A PRELIMINARY EVALUATION OF BANDWIDTH ALLOCATION MODEL DYNAMIC SWITCHING

Rafael F. Reale\textsuperscript{1}, Romildo Martins S. Bezerra\textsuperscript{2} and Joberto Sérgio B. Martins\textsuperscript{3}

\textsuperscript{1}DMCC, UFBA, Salvador, Bahia, Brazil
\textsuperscript{2}GSORT, IFBA, Salvador, Bahia, Brazil
\textsuperscript{3}NUPERC, UNIFACS, Salvador, Bahia, Brazil

ABSTRACT

Bandwidth Allocation Models (BAMs) are used in order to define Bandwidth Constraints (BCs) in a per-class basis for MPLS/DS-TE networks and effectively define how network resources like bandwidth are obtained and shared by applications. The BAMs proposed (MAM – Maximum Allocation Model, RDM – Russian Dolls Model, G-RDM – Generic RDM and AllocTC-Sharing) attempt to optimize the use of bandwidth resources on a per-link basis with different allocation and resource sharing characteristics. As such, the adoption of distinct BAMs and/or changes in network resource demands (network traffic profile) may result in different network traffic allocation and operational behavior for distinct BAMs. This paper evaluates the resulting network characteristics (link utilization, preemption and flows blocking) of using BAMs dynamically with different traffic scenarios. In brief, it is investigated the dynamics of BAM switching with distinct traffic scenarios. The paper presents initially the investigated BAMs in relation to their behavior and resource allocation characteristics. Then, distinct BAMs are compared using different traffic scenarios in order to investigate the impact of a dynamic change of the BAM configured in the network. Finally, the paper shows that the adoption of a dynamic BAM allocation strategy may result in benefits for network operation in terms of link utilization, preemption and flows blocking.

KEYWORDS

BAM, Management Dynamic, RDM, AllocTC

1. INTRODUCTION AND MOTIVATION

Bandwidth Allocation Models (BAMs) are used in order to define Bandwidth Constraints (BCs) in a per-class basis for MPLS/DS-TE networks [3] and effectively define how network resources like bandwidth are obtained and shared by applications.

The adoption and configuration of a specific BAM (MAM or RDM or AllocTC-Sharing) for a network is dependent on an evaluation process done typically by the manager. This evaluation process considers aspects such as the set of applications mapped to traffic classes (TCs), the priorities allocated for the classes of applications (TCs) and the SLA (Service Level Agreement)/QoS (Quality of Service) requirements for these applications (SLA/QoS dependencies).

Once a specific BAM is defined and configured for a network, it will have a static behavior for the set of running applications in terms of link utilization, preemption and LSP (calls) blocking. That is so, since BAMs attempt to optimize the use of bandwidth resources on a per-link basis with different allocation and resource sharing characteristics.
In most cases, the evaluation and definition of a BAM to be used in a network is a non-trivial evaluation process. In effect, from the management point of view the evaluation and configuration processes might be eventually done by an autonomic framework capable of analyzing the network current state, its SLA/QoS requirements and, based on that, to infer on-the-fly the most adequate BAM to be used [10][11].

The main motivation addressed by this paper is to preliminarily investigate the feasibility of adopting a dynamic BAM utilization strategy based on different traffic scenarios. The focus is on the identification of eventual advantages and disadvantages by using BAMs dynamically under different traffic scenarios.

The paper initially reviews MAM, RDM, G-RDM and AllocTC-Sharing (BAMs) in relation to their behavior and resource allocation characteristics. In sequence, these BAMs are compared using different traffic scenarios in order to investigate the impact of a dynamic BAM configuration in networks. Finally, the paper shows that the adoption of a dynamic BAM allocation strategy may result in benefits for network operation in terms of link utilization, preemption and flows blocking.

2. BANDWIDTH ALLOCATION MODELS – A BRIEF REVIEW

Current research on bandwidth allocation models has been mostly focused on finding new models with distinct strategies for resource allocation and variations on these alternatives [1][2][4][6][7][8].

The Maximum Allocation Model (MAM) is discussed in [4][7] and, in this model, there are multiple traffic classes (TCs) and each TC is configured by the network administrator (network manager) to use a certain amount of link bandwidth (resource). This resource is allocated on demand to the applications belonging to traffic classes (TCs). MAM effectively isolates traffic classes (TCs) and there is no resource (bandwidth) sharing among applications belonging to different classes.

The Russian Dolls Model (RDM) is presented in [2][8] and, in this model, TCs with higher values are hierarchically superior to TCs with lower values. As such, all LSPs associated with TC2 do not use a bandwidth greater than BC2, all LSPs associated with TC1 and TC2 do not use bandwidth greater than BC1 and all LSPs associated with TC0, TC1 and TC2 do not use bandwidth greater than BC0 (Figure. 1).

RDM is an effective evolution of MAM model and introduced the first attempt to share resources among TCs and, consequently, LSPs. In effect, RDM allows the sharing of unallocated bandwidth resources unused by high-priority applications by low-priority applications.

The G-RDM model is presented in [1] and is basically a variation of the RDM model in which TCs have both shared resources (like in RDM) and private resources. TC private resources are not shared in any circumstance (like MAM). The overall operation results in having a hybrid MAM/RDM with a reduced volume of resource sharing for applications allocated in distinct traffic classes.
The AllocTC-Sharing model is presented in [6] and, in this model, an opportunistic strategy for resource (bandwidth) allocation is used. AllocTC-Sharing allows two different styles for resource (bandwidth) sharing concomitantly: a “high-to-low” (HTL) bandwidth allocation and a “low-to-high” (LTH) bandwidth allocation. The “high-to-low” bandwidth allocation style is equivalent to RDM model. The “low-to-high” bandwidth allocation style allows high priority classes temporarily allocate non-used bandwidth primarily reserved for low priority classes.

3. BANDWIDTH ALLOCATION MODELS CHARACTERISTICS

The characteristics of the Bandwidth Allocation Models currently available at the literature are distinct and, as such, imply in distinct network overall behaviour for different traffic matrixes and/or different types and classes of applications. For the sake of a preliminary evaluation of these characteristics, this paper will consider three BAMs: MAM, RDM and AllocTC-Sharing that approximately represent the basic operational aspects and evolution of BAMs.

The basic characteristics of these BAMs are illustrated in Table 1 and the impact of these characteristics is considered in relation to the set of applications grouped as TCs and running on the network.

<table>
<thead>
<tr>
<th>BAM Operational Characteristics</th>
<th>MAM</th>
<th>RDM</th>
<th>AllocTC-Sharing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sharing from “high to low”</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Sharing from “low to high”</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Efficient bandwidth utilization with high volume of traffic (low priority)</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Efficient bandwidth utilization with high volume of traffic (high priority)</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Isolation between TCs</td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
</tr>
</tbody>
</table>

The MAM model does not admit bandwidth sharing between traffic classes (TCs) (either "high to low" or "low to high") and, as such, is indicated when we have a network traffic profile where “application classes” (Traffic Classes – TCs) do not affect each other significantly [1]. MAM model utilization in a network does not imply in efficient bandwidth utilization since unallocated resources cannot be used among different TCs (class of applications).

By using the RDM model low-priority applications can benefit from idle bandwidth that would be normally allocated for high-priority applications. In relation to high-priority applications the
opposite effect occurs. In fact, in case of exceeding high-priority applications bandwidth demands (beyond configured BC) there will be flows blocking since they cannot benefit from idle bandwidth allocated primarily for low-priority applications. The overall impact for this model is that link utilization is not maximized with this BAM [6].

The AllocTC-Sharing model supports both high-priority and low-priority traffic classes (TCs) bandwidth sharing. The model allows the use of available bandwidth in both directions (“high to low” and “low to high”).

The “cost” of allowing bandwidth sharing in both directions is that preemptions may also occur in both directions (conventional preemption of low-priority applications by high-priority applications and “devolutions” of bandwidth which are, in effect, the preemption of high-priority applications by low-priority applications).

As discussed in [6], AllocTC-Sharing tends to opportunistically maximize link utilization but has also to consider the impact of “returning” bandwidth borrowed from low-priority applications by high-priority applications. As such, this model is most suitable for elastic and bandwidth eager high-priority applications.

4. Dynamic BAM Utilization

Computer networks have a dynamic traffic profile (traffic matrixes) and do have fixed and, normally, limited resources in terms of bandwidth. Different Bandwidth Allocation Models (BAMs) operate using distinct resources (bandwidth) allocation strategies and, as such, the overall network behavior is different for distinct BAMs under identical traffic profiles.

From the management point of view it would be more effective to adopt and configure BAMs according with current network traffic profile considering SLA/QoS application requirements (indirectly TCs SLA/QoS requirements, since applications are grouped in traffic classes – TCs). The first step towards this management approach is to investigate if, effectively, to switch among BAMs (MAM, RDM, G-RDM or AllocTC-Sharing) in order to reflect changes in the application’s (TCs) resource demand is advantageous for network operation. This is the focus of the next sections on this paper.

The step following the certification of BAM switching adequateness is to consider how to effectively realize it. There is always the possibility to switch BAMs in a “manual style” by, for instance, using configuration scripts for time periods during the day. These time periods would have, typically, a certain profile of traffic which could be estimated using methods and tools currently available (offline approach).

Another possible alternative is to develop a framework with the capability of dynamically (on-the-fly) to decide which bandwidth allocation model is currently adequate for a given network traffic state [9]. This corresponds to a more “on-the-fly” alternative and comprises a challenge in terms of computing a great deal of parameters (SLAs, QoS, LSP routing, other) in order to infer the best possible configuration outcome for the network. As a last aspect in relation to the alternatives for BAM switching, the referred framework could eventually incorporate some autonomic characteristics in order to better support the management decision-making process.

In the following sections we focus on the evaluation of BAM switching adequateness with distinct traffic scenarios.
5. **Dynamic BAM Utilization**

Computer networks have a dynamic traffic profile (traffic matrixes) and do have fixed and, normally, limited resources in terms of bandwidth. Different Bandwidth Allocation Models (BAMs) operate using distinct resources (bandwidth) allocation strategies and, as such, the overall network behavior is different for distinct BAMs under identical traffic profiles.

From the management point of view it would be more effective to adopt and configure BAMs according with current network traffic profile considering SLA/QoS application requirements (indirectly TCs SLA/QoS requirements, since applications are grouped in traffic classes – TCs). The first step towards this management approach is to investigate if, effectively, to switch among BAMs (MAM, RDM, G-RDM or AllocTC-Sharing) in order to reflect changes in the application´s (TCs) resource demand is advantageous for network operation. This is the focus of the next sections on this paper.

The step following the certification of BAM switching adequateness is to consider how to effectively realize it. There is always the possibility to switch BAMs in a “manual style” by, for instance, using configuration scripts for time periods during the day. These time periods would have, typically, a certain profile of traffic which could be estimated using methods and tools currently available (offline approach).

Another possible alternative is to develop a framework with the capability of dynamically (on-the-fly) to decide which bandwidth allocation model is currently adequate for a given network traffic state [9]. This corresponds to a more “on-the-fly” alternative and comprises a challenge in terms of computing a great deal of parameters (SLAs, QoS, LSP routing, other) in order to infer the best possible configuration outcome for the network. As a last aspect in relation to the alternatives for BAM switching, the referred framework could eventually incorporate some autonomic characteristics in order to better support the management decision-making process. In the following sections we focus on the evaluation of BAM switching adequateness with distinct traffic scenarios.

6. **Evaluation Scenario**

The main objective of this preliminary evaluation of BAM dynamic utilization (BAM switching) in networks will be the identification of some potential advantages and disadvantages resulting from adopting this approach. As such, the evaluation scenarios discussed in this paper are a specific subset of network’s full operation scenarios and is focused on more extreme traffic profiles where the network is either stressed or alleviated with combinations of high and low volume of traffic (TCs).

Two BAMs are used in the simulation (RDM and AllocTC-Sharing). Two traffic scenarios were adopted and configured with 03 traffic classes:

- TC0 – low priority applications;
- TC1 – intermediate priority applications and
- TC2 – high priority applications.

TC1 traffic is intended to be interference traffic and will be kept stable in terms of traffic and network resource demands for both scenarios.
6.1 Scenario 01 - BAM Switching Analysis

In the first scenario, the combination of traffic used to identify BAM switching characteristics will have two distinct phases of traffic as follows:

- Phase 01 simulation run has a high volume of traffic (bandwidth demand) for TC2 (high-priority applications) and low volume of traffic for TC0 (low-priority applications).
- Phase 02 simulation run maintains a high volume of traffic (bandwidth demand) for TC2 (high-priority applications) and enforces a high volume of traffic for TC0 (low-priority applications).

6.2 Scenario 02 - BAM Switching Analysis with Inverted Traffic Demand

In scenario 02, the traffic demands are inverted as follows:

- Phase 1 simulation run has a high volume of traffic (bandwidth demand) for TC2 (high-priority applications) and high volume of traffic for TC0 (low-priority applications)
- Phase 2 simulation run has a high volume of traffic (bandwidth demand) for TC2 (high-priority applications) and low volume of traffic for TC0 (low-priority applications).

In effect, the scenarios and phases were defined considering that the characteristics of BAMs (RDM and AllocTC-Sharing) differ when traffic profiles are on more extreme situations.

6.3 BAM Switching Analysis Details

The parameters evaluated in the simulation were link utilization (by TC and by “link”), preemptions and LSP (bandwidth demand) blocking.

The performance evaluation used a network topology with one traffic source (S1), two traffic interference (S2 and S3) and one destination (D) (Figure 2). CSPF (Constrained Shortest Path First) algorithm was used for path selection.

![Figure 2. Network Topology](image)

Each scenario was simulated with 05 random seeds and the results are presented in terms of the mean value obtained (confidence interval of 95%) [5]. The configuration parameters of the simulation scenarios are as follows:
• Links: 622 Mbps (STM-4 – SDH)
• Bandwidth Constraint (BCs) according Table 2

Table 2 - Bandwidth Constraint (BCs) per Traffic Class (TCs)

<table>
<thead>
<tr>
<th>BC</th>
<th>Max BC (%)</th>
<th>MAX BC (Mbps)</th>
<th>TC per BC</th>
</tr>
</thead>
<tbody>
<tr>
<td>BC_0</td>
<td>100</td>
<td>622</td>
<td>TC_0+TC_1+TC_2</td>
</tr>
<tr>
<td>BC_1</td>
<td>80</td>
<td>497.6</td>
<td>TC_1+TC_2</td>
</tr>
<tr>
<td>BC_2</td>
<td>45</td>
<td>279.9</td>
<td>TC_2</td>
</tr>
</tbody>
</table>

The switching between BAMs (AllocTC-Sharing ↔ RDM) is another simulation parameter definition and, for this preliminary evaluation, is enforced a specific condition when TC0 bandwidth utilization reaches 80%. As such, when CT0 utilization is below 80%, AllocTC-Sharing is used, otherwise RDM is adopted. With this condition we attempt to use the opportunistic characteristics of AllocTC-Sharing of borrowing bandwidth from low-priority applications.

The choice (80% TC0) considers that a traffic stress situation is being reached and, as such, the BAM switching evaluation becomes relevant. In a full range evaluation, the switching between BAMs may vary depending on parameters defined by the manager such as link utilization, link utilization by TC and number of preemptions, among others. In effect, in case BAM switching is configured, the switching point must also be investigated and evaluated against management parameters. Other basic simulation parameters are:

• Flow – flow setup duration modeled exponentially – mean of 220 seconds; leads to link saturation;
• Flow bandwidth – uniform distribution between 05 Mbps e 25 Mbps;
• Number of flows – 2.000 flows.

6. BAM DYNAMIC CHARACTERISTIC – PRELIMINARY EVALUATION

Figure 3 illustrates the AllocTC-Sharing to RDM BAM enforced switching and link utilization by TCs for the scenario 01. It is observed that link utilization by TCs is maximized at phase 01 (first part of the run) with TC2 (high-priority applications) borrowing bandwidth from TC0 (low-priority application) and this reflects AllocTC-Sharing behaviour for this traffic pattern since there is a low demand for low-priority traffic. At phase 02, the link tends to saturation (high volume of traffic for TC0 and TC2) and, as such, both AllocTC-Sharing and RDM have a similar behaviour with the bandwidth usage reaching the configured limits (BCs). As such, simulation suggests that link utilization by TCs (Figure 3) and link utilization (Figure 4) might be improved by using AllocTC-Sharing for phase 01 and phase 02 traffic scenarios or a combination of AllocTC-Sharing (1st phase) followed by RDM (2nd phase).
The BAMs (RDM and AllocTC-Sharing) have an operational cost since they may temporarily “borrow” bandwidth from other classes and with either TC or link saturation (high volume of traffic) bandwidth has to be returned. Two parameters reflect this behavior: the number of preemptions and the number of LSPs blocked (bandwidth demand not granted).

Figure 3. Link utilization by TC. The vertical line indicates the BAM switching point, as mentioned in section V. BAM switching point happens in traffic stress situation (80% TC0).

Figure 4. Link utilization

Figure 5 illustrates the number of “devolutions” by preemption [6] and suggests that switching from AllocTC-Sharing to RDM would improve the behaviour for high-priority applications (TC2) since it would reduce the number of LSP teardown in order to return the “loans” taken by AllocTC-Sharing model. In brief, in a traffic scenario starting with low-priority applications with low traffic and evolving to high volume of traffic for all classes (TCs) the dynamic switching would improve network operation. In this context, it is considered that high-priority applications preemption, depending on the manager decision and application SLA/QoS requirements, is an eventually undesired situation.
Figure 5. “Devolution” by preemption of high-priority traffic by low-priority traffic.

Figure 6 illustrates the number of preemptions of low-priority traffic by high-priority traffic in scenario 01 and firstly suggests that AllocTC-Sharing has an improved behaviour in relation to RDM utilization. Figure 05 also indicates that BAM switching (AllocTC-Sharing to RDM) would imply in having more preemptions in relation to keeping AllocTC-Sharing utilization for scenario 01. In effect, Figures 04 and 05 indicate different behaviours for preemption and devolution and an effective decision on the most adequate approach (switch BAM or not) could be dictated by the SLA/QoS application requirements.

Figure 6. Preemption of low-priority traffic by high-priority traffic.

The blocking behaviour for the scenario 01 is illustrated in Figure 7. In this case, LSP blocking and preemption of low-priority traffic by high-priority traffic have similar behaviour and the same considerations apply.

Figure 8 illustrate link utilization by TCs in scenario 2. For phase 1 the traffic is high for all TCs (classes) and, as such, each traffic class tries to use the reserved bandwidth till the limit defined by its BC. At the beginning of phase 1, AllocTC-Sharing is used first (less than 80% of TC0 capacity) and, using its opportunistic characteristics, TC2 “borrows” bandwidth from TC0 keeping link load by TC (Figure 8) and link utilization (Figure 9) at its maximum. Switching to RDM occurs approximately at 250 LSPs being setup and from there on link utilization by TC and link utilization reach their maximums leading to a standard behavior independent of any switching among BAMs.
A negative aspect to be observed in phase 1 is that, since TC0 (low-priority application) has a high volume of traffic, the number of devolutions (preemption of high-priority class by low-priority class) is potentially high during phase 1. Till approximately 240 LSPs setup the dynamic BAM follows AllocTC-Sharing (Figure 10). During the remaining of phase 1 (240 LSPs to 1000 LSPs approximately), the number of devolutions remains stable and it may also be observed that the dynamic BAM behavior (AllocTC-Sharing □□ RDM switching) has much less devolutions than keeping AllocTC-Sharing as the BAM for the entire phase 1.

As this point it is important to consider that the adoption of AllocTC-Sharing first has, additionally, to consider if high-priority applications at TC2 can afford “devolutions”. We consider that high-priority elastic multimedia applications, which are common, could be an example of a class of application that could deal with this BAM behavior. Any framework considering the effective implementation of this strategy should also consider what to do with the applications that are elected to “return” bandwidth. At phase 2, the switching back to AllocTC-Sharing restores BCs limits and also keeps the overall link utilization at its maximum.
In relation to parameter “preemption” (Figure 11), we observe that the dynamic switching approach follows AllocTC-Sharing behavior till switching to RDM at phase 1. Till the end of phase 1 (RDM BAM), it follows the RDM with a behavior less efficient than AllocTC-Sharing. For phase 2 and after AllocTC-Sharing is switched back, the number of preemptions follows the AllocTC-Sharing behavior, leading to a better situation when compared to RDM kept till the end of the phase. In general, the dynamic model may achieve a better performance by traffic profile and this should also be considered by the network manager whenever adopting the proposed strategy.
7. FINAL CONSIDERATIONS

Current BAMs (MAM, RDM, G-RDM, AllocTC-Sharing) have different behaviors for enforcing bandwidth allocation for traffic classes (TCs). As such, when the traffic profile of a network changes, distinct BAM utilization imply in distinct responses for applications demands for resources (bandwidth). In this context, it is argued that a strategy of switching among BAMs might be relevant for network operation.

This paper considered a preliminary evaluation of the switching characteristics for 02 specific BAMs (AllocTC-Sharing and RDM) and the simulation results obtained for a specific set of network operation and simulation parameters suggest that BAM dynamic switching may optimize network resources like link utilization and minimize preemptions and LSPs blocking. In effect, BAM operation disadvantages like preemptions, devolutions and LSP blocking might be minimized with the dynamic BAM switching strategy and the network traffic matrix may be used to dictate which parameters should be enforced according with the applications or TCs profile supported by the network.

The BAM dynamic switching may be triggered by a number of parameters such as preemptions, blocking, devolutions, and link utilization, among others. This leads to a set of possibilities for configuring BAM utilization in the network management context.

ACKNOWLEDGEMENTS

We gratefully acknowledge financial support from FAPESB (Fundo de Amparo à Pesquisa do Estado da Bahia) - Brazil to this project.

REFERENCES

Authors

PhD student in Computer Science by DMCC (UFBA/UNIFACS/UEFS), Masters in Computer and Systems by Salvador University - UNIFACS (2011) and graduation in Informatics by Universidade Católica do Salvador (2005). Professor at Instituto Federal da Bahia (IFBA). Has been working in Computer Science with emphasis in computer networks in topics like: Bandwidth Allocation Model (MAM, RDM, G-RDM, AllocTC-Sharing), MPLS, MPLS-TE, DS-TE, Autonomy, QoS, Future Internet and Software Defined Networks.


Graduate at Federal University of Paraiba (1977) - Brazil, Masters from Technische Universiteit Eindhoven (TUE)/ NUFFIC (1979) - Holland and Ph.D. from Université Pierre et Marie Curie – UPMC (1986) - France. Has extensive experience in computer science, teaching and researching on innovative areas such as Future Internet, OpenFlow & Virtualization, Quality of Service, Ubiquitous Computing and Autonomic Management. Currently, invited Professor at HTW - Hochschule für Technik und Wirtschaft des Saarlandes – Germany and Université d’Evry – France.