Network Working Group
Request for Comments: 4126
Category: Experimental

J. Ash
AT&T

June 2005

Max Allocation with Reservation Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering & Performance Comparisons

Status of This Memo

Copyright Notice

Abstract

This document complements the Diffserv-aware MPLS Traffic Engineering (DS-TE) requirements document by giving a functional specification for the Maximum Allocation with Reservation (MAR) Bandwidth Constraints Model. Assumptions, applicability, and examples of the operation of the MAR Bandwidth Constraints Model are presented. MAR performance is analyzed relative to the criteria for selecting a Bandwidth Constraints Model, in order to provide guidance to user implementation of the model in their networks.

Table of Contents

1. Introduction

Diffserv-aware MPLS traffic engineering (DS-TE) requirements and protocol extensions are specified in [DSTE-REQ, DSTE-PROTO]. A requirement for DS-TE implementation is the specification of Bandwidth Constraints Models for use with DS-TE. The Bandwidth Constraints Model provides the 'rules' to support the allocation of bandwidth to individual class types (CTs). CTs are groupings of service classes in the DS-TE model, which are provided separate bandwidth allocations, priorities, and QoS objectives. Several CTs can share a common bandwidth pool on an integrated, multiservice MPLS/Diffserv network.

This document is intended to complement the DS-TE requirements document [DSTE-REQ] by giving a functional specification for the Maximum Allocation with Reservation (MAR) Bandwidth Constraints Model. Examples of the operation of the MAR Bandwidth Constraints Model are presented. MAR performance is analyzed relative to the criteria for selecting a Bandwidth Constraints Model, in order to provide guidance to user implementation of the model in their networks.

Two other Bandwidth Constraints Models are being specified for use in DS-TE:

1. Maximum Allocation Model (MAM) [MAM] - the maximum allowable bandwidth usage of each CT is explicitly specified.

2. Russian Doll Model (RDM) [RDM] - the maximum allowable bandwidth usage is done cumulatively by grouping successive CTs according to priority classes.

MAR is similar to MAM in that a maximum bandwidth allocation is given to each CT. However, through the use of bandwidth reservation and protection mechanisms, CTs are allowed to exceed their bandwidth allocations under conditions of no congestion but revert to their allocated bandwidths when overload and congestion occurs.

All Bandwidth Constraints Models should meet these objectives:

1. applies equally when preemption is either enabled or disabled (when preemption is disabled, the model still works 'reasonably' well),

2. bandwidth efficiency, i.e., good bandwidth sharing among CTs under both normal and overload conditions,

3. bandwidth isolation, i.e., a CT cannot hog the bandwidth of another CT under overload conditions,

4. protection against QoS degradation, at least of the high-priority CTs (e.g., high-priority voice, high-priority data, etc.), and

5. reasonably simple, i.e., does not require additional IGP extensions and minimizes signaling load processing requirements.

In Appendix A, modeling analysis is presented that shows the MAR Model meets all of these objectives and provides good network performance, relative to MAM and full-sharing models, under normal and abnormal operating conditions. It is demonstrated that MAR simultaneously achieves bandwidth efficiency, bandwidth isolation, and protection against QoS degradation without preemption.

In Section 3 we give the assumptions and applicability; in Section 4 a functional specification of the MAR Bandwidth Constraints Model; and in Section 5 we give examples of its operation. In Appendix A, MAR performance is analyzed relative to the criteria for selecting a Bandwidth Constraints Model, in order to provide guidance to user implementation of the model in their networks. In Appendix B, bandwidth prediction for path computation is discussed.

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, DSTE-PROTO] are repeated here:

Traffic Trunk: an aggregation of traffic flows of the same class (i.e., treated equivalently from the DS-TE perspective), which is placed inside a Label Switched Path (LSP).

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.

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 BCc, 0 <= c <= MaxBC-1 = 7.

TE-Class: A pair of: a) a CT, and b) a preemption priority allowed for that CT. This means that an LSP, transporting a Traffic Trunk from that CT, can use that preemption priority as the set-up priority, the holding priority, or both.

MAX_RESERVABLE_BWk: maximum reservable bandwidth on link k specifies the maximum bandwidth that may be reserved; this may be greater than the maximum link bandwidth, in which case the link may be oversubscribed [OSPF-TE].

BCck: bandwidth constraint for CTc on link k = allocated (minimum guaranteed) bandwidth for CTc on link k (see Section 4).

RBW_THRESk: reservation bandwidth threshold for link k (see Section 4).

RESERVED_BWck: reserved bandwidth-in-progress on CTc on link k (0 <= c <= MaxCT-1), RESERVED_BWck = total amount of the bandwidth reserved by all the established LSPs that belong to CTc.

UNRESERVED_BWk: unreserved link bandwidth on link k specifies the amount of bandwidth not yet reserved for any CT, UNRESERVED_BWk = MAX_RESERVABLE_BWk - sum [RESERVED_BWck (0 <= c <= MaxCT-1)].

UNRESERVED_BWck: unreserved link bandwidth on CTc on link k specifies the amount of bandwidth not yet reserved for CTc, UNRESERVED_BWck = UNRESERVED_BWk - delta0/1(CTck) * RBW-THRESk where

delta0/1(CTck) = 0 if RESERVED_BWck < BCck
delta0/1(CTck) = 1 if RESERVED_BWck >= BCck

A number of recovery mechanisms under investigation 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_BWck introduced in [DSTE-REQ] is compatible with such a concept of bandwidth sharing across multiple LSPs, the wording of the definition provided in [DSTE-REQ] is generalized. With this generalization, the definition is compatible with Shared Mesh Restoration defined in [GMPLS-RECOV], so that DS-TE and Shared Mesh Protection can operate simultaneously, under the assumption 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 CTc also need to belong to CTc; excess traffic LSPs that share bandwidth with backup LSPs of CTc also need to belong to CTc.

3. Assumptions & Applicability

In general, DS-TE is a bandwidth allocation mechanism for different classes of traffic allocated to various CTs (e.g., voice, normal data, best-effort data). Network operation functions such as capacity design, bandwidth allocation, routing design, and network planning are normally based on traffic-measured load and forecast [ASH1].

As such, the following assumptions are made according to the operation of MAR:

1. Connection admission control (CAC) allocates bandwidth for network flows/LSPs according to the traffic load assigned to each CT, based on traffic measurement and forecast.

2. CAC could allocate bandwidth per flow, per LSP, per traffic trunk, or otherwise. That is, no specific assumption is made about a specific CAC method, except that CT bandwidth allocation is related to the measured/forecasted traffic load, as per assumption #1.

3. CT bandwidth allocation is adjusted up or down according to measured/forecast traffic load. No specific time period is assumed for this adjustment, it could be short term (seconds, minutes, hours), daily, weekly, monthly, or otherwise.

4. Capacity management and CT bandwidth allocation thresholds (e.g., BCc) are designed according to traffic load, and are based on traffic measurement and forecast. Again, no specific time period is assumed for this adjustment, it could be short term (hours), daily, weekly, monthly, or otherwise.

5. No assumption is made on the order in which traffic is allocated to various CTs; again traffic allocation is assumed to be based only on traffic load as it is measured and/or forecast.

6. If link bandwidth is exhausted on a given path for a flow/LSP/traffic trunk, alternate paths may be attempted to satisfy CT bandwidth allocation.

Note that the above assumptions are not unique to MAR, but are generic, common assumptions for all BC Models.

4. Functional Specification of the MAR Bandwidth Constraints Model

A DS-TE Label Switching Router (LSR) that implements MAR MUST support enforcement of bandwidth constraints, in compliance with the specifications in this section.

In the MAR Bandwidth Constraints Model, the bandwidth allocation control for each CT is based on estimated bandwidth needs, bandwidth use, and status of links. The Label Edge Router (LER) makes needed bandwidth allocation changes, and uses [RSVP-TE], for example, to determine if link bandwidth can be allocated to a CT. Bandwidth allocated to individual CTs is protected as needed, but otherwise it is shared. Under normal, non-congested network conditions, all CTs/services fully share all available bandwidth. When congestion occurs for a particular CTc, bandwidth reservation prohibits traffic from other CTs from seizing the allocated capacity for CTc.

On a given link k, a small amount of bandwidth RBW_THRESk (the reservation bandwidth threshold for link k) is reserved and governs the admission control on link k. Also associated with each CTc on link k are the allocated bandwidth constraints BCck to govern bandwidth allocation and protection. The reservation bandwidth on a link (RBW_THRESk) can be accessed when a given CTc has bandwidth-in-use (RESERVED_BWck) below its allocated bandwidth constraint (BCck). However, if RESERVED_BWck exceeds its allocated bandwidth constraint (BCck), then the reservation bandwidth (RBW_THRESk) cannot be accessed. In this way, bandwidth can be fully shared among CTs if available, but is otherwise protected by bandwidth reservation methods.

Bandwidth can be accessed for a bandwidth request = DBW for CTc on a given link k based on the following rules:

Table 1: Rules for Admitting LSP Bandwidth Request = DBW on Link k

For LSP on a high priority or normal priority CTc:

If RESERVED_BWck <= BCck: admit if DBW <= UNRESERVED_BWk
If RESERVED_BWck > BCck: admit if DBW <= UNRESERVED_BWk - RBW_THRESk;

or, equivalently:

If DBW <= UNRESERVED_BWck, admit the LSP.

For LSP on a best-effort priority CTc:
allocated bandwidth BCck = 0;
Diffserv queuing admits BE packets only if there is available link bandwidth.

The normal semantics of setup and holding priority are applied in the MAR Bandwidth Constraints Model, and cross-CT preemption is permitted when preemption is enabled.

The bandwidth allocation rules defined in Table 1 are illustrated with an example in Section 6 and simulation analysis in Appendix A.

5. Setting Bandwidth Constraints

For a normal priority CTc, the bandwidth constraints BCck on link k are set by allocating the maximum reservable bandwidth (MAX_RESERVABLE_BWk) in proportion to the forecast or measured traffic load bandwidth (TRAF_LOAD_BWck) for CTc on link k. That is:

```
PROPORTIONAL_BWck = TRAF_LOAD_BWck/[sum {TRAF_LOAD_BWck, c=0, MaxCT-1}]
                    X MAX_RESERVABLE_BWk
```

For normal priority CTc:
BCck = PROPORTIONAL_BWck

For a high priority CT, the bandwidth constraint BCck is set to a multiple of the proportional bandwidth. That is:

For high priority CTc:
BCck = FACTOR X PROPORTIONAL_BWck

where FACTOR is set to a multiple of the proportional bandwidth (e.g., FACTOR = 2 or 3 is typical). This results in some 'over-allocation' of the maximum reservable bandwidth, and gives priority

to the high priority CTs. Normally the bandwidth allocated to high priority CTs should be a relatively small fraction of the total link bandwidth, with a maximum of 10-15 percent being a reasonable guideline.

As stated in Section 4, the bandwidth allocated to a best-effort priority CTc should be set to zero. That is:

For best-effort priority CTc:
BCck = 0

6. Example of MAR Operation

In the example, assume there are three class-types: CT0, CT1, CT2. We consider a particular link with

MAX-RESERVABLE_BW = 100

And with the allocated bandwidth constraints set as follows:

BC0 = 30
BC1 = 20
BC2 = 20

These bandwidth constraints are based on the normal traffic loads, as discussed in Section 5. With MAR, any of the CTs is allowed to exceed its bandwidth constraint (BCc) as long a there are at least RBW_THRES (reservation bandwidth threshold on the link) units of spare bandwidth remaining. Let’s assume

RBW_THRES = 10

So under overload, if

RESERVED_BW0 = 50
RESERVED_BW1 = 30
RESERVED_BW2 = 10

Therefore, for this loading

UNRESERVED_BW = 100 - 50 - 30 - 10 = 10

CT0 and CT1 can no longer increase their bandwidth on the link, because they are above their BC values and there is only RBW_THRES=10 units of spare bandwidth left on the link. But CT2 can take the additional bandwidth (up to 10 units) if the demand arrives, because it is below its BC value.

As also discussed in Section 4, if best effort traffic is present, it can always seize whatever spare bandwidth is available on the link at the moment, but is subject to being lost at the queues in favor of the higher priority traffic.

Let’s say an LSP arrives for CT0 needing 5 units of bandwidth (i.e., DBW = 5). We need to decide, based on Table 1, whether to admit this LSP or not. Since for CT0

RESERVED_BW0 > BC0 (50 > 30), and
DBW > UNRESERVED_BW - RBW_THRES (i.e., 5 > 10 - 10)

Table 1 says the LSP is rejected/blocked.

Now let’s say an LSP arrives for CT2 needing 5 units of bandwidth (i.e., DBW = 5). We need to decide based on Table 1 whether to admit this LSP or not. Since for CT2

RESERVED_BW2 < BC2 (10 < 20), and
DBW < UNRESERVED_BW (i.e., 5 < 10)

Table 1 says to admit the LSP.

Hence, in the above example, in the current state of the link and in the current CT loading, CT0 and CT1 can no longer increase their bandwidth on the link, because they are above their BCc values and there is only RBW_THRES=10 units of spare bandwidth left on the link. But CT2 can take the additional bandwidth (up to 10 units) if the demand arrives, because it is below its BCc value.

7. Summary

The proposed MAR Bandwidth Constraints Model includes the following:

1. allocation of bandwidth to individual CTs,

2. protection of allocated bandwidth by bandwidth reservation methods, as needed, but otherwise full sharing of bandwidth,

3. differentiation between high-priority, normal-priority, and best-effort priority services, and

4. provision of admission control to reject connection requests, when needed, in order to meet performance objectives.

The modeling results presented in Appendix A show that MAR bandwidth allocation achieves a) greater efficiency in bandwidth sharing while still providing bandwidth isolation and protection against QoS

degradation, and b) service differentiation for high-priority, normal-priority, and best-effort priority services.

8. Security Considerations

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

9. 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, the IANA has assigned a Bandwidth Constraints Model Id for MAR from the range 0-239 (which is to be managed as per the "Specification Required" policy defined in [IANA-CONS]).

Bandwidth Constraints Model Id 2 was allocated by IANA to MAR.

10. Acknowledgements

DS-TE and Bandwidth Constraints Models have been an active area of discussion in the TEWG. I would like to thank Wai Sum Lai for his support and review of this document. I also appreciate helpful discussions with Francois Le Faucheur.

Appendix A. MAR Operation & Performance Analysis

A.1. MAR Operation

In the MAR Bandwidth Constraints Model, the bandwidth allocation control for each CT is based on estimated bandwidth needs, bandwidth use, and status of links. The LER makes needed bandwidth allocation changes, and uses [RSVP-TE], for example, to determine if link bandwidth can be allocated to a CT. Bandwidth allocated to individual CTs is protected as needed, but otherwise it is shared. Under normal, non-congested network conditions, all CTs/services fully share all available bandwidth. When congestion occurs for a particular CTc, bandwidth reservation acts to prohibit traffic from other CTs from seizing the allocated capacity for CTc. Associated with each CT is the allocated bandwidth constraint (BCc) which governs bandwidth allocation and protection; these parameters are illustrated with examples in this Appendix.

In performing MAR bandwidth allocation for a given flow/LSP, the LER first determines the egress LSR address, service-identity, and CT. The connection request is allocated an equivalent bandwidth to be routed on a particular CT. The LER then accesses the CT priority, QoS/traffic parameters, and routing table between the LER and egress LSR, and sets up the connection request using the MAR bandwidth allocation rules. The LER selects a first-choice path and determines if bandwidth can be allocated on the path based on the MAR bandwidth allocation rules given in Section 4. If the first choice path has insufficient bandwidth, the LER may then try alternate paths, and again applies the MAR bandwidth allocation rules now described.

MAR bandwidth allocation is done on a per-CT basis, in which aggregated CT bandwidth is managed to meet the overall bandwidth requirements of CT service needs. Individual flows/LSPs are allocated bandwidth in the corresponding CT according to CT bandwidth availability. A fundamental principle applied in MAR bandwidth allocation methods is the use of bandwidth reservation techniques.

Bandwidth reservation gives preference to the preferred traffic by allowing it to seize idle bandwidth on a link more easily than the non-preferred traffic. Burke [BUR] first analyzed bandwidth reservation behavior from the solution of the birth-death equations for the bandwidth reservation model. Burke’s model showed the relative lost-traffic level for preferred traffic, which is not subject to bandwidth reservation restrictions, as compared to non-preferred traffic, which is subject to the restrictions. Bandwidth reservation protection is robust to traffic variations and provides

significant dynamic protection of particular streams of traffic. It is widely used in large-scale network applications [ASH1, MUM, AKI, KRU, NAK].

Bandwidth reservation is used in MAR bandwidth allocation to control sharing of link bandwidth across different CTs. On a given link, a small amount of bandwidth (RBW_THRES) is reserved (perhaps 1% of the total link bandwidth), and the reservation bandwidth can be accessed when a given CT has reserved bandwidth-in-progress (RESERVED_BW) below its allocated bandwidth (BC). That is, if the available link bandwidth (unreserved idle link bandwidth UNRESERVED_BW) exceeds RBW_THRES, then any CT is free to access the available bandwidth on the link. However, if UNRESERVED_BW is less than RBW_THRES, then the CT can utilize the available bandwidth only if its current bandwidth usage is below the allocated amount (BC). In this way, bandwidth can be fully shared among CTs if available, but it is protected by bandwidth reservation if below the reservation level.

Through the bandwidth reservation mechanism, MAR bandwidth allocation also gives preference to high-priority CTs, in comparison to normal-priority and best-effort priority CTs.

Hence, bandwidth allocated to each CT is protected by bandwidth reservation methods, as needed, but otherwise shared. Each LER monitors CT bandwidth use on each CT, and determines if connection requests can be allocated to the CT bandwidth. For example, for a bandwidth request of DBW on a given flow/LSP, the LER determines the CT priority (high, normal, or best-effort), CT bandwidth-in-use, and CT bandwidth allocation thresholds, and uses these parameters to determine the allowed load state threshold to which capacity can be allocated. In allocating bandwidth DBW to a CT on given LSP (for example, A-B-E), each link in the path is checked for available bandwidth in comparison to the allowed load state. If bandwidth is unavailable on any link in path A-B-E, another LSP could be tried, such as A-C-D-E. Hence, determination of the link load state is necessary for MAR bandwidth allocation, and two link load states are distinguished: available (non-reserved) bandwidth (ABW_STATE), and reserved-bandwidth (RBW_STATE). Management of CT capacity uses the link state and the allowed load state threshold to determine if a bandwidth allocation request can be accepted on a given CT.

A.2. Analysis of MAR Performance

In this Appendix, modeling analysis is presented in which MAR bandwidth allocation is shown to provide good network performance, relative to full sharing models, under normal and abnormal operating conditions. A large-scale Diffserv-aware MPLS traffic engineering simulation model is used, in which several CTs with different

priority classes share the pool of bandwidth on a multiservice, integrated voice/data network. MAR methods have also been analyzed in practice for networks that use time division multiplexing (i.e., TDM-based networks) [ASH1], and in modeling studies for IP-based networks [ASH2, ASH3, E.360].

All Bandwidth Constraints Models should meet these objectives:

1. applies equally when preemption is either enabled or disabled (when preemption is disabled, the model still works 'reasonably' well),

2. bandwidth efficiency, i.e., good bandwidth sharing among CTs under both normal and overload conditions,

3. bandwidth isolation, i.e., a CT cannot hog the bandwidth of another CT under overload conditions,

4. protection against QoS degradation, at least of the high-priority CTs (e.g., high-priority voice, high-priority data, etc.), and

5. reasonably simple, i.e., does not require additional IGP extensions and minimizes signaling load processing requirements.

The use of any given Bandwidth Constraints Model has significant impacts on the performance of a network, as explained later. Therefore, the criteria used to select a model need to enable us to evaluate how a particular model delivers its performance, relative to other models. Lai [LAI, DSTE-PERF] has analyzed the MAM and RDM Models and provided valuable insights into the relative performance of these models under various network conditions.

In environments where preemption is not used, MAM is attractive because a) it is good at achieving isolation, and b) it achieves reasonable bandwidth efficiency with some QoS degradation of lower classes. When preemption is used, RDM is attractive because it can achieve bandwidth efficiency under normal load. However, RDM cannot provide service isolation under high load or when preemption is not used.

Our performance analysis of MAR bandwidth allocation methods is based on a full-scale, 135-node simulation model of a national network, combined with a multiservice traffic demand model to study various scenarios and tradeoffs [ASH3, E.360]. Three levels of traffic priority -- high, normal, and best effort -- are given across 5 CTs: normal priority voice, high priority voice, normal priority data, high priority data, and best effort data.

The performance analyses for overloads and failures include a) the MAR Bandwidth Constraints Model, as specified in Section 4, b) the MAM Bandwidth Constraints Model, and c) the No-DSTE Bandwidth Constraints Model.

The allocated bandwidth constraints for MAR are described in Section 5 as:

```
Normal priority CTs:      BCck = PROPORTIONAL_BWk,
High priority CTs:        BCck = FACTOR X PROPORTIONAL_BWk
Best-effort priority CTs: BCck = 0
```

In the MAM Bandwidth Constraints Model, the bandwidth constraints for each CT are set to a multiple of the proportional bandwidth allocation:

```
Normal priority CTs:      BCck = FACTOR1 X PROPORTIONAL_BWk,
High priority CTs:        BCck = FACTOR2 X PROPORTIONAL_BWk
Best-effort priority CTs: BCck = 0
```

Simulations show that for MAM, the sum (BCc) should exceed MAX_RESERVABLE_BWk for better efficiency, as follows:

1. The normal priority CTs and the BCc values need to be over-allocated to get reasonable performance. It was found that over-allocating by 100% (i.e., setting FACTOR1 = 2), gave reasonable performance.

2. The high priority CTs can be over-allocated by a larger multiple FACTOR2 in MAM and this gives better performance.

The rather large amount of over-allocation improves efficiency, but somewhat defeats the 'bandwidth protection/isolation' needed with a BC Model, because one CT can now invade the bandwidth allocated to another CT. Each CT is restricted to its allocated bandwidth constraint BCck, which is the maximum level of bandwidth allocated to each CT on each link, as in normal operation of MAM.

In the No-DSTE Bandwidth Constraints Model, no reservation or protection of CT bandwidth is applied, and bandwidth allocation requests are admitted if bandwidth is available. Furthermore, no queuing priority is applied to any of the CTs in the No-DSTE Bandwidth Constraints Model.

Table 2 gives performance results for a six-times overload on a single network node at Oakbrook, Illinois. The numbers given in the table are the total network percent lost (i.e., blocked) or delayed

traffic. Note that in the focused overload scenario studied here, the percentage of lost/delayed traffic on the Oakbrook node is much higher than the network-wide average values given.

Table 2
Performance Comparison for MAR, MAM, & No-DSTE
Bandwidth Constraints (BC) Models
6X Focused Overload on Oakbrook
(Total Network % Lost/Delayed Traffic)

Class Type	MAR BC Model	MAM BC Model	No-DSTE BC Model
NORMAL PRIORITY VOICE	0.00	1.97	10.30
HIGH PRIORITY VOICE	0.00	0.00	7.05
NORMAL PRIORITY DATA	0.00	6.63	13.30
HIGH PRIORITY DATA	0.00	0.00	7.05
BEST EFFORT PRIORITY DATA	12.33	11.92	9.65

Clearly the performance is better with MAR bandwidth allocation, and the results show that performance improves when bandwidth reservation is used. The reason for the poor performance of the No-DSTE Model, without bandwidth reservation, is due to the lack of protection of allocated bandwidth. If we add the bandwidth reservation mechanism, then performance of the network is greatly improved.

The simulations showed that the performance of MAM is quite sensitive to the over-allocation factors discussed above. For example, if the BCc values are proportionally allocated with FACTOR1 = 1, then the results are much worse, as shown in Table 3:

Table 3
Performance Comparison for MAM Bandwidth Constraints Model
with Different Over-allocation Factors
6X Focused Overload on Oakbrook
(Total Network % Lost/Delayed Traffic)

Class Type	(FACTOR1 = 1)	(FACTOR1 = 2)
NORMAL PRIORITY VOICE	31.69	1.97
HIGH PRIORITY VOICE	0.00	0.00
NORMAL PRIORITY DATA	31.22	6.63
HIGH PRIORITY DATA	0.00	0.00
BEST EFFORT PRIORITY DATA	8.76	11.92

Table 4 illustrates the performance of the MAR, MAM, and No-DSTE Bandwidth Constraints Models for a high-day network load pattern with a 50% general overload. The numbers given in the table are the total network percent lost (i.e., blocked) or delayed traffic.

Table 4
Performance Comparison for MAR, MAM, & No-DSTE
Bandwidth Constraints (BC) Models
50% General Overload (Total Network % Lost/Delayed Traffic)

Class Type	MAR BC Model	MAM BC Model	No-DSTE BC Model
NORMAL PRIORITY VOICE	0.02	0.13	7.98
HIGH PRIORITY VOICE	0.00	0.00	8.94
NORMAL PRIORITY DATA	0.00	0.26	6.93
HIGH PRIORITY DATA	0.00	0.00	8.94
BEST EFFORT PRIORITY DATA	10.41	10.39	8.40

Again, we can see the performance is always better when MAR bandwidth allocation and reservation is used.

Table 5 illustrates the performance of the MAR, MAM, and No-DSTE Bandwidth Constraints Models for a single link failure scenario (3 OC-48). The numbers given in the table are the total network percent lost (blocked) or delayed traffic.

Table 5
Performance Comparison for MAR, MAM, & No-DSTE
Bandwidth Constraints (BC) Models
Single Link Failure (2 OC-48)
(Total Network % Lost/Delayed Traffic)

Class Type	MAR BC Model	MAM BC Model	No-DSTE BC Model
NORMAL PRIORITY VOICE	0.00	0.62	0.63
HIGH PRIORITY VOICE	0.00	0.31	0.32
NORMAL PRIORITY DATA	0.00	0.48	0.50
HIGH PRIORITY DATA	0.00	0.31	0.32
BEST EFFORT PRIORITY DATA	0.12	0.72	0.63

Again, we can see the performance is always better when MAR bandwidth allocation and reservation is used.

Table 6 illustrates the performance of the MAR, MAM, and No-DSTE Bandwidth Constraints Models for a multiple link failure scenario (3 links with 3 OC-48, 3 OC-3, 4 OC-3 capacity, respectively). The numbers given in the table are the total network percent lost (blocked) or delayed traffic.

Table 6
Performance Comparison for MAR, MAM, & No-DSTE
Bandwidth Constraints (BC) Models
Multiple Link Failure
(3 Links with 2 OC-48, 2 OC-12, 1 OC-12, Respectively)
(Total Network % Lost/Delayed Traffic)

Class Type	MAR BC Model	MAM BC Model	No-DSTE BC Model
NORMAL PRIORITY VOICE	0.00	0.91	0.92
HIGH PRIORITY VOICE	0.00	0.44	0.44
NORMAL PRIORITY DATA	0.00	0.70	0.72
HIGH PRIORITY DATA	0.00	0.44	0.44
BEST EFFORT PRIORITY DATA	0.14	1.03	1.04

Again, we can see the performance is always better when MAR bandwidth allocation and reservation is used.

Lai's results [LAI, DSTE-PERF] show the trade-off between bandwidth sharing and service protection/isolation, using an analytic model of a single link. He shows that RDM has a higher degree of sharing than MAM. Furthermore, for a single link, the overall loss probability is the smallest under full sharing and largest under MAM, with RDM being intermediate. Hence, on a single link, Lai shows that the full sharing model yields the highest link efficiency, while MAM yields the lowest; and that full sharing has the poorest service protection capability.

The results of the present study show that, when considering a network context in which there are many links and multiple-link routing paths are used, full sharing does not necessarily lead to maximum, network-wide bandwidth efficiency. In fact, the results in Table 4 show that the No-DSTE Model not only degrades total network throughput, but also degrades the performance of every CT that should be protected. Allowing more bandwidth sharing may improve performance up to a point, but it can severely degrade performance if care is not taken to protect allocated bandwidth under congestion.

Both Lai's study and this study show that increasing the degree of bandwidth sharing among the different CTs leads to a tighter coupling between CTs. Under normal loading conditions, there is adequate capacity for each CT, which minimizes the effect of such coupling.

Under overload conditions, when there is a scarcity of capacity, such coupling can cause severe degradation of service, especially for the lower priority CTs.

Thus, the objective of maximizing efficient bandwidth usage, as stated in Bandwidth Constraints Model objectives, needs to be exercised with care. Due consideration also needs to be given to achieving bandwidth isolation under overload, in order to minimize the effect of interactions among the different CTs. The proper tradeoff of bandwidth sharing and bandwidth isolation needs to be achieved in the selection of a Bandwidth Constraints Model. Bandwidth reservation supports greater efficiency in bandwidth sharing, while still providing bandwidth isolation and protection against QoS degradation.

In summary, the proposed MAR Bandwidth Constraints Model includes the following: a) allocation of bandwidth to individual CTs, b) protection of allocated bandwidth by bandwidth reservation methods, as needed, but otherwise full sharing of bandwidth, c) differentiation between high-priority, normal-priority, and best-effort priority services, and d) provision of admission control to reject connection requests, when needed, in order to meet performance objectives.

In the modeling results, the MAR Bandwidth Constraints Model compares favorably with methods that do not use bandwidth reservation. In particular, some of the conclusions from the modeling are as follows:

- MAR bandwidth allocation is effective in improving performance over methods that lack bandwidth reservation; this allows more bandwidth sharing under congestion.

- MAR achieves service differentiation for high-priority, normal-priority, and best-effort priority services.

- Bandwidth reservation supports greater efficiency in bandwidth sharing while still providing bandwidth isolation and protection against QoS degradation, and is critical to stable and efficient network performance.

Appendix B. Bandwidth Prediction for Path Computation

As discussed in [DSTE-PROTO], there are potential advantages for a Head-end when predicting the impact of an LSP on the unreserved bandwidth for computing the path of the LSP. One example would be to perform better load-distribution of multiple LSPs across multiple paths. Another example would be to avoid CAC rejection when the LSP no longer fits on a link after establishment.

Where such predictions are used on Head-ends, the optional Bandwidth Constraints sub-TLV and the optional Maximum Reservable Bandwidth sub-TLV MAY be advertised in the IGP. This can be used by Head-ends to predict how an LSP affects unreserved bandwidth values. Such predictions can be made with MAR by using the unreserved bandwidth values advertised by the IGP, as discussed in Sections 2 and 4:

```
UNRESERVED_BWck = MAX_RESERVABLE_BWk - UNRESERVED_BWk -
                  delta0/1(CTck) * RBW-THRESk
```

where

```
delta0/1(CTck) = 0 if RESERVED_BWck < BCck
delta0/1(CTck) = 1 if RESERVED_BWck >= BCck
```

Furthermore, the following estimate can be made for RBW_THRESk:

```
RBW_THRESk = RBW_% * MAX_RESERVABLE_BWk,
```

where RBW_% is a locally configured variable, which could take on different values for different link speeds. This information could be used in conjunction with the BC sub-TLV, MAX_RESERVABLE_BW sub-TLV, and UNRESERVED_BW sub-TLV to make predictions of available bandwidth on each link for each CT. Because admission control algorithms are left for vendor differentiation, predictions can only be performed effectively when the Head-end LSR predictions are based on the same (or a very close) admission control algorithm used by other LSRs.

LSPs may occasionally be rejected when head-ends are establishing LSPs through a common link. As an example, consider some link L, and two head-ends H1 and H2. If only H1 or only H2 is establishing LSPs through L, then the prediction is accurate. But if both H1 and H2 are establishing LSPs through L at the same time, the prediction would not work perfectly. In other words, the CAC will occasionally run into a rejected LSP on a link with such 'race' conditions. Also, as mentioned in Appendix A, such a prediction is optional and outside the scope of the document.

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

[AKI] Akinpelu, J. M., "The Overload Performance of Engineered Networks with Nonhierarchical & Hierarchical Routing," BSTJ, Vol. 63, 1984.

[ASH1] Ash, G. R., "Dynamic Routing in Telecommunications Networks," McGraw-Hill, 1998.

[ASH2] Ash, G. R., et al., "Routing Evolution in Multiservice Integrated Voice/Data Networks," Proceeding of ITC-16, Edinburgh, June 1999.

[ASH3] Ash, G. R., "Performance Evaluation of QoS-Routing Methods for IP-Based Multiservice Networks," Computer Communications Magazine, May 2003.

[BUR] Burke, P. J., Blocking Probabilities Associated with Directional Reservation, unpublished memorandum, 1961.

[DSTE-PERF] Lai, W., "Bandwidth Constraints Models for Differentiated Services-aware MPLS Traffic Engineering: Performance Evaluation", RFC 4128, June 2005.

[E.360] ITU-T Recommendations E.360.1 - E.360.7, "QoS Routing & Related Traffic Engineering Methods for Multiservice TDM-, ATM-, & IP-Based Networks".

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

[KRU] Krupp, R. S., "Stabilization of Alternate Routing Networks", Proceedings of ICC, Philadelphia, 1982.

[LAI] Lai, W., "Traffic Engineering for MPLS, Internet Performance and Control of Network Systems III Conference", SPIE Proceedings Vol. 4865, pp. 256-267, Boston, Massachusetts, USA, 29 July-1 August 2002.

[MAM] Le Faucheur, F., Lai, W., "Maximum Allocation Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering", RFC 4125, June 2005.

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

[MUM] Mummert, V. S., "Network Management and Its Implementation on the No. 4ESS, International Switching Symposium", Japan, 1976.

[NAK] Nakagome, Y., Mori, H., Flexible Routing in the Global Communication Network, Proceedings of ITC-7, Stockholm, 1973.

[OSPF-TE] Katz, D., Kompella, K. and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

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

[RSVP-TE] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V. and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

Author’s Address

Jerry Ash
AT&T
Room MT D5-2A01
200 Laurel Avenue
Middletown, NJ 07748, USA

Phone: +1 732-420-4578
EMail: gash@att.com

Full Copyright Statement

Intellectual Property

Acknowledgement

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