                    Bandwidth Constraints Models for
          Differentiated Services-aware MPLS Traffic Engineering:
                        Performance Evaluation

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

    The Differentiated Services (Diffserv)-aware MPLS Traffic
    Engineering Requirements RFC 3564 specifies the requirements and
    selection criteria for Bandwidth Constraints Models. Two such
    models, the Maximum Allocation and the Russian Dolls, are described
    therein. This document complements RFC 3564 by presenting the
    results of a performance evaluation of these two models under
    various operational conditions: normal load, overload, preemption
    fully or partially enabled, pure blocking, or complete sharing.

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

   Differentiated Services (Diffserv)-aware MPLS Traffic Engineering
   (DS-TE) mechanisms operate on the basis of different Diffserv
   classes of traffic to improve network performance. Requirements for
   DS-TE and the associated protocol extensions are specified in
   references [1, 2], respectively.

   To achieve per-class traffic engineering, rather than on an
   aggregate basis across all classes, DS-TE enforces different
   Bandwidth Constraints (BCs) on different classes. Reference [1]
   specifies the requirements and selection criteria for Bandwidth
   Constraints Models (BCM) for the purpose of allocating bandwidth to
   individual classes.

   This document presents a performance analysis for the two BCMs
   described in [1]:

   (1) Maximum Allocation Model (MAM) - the maximum allowable bandwidth
   usage of each class, together with the aggregate usage across all
   classes, are explicitly specified.
   (2) Russian Dolls Model (RDM) - specification of maximum allowable
   usage is done cumulatively by grouping successive priority classes
   recursively.

The following criteria are also listed in [1] for investigating the
performance and trade-offs of different operational aspects of BCMs:

(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

The use of any given BCM has significant impacts on the capability
of a network to provide protection for different classes of traffic,
particularly under high load, so that performance objectives can be
met [3]. This document complements [1] by presenting the results of
a performance evaluation of the above two BCMs under various
operational conditions: normal load, overload, preemption fully or
partially enabled, pure blocking, or complete sharing. Thus, our
focus is only on the performance-oriented criteria and their
implications for a network implementation. In other words, we are
only concerned with criteria (2), (3), and (5); we will not address
criteria (1), (4), or (6).

Related documents in this area include [4, 5, 6, 7, 8].

In the rest of this document, the following acronyms are used:

BC	Bandwidth Constraints
BCM	Bandwidth Constraints Model
MAM	Maximum Allocation Model
RDM	Russian Dolls Model

2. Bandwidth Constraints Models

To simplify our presentation, we use the informal name "class of
traffic" for the terms Class-Type and TE-Class defined in [1]. We
assume that (1) there are only three classes of traffic, and (2) all
label-switched paths (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.

The concept of reserved bandwidth is also defined in [1] to account
for the possible use of overbooking. Rather than getting into these
details, we assume that each LSP is allocated 1 unit of bandwidth on
a given link after establishment. This allows us to express link
bandwidth usage simply in terms of the *number of simultaneously
established LSPs*. Link capacity can then be used as the aggregate
constraint on bandwidth usage across all classes.

Suppose that the three classes of traffic assumed above for the
purpose of this document are denoted by class 1 (highest priority),
class 2, and class 3 (lowest priority). When preemption is enabled,

these are the preemption priorities. To define a generic class of
BCMs for the purpose of our analysis in accordance with the above
assumptions, let

Nmax = link capacity, i.e., the maximum number of simultaneously
      established LSPs for all classes together,
Nc = the number of simultaneously established class c LSPs, for c =
      1, 2, and 3, respectively.

For MAM, let

Bc = maximum number of simultaneously established class c LSPs.

Then, Bc is the Bandwidth Constraint for class c, and we have

Nc <= Bc <= Nmax, for c = 1, 2, and 3,
N1 + N2 + N3 <= Nmax,
B1 + B2 + B3 >= Nmax.

For RDM, the BCs are specified as:

B1 = maximum number of simultaneously established class 1 LSPs,
B2 = maximum number of simultaneously established LSPs for classes 1
      and 2 together,
B3 = maximum number of simultaneously established LSPs for classes
      1, 2, and 3 together.

Then, we have the following relationships:

N1 <= B1,
N1 + N2 <= B2,
N1 + N2 + N3 <= B3,
B1 < B2 < B3 = Nmax.

3. Performance Model

   Reference [8] presents a 3-class Markov-chain performance model to
   analyze a general class of BCMs. The BCMs that can be analyzed
   include, besides MAM and RDM, also BCMs with privately reserved
   bandwidth that cannot be preempted by other classes.

   The Markov-chain performance model in [8] assumes Poisson arrivals
   for LSP requests with exponentially distributed lifetime. The
   Poisson assumption for LSP requests is relevant since we are not
   dealing with the arrivals of individual packet within an LSP. Also,
   LSP lifetime may exhibit heavy-tail characteristics. This effect
   should be accounted for when the performance of a particular BCM by
   itself is evaluated. As the effect would be common for all BCMs, we
   ignore it for simplicity in the comparative analysis of the relative
   performance of different BCMs. In principle, a suitably chosen
   hyperexponential distribution may be used to capture some aspects of

heavy tail. However, this will significantly increase the
complexity of the non-product-form preemption model in [8].

The model in [8] assumes the use of admission control to allocate
link bandwidth to LSPs of different classes in accordance with their
respective BCs. Thus, the model accepts as input the link capacity
and offered load from different classes. The blocking and
preemption probabilities for different classes under different BCs
are generated as output. Thus, from a service provider's
perspective, given the desired level of blocking and preemption
performance, the model can be used iteratively to determine the
corresponding set of BCs.

To understand the implications of using criteria (2), (3), and (5)
in the Introduction Section to select a BCM, we present some
numerical results of the analysis in [8]. This is to gain some
insight to facilitate the discussion of the issues that can arise.
The major performance objective is to achieve a balance between the
need for bandwidth sharing so as to gain bandwidth efficiency, and
the need for bandwidth isolation so as to protect bandwidth access
by different classes.

3.1 LSP Blocking and Preemption

As described in Section 2, the three classes of traffic used as an
example are class 1 (highest priority), class 2, and class 3 (lowest
priority). Preemption may or may not be used and we will examine
the performance of each scenario. When preemption is used, the
priorities are the preemption priorities. We consider cross-class
preemption only, with no within-class preemption. In other words,
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. (In
packet-based networks, traffic volume is usually measured by
counting the number of bytes and/or packets that are sent or
received over an interface, during a measurement period. Here we
are only concerned with bandwidth allocation and usage at the LSP
level. Hence, the erlang as a measure of resource utilization in a
link-speed independent manner is an appropriate unit for our purpose
[9].)

To prevent Diffserv QoS degradation at the packet level, the
expected number of established LSPs for a given class should be kept
in line with the average service rate that the Diffserv scheduler
can provide to that class. Because of the use of overbooking, the
actual traffic carried by a link may be higher than expected, and
hence QoS degradation may not be totally avoidable.

However, the use of admission control at the LSP level helps to
minimize QoS degradation by enforcing the BCs established for the

different classes, according to the rules of the BCM adopted. That
is, the BCs are used to determine the number of LSPs that can be
simultaneously established for different classes under various
operational conditions. By controlling the number of LSPs admitted
from different classes, this in turn ensures that the amount of
traffic submitted to the Diffserv scheduler is compatible with the
targeted packet-level QoS objectives.

The performance of a BCM can therefore be measured by how well the
given BCM handles the offered traffic, under normal or overload
conditions, while maintaining packet-level service objectives.
Thus, assuming the enforcement of Diffserv QoS objectives by
admission control as a given, the performance of a BCM can be
expressed in terms of *LSP blocking and preemption probabilities*.

Different BCMs have different strengths and weaknesses. Depending
on the BCs chosen for a given load, a BCM may perform well in one
operating region and poorly in another region. Service providers
are mainly concerned with the utility of a BCM to meet their
operational needs. Regardless of which BCM is deployed, the
foremost consideration is that the BCM works well under the
engineered load, such as the ability to deliver service-level
objectives for LSP blocking probabilities. It is also expected that
the BCM handles overload "reasonably" well. Thus, for comparison,
the common operating point we choose for each BCM is that they meet
specified performance objectives in terms of blocking/preemption
under given normal load. We then observe how their performance
varies under overload. More will be said about this aspect later in
Section 4.2.

3.2 Example Link Traffic Model

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.

We now consider a scenario whereby the blocking/preemption
performance objectives for the three classes are desired to be
comparable under normal conditions (other scenarios are covered in
later sections). To meet this service requirement under the above
given load, the BCs are selected as follows:

For MAM:
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 RDM:
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.

Note that the driver is service requirement, independent of BCM.
The above BCs are not arbitrarily picked; they are chosen to meet
specific performance objectives in terms of blocking/preemption
(detailed in the next section).

An intuitive "explanation" for the above set of BCs may be as
follows. Class 1 BC is the same (6) for both models, as class 1 is
treated the same way under either model with preemption. However,
MAM and RDM operate in fundamentally different ways and give
different treatments to classes with lower preemption priorities.
In can be seen from Section 2 that while RDM imposes a strict
ordering of the different BCs (B1 < B2 < B3) and a hard boundary (B3
= Nmax), MAM uses a soft boundary (B1+B2+B3 >= Nmax) with no
specific ordering. As to be explained in Section 4.3, this allows
RDM to have a higher degree of sharing among different classes.
Such a higher degree of coupling means that the numerical values of
the BCs can be relatively smaller when compared with those for MAM,
to meet given performance requirements under normal load.

Thus, in the above example, the RDM BCs of (6, 11, 15) may be
thought of as roughly corresponding to the MAM BCs of (6, 6+7,
6+7+15). (The intent here is just to point out that the design
parameters for the two BCMs need to be different as they operate
differently - strictly speaking, the numerical correspondence is
incorrect.) Of course, both BCMs are bounded by the same aggregate
constraint of the link capacity (15).

The BCs chosen in the above example are not intended to be regarded
as typical values used by any 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.

3.3 Performance Under Normal Load

In the example above, based on the BCs chosen, the blocking and
preemption probabilities for LSP setup requests under normal
conditions for the two BCMs are given in Table 1. Remember that the
BCs have been selected for this scenario to address the service
requirement to offer comparable blocking/preemption objectives for
the three classes.

Table 1.   Blocking and preemption probabilities

BCM	PB1	PB2	PB3	PP2	PP3	PB2+PP2	PB3+PP3
MAM	0.03692	0.03961	0.02384	0	0.02275	0.03961	0.04659
RDM	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

    First, we observe that, indeed, the values for (PB1, PB2+PP2,
    PB3+PP3) are very similar one to another. This confirms that the
    service requirement (of comparable blocking/preemption objectives
    for the three classes) has been met for both BCMs.

    Then, we observe that the (PB1, PB2+PP2, PB3+PP3) values for MAM are
    very similar to the (PB1, PB2+PP2, PB3+PP3) values for RDM. This
    indicates that, in this scenario, both BCMs offer very similar
    performance under normal load.

    From column 2 of the above table, it can be seen that class 1 sees
    exactly the same blocking under both BCMs. This should be obvious
    since both allocate up to 6 simultaneous LSPs for use by class 1
    only. Slightly better results are obtained from RDM, as shown by
    the last two columns in Table 1. This comes about because the
    cascaded bandwidth separation in RDM 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 BCs for MAM 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 preempting class 3 alone.) In general,
    this will not be the case.

    It is interesting to compare these results with those 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
    BCM 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. This aspect will be examined in more details later in the
    section on Complete Sharing.

4. Performance Under Overload

    Overload occurs when the traffic on a system is greater than the
    traffic capacity of the system. To investigate the performance
    under overload conditions, the load of each class is varied
    separately. Blocking and preemption probabilities for each case are

   not shown separately: they are added together to yield a combined
   blocking/preemption probability.

4.1 Bandwidth Sharing Versus Isolation

   Figures 1 and 2 show the relative performance when the load of each
   class in the example of Section 3.2 is varied separately. 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.

   The following observations apply to both BCMs:

   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 BCMs: being negligible in MAM
   and significant in RDM.

   1. While class 2 sees little improvement (no improvement in this
      particular example) in performance when class 1 traffic is below
      normal when MAM is used, it sees better than normal performance
      under RDM.
   2. Class 2 sees no degradation in performance when class 1 traffic is
      above normal when MAM is used. In this example, with BCs 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 MAM.
   3. This is not the case in RDM. 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 2 traffic is sharing a BC

with class 1. Consequently, class 2 suffers when class 1 traffic
increases.

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

4.2 Improving Class 2 Performance at the Expense of Class 3

We now consider a scenario in which the service requirement is to
give better blocking/preemption performance to class 2 than to class
3, while maintaining class 1 performance at the same level as in the
previous scenario. (The use of minimum deterministic guarantee for
class 3 is to be considered in the next section.) So that the
specified class 2 performance objective can be met, class 2 BC is
appropriately increased. As an example, BCs (6, 9, 15) are now used
for MAM, and (6, 13, 15) for RDM. For both BCMs, as shown in
Figures 1bis and 2bis, while class 1 performance remains unchanged,
class 2 now receives better performance, at the expense of class 3.
This is of course due to the increased access of bandwidth by class
2 over class 3. Under normal conditions, the performance of the two
BCMs is similar in terms of their blocking and preemption
probabilities for LSP setup requests, as shown in Table 2.

Table 2.   Blocking and preemption probabilities

BCM	PB1	PB2	PB3	PP2	PP3	PB2+PP2	PB3+PP3
MAM	0.03692	0.00658	0.02733	0	0.02709	0.00658	0.05441
RDM	0.03692	0.00449	0.02759	0.00272	0.02436	0.00721	0.05195

Under overload, the observations in Section 4.1 regarding the
difference in the general behavior between the two BCMs still apply,
as shown in Figures 1bis and 2bis.

Some frequently asked questions about the operation of BCMs are as
follows. For a link capacity of 15, would a class 1 BC of 6 and a
class 2 BC of 9 in MAM result in the possibility of a total lockout
for class 3? This will certainly be the case when there are 6 class
1 and 9 class 2 LSPs being simultaneously established. Such an
offered load (with 6 class 1 and 9 class 2 LSP requests) will not
cause a lockout of class 3 with RDM having a BC of 13 for classes 1
and 2 combined, but will result in class 2 LSPs being rejected. If
class 2 traffic were considered relatively more important then class
3 traffic, then RDM would perform very poorly when compared with MAM
with BCs of (6, 9, 15). Should MAM with BCs of (6, 7, 15) be used
instead so as to make the performance of RDM look comparable?

The answer is that the above scenario is not very realistic when the
offered load is assumed to be (2.7, 3.5, 3.5) for the three classes,
as stated in Section 3.2. Treating an overload of (6, 9, x) as
normal operating condition is incompatible with the engineering of
BCs according to needed bandwidth from different classes. It would

be rare for a given class to need so much more than its engineered
bandwidth level. But if the class did, the expectation based on
design and normal traffic fluctuations is that this class would
quickly release unneeded bandwidth toward its engineered level,
freeing up bandwidth for other classes.

Service providers engineer their networks based on traffic
projections to determine network configurations and needed capacity.
All BCMs should be designed to operate under realistic network
conditions. For any BCM to work properly, the selection of values
for different BCs must therefore be based on the projected bandwidth
needs of each class, as well as the bandwidth allocation rules of
the BCM itself. This is to ensure that the BCM works as expected
under the intended design conditions. In operation, the actual load
may well turn out to be different from the design. Thus, an
assessment of the performance of a BCM under overload is essential
to see how well the BCM can cope with traffic surges or network
failures. Reflecting this view, the basis for comparison of two
BCMs is that they meet the same or similar performance requirements
under normal conditions, and how they withstand overload.

In operational practice, load measurement and forecast would be
useful to calibrate and fine-tune the BCs so that traffic from
different classes could be redistributed accordingly. Dynamic
adjustment of the Diffserv scheduler could also be used to minimize
QoS degradation.

4.3 Comparing Bandwidth Constraints of Different Models

As pointed out in Section 3.2, the higher degree of sharing among
the different classes in RDM means that the numerical values of the
BCs could be relatively smaller, when compared with those for MAM.
We now examine this aspect in more details by considering the
following scenario. We set the BCs so that, (1) for both BCMs, the
same value is used for class 1, (2) the same minimum *deterministic*
guarantee of bandwidth for class 3 is offered by both BCMs, and (3)
the blocking/preemption probability is minimized for class 2. We
want to emphasize that this may not be the way service providers
select BCs. It is done here to investigate the *statistical*
behavior of such a deterministic mechanism.

For illustration, we use BCs (6, 7, 15) for MAM, and (6, 13, 15) for
RDM. In this case, both BCMs have 13 units of bandwidth for classes
1 and 2 together, and dedicate 2 units of bandwidth for use by class
3 only. The performance of the two BCMs under normal conditions is
shown in Table 3. It is clear that MAM with (6, 7, 15) gives fairly
comparable performance objectives across the three classes, while
RDM with (6, 13, 15) strongly favors class 2 at the expense of class
3. They therefore cater to different service requirements.

Table 3. Blocking and preemption probabilities

BCM	PB1	PB2	PB3	PP2	PP3	PB2+PP2	PB3+PP3

MAM	0.03692	0.03961	0.02384	0	0.02275	0.03961	0.04659
RDM	0.03692	0.00449	0.02759	0.00272	0.02436	0.00721	0.05195

By comparing Figures 1 and 2bis, it can be seen that, when being
subjected to the same set of BCs, RDM gives class 2 much better
performance than MAM, with class 3 being only slightly worse.

This confirms the observation in Section 3.2 that, when the same
service requirements under normal conditions are to be met, the
numerical values of the BCs for RDM can be relatively smaller than
those for MAM. This should not be surprising in view of the hard
boundary (B3 = Nmax) in RDM versus the soft boundary (B1+B2+B3 >=
Nmax) in MAM. The strict ordering of BCs (B1 < B2 < B3) gives RDM
the advantage of a higher degree of sharing among the different
classes, i.e., the ability to reallocate the unused bandwidth of
higher-priority classes to lower-priority ones, if needed.
Consequently, this leads to better performance when an identical set
of BCs is used as exemplified above. Such a higher degree of
sharing may necessitate the use of minimum deterministic bandwidth
guarantee to offer some protection for lower-priority traffic from
preemption. The explicit lack of ordering of BCs in MAM together
with its soft boundary implies that the use of minimum deterministic
guarantees for lower-priority classes may not need to be enforced
when there is a lesser degree of sharing. This is demonstrated by
the example in Section 4.2 with BCs (6, 9, 15) for MAM.

For illustration, Table 4 shows the performance under normal
conditions of RDM with BCs (6, 15, 15).

Table 4. Blocking and preemption probabilities

BCM	PB1	PB2	PB3	PP2	PP3	PB2+PP2	PB3+PP3
RDM	0.03692	0.00060	0.02800	0.00032	0.02740	0.00092	0.05540

Regardless of whether deterministic guarantees are used or not, both
BCMs are bounded by the same aggregate constraint of the link
capacity. Also, in both BCMs, bandwidth access guarantees are
necessarily achieved statistically because of traffic fluctuations,
as explained in Section 4.2. (As a result, service-level objectives
are typically specified as monthly averages, under the use of
statistical guarantees, rather than deterministic guarantees.)
Thus, given the fundamentally different operating principles of the
two BCMs (ordering, hard versus soft boundary), the dimensions of
one BCM should not be adopted to design for the other. Rather, it
is the service requirements, and perhaps also the operational needs,
of a service provider that should be used to drive how the BCs of a
BCM are selected.

5. Performance Under Partial Preemption

In the previous two sections, preemption is *fully enabled* in the
sense that class 1 can preempt class 3 or class 2 (in that order),

and class 2 can preempt class 3. That is, both classes 1 and 2 are
preemptor-enabled, while classes 2 and 3 are preemptable. A class
that is preemptor-enabled can preempt lower-priority classes
designated as preemptable. A class not designated as preemptable
cannot be preempted by any other classes, regardless of relative
priorities.

We now consider the three cases shown in Table 5 when preemption is
only partially enabled.

Table 5. Partial preemption modes

preemption modes	preemptor-enabled	preemptable
"1+2 on 3" (Fig. 3, 6)	class 1, class 2	class 3
"1 on 3" (Fig. 4, 7)	class 1	class 3
"1 on 2+3" (Fig. 5, 8)	class 1	class 3, class 2

In this section, we evaluate how these preemption modes affect the
performance of a particular BCM. Thus, we are comparing how a given
BCM performs when preemption is fully enabled versus how the same
BCM performs when preemption is partially enabled. The performance
of these preemption modes is shown in Figures 3 to 5 for RDM, and
Figures 6 to 8 for MAM, respectively. In all of these figures, the
BCs of Section 3.2 are used for illustration, i.e., (6, 7, 15) for
MAM and (6, 11, 15) for RDM. However, the general behavior is
similar when the BCs are changed to those in Sections 4.2 and 4.3,
i.e., (6, 9, 15) and (6, 13, 15), respectively.

5.1 Russian Dolls Model

Let us first examine the performance under RDM. There are two sets
of results, depending on whether class 2 is preemptable or not: (1)
Figures 3 and 4 for the two modes when only class 3 is preemptable,
and (2) Figure 2 in the previous section and Figure 5 for the two
modes when both classes 2 and 3 are preemptable. By comparing these
two sets of results, the following impacts can be observed.
Specifically, when class 2 is non-preemptable, and when compared
with the case of class 2 being preemptable, then the behavior of
each class is:

1. Class 1 generally sees a higher blocking probability when class 2
   is non-preemptable. As the class 1 space allocated by the class 1
   BC is shared with class 2, which is now non-preemptable, class 1
   cannot reclaim any such space occupied by class 2 when needed.
   Also, class 1 has less opportunity to preempt - being able to
   preempt class 3 only.
2. Class 3 also sees higher blocking/preemption when its own load is
   increased, as it is being preempted more frequently by class 1,
   when class 1 cannot preempt class 2. (See the last set of four
   points in the series for class 3 shown in Figures 3 and 4, when
   comparing with Figures 2 and 5.)
3. Class 2 blocking/preemption is reduced even when its own load is
   increased, since it is not being preempted by class 1. (See the

        middle set of four points in the series for class 2 shown in
     Figures 3 and 4, when comparing with Figures 2 and 5.)

   Another two sets of results are related to whether class 2 is
   preemptor-enabled or not. In this case, when class 2 is not
   preemptor-enabled, class 2 blocking/preemption is increased when
   class 3 load is increased (the last set of four points in the series
   for class 2 shown in Figures 4 and 5, when comparing with Figures 2
   and 3). This is because both classes 2 and 3 are now competing
   independently with each other for resources.

5.2 Maximum Allocation Model

   Turning now to MAM, the significant impact appears to be only on
   class 2, when it cannot preempt class 3, thereby causing its
   blocking/preemption to increase in two situations.

   1. When class 1 load is increased (the first set of four points in
      the series for class 2 shown in Figures 7 and 8, when comparing
      with Figures 1 and 6).
   2. When class 3 load is increased (the last set of four points in the
      series for class 2 shown in Figures 7 and 8, when comparing with
      Figures 1 and 6). This is similar to RDM, i.e., class 2 and class
      3 are now competing with each other.

   When comparing Figure 1 (for the case of fully enabled preemption)
   with Figures 6 to 8 (for partially enabled preemption), it can be
   seen that the performance of MAM is relatively insensitive to the
   different preemption modes. This is because when each class has its
   own bandwidth access limits, the degree of interference among the
   different classes is reduced.

   This is in contrast with RDM, whose behavior is more dependent on
   the preemption mode in use.

6. Performance Under Pure Blocking

   This section covers the case when preemption is completely disabled.
   We continue with the numerical example used in the previous sections
   with the same link capacity and offered load.

6.1 Russian Dolls Model

   For RDM, we consider two different settings:

   "Russian Dolls (1)" BCs:
   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.

   "Russian Dolls (2)" BCs:
   up to 9 simultaneous LSPs for class 3 by itself,

up to 14 simultaneous LSPs for classes 3 and 2 together, and
up to 15 simultaneous LSPs for all three classes together.

Note that the "Russian Dolls (1)" set of BCs is the same as
previously with preemption enabled, while the "Russian Dolls (2)"
has the cascade of bandwidth arranged in *reverse* order of the
classes.

As observed in Section 4, the cascaded bandwidth arrangement is
intended to offer lower priority traffic some protection from
preemption by higher priority traffic. This is to avoid starvation.
In a pure blocking environment, such protection is no longer
necessary. As depicted in Figure 9, it actually produces the
opposite, undesirable, effect: higher priority traffic sees higher
blocking than lower priority traffic. With no preemption, higher
priority traffic should be protected instead to ensure that they
could get through when under high load. Indeed, when the reverse
cascade is used in "Russian Dolls (2)," the required performance of
lower blocking for higher priority traffic is achieved as shown in
Figure 10. In this specific example, there is very little
difference among the performance of the three classes in the first
eight data points for each of the three series. However, the BCs
can be tuned to get a bigger differentiation.

6.2 Maximum Allocation Model

   For MAM, we also consider two different settings:

   "Exp. Max. Alloc. (1)" BCs:
   up to 7 simultaneous LSPs for class 1,
   up to 8 simultaneous LSPs for class 2, and
   up to 8 simultaneous LSPs for class 3.

   "Exp. Max. Alloc. (2)" BCs:
   up to 7 simultaneous LSPs for class 1, with additional bandwidth for
      1 LSP privately reserved
   up to 8 simultaneous LSPs for class 2, and
   up to 8 simultaneous LSPs for class 3.

   These BCs are chosen so that, under normal conditions, the blocking
   performance is similar to all the previous scenarios. The only
   difference between these two sets of values is that the "Exp. Max.
   Alloc. (2)" algorithm gives class 1 a private pool of 1 server for
   class protection. As a result, class 1 has a relatively lower
   blocking especially when its traffic is above normal, as can be seen
   by comparing Figures 11 and 12. This is of course at the expense of
   a slight increase in the blocking of classes 2 and 3 traffic.

   When comparing the "Russian Dolls (2)" in Figure 10 with MAM in
   Figures 11 or 12, the difference between their behavior and the
   associated explanation are again similar to the case when preemption
   is used. The higher degree of sharing in the cascaded bandwidth
   arrangement of RDM leads to a tighter coupling between the different

classes of traffic when under overload. Their performance therefore
tends to degrade together when the load of any one class is
increased. By imposing explicit maximum bandwidth usage on each
class individually, better class isolation is achieved. The trade-
off is that, generally, blocking performance in MAM is somewhat
higher than RDM, because of reduced sharing.

The difference in the behavior of RDM with or without preemption has
already been discussed at the beginning of this section. For MAM,
some notable difference can also be observed from a comparison of
Figures 1 and 11. If preemption is used, higher-priority traffic
tends to be able to maintain their performance despite the
overloading of other classes. This is not so if preemption is not
allowed. The trade-off is that, generally, the overloaded class
sees a relatively higher blocking/preemption when preemption is
enabled, than the case when preemption is disabled.

7. Performance Under Complete Sharing

As observed towards the end of Section 3, 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 full
sharing of the total bandwidth among all the classes. This is
referred to as the Complete Sharing Model.

For MAM, this means that the BCs are such that up to 15 simultaneous
LSPs are allowed for any class.

Similarly, for RDM, the BCs 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 MAM and RDM.
Figure 13 shows the performance when all classes have equal access
to link bandwidth under Complete Sharing.

With preemption being fully 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 13 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 MAM or
RDM. In a sense, class 3 is starved under overload as no protection
of its traffic is being provided under Complete Sharing.

8. Implications on Performance Criteria

   Based on the previous results, a general theme is shown to be the
   trade-off between bandwidth sharing and class protection/isolation.
   To show this more concretely, let us compare the different BCMs 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,
   for a given level of offered load.

   As noted from the previous sections, while RDM has a higher degree
   of sharing then MAM, both converge ultimately to the Complete
   Sharing Model as the degree of sharing in each of them is increased.
   Figure 14 shows that, for a single link, the overall loss
   probability is the smallest under Complete Sharing and the largest
   under MAM, with RDM being intermediate. Expressed differently,
   Complete Sharing yields the highest link efficiency and MAM 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 class 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 criterion
   (5) of Section 1 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 Model, as shown in Figure 14.

   The intention of criterion (2) in judging the effectiveness of
   different BCMs 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 BCM 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.

9. Conclusions

BCMs are used in DS-TE for path computation and admission control of
LSPs by enforcing different BCs for different classes of traffic so
that Diffserv QoS performance can be maximized. Therefore, it is of
interest to measure the performance of a BCM by the LSP
blocking/preemption probabilities under various operational
conditions. Based on this, the performance of RDM and MAM for LSP
establishment has been analyzed and compared. In particular, three
different scenarios have been examined: (1) all three classes have
comparable performance objectives in terms of LSP
blocking/preemption under normal conditions, (2) class 2 is given
better performance at the expense of class 3, and (3) class 3
receives some minimum deterministic guarantee.

A general theme is shown to be the trade-off between bandwidth
sharing to achieve greater efficiency under normal conditions, and
robust class protection/isolation under overload. The general
properties of the two BCMs are:

RDM
. allows greater sharing of bandwidth among different classes
. performs somewhat better under normal conditions
. works well when preemption is fully enabled; under partial
  preemption, not all preemption modes work equally well

MAM
. does not depend on the use of preemption
. is relatively insensitive to the different preemption modes when
  preemption is used
. provides more robust class isolation under overload

Generally, the use of preemption gives higher-priority traffic some
degree of immunity against the overloading of other classes. This
results in a higher blocking/preemption for the overloaded class,
when compared with a pure blocking environment.

10. Security Considerations

No new security considerations are raised by the BCMs presented in
this document; they are the same as in the DS-TE Requirements
document [1].

11. References

Normative References

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

Informative References

2 F. Le Faucheur (Editor), "Protocol extensions for support of
   Diff-Serv-aware MPLS Traffic Engineering," Internet-Draft, Work
   in Progress.
3 J. Boyle, V. Gill, A. Hannan, D. Cooper, D. Awduche, B.
   Christian, and W.S. Lai, "Applicability Statement for Traffic
   Engineering with MPLS," RFC 3346, July 2002.
4 F. Le Faucheur and W.S. Lai, "Maximum Allocation Bandwidth
   Constraints Model for Diff-Serv-aware MPLS Traffic Engineering,"
   Internet-Draft, Work in Progress.
5 F. Le Faucheur (Editor), "Russian Dolls Bandwidth Constraints
   Model for Diff-Serv-aware MPLS Traffic Engineering," Internet-
   Draft, Work in Progress.
6 J. Ash, "Max Allocation with Reservation Bandwidth Constraint
   Model for MPLS/DiffServ TE & Performance Comparisons," Internet-
   Draft, Work in Progress.
7 F. Le Faucheur, "Considerations on Bandwidth Constraints Models
   for DS-TE," Internet-Draft, Work in Progress.
8 W.S. Lai, "Traffic Engineering for MPLS," Internet Performance
   and Control of Network Systems III Conference, SPIE Proceedings
   Vol. 4865, Boston, Massachusetts, USA, 30-31 July 2002, pp. 256-
   267.
9 W.S. Lai, "Traffic Measurement for Dimensioning and Control of IP
   Networks," Internet Performance and Control of Network Systems II
   Conference, SPIE Proceedings Vol. 4523, Denver, Colorado, USA,
   21-22 August 2001, pp. 359-367.

12. Acknowledgments

13. Author's Address

   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

Explicit Maximum Allocation, With Preemption


Figure 1. Maximum Allocation (6, 7, 15), with full preemption.

Russian Doll, With Preemption
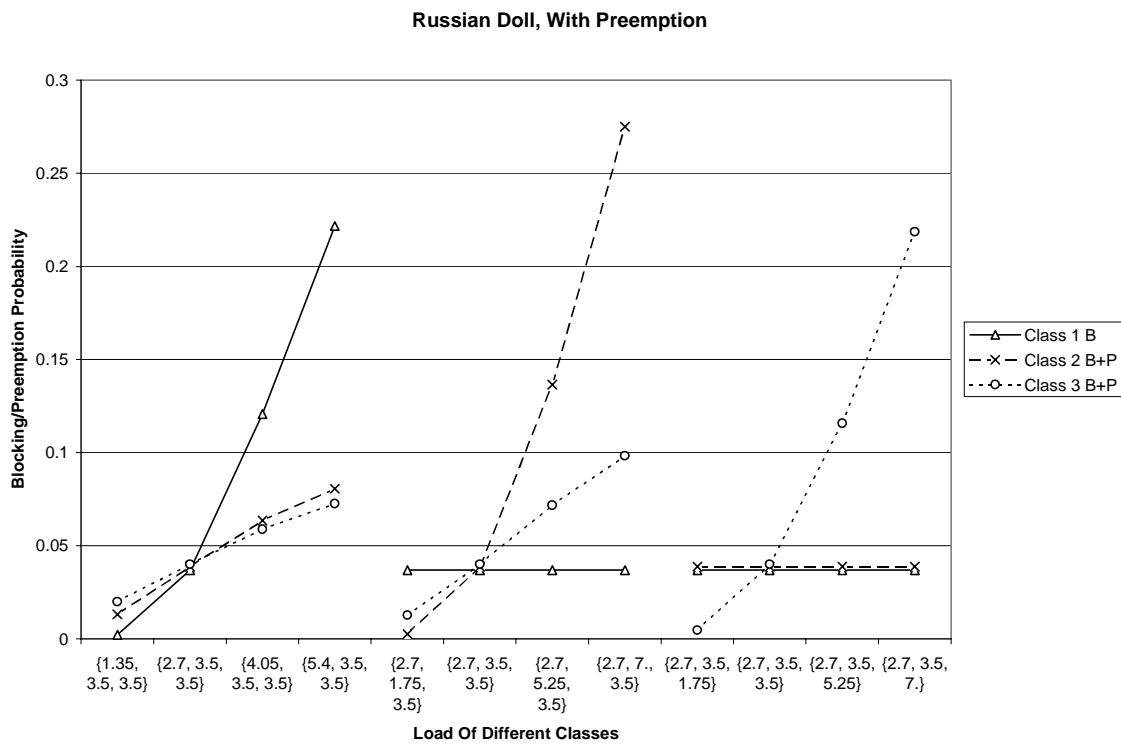

Figure 2. Russian Doll (6, 11, 15), with full preemption.

Explicit Maximum Allocation, With Preemption
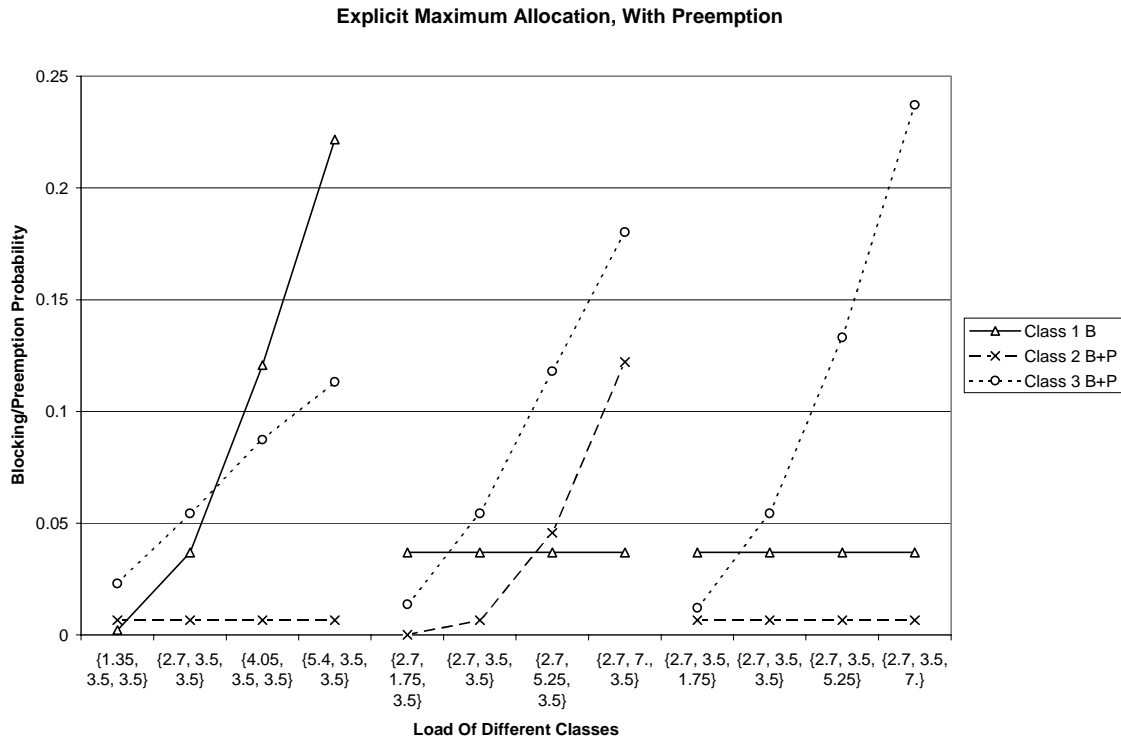
Figure 1bis. Maximum Allocation (6, 9, 15), with full preemption.

Russian Doll, With Preemption
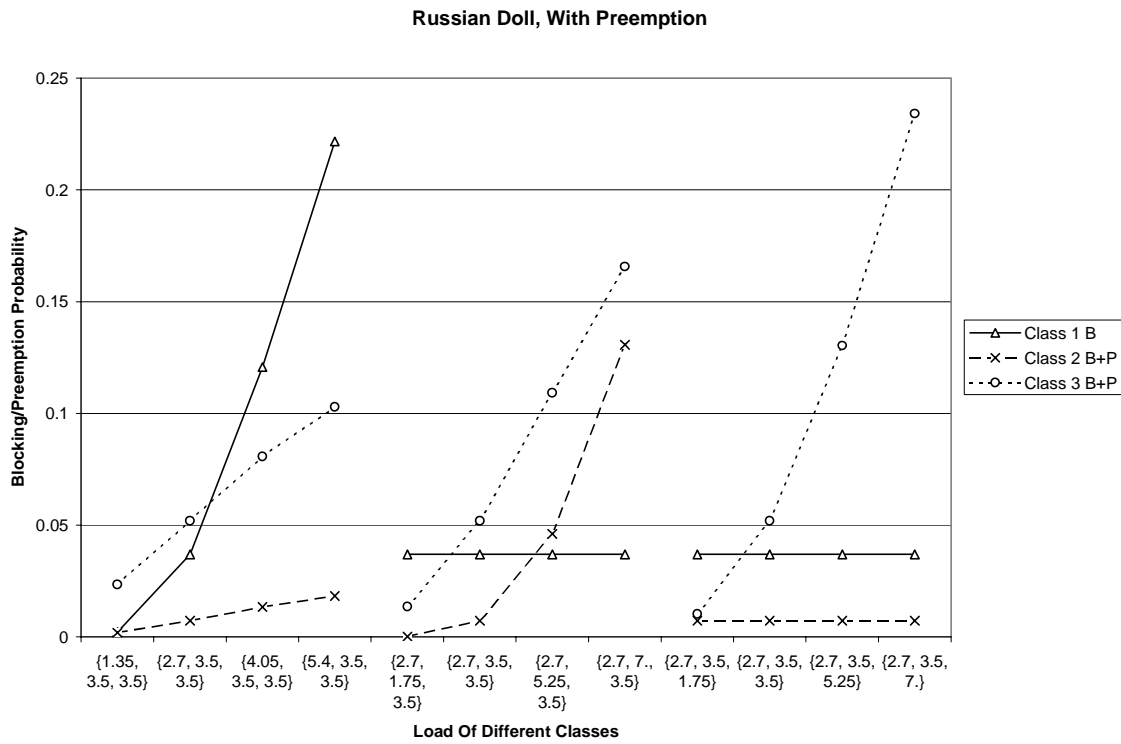
Figure 2bis. Russian Doll (6, 13, 15), with full preemption.

Russian Doll, With Partial Preemption (1+2 on 3)
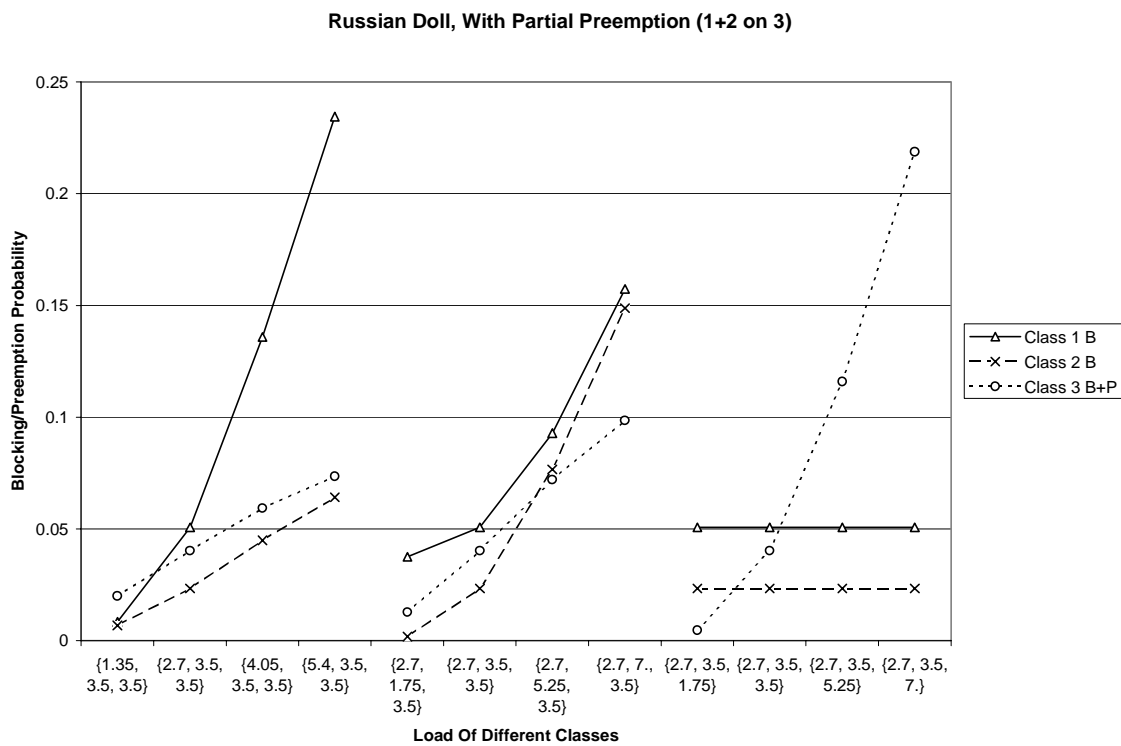
Figure 3. Russian Doll, with partial preemption (1+2 on 3).

Russian Doll, With Partial Preemption (1 on 3)
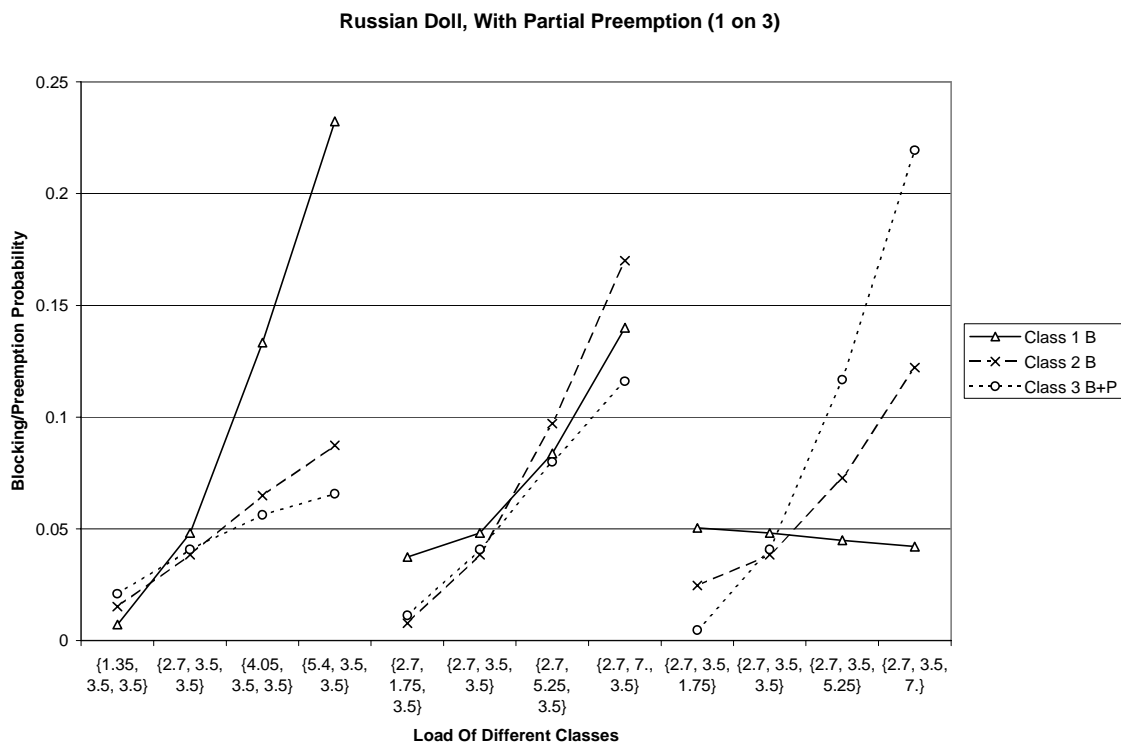
Figure 4. Russian Doll, with partial preemption (1 on 3).

Russian Doll, With Partial Preemption (1 on 2+3)
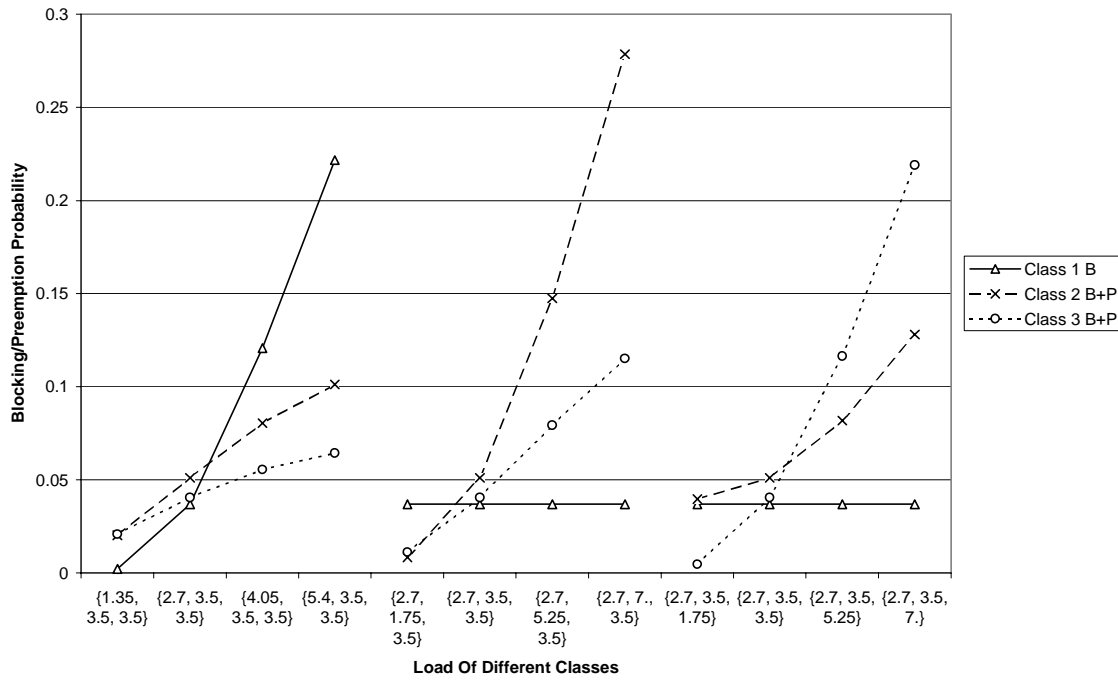
Figure 5. Russian Doll, with partial preemption (1 on 2+3).

Explicit Maximum Allocation, With Partial Preemption (1+2 on 3)
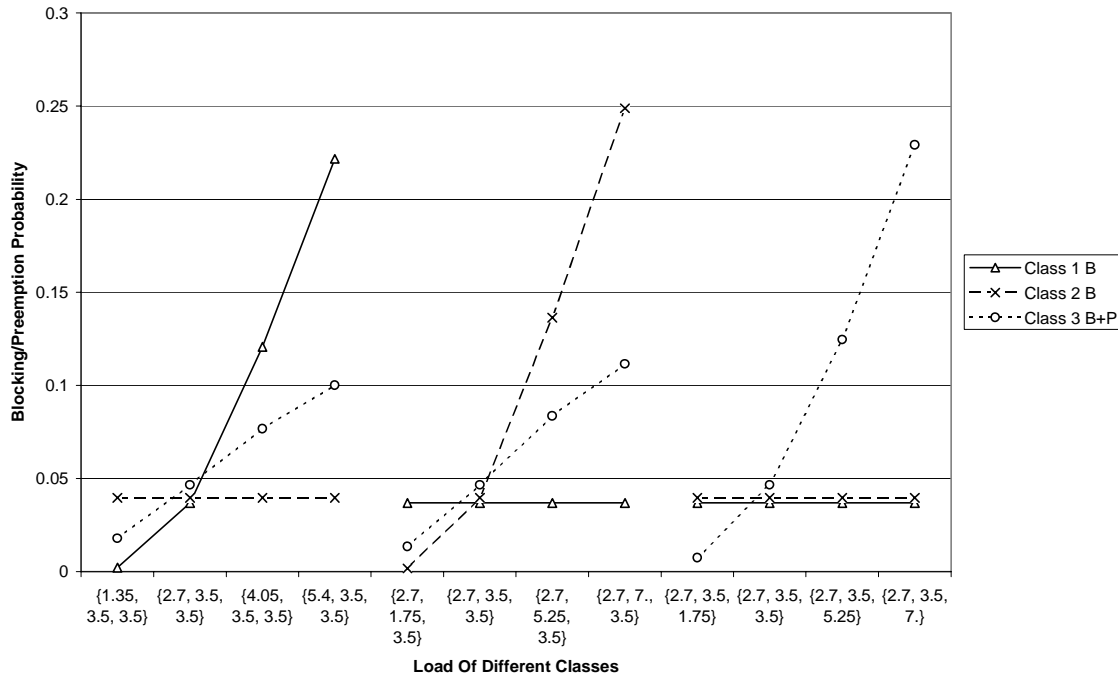
Figure 6. Maximum Allocation, with partial preemption (1+2 on 3).

Explicit Maximum Allocation, With Partial Preemption (1 on 3)
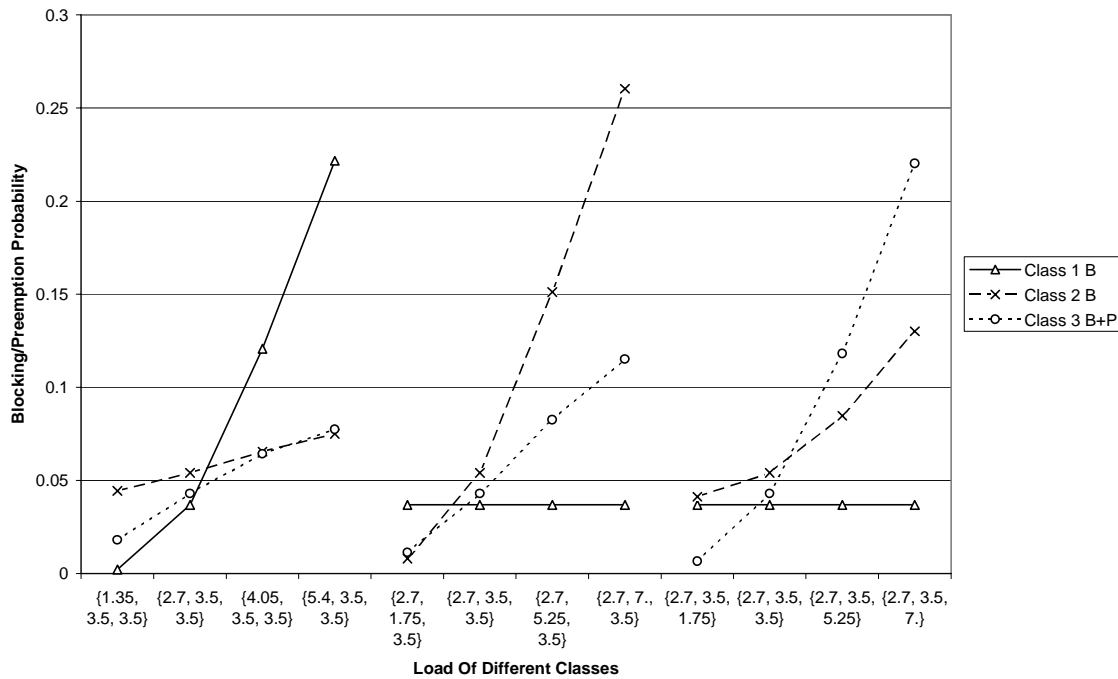
Figure 7. Maximum Allocation, with partial preemption (1 on 3).

Explicit Maximum Allocation, With Partial Preemption (1 on 2+3)
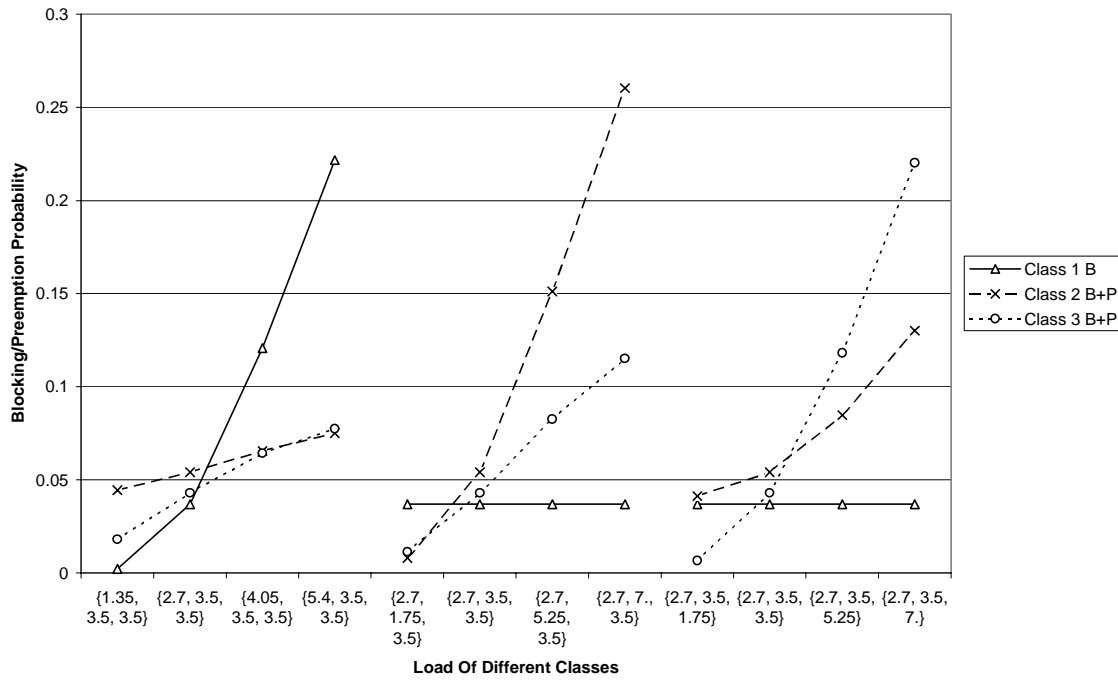
Figure 8. Maximum Allocation, with partial preemption (1 on 2+3).

Russian Doll (1), No Preemption
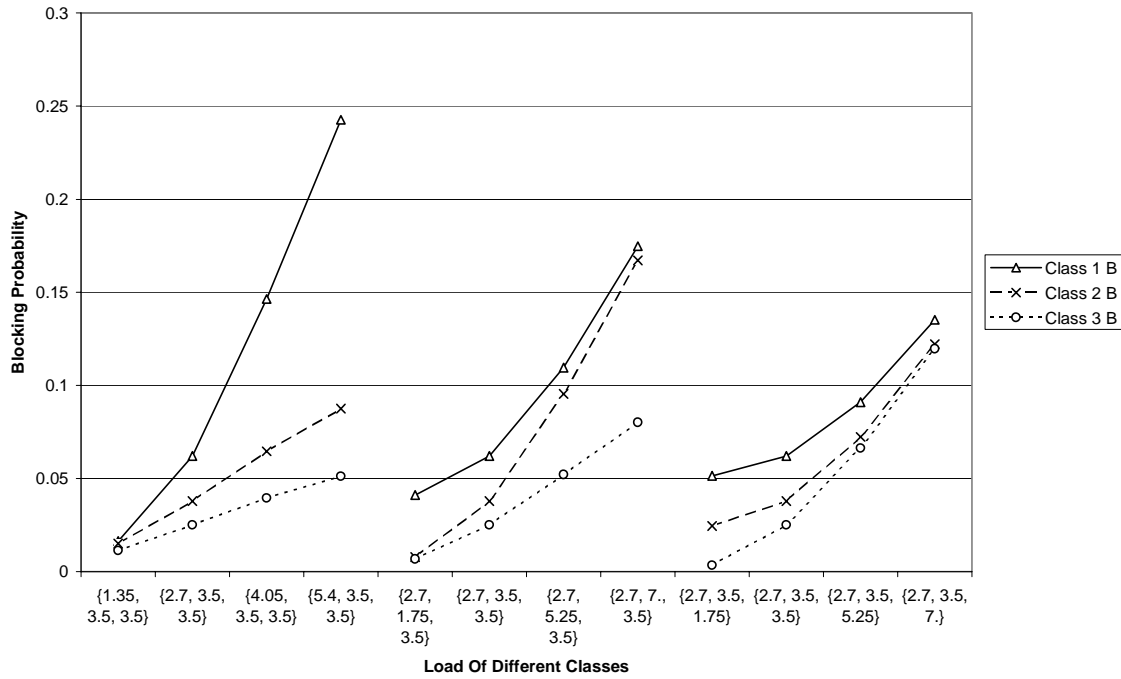
Figure 9. "Russian Doll (1)", with no preemption.

Russian Doll (2), No Preemption
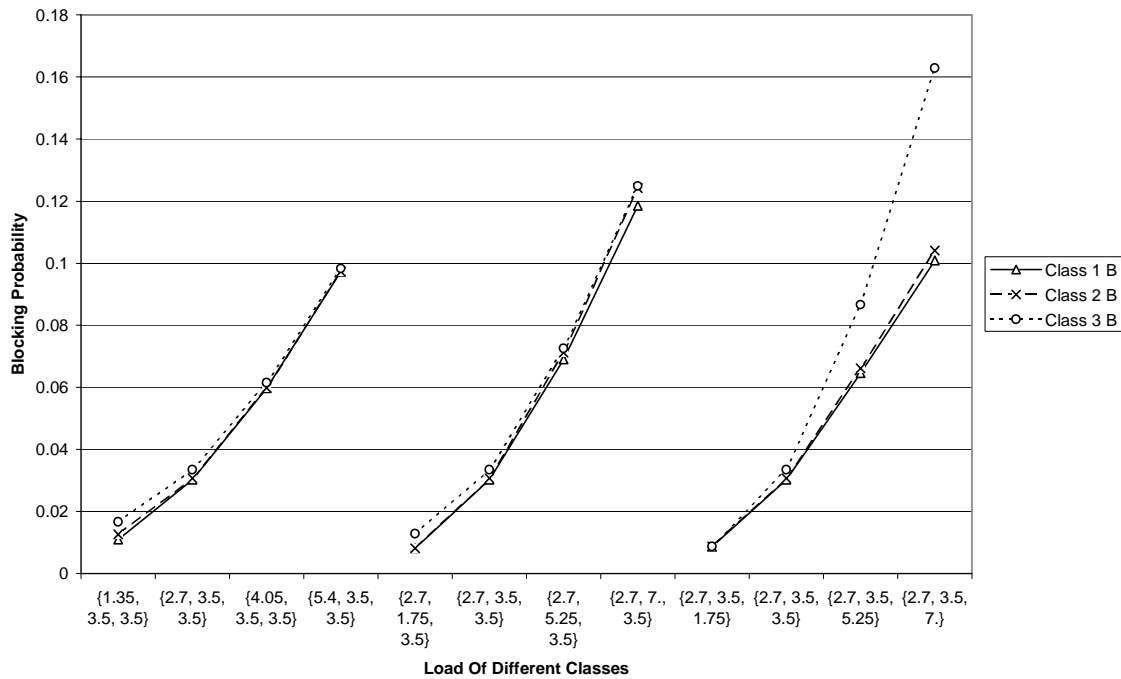
Figure 10. "Russian Doll (2)", with no preemption.

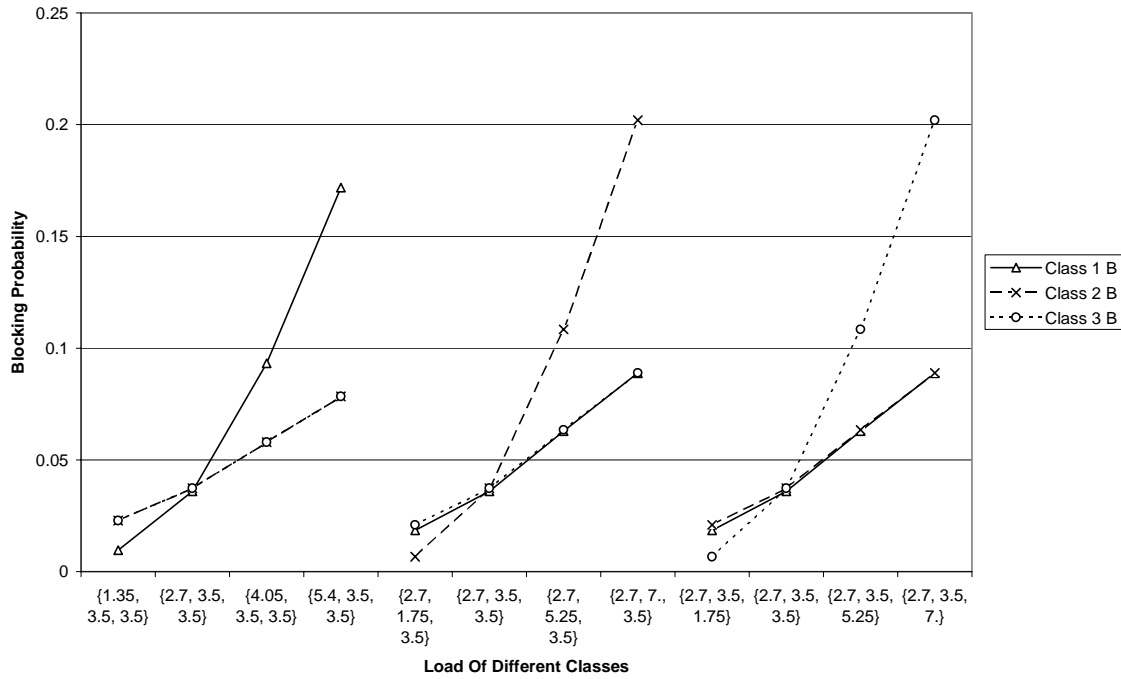**Explicit Maximum Allocation (1), No Preemption, No Min Alloc**

Figure 11. "Maximum Allocation (1)", with no preemption.

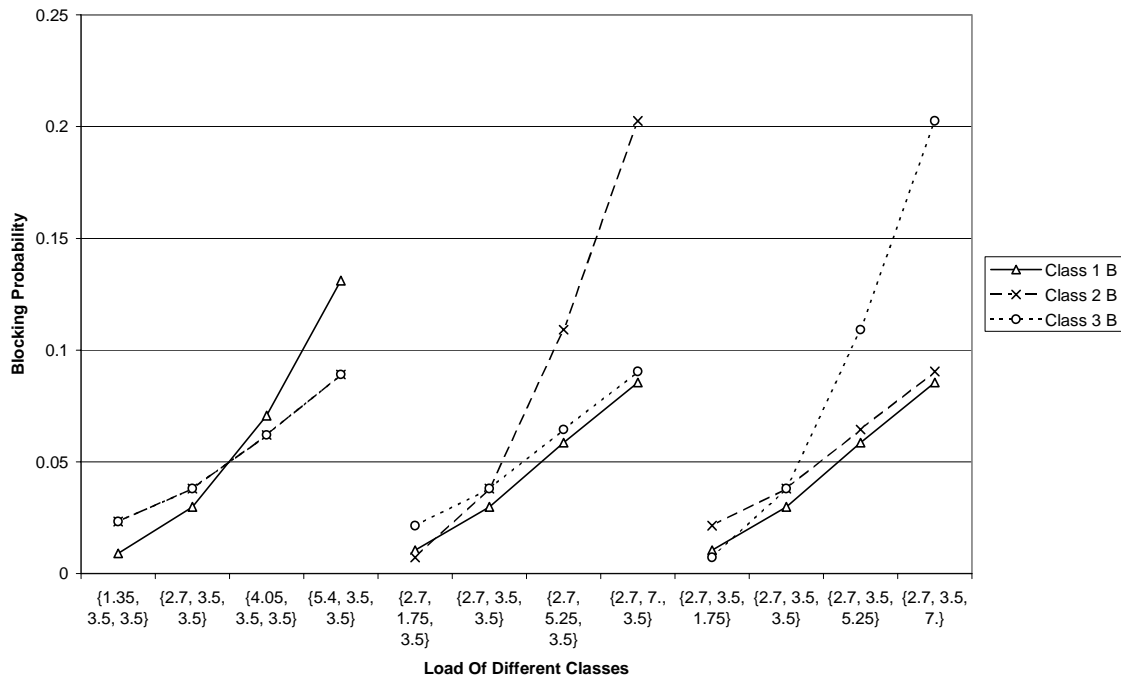**Explicit Maximum Allocation (2), No Preemption, Min Alloc For Class 1**

Figure 12. "Maximum Allocation (2)", with no preemption.
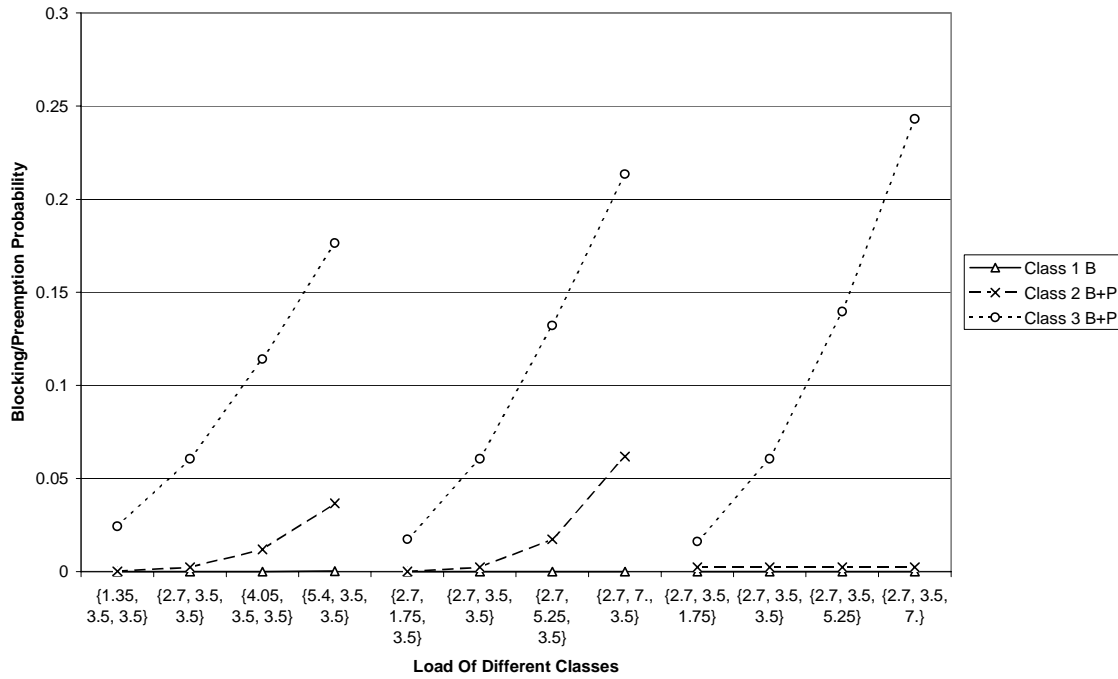
Complete Sharing, With Preemption

Figure 13.   Complete Sharing, with full preemption.
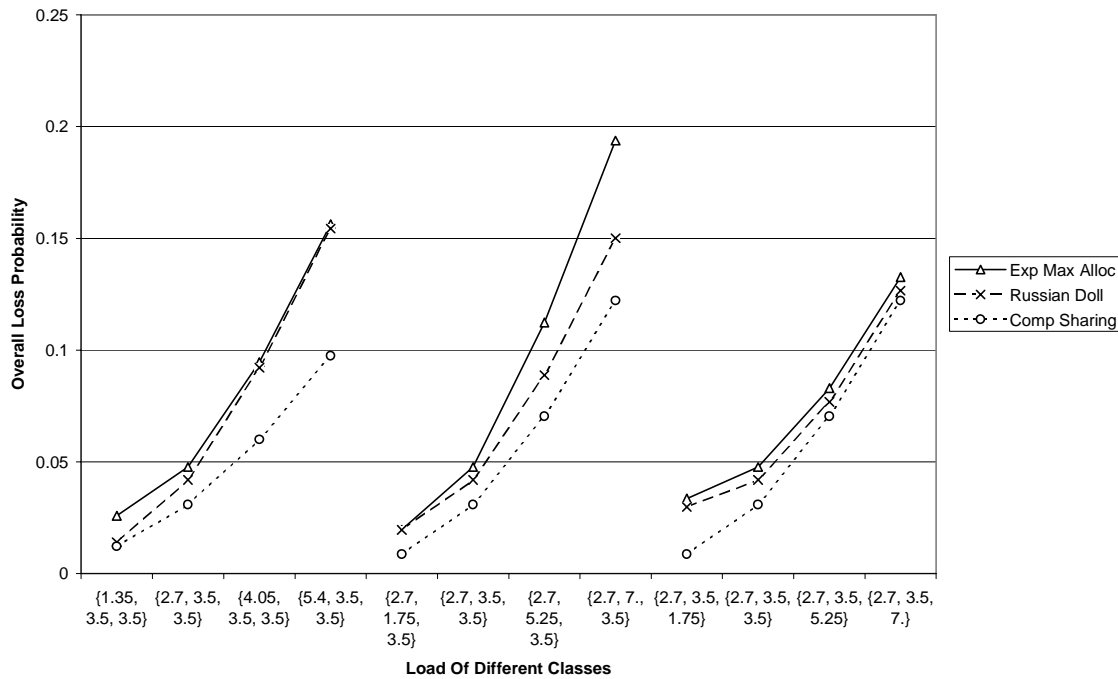
Total Loss Over All Classes

Figure 14.   Total loss over all classes, with full preemption.