Traffic Engineering Working Group
Internet Draft
Document: draft-wlai-tewg-bcmodel-00.txt
Expires: December 2002

Wai Sum Lai
AT&T

June 2002

Bandwidth Constraint Models for Diffserv-aware MPLS Traffic Engineering

Status of this Memo

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 describing the implications of some of the criteria for selecting a default bandwidth constraint model. Properties of candidate models are also presented to provide guidance to the corresponding Solution document for this selection.

Contributions are welcome. Please send comments to the mailing list te-wg@ops.ietf.org

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 [2].

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) [3, 4]. A major item is the specification of bandwidth constraint models for use with DS-TE. This document is intended to complement the Requirements document [3] by describing the implications of some of the criteria for selecting a default model. Properties of different models are also presented to provide guidance to the Solution document [4] for this selection.

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

(1) addresses the scenarios in Section 2 (of [3])
(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

Also, two bandwidth constraint models are described in the Requirements document:
(1) explicit maximum allocation - the maximum allowable bandwidth usage of each class is being explicitly specified
(2) Russian Doll - specification of maximum allowable usage is being done cumulatively by grouping successive priority classes

The use of any given bandwidth constraint 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) and (5) only. Criteria (3) and (4) are to be included in the next version. Criterion (1) relates mainly to the Requirements document and will not be further discussed.

2. Performance Under Normal Load

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

To simplify our presentation, we 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. 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 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 Doll model, we assume that the 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 Doll model, as shown by the last two columns in Table 1. This comes about because the cascaded bandwidth separation in the Russian Doll design effectively gives class 3 some form of protection from being preempted by higher priority classes.

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.

3. 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 none in one case and significant in the other.

1. While class 2 sees no 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 Doll 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 Doll 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 Doll algorithm works better under normal conditions, such interaction makes it less effective to provide service isolation under overload conditions.

4. 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 Doll 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 Doll models. In a sense, class 3 is starved under overload as no protection of its service is being provided under complete sharing.

5. 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 Doll 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 the overall loss probability is the smallest under complete sharing and the largest under explicit maximum allocation, with Russian Doll 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.

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.

Security Considerations

No new security considerations are raised by this document, as they are the same as the DS-TE requirements document [3].

References

1 Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC 2026, October 1996.

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

3 F. Le Faucheur, T. Nadeau, M. Tatham, T. Telkamp, D. Cooper, J. Boyle, W. Lai, L. Fang, G. Ash, P. Hicks, A. Chiu, W. Townsend, and D. Skalecki, "Requirements for Support of Diff-Serv-aware MPLS Traffic Engineering," Internet-Draft, Work in Progress, April 2002.

4 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, February 2002.

Acknowledgments

To be added.

Author's Addresses

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