Maximum Allocation Bandwidth Constraints Model for
Diffserv-TE & Performance Comparisons

Status of this Memo

   This document is an Internet-Draft and is in full conformance with
   all provisions of Section 10 of RFC2026.

   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 document is available in both .txt and .pdf formats.

Abstract

   This document is intended to complement the Diffserv-aware MPLS TE
   Requirements document by giving a functional specification for the
   Maximum Allocation bandwidth constraint model. We also provide a
   performance comparison of the Maximum Allocation and the Russian
   Dolls models to provide guidance to user implementation of the
   models in their networks.

Conventions used in this document

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in
   this document are to be interpreted as described in RFC-2119.

Table of Contents

1. **Introduction**

   Work is currently ongoing in the Traffic Engineering Working Group
   to provide the capability for Diffserv-aware MPLS traffic
   engineering (DS-TE) [1, 2]. A major item is the specification of
   bandwidth constraints models for use with DS-TE. This document is
   intended to complement the Requirements document [1] by describing
   the implications of some of the criteria for selecting a model for
   use in a network implementation. Related documents in this area
   include [3, 4, 5, 6].

   The following selection criteria are currently listed in the
   Requirements document:

   (1) addresses the scenarios in Section 2 (of [1])
   (2) works well under both normal and overload conditions
   (3) applies equally when preemption is either enabled or disabled
   (4) minimizes signaling load processing requirements
   (5) maximizes efficient use of the network
   (6) minimizes implementation and deployment complexity

   Also, two bandwidth constraints models are described in the
   Requirements document:

   (1) Maximum Allocation model (MAM) - the maximum allowable bandwidth
   usage of each class is being explicitly specified
   (2) Russian Dolls model (RDM) - specification of maximum allowable
   usage is being done cumulatively by grouping successive priority
   classes

   The use of any given bandwidth constraints model has significant
   impacts on the performance of a network, as to be explained later.
   Therefore, the criteria used to select a model must enable us to

evaluate how a particular model delivers its performance, relative
to other models. This version of the present document deals with
criteria (2), (3), and (5). Criteria (4) and (6) are to be included

in the next version. Criterion (1) relates mainly to the
Requirements document and will not be further discussed.

2. **Definitions of Bandwidth Constraints Models**

The Requirements document defines the concepts of Class Type and
Reserved Bandwidth as follows.

Class-Type (CT) is 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.

Up to 8 CTs (MaxCT = 8) are supported. They are referred to as CTc,
$0 <= c <= MaxCT-1 = 7$.

Each CT is assigned either a Bandwidth Constraint, or a set of
Bandwidth Constraints. Up to 8 Bandwidth Constraints (MaxBC = 8)
are supported and they are referred to as BCb, $0 <= b <= MaxBC-1 =
7$.

For a given Class-Type CTc, its Reserved Bandwidth "Reserved(CTc)"
is defined as the sum of the bandwidth reserved by all established
label switched paths (LSPs) which belong to CTc.

The Requirements document also describes the concept of overbooking.
This aspect has significant impact on performance and will be
further discussed in later sections of this document.

 3.0 Functional Specification of Maximum Allocation Model (MAM)

MAM is defined in [1] as a model with one separate Bandwidth
Constraint per CT:
     - MaxBC = MaxCT = 8
     - for each value of b in the range $0 <= b <= (MaxCT - 1)$:
             Reserved (CTb) <= BCb

For illustration purposes, on a link of 100 units of bandwidth where
three CTs are used with no overbooking, a network administrator
might configure BC0 = 30, BC1 = 50, and BC2 = 20 such that:
-     All LSPs supporting Traffic Trunks from CT0 use no more than 30
   (e.g. Voice <= 30)
-     All LSPs supporting Traffic Trunks from CT1 use no more than 50
   (e.g. Premium Data <= 50)
-     All LSPs supporting Traffic Trunks from CT2 use no more than 20
   (e.g. Best Effort <= 20)

ANNEX A û Performance Comparisons of MAM bandwidth constraint model &
Russian Dolls bandwidth constraint model

In this annex we define the Russian Dolls bandwidth constraint model
and provide a performance comparison to MAM. This will provide
guidance to user implementation of the models in their networks.

A.1 Russian Dolls Model (RDM)

RDM is defined in [1] as follows:
- MaxBC = MaxCT = 8
- for each value of b in the range 0 <= b <= (MaxCT - 1):
        SUM (Reserved (CTc)) <= BCb,
        for all "c" in the range b <= c <= (MaxCT - 1)

For illustration purposes, on a link of 100 units of bandwidth where
three CTs are used with no overbooking, a network administrator
might configure BC0 = 100, BC1 = 80, BC2 = 60 such that:
-    All LSPs supporting Traffic Trunks from CT2 use no more than 60
  (e.g. Voice <= 60)
-    All LSPs supporting Traffic Trunks from CT1 or CT2 use no more
  than 80 (e.g. Voice + Premium Data <= 80)
-    All LSPs supporting Traffic Trunks from CT0 or CT1 or CT2 use no
  more than 100 (e.g. Voice + Premium Data + Best Effort <= 100).

A.2 Other Bandwidth Constraints Models

Currently, the Maximum Allocation with Reservation model [6] is
under consideration for use as an another candidate bandwidth
constraint model. However, this model is not further discussed
here.

A.3. Performance Under Normal Load

To understand the implications of using criteria (2), (3), and (5)
to select a bandwidth constraint model, we first present some
numerical results of our analysis [7]. This is to gain some insight
to facilitate the discussion of the issues that can arise.

To simplify our presentation, we use the informal name "class of
traffic" for Class-Type and assume that (1) there are only three
classes of traffic, and (2) all LSPs, regardless of class, require
the same amount of bandwidth. Furthermore, the focus is on the
bandwidth usage of an individual link with a given capacity; routing
aspects of LSP setup are not considered.

Let the three classes of traffic be denoted as class 1 (highest
priority), class 2, and class 3 (lowest priority). Preemption is
enabled so that, when necessary, class 1 can preempt class 3 or
class 2 (in that order), and class 2 can preempt class 3. Each
class offers a load of traffic to the network that is expressed in

terms of the arrival rate of its LSP requests and the average
lifetime of an LSP. A unit of such a load is an erlang.

As an example, consider a link with a capacity that allows a maximum
of 15 LSPs from different classes to be established simultaneously.
Overbooking is allowed, as is to be described below. All LSPs are
assumed to have an average lifetime of 1 time unit. Suppose that
this link is being offered a load of
2.7 erlangs from class 1,
3.5 erlangs from class 2, and
3.5 erlangs from class 3.

For the explicit maximum allocation model, we assume that the
bandwidth constraints are:
up to 6 simultaneous LSPs for class 1,
up to 7 simultaneous LSPs for class 2, and
up to 15 simultaneous LSPs for class 3.

For the Russian Dolls model, we assume that the bandwidth
constraints are:
up to 6 simultaneous LSPs for class 1 by itself,
up to 11 simultaneous LSPs for classes 1 and 2 together, and
up to 15 simultaneous LSPs for all three classes together.

Obviously, these should not be regarded as typical values used by
any Internet service provider. They are used here mainly for
illustrative purposes. The method we used for analysis can easily
accommodate another set of parameter values as input.

In the example here, the values of these parameters are chosen so
that, under normal conditions, the performance of the two models is
similar in terms of their blocking and preemption behavior for LSP
setup requests. Specifically, the following table shows their
relative performance.

Table 1. Blocking and preemption probabilities

Model	PB1	PB2	PB3	PP2	PP3	PB2+PP2	PB3+PP3
MaxAll	0.03692	0.03961	0.02384	0	0.02275	0.03961	0.04659
RussDoll	0.03692	0.02296	0.02402	0.01578	0.01611	0.03874	0.04013

In the above table,

PB1 = blocking probability of class 1
PB2 = blocking probability of class 2
PB3 = blocking probability of class 3

PP2 = preemption probability of class 2
PP3 = preemption probability of class 3

```
   PB2+PP2 = combined blocking/preemption probability of class 2
   PB3+PP3 = combined blocking/preemption probability of class 3
```

   From column 2 of the above table, it can be seen that class 1 sees
   the same blocking under both models. This should be obvious since
   both allocate up to 6 simultaneous LSPs for use by class 1 only.

Slightly better results are obtained from the Russian Dolls model,
as shown by the last two columns in Table 1. This comes about
because the cascaded bandwidth separation in the Russian Dolls
design effectively gives class 3 some form of protection from being
preempted by higher priority classes.

Also, note that PP2 is zero in this particular case, simply because
the parameters for the explicit maximum allocation algorithm happen
to have been chosen in such a way that Class 1 never has to preempt
Class 2 for any of the bandwidth that Class 1 needs. (This is
because Class 1 can, in the worst case, get all the bandwidth it
needs simply by pre-empting Class 3 alone.) In general, this will
not be the case.

It is interesting to compare these results with that for the case of
a single class. Based on the Erlang loss formula, a capacity of 15
servers can support an offered load of 10 erlangs with a blocking
probability of 0.0364969. Whereas the total load for the 3-class
model is less with 2.7 + 3.5 + 3.5 = 9.7 erlangs, the probabilities
of blocking/preemption are higher. Thus, there is some loss of
efficiency due to the link bandwidth being partitioned to
accommodate for different traffic classes, thereby resulting in less
sharing.

A.4. Performance Under Overload

To investigate the performance under overload conditions, the load
of each class in the above example is varied separately. Figures 1
and 2 show their relative performance. The three series of data in
each of these figures are, respectively, class 1 blocking
probability ("Class 1 B"), class 2 blocking/preemption probability
("Class 2 B+P"), and class 3 blocking/preemption probability ("Class
3 B+P"). For each of these series, the first set of four points is
for the performance when class 1 load is increased from half of its
normal load to twice its normal. Similarly, the next and the last
sets of four points are when class 2 and class 3 loads are
correspondingly increased.

Here is something common to both algorithms:

   1.   The performance of any class generally degrades as its load
       increases.
   2.   The performance of class 1 is not affected by any changes
       (increases or decreases) in either class 2 or class 3 traffic,
       because class 1 can always preempt others.
   3.   Similarly, the performance of class 2 is not affected by any
       changes in class 3 traffic.

4.   Class 3 sees better (worse) than normal performance when either
   class 1 or class 2 traffic is below (above) normal.

In contrast, the impact of the changes in class 1 traffic on class 2
performance is different for the two algorithms: being negligible in
one case and significant in the other.

1.   While class 2 sees little improvement in performance when class
    1 traffic is below normal when the explicit maximum allocation
    algorithm is used, it sees better than normal performance under
    the Russian Dolls algorithm.

2.   Class 2 sees no degradation in performance when class 1 traffic
    is above normal when the explicit maximum allocation algorithm
    is used. In this example, with bandwidth constraints 6 + 7 <
    15, class 1 and class 2 traffic are effectively being served by
    separate pools. Therefore, class 2 sees no preemption, and
    only class 3 is being preempted whenever necessary. This fact
    is confirmed by the Erlang loss formula: a load of 2.7 erlangs
    offered to 6 servers sees a 0.03692 blocking, a load of 3.5
    erlangs offered to 7 servers sees a 0.03961 blocking. These
    blocking probabilities are exactly the same as the
    corresponding entries in Table 1: PB1 and PB2 for MaxAll.

3.   This is not the case in the Russian Dolls algorithm. Here, the
    probability for class 2 to be preempted by class 1 is nonzero
    because of two effects. (1) Through the cascaded bandwidth
    arrangement, class 3 is protected somewhat from preemption.
    (2) Class 1 and class 2 traffic are sharing their bandwidth
    allocations to some extent. Consequently, class 2 suffers when
    class 1 traffic increases.

Thus, it appears that while the cascaded bandwidth arrangement and
the resulting bandwidth sharing makes the Russian Dolls algorithm
works better under normal conditions, such interaction makes it less
effective to provide service isolation under overload conditions.

A.5. **Performance Under Complete Sharing**

As observed towards the end of Section 2, the partitioning of
bandwidth capacity for access by different traffic classes tends to
reduce the maximum link efficiency achievable. We now consider the
case where there is no such partitioning, thereby resulting in
complete sharing of the total bandwidth among all the classes.

For the explicit maximum allocation model, this means that the
constraints are such that up to 15 simultaneous LSPs are allowed for
any class.

Similarly, for the Russian Dolls model, the constraints are
up to 15 simultaneous LSPs for class 1 by itself,
up to 15 simultaneous LSPs for classes 1 and 2 together, and

up to 15 simultaneous LSPs for all three classes together.

Effectively, there is now no distinction between the two models.
Figure 3 shows the performance when all classes have equal access to
link bandwidth under the complete sharing scheme.

With preemption being enabled, it can be seen that class 1 virtually sees no blocking, regardless of the loading conditions of the link. Since class 2 can only preempt class 3, class 2 sees some blocking and/or preemption when either class 1 load or its own load is above normal; otherwise, class 2 is unaffected by increases of class 3 load. As higher priority classes always preempt class 3 when the link is full, class 3 suffers the most with high blocking/preemption when there is any load increase from any class. A comparison of Figures 1, 2, and 3 shows that, while the performance of both classes 1 and 2 is far superior under complete sharing, class 3 performance is much better off under either the explicit maximum allocation or Russian Dolls models. In a sense, class 3 is starved under overload as no protection of its service is being provided under complete sharing.

A.6. Performance Under Pure Blocking

This section is to cover the case when preemption is disabled. It will be discussed in the next version of this document.

A.7. Implications on Selection Criteria

Based on the previous results, a general theme is shown to be the trade-off between bandwidth sharing and service protection/isolation. To show this more concretely, let us compare the different models in terms of the *overall loss probability*. This quantity is defined as the long-term proportion of LSP requests from all classes combined that are lost as a result of either blocking or preemption.

As noted from the previous sections, while the Russian Dolls model has a higher degree of sharing then explicit maximum allocation, both converge ultimately to the complete sharing model as the degree of sharing in each of them is increased. Figure 4 shows that, for a single link, the overall loss probability is the smallest under complete sharing and the largest under explicit maximum allocation, with Russian Dolls being intermediate. Expressed differently, complete sharing yields the highest link efficiency and explicit maximum allocation the lowest. As a matter of fact, the overall loss probability of complete sharing is identical to loss probability of a single class as computed by the Erlang loss formula. Yet complete sharing has the poorest service protection capability. (We want to point out that, in a network with many links and multiple-link routing paths, analysis in [6] showed that complete sharing does not necessarily lead to maximum network-wide bandwidth efficiency.)

Increasing the degree of bandwidth sharing among the different
traffic classes helps to increase link efficiency. Such increase,
however, will lead to a tighter coupling between different classes.
Under normal loading conditions, proper dimensioning of the link so

that there is adequate capacity for each class can minimize the
effect of such coupling. Under overload conditions, when there is a
scarcity of capacity, such coupling will be unavoidable and can
cause severe degradation of service to the lower priority classes.
Thus, the objective of maximizing link usage as stated in selection
criterion (5) must be exercised with care, with due consideration to
the effect of interactions among the different classes. Otherwise,
use of this criterion alone will lead to the selection of the
complete sharing scheme, as shown in Figure 4.

The intention of criterion (2) in judging the effectiveness of
different models is to evaluate how they help the network to achieve
the expected performance. This can be expressed in terms of the
blocking and/or preemption behavior as seen by different classes
under various loading conditions. For example, the relative
strength of a model can be demonstrated by examining how many times
the per-class blocking or preemption probability under overload is
worse off than the corresponding probability under normal load.

(end of ANNEX A)

4. Security Considerations

No new security considerations are raised the Bandwidth Constraints
models presented in this document, as they are the same as the DS-TE
Requirements document [1].

5. References

Normative References

1 F. Le Faucheur (Editor), W.S. Lai (Co-editor), "Requirements for
   Support of Diff-Serv-aware MPLS Traffic Engineering," Internet-
   Draft, Work in Progress.
2 F. Le Faucheur, T. Nadeau, J. Boyle, K. Kompella, W. Townsend,
   and D. Skalecki, "Protocol extensions for support of Diff-Serv-
   aware MPLS Traffic Engineering," Internet-Draft, Work in
   Progress.

Informative References

3 F. Le Faucheur, "Maximum Allocation Bandwidth Constraints Model
   for Diff-Serv-aware MPLS Traffic Engineering", Internet-Draft,
   Work in Progress.
4 F. Le Faucheur, "Russian Dolls Bandwidth Constraints Model for
   Diff-Serv-aware MPLS Traffic Engineering", Internet-Draft, Work
   in Progress.
5 F. Le Faucheur, "Considerations on Bandwidth Constraints Models

for DS-TE", Internet-Draft, Work in Progress.

6   J. Ash, "Max Allocation with Reservation BW Constraint Model for
    MPLS/DiffServ TE", Internet-Draft, Work in Progress.

    7 W.S. Lai, "Traffic Engineering for MPLS," Internet Performance
       and Control of Network Systems III Conference, SPIE Proceedings
       Vol. 4865, Boston, Massachusetts, USA, 29 July-1 August 2002.
       (URL: http://www.columbia.edu/~ffl5/waisum/bcmodel.pdf)

6.. Acknowledgments

   DS-TE has been an active area within the TEWG. Inputs from Jerry
   Ash, Jim Boyle, Francois Le Faucheur, and Vishal Sharma are much
   appreciated.

7. Author's Addresses

   Wai Sum Lai
   AT&T Labs
   Room D5-3D18
   200 Laurel Avenue
   Middletown, NJ 07748, USA
   Phone: +1 732-420-3712
   Email: wlai@att.com

Full Copyright Statement