

5.3. Example Usage of LOM

To illustrate usage of the local overbooking method with MAM, let's consider a DS-TE deployment where two CTs (CT0 for data and CT1 for voice) and a single preemption priority are used.

The TE-Class mapping is the following:

```
     TE-Class  <-->  CT, preemption
     ==============================
         0           CT0, 0
         1           CT1, 0
         rest         unused
```

Let's assume that on a given link, BCs and LOMs are configured in the following way:

```
     BC0 = 200
     BC1 = 100
     LOM(0) = 4  (i.e. = 400%)
     LOM(1) = 2  (i.e. = 200%)
```

Let's further assume that the DS-TE LSR uses the example formulas presented above for computing unreserved bandwidth values.

If there is no established LSP on the considered link, the LSR will advertise for that link in IGP :

```
        Unreserved TE-Class [0] = 4 x (200 - 0/4)= 800
        Unreserved TE-Class [1] = 2 x (100- 0/2) = 200
```

Note again that these values advertised for Unreserved Bandwidth are larger than BC1 and BC0.

If there is only a single established LSP, with CT=CT0 and BW=100, the LSR will advertise:

```
        Unreserved TE-Class [0] = 4 x (200 - 100/4)= 700
        Unreserved TE-Class [1] = 2 x (100- 0/2)   = 200
```

If there is only a single established LSP, with CT=CT1 and BW=100, the LSR will advertise:

```
        Unreserved TE-Class [0] = 4 x (200 - 0/4)  = 800
        Unreserved TE-Class [1] = 2 x (100- 100/2) = 100
```

6. Security Considerations

Security considerations related to the use of DS-TE are discussed in [DSTE-PROTO]. Those apply independently of the Bandwidth Constraints model, including MAM specified in this document.

7. Acknowledgments

A lot of the material in this document has been derived from ongoing discussions within the TEWG work. This involved many people including Jerry Ash, Waisum Lai and Dimitry Haskin.

8. Normative References

[DSTE-REQ] Le Faucheur et al, Requirements for support of Diff-Serv-aware MPLS Traffic Engineering, draft-ietf-tewg-diff-te-reqts-07.txt, February 2003.

[DSTE-PROTO] Le Faucheur et al, Protocol extensions for support of Diff-Serv-aware MPLS Traffic Engineering, draft-ietf-tewg-diff-te-proto-03.txt, February 2003

[RFC2119] S. Bradner, Key words for use in RFCs to Indicate Requirement Levels, RFC2119, March 1997.

9. Informative References

[BC-CONS] Le Faucheur, "Considerations on Bandwidth Constraints Model for DS-TE", draft-lefaucheur-tewg-russian-dolls-00.txt, June 2002.

[BC-MODEL] Lai, "Bandwidth Constraints Models for DS-TE", draft-wlai-tewg-bcmodel-00.txt, June 2002.

[RDM] Le Faucheur et al., "Russian Dolls Bandwidth Constraints Model for Diff-Serv-aware MPLS Traffic Engineering", draft-ietf-tewg-diff-te-russian-01.txt, February 2003

[OSPF-TE] Katz et al., Traffic Engineering Extensions to OSPF, draft-katz-yeung-ospf-traffic-09.txt, October 2002.

[ISIS-TE] Smit et al., IS-IS extensions for Traffic Engineering, draft-ietf-isis-traffic-04.txt, December 2002.

[RSVP-TE] Awduche et al, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[DIFF-MPLS] Le Faucheur et al, "MPLS Support of Diff-Serv", RFC3270, May 2002.

10. Author'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