Network Working Group
Request for Comments: 4125
Category: Experimental

F. Le Faucheur
Cisco Systems, Inc.
W. Lai
AT&T Labs

June 2005

Maximum Allocation Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering

Status of This Memo

Copyright Notice

Abstract

This document provides specifications for one Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering, which is referred to as the Maximum Allocation Model.

Table of Contents

1. Introduction

[DSTE-REQ] presents the Service Providers requirements for support of Diffserv-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 Constraints Model for DS-TE and states that "The DS-TE technical solution MUST specify at least one Bandwidth Constraints Model and MAY specify multiple Bandwidth Constraints Models."

This document provides a detailed description of one particular Bandwidth Constraints 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.

1.1. 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].

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.

A number of recovery mechanisms, under investigation or specification in the IETF, take advantage of the concept of bandwidth sharing across particular sets of LSPs. "Shared Mesh Restoration" in [GMPLS-RECOV] and "Facility-based Computation Model" in [MPLS-BACKUP] are example mechanisms that increase bandwidth efficiency by sharing bandwidth across backup LSPs protecting against independent failures. To ensure that the notion of "Reserved (CTc)" introduced in [DSTE-REQ] is compatible with such a concept of bandwidth sharing across multiple LSPs, the wording of the "Reserved (CTc)" definition provided in [DSTE-REQ] is generalized into the following:

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

With this generalization, the Maximum Allocation Model definition provided in this document is compatible with Shared Mesh Restoration defined in [GMPLS-RECOV], so that DS-TE and Shared Mesh Protection can operate simultaneously. This assumes that Shared Mesh Restoration operates independently within each DS-TE Class-Type and does not operate across Class-Types (for example, backup LSPs protecting Primary LSPs of CTx also need to belong to CTx; Excess Traffic LSPs sharing bandwidth with Backup LSPs of CTx also need to belong to CTx).

We also introduce the following definition:

Reserved(CTb,q): Let us define "Reserved(CTb,q)" as the total amount of the bandwidth reserved by all the established LSPs that 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 cannot be any established LSPs that belongs to CTb and has a holding priority of q; therefore, in this case, Reserved(CTb,q) = 0.

3. Maximum Allocation Model Definition

MAM is defined in the following manner:

- o Maximum Number of Bandwidth Constraints (MaxBC) =
  Maximum Number of Class-Types (MaxCT) = 8

- o for each value of c in the range 0 <= c <= (MaxCT - 1):
  Reserved (CTc) <= BCc <= Max-Reservable-Bandwidth,

o SUM (Reserved(CTc)) <= Max-Reservable-Bandwidth
  where the SUM is across all values of c in the range
  0 <= c <= (MaxCT - 1)

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

To increase the degree of bandwidth sharing among the different CTs, the sum of Bandwidth Constraints may exceed the Maximum Reservable Bandwidth, so that the following relationship may hold true:

o SUM (BCc) > Max-Reservable-Bandwidth,
  where the SUM is across all values of c in the range
  0 <= c <= (MaxCT - 1)

The sum of Bandwidth Constraints may also be equal to (or below) the Maximum Reservable Bandwidth. In that case, the Maximum Reservable Bandwidth does not actually constrain CT bandwidth reservations (in other words, the 3rd bullet item of the MAM definition above will never effectively come into play). This is because the 2nd bullet item of the MAM definition above implies that:

SUM (reserved(CTc)) <= SUM (BCc)

and we assume here that

SUM (BCc) <= Maximum Reservable Bandwidth.

Therefore, it will always be true that:

SUM (Reserved(CTc)) <= Max-Reservable-Bandwidth.

Both preemption within a CT and across CTs is allowed.

Where 8 CTs 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

- All LSPs from all CTs collectively use no more than the Maximum Reservable Bandwidth

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

```
        I----------------------------I
        <---BC0--->                  I
        I---------I                  I
        I         I                  I
        I   CT0   I                  I
        I         I                  I
        I---------I                  I
        I                            I
        I                            I
        <-------BC1------->          I
        I-----------------I          I
        I                 I          I
        I       CT1       I          I
        I                 I          I
        I-----------------I          I
        I                            I
        I                            I
        <-----BC2----->              I
        I-------------I              I
        I             I              I
        I     CT2     I              I
        I             I              I
        I-------------I              I
        I                            I
        I        CT0+CT1+CT2         I
        I                            I
        I----------------------------I

        <--Max Reservable Bandwidth-->
```

(Note that, in this illustration, the sum BC0 + BC1 + BC2 exceeds the Max Reservable Bandwidth.)

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

- Network administrators generally find MAM simple and intuitive

- MAM matches simple bandwidth control policies that Network Administrators may want to enforce, such as setting an individual Bandwidth Constraint for a given type of traffic (a.k.a. Class-Type) and simultaneously limit the aggregate of reserved bandwidth across all types of traffic.

- MAM can be used in a way which ensures isolation across Class-Types, 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, and its comparison with RDM, can be found in [BC-CONS] and [BC-MODEL].

As a very simple example of usage of MAM, 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 = 2 Gb/s (i.e., Data is limited to 2 Gb/s)

- BC1 = 1 Gb/s (i.e., Voice is limited to 1 Gb/s)

- Maximum Reservable Bandwidth = 2.5 Gb/s (i.e., aggregate Data + Voice is limited to 2.5 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 when computing "Unreserved TE-Class [i]".

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

Then:

```
  "Unreserved TE-Class [i]" =

   MIN  [
  [ BCc - SUM ( Reserved(CTc,q) ) ] for q <= p  ,
  [ Max-Res-Bw - SUM (Reserved(CTb,q)) ] for q <= p and 0 <= b <= 7,
         ]

  where:
       TE-Class [i] <--> < CTc , preemption p>
       in the configured TE-Class mapping.
```

5. 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.

6. IANA Considerations

[DSTE-PROTO] defines a new name space for "Bandwidth Constraints Model Id". The guidelines for allocation of values in that name space are detailed in section 13.1 of [DSTE-PROTO]. In accordance with these guidelines, IANA has assigned a Bandwidth Constraints Model Id for MAM from the range 0-239 (which is to be managed as per the "Specification Required" policy defined in [IANA-CONS]).

Bandwidth Constraints Model Id 1 was allocated by IANA to MAM.

7. Acknowledgements

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 and Dimitry Haskin.

Appendix A: Addressing [DSTE-REQ] Scenarios

This Appendix provides examples of how the Maximum Allocation Bandwidth Constraints Model can be used to support each of the scenarios described in [DSTE-REQ].

A.1. Scenario 1: Limiting Amount of Voice

By configuring on every link:

- Bandwidth Constraint 1 (for CT1 = Voice) = "certain percentage" of link capacity

- Bandwidth Constraint 0 (for CT0 = Data) = link capacity (or a constraint specific to data traffic)

- Max Reservable Bandwidth = link capacity

By configuring:

- every CT1/Voice TE-LSP with preemption = 0

- every CT0/Data TE-LSP with preemption = 1

DS-TE with the Maximum Allocation Model will address all the requirements:

- amount of Voice traffic limited to desired percentage on every link

- data traffic capable of using all remaining link capacity (or up to its own specific constraint)

- voice traffic capable of preempting other traffic

A.2. Scenario 2: Maintain Relative Proportion of Traffic Classes

By configuring on every link:

- BC2 (for CT2) = e.g., 45% of link capacity

- BC1 (for CT1) = e.g., 35% of link capacity

- BC0 (for CT0) = e.g., 100% of link capacity

- Max Reservable Bandwidth = link capacity

DS-TE with the Maximum Allocation Model will ensure that the amount of traffic of each CT established on a link is within acceptable levels as compared to the resources allocated to the corresponding Diffserv Per Hop Behaviors (PHBs) regardless of which order the LSPs are routed in, regardless of which preemption priorities are used by which LSPs and regardless of failure situations.

By also configuring:

- every CT2/Voice TE-LSP with preemption = 0

- every CT1/Premium Data TE-LSP with preemption = 1

- every CT0/Best-Effort TE-LSP with preemption = 2

DS-TE with the Maximum Allocation Model will also ensure that:

- CT2 Voice LSPs always have first preemption priority in order to use the CT2 capacity

- CT1 Premium Data LSPs always have second preemption priority in order to use the CT1 capacity

- Best-Effort can use up to link capacity of what is left by CT2 and CT1.

Optional automatic adjustment of Diffserv scheduling configuration could be used for maintaining very strict relationships between the amounts of established traffic of each CT and corresponding Diffserv resources.

A.3. Scenario 3: Guaranteed Bandwidth Services

By configuring on every link:

- BC1 (for CT1) = "given" percentage of link bandwidth (appropriate to achieve the QoS objectives of the Guaranteed Bandwidth service)

- BC0 (for CT0 = Data) = link capacity (or a constraint specific to data traffic)

- Max Reservable Bandwidth = link capacity

DS-TE with the Maximum Allocation Model will ensure that the amount of Guaranteed Bandwidth Traffic established on every link remains below the given percentage so that it will always meet its QoS objectives. At the same time, it will allow traffic engineering of the rest of the traffic such that links can be filled up (or limited to the specific constraint for such traffic).

Normative References

[DSTE-REQ] Le Faucheur, F. and W. Lai, "Requirements for Support of Differentiated Services-aware MPLS Traffic Engineering", RFC 3564, July 2003.

[DSTE-PROTO] Le Faucheur, F., Ed., "Protocol Extensions for Support of Diffserv-aware MPLS Traffic Engineering", RFC 4124, June 2005.

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

[IANA-CONS] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 2434, October 1998.

Informative References

[BC-CONS] Le Faucheur, F., "Considerations on Bandwidth Constraints Model for DS-TE", Work in Progress, June 2002.

[BC-MODEL] Lai, W., "Bandwidth Constraints Models for Differentiated Services (Diffserv)-aware MPLS Traffic Engineering: Performance Evaluation", RFC 4128, June 2005.

[DSTE-RDM] Le Faucheur, F., Ed., "Russian Dolls Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering", RFC 4127, June 2005.

[GMPLS-RECOV] Lang, et al., "Generalized MPLS Recovery Functional Specification", Work in Progress.

[MPLS-BACKUP] Vasseur, et al., "MPLS Traffic Engineering Fast reroute: Bypass Tunnel Path Computation for Bandwidth Protection", Work in Progress.

Authors' Addresses

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

Wai Sum Lai
AT&T Labs
200 Laurel Avenue
Middletown, New Jersey 07748, USA

Phone: (732) 420-3712
EMail: wlai@att.com

Intellectual Property

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.

Acknowledgement

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