Performance Evaluation of MBAC solutions

Doreid Ammar, Thomas Begin, Isabelle Guérin Lassous, Ludovic Noirie

Performance Evaluation of MBAC solutions

Doreid Ammar, Thomas Begin , Isabelle Guérin-Lassous , Ludovic Noirie

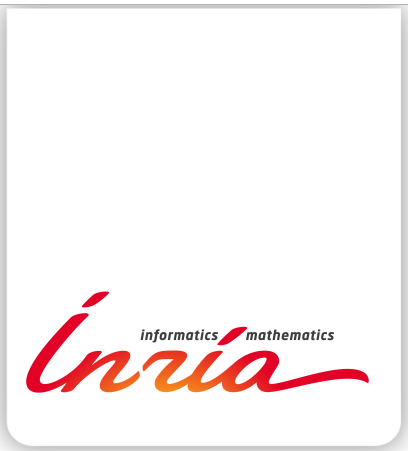
RESEARCH REPORT

N° 8080

May 2012

Project-Team RESO

ISSN 0249-6399 ISRN INRIA/RR--8080--FR+ENG



Performance Evaluation of MBAC solutions

Doreid Ammar *, Thomas Begin *, Isabelle Guérin-Lassous *,
Ludovic Noirie †

Project-Team RESO

Research Report n° 8080 — May 2012 — 19 pages

Abstract: Admission control prevents certain flows from accessing a network with regard to the current utilization level of its resources with the ultimate goal of avoiding congestion and performance collapses, so that, accepted flows receive a sufficient level of Quality of Service (QoS). In this paper, we evaluate the performance of three measurement-based admission control (MBAC) solutions in the context of *semantic networks*, which autonomously acquire a knowledge on the on-going traffic. Each MBAC solution is carefully parameterized to have identical performance target expressed in terms of maximum tolerable loss rate or either queueing delay. We also include the results that would be obtainable by an ideal admission control so as to benchmark the performance of existing MBAC. An extensive set of simulations, using different methods for representing the background traffic, viz. a Poisson process, a PPBP process and a real trace, were carried out to evaluate the performance of each solution. Our results tend to show that, in case of a target loss rate, when the characteristics of background traffic deviate significantly from a "regular" source (e.g., a Poisson process) to more variable sources (e.g., a PPBP process, a real trace), the *Aggregate Traffic Envelopes* and *Equivalent Capacity* solutions brings satisfactory results, much better than those obtained by the *Measured Sum* solution. However, if one deals with a queuing delay target, the outcomes of any MBAC solution are generally less successful, as they typically tend to deviate further from the ideal admission control, often in an overly conservative way.

Key-words: admission control, performance evaluation, QoS, measurement

This work has been partly supported by the project *Semantic Networking* within the common laboratory INRIA - Alcatel Lucent-Bell Labs.

* Université Lyon 1 / LIP (UMR ENS Lyon - INRIA - CNRS - UCBL) - Email: firstname.lastname@ens-lyon.fr
† Alcatel-Lucent Bell Labs, Nozay, France - Email: firstname.lastname@alcatel-lucent.com

RESEARCH CENTRE
GRENOBLE – RHÔNE-ALPES

Inovallée
655 avenue de l'Europe Montbonnot
38334 Saint Ismier Cedex

Evaluation des performances des solutions MBAC

Résumé : Le contrôle d'admission est un mécanisme destiné à prévenir la congestion des réseaux informatiques et à assurer ainsi à tous les flux du réseau un niveau de performances suffisant. Ce travail vise à présenter une évaluation pratique de trois solutions existantes pour le contrôle d'admission basé sur les mesures (MBAC) dans le cadre des réseaux sémantiques. Nous comparons ces trois solutions en les paramétrant de telle sorte qu'elles aient un objectif identique en termes de performances. Nous avons évalué les performances de chacune de ces solutions par simulation en les rapportant à l'oracle afin d'apprécier leurs performances. Par soucis de généralité, nous avons considéré plusieurs possibilités pour modéliser le trafic de fond : un processus de Poisson, un processus PPBP et une trace collectée sur un réseau réel. Les résultats obtenus montrent que, dans le cas d'un taux de perte toléré, lorsque les caractéristiques du trafic de fond varient, les solutions *Aggregate Traffic Envelopes* et *Equivalent Capacity* semblent être plus satisfaisantes que la solution *Measured Sum*. Cependant, dans le cas d'un délai d'attente toléré, les résultats des solutions MBAC sont généralement moins satisfaisants car ils ont tendance à s'écarter de l'oracle et sont souvent très conservateurs.

Mots-clés : contrôle d'admission, évaluation de performances, QoS, mesures

1 Introduction

Over the last few years, network operators, both wireless and wireline, have faced a rapid growth in the volume of data that are being exchanged over their network infrastructure. New applications, which may be network-intensive, time-sensitive, and may benefit from intensive use, have continued to flourish. It seems very unlikely that this growth will slow down anytime soon. The "Fall 2011 Global Internet Phenomena Report" recenly reflects that, within wired networks in the United States, Real-Time Entertainment applications are the primary drivers of network capacity requirements, accounting for 60% of peak downstream traffic, up from 50% in 2010 [**?**]. In particular, this report revealed that Netflix traffic accounted for nearly 30% of peak downstream traffic during peak periods, a relative increase of more than 10% since Spring 2011.

However, this steady increase in the amount of data being uploaded or downloaded has significantly affected the utilization of networking resources, bringing them to possibly too high levels. This surge might, ultimately, be a major factor for significant network congestions and performance disruptions. A case in the point is the traffic collapse that occurred during summer 2010 on AT&T wireless access networks [**?**], where available bandwidth is known to be a scarce resource.

To cope with this increasingly demand for bandwidth, operators can either deploy new resources or improve the use of their existing resources. As operators may be reluctant to grant additional resources, they may be more inclined to improve the actual configuration of their network and implement new management strategies. Network management options include a large choice of possible policies such as congestion control, scheduling algorithms, traffic shaping, admission control. In this paper, we focus on admission control.

Admission control is a mechanism used to prevent certain flows from accessing a computer network with regard to the current utilization level of the network resource. By regulating the number of on-going flows, admission control aims at preventing overloading, congestion and performance collapses, so that, accepted flows receive a sufficient level of Quality of Service (QoS), which is of utmost importance for delay-sensitive applications (*e.g.*, Telephony over IP) and resource-intensive applications (*e.g.*, streaming video). There are different approaches to perform admission control. First, endpoint admission control solutions make use of probing packets that aim at reproducing the traffic pattern that the source is on the verge to transmit through the network [6]. This approach is referred to as an active technique since artificial traffic is injected into the network to perform admission control. Network operators usually discard this approach for several reasons: (i) Generating traffic that shares close characteristics with the original is everything but an easy task for an operator; (ii) Analyzing the distortion on the injected traffic reflects the state of the network over a very limited length of time, corresponding to the time during which the probe packets were sent. Second, admission control solutions can be based on the use of traffic descriptors. The underlying idea primarily consists in theoretically assessing the current network workload using traffic descriptors. Then, the admission control uses the found value to decide, given the incoming flow traffic descriptor, whether or not to let it come into the network. Clearly, such an approach requires to know traffic descriptors for every on-going (accepted) flow as well as for any incoming flow [12]. Operators generally deprecate this approach because acquiring and maintaining the knowledge on traffic descriptors represent both a costly and difficult task. Third, measurement-based admission control (MBAC) solutions rely exclusively on measurements to assess the workload of on-going traffic over each communication link. Unlike the first type of solutions, these solutions are categorized as passive techniques. MBAC solutions differ from the second type of solutions since they do not require any explicit knowledge on the traffic descriptors of on-going flows. Because of its simplicity and its neutral impact, MBAC appears as an attractive approach for network operator. It is therefore this

approach that we consider in the remainder of this paper.

Several MBAC solutions have been proposed in the literature. These solutions are generally thought to operate on a single communication link, and the admission control must be repeated for each link along the path of the flow. These solutions are basically made up of two parts. First, they perform measurements on the on-going traffic, and deliver measured metrics (*e.g.*, the residual capacity of the link). Second, they rely on an algorithm that includes a test operation, whose outcome decides whether or not to let a new flow requesting admission come into the network. In its simplest form, the algorithm can simply check that the rate requested by the incoming flow is less than the residual link capacity. Existing MBAC solutions mainly differ by their measurement operations and by the theoretical assumptions made on the on-going traffic.

In this paper, we focus our study on admission control in the context of semantic networks [**?**]. *Semantic networks* refer to computer networks that autonomously acquire a knowledge on the on-going traffic. They analyze the features of the transmitted traffic at the flow granularity and exploit this knowledge to dynamically adjust their behavior. In the context of admission control, not only does the network acquire knowledge on the characteristics of on-going traffic, as it would be the case for any MBAC, but it also gets knowledge on any new incoming flow requesting admission (thanks to the inspection of its first packets). It is this paper goal to evaluate and to compare the performance of three existing MBAC solutions in the context of semantic networks.

The originality of this work is twofold. First, as opposed to previous comparison studies [12, 5, 14, 13], we do not assume any explicit knowledge, neither on incoming flows nor on on-going traffic. To this end, we introduce a method to estimate the peak rate of an incoming flow based on its first transmitted packets[1]. Second, we carefully parameterize each MBAC solution in a way that leads to identical target in terms of performance and we compare their efficiency. The selected target is alternatively the maximum tolerable packet loss rate or the maximum acceptable packet queueing delay.

The remainder of this paper is organized as follows. Section 2 relates the state-of-the-art on MBAC solutions. In Section 3, we detail the admission control solutions which are investigated in our study. Our experimental framework is presented in Section 4. Section 5 is devoted to the numerical results. Finally, Section 6 concludes this paper.

2 State-of-the-Art

This state-of-the-art is restricted to measurement-based admission control solutions (MBAC) since we consider only this approach in this paper. *Guerin et al.* were the first to introduce in [9] the approach of *Equivalent Capacity* used in several MBAC solutions. The Equivalent Capacity of aggregated traffic over a communication link, $C(\epsilon)$, is defined as being such that the probability for the arrival data rate of aggregated traffic to exceed $C(\epsilon)$ is at most ϵ. Basically, any MBAC solution based on Equivalent Capacity attempts to ensure that, for any link on the path between the source and the destination, the rate of the flow requesting admission summed to the actual Equivalent Capacity keeps below the nominal link capacity. The formula for the Equivalent Capacity given in [9] assumes a buffer-less model and an aggregate arrival rate that follows a Normal distribution. Floyd proposed in [7] an alternative formula for the measurement of Equivalent Capacity based on Hoeffding bounds. In [8], *Georgoulas et al.* used the formula of the Equivalent Capacity given in [9], but they include an Admission Policy Factor in their admission control algorithm that allows the operator to tune its degree of conservativeness in terms of packet loss rate. These three latter solutions require measurements only on the utilization rate of each

[1] This property implies that a flow can be rejected even though its first packets were transmitted.

communication link to be run. In [11], *Jamin et al.* were the first to integrate in their admission control the queueing delay constraint. To be performed, this solution requires, in addition to a measurement on the actual utilization rate of the link, a measurement on the waiting time being spent in the queue (buffer). Specifically, their admission control algorithm consists of two tests: a test on utilization and a test on delay. *Qiu and Knightly* propose to improve in [15] the works of *Jamin et al.* by proposing an alternative measurement of the utilization rate of the link in order to have a better traffic characterization over this link. To do this, the authors introduce the notion of aggregate traffic envelopes. It is worth noting that all the solutions mentioned above were designed and evaluated assuming an explicit knowledge on the peak rate of incoming flows requesting admission. In some cases, this peak rate is derived assuming the existence of a token bucket mechanism.

Former studies compare these different solutions. In [12], a comparison of three MBAC solutions was performed using simulation. The results show that a simplified version of *Jamin et al.* solution [11] (without incorporating the delay constraint) is more likely to achieve a higher link utilization inducing a small loss rate of packets, as opposed to *Floyd* solution [7] which leads to a lower utilization rate of the link but with no loss of packets. *Breslau et al.* compare in [5] several MBAC solutions performed under NS-2. By considering a range of values for the parameters of the investigated admission controls, the authors show that all these solutions yield to the same performance frontier, which defines the trade-off between utilization and loss rate. The authors conclude that the main difficulty associated to MBAC consists in calibrating correctly their parameters so as to achieve the target trade-off between utilization and loss rate. *Nevin et al.* compare in [14] three admission control solutions to an ideal admission control using various traffic conditions. They found that none of the three measurement-based admission control solutions is able to meet the QoS targets. *Moore* compares in [13] a subset of MBAC solutions under a purpose-built test environment. He highlights the impact of the admission decision upon heterogeneous traffic systems. The results show that only the aggregate traffic envelopes solution [15] achieves acceptable results.

Despite the number and the variety of tested scenarios from the previously mentioned works, virtually all of them, if not all, assume that the incoming traffic requesting admission is smoothed using a fully described token bucket. In our study, we assume no explicit knowledge on incoming flows requesting admission. Moreover, very few works explain how to parameterize the different admission control solutions. In our study, we carefully tune the parameters of the different tested solutions with the goal of providing a given QoS to each accepted flow.

3 Investigated MBAC Solutions

In our study, we investigate three MBAC solutions. Note that all these solutions assume to know the peak rate of each new flow requesting admission. We denote by r the peak rate of an incoming flow and by C the nominal capacity (transmission speed) of a communication link. In Section 4.5, we detail a simple technique to evaluate r.

3.1 Measured Sum (*M.S.*)

Jamin et al. present in [12] a MBAC solution based on the measured load of existing traffic over the link, denoted by R. This solution admits a new flow requesting admission, with a peak rate r, if and only if:

$$R + r \leq \nu C, \tag{1}$$

where ν is a parameter that defines the targeted link utilization.
The measured load of existing traffic is updated every measurement window of length T. This time window is split into smaller sampling periods of equal length. The average rate of existing flows is then computed on every sampling period. At the end of a measurement window, R is defined as the highest average rate seen in the sampling periods constituting this time window. However, the value of the measured link load may be updated within a measurement window for two reasons: whenever an individual average rate on a sampling period exceeds the current link load of the measurement window or whenever a new flow is admitted, the value of the measured network load is then updated with the value of the average rate of the sampling period or with the peak rate of the new admitted flow added to the current load respectively. Note that the measured loads on the sampling periods are always stored and used to compute the average load at the end of a measurement window.

Jamin et al. introduce a delay test to their admission control solution. The measured delay, denoted by $\widehat{D}$, tracks the maximum queueing delay of every packet computed over a time window of length T. The solution rejects an incoming flow requesting admission if admitting this new flow violates the following constraint:

$$D > \widehat{D} + \frac{b_i}{C}, \tag{2}$$

where D is the delay bound and b_i is the burstiness of the flow (see details in [12]). The value of $\widehat{D}$ is updated at the end of each measurement window. Whenever an individual delay measurement exceeds the estimated maximum queueing delay, the value of $\widehat{D}$ is also updated to be λ times this sampled delay. Finally, we update the measured delay to the right side of (2), whenever a new flow is admitted.

3.2 Equivalent Capacity (*E.C.*)

In [7], Floyd presents an admission control solution based on the estimation of the *Equivalent Capacity* of the link for a set of aggregated flows. A new flow is accepted if the sum of the peak rate r, requested by this flow, and the *Equivalent Capacity* of the link, $C(\epsilon)$, is less than or equal to the capacity of the link C. More formally, this condition is expressed as:

$$C(\epsilon) + r \leq C \tag{3}$$

The critical point of this method relies on the estimation of the *Equivalent Capacity*, $C(\epsilon)$. In our case study, we chose the formula given in [9] because it is easier to use in the context of *semantic networks*. The *Equivalent Capacity* proposed in [9] is a linear function of the average rate of aggregate traffic and its standard-deviation, denoted by $\widehat{r}$ and σ, respectively. This function is given by:

$$C(\epsilon) = \widehat{r} + \alpha.\sigma, \text{with } \alpha = \sqrt{2\ln\frac{1}{\epsilon} + \ln\frac{1}{2\pi}}, \tag{4}$$

where ϵ is the probability that the arrival rate exceeds the expected *Equivalent Capacity*.
In order to compute the average arrival rate of aggregated traffic, $\widehat{r}$, *Floyd* suggests to define it as an exponential-weighted moving average with a weight ω updated after each measurement window T. The average arrival rate could then be calculated using: $\widehat{r} \leftarrow (1-\omega).\widehat{r} + \omega.R$, where R is the average rate of the aggregated traffic measured every measurement window T and ω is a real number between 0 and 1. Since nothing was recommended by the authors about the computation of the standard-deviation σ, we chose to compute the value of σ from the M previous measured values of R.

3.3 Aggregate Traffic Envelopes (*Env.*)

Qiu and Knightly present in [15] a MBAC solution that aims to characterize the aggregate traffic rate by the maximal rate envelope. To do this, they consider a time window of length T divided into t sampling periods of equal length. Within a time window, maximal rate measurements are done on different time scales. R_k^m represents the maximal observed rate in the time scale k. This time scale is equal to k sampling periods in the m^{th} measurement window. The rate of the aggregate traffic and its standard-deviation are estimated over the last M measurement windows as follows:

$$\overline{R}_k = \sum_{m=1}^{M} \frac{R_k^m}{M} \text{ and } \sigma_k^2 = \frac{1}{M-1} \sum_{m=1}^{M} (R_k^m - \overline{R}_k)^2. \quad (5)$$

This measurement-based admission control solution consists of two parts: a short time scale test that ensures that no packet is too long delayed, and a long time scale test that checks that the flow requesting admission does not exceed the link capacity. Note that envelopes are used only to check the first condition. A new flow requesting admission with a peak rate r is accepted if and only if:

$$\max_{k=1,\ldots,t} \{k\tau(\overline{R}_k + r + \alpha_E \sigma_k - C)\} \leq C \times D \quad (6)$$

and

$$\overline{R}_t + r + \alpha_E \sigma_t \leq C \quad (7)$$

where D is the maximum delay requirement and α_E is a constant specifying the confidence level, $\Phi(\alpha_E)$, that on-going flows do not experience any packet loss. $\Phi(\alpha_E)$ is defined as:

$$\Phi(\alpha_E) \approx \frac{1}{\sqrt{2\pi}\sigma_k} \int_{-\infty}^{\overline{R}_k + \alpha_E \sigma_k} \exp\left(-\frac{(r-\overline{R}_k)^2}{2\sigma_k^2}\right) \mathrm{d}r. \quad (8)$$

4 Methodology

The performance evaluation of an admission control can be handled through different aspects. One can consider the overhead costs for network nodes in terms of CPU time or memory consumption, the ease of configuration, the quality of the decision, etc. Many studies comparing admission controls ([12, 5, 14]) aimed to quantify, for a given experimental scenario, the attained level of utilization of the link versus the packet loss rate. As stated in [5], the results tend to show that the different tested admission control solutions achieve nearly the same behaviors. This can be explained by representing the performance of the link as being those of a queueing model with a single server and a finite buffer in which the inter-packets arrival times and the service times follow arbitrary processes (*i.e.*, a *G/G/1/K* queue). Then, the relation between the attained level of utilization and the loss probability is necessary to be the same for any admission control.

Therefore, in this paper, we focus our work on the tuning of the parameters of the admission control with the goal of providing a given QoS to each accepted flow. More specifically, we aim at highlighting the ability for each of the three MBAC presented above to achieve the maximum level of utilization of the link, while respecting a given target in terms of performance. In this work, we choose alternatively the loss rate and the queueing delay as target. We consider the following values for the target loss rate, Pr, and for the target queueing delay, D, namely $Pr = 10^{-2}$ and $D = 10$ms. Note that we did not consider the case where a target loss rate and a target queueing delay are both specified since, following the queueing theory analysis given above, meeting a target loss rate implies an unknown but fixed target queueing delay, and vice versa.

4.1 Scenarios

We consider a communication link of capacity, C, set to 10 Mb/s. The size of the buffer is set to 20ms (corresponding to a queue size of nearly 130 packets of size 190 bytes) when we deal with the target loss rate, and to 60ms when we study the queueing delay target. The queueing discipline is FIFO (*First In First Out*) and the queue management algorithm is Drop-Tail.

Each incoming flow that requests access to the communication link will generate variable bit rate (VBR) traffic. Departures times of its packets are determined as follows: with a probability p, the next packet departure is scheduled t_p milliseconds later after the previous packet, and with a probability q = 1 - p, the next packet departure occurs t_q milliseconds later. Overall, the average sending rate of each VBR flow is given by:

$$\overline{r} = \frac{p}{t_p} + \frac{q}{t_q} \tag{9}$$

In our experiments, we select $p = 0.95$, $t_q = 28 \times t_p$ and a constant packet size equal to 190 bytes. Hence, each VBR flows will generate packets with an average sending rate $\overline{r}$ of 64 kb/s and a coefficient of variation equal to 2.5 (remind that it is 0 for a CBR flow and 1 for a Poisson source).

The VBR flows arrive randomly to the communication link according to a Poisson process with a constant rate, denoted by γ. Their durations are decided by an exponentially distributed random variable with mean l_{vbr}. Then, if no admission control were to be performed, the cumulated sending rate of VBR flows would be equal to:

$$\Lambda_{vbr} = \overline{n}.\overline{r} \tag{10}$$

where $\overline{n} = l_{vbr}.\gamma$ (see Little's law [1]) represents the average number of VBR flows over the communication link (without admission control policy). We choose l_{vbr} = 120 s and γ = 0.717 arrivals per second. Hence, we have: Λ_{vbr} = 5.5 Mb/s.

Clearly, the outcome of any admission control regarding an arriving flow requesting access is highly tied to the characteristics of the aggregate traffic currently traversing the link. Hereafter, we refer to this aggregate traffic as the *background traffic*. As it is well-known, the statistical properties of flows may vary considerably according to the type of networks, to the observation location and to the nature of transmitted application. Thus, when it comes to the question of how representing the background traffic, no simple and universal answer should be expected. In the related work ([12, 5, 14, 13]), authors often consider that the background traffic is simply determined as the aggregation of the previously accepted flows, each of them being usually identically distributed. In their work [5], *Breslau et al.* consider an aggregation of flows with heterogeneous characteristics, but all these flows are smoothed by the same token bucket. On the other hand, in this paper (i) we deal with various conditions for the background traffic, and (ii) we define it as a two-layered process. More specifically, the background traffic consists of an initial background traffic, to which is summed up the aggregation of VBR flows accepted by the admission control (see Figure 1). This initial background traffic sent without admission control can represent, for instance, priority traffic or VPN traffic under no or limited control. We choose to represent the initial background traffic by three different processes:

1) Initial background traffic based on a Poisson process

Initially, the background traffic delivers packets of length 190 bytes according to a Poisson process with a rate of 4.5 Mb/s.

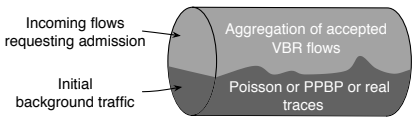


Figure 1: Background traffic conditions over the communication link

2) Initial background traffic based on a PPBP process

The PPBP (*Poisson Pareto Burst Process*) process [**?**], represents the cumulated behavior of infinite and independent heavy-tailed On/Off sources. The On durations are chosen from a Pareto distribution with mean 200 ms and a *Hurst parameter*, $H = 0.7$. While being active, each source delivers a CBR traffic with a fixed rate of 1 Mb/s and a packet size of 190 bytes. Hence, the PPBP process also generates packets at a rate of 4.5 Mb/s.

3) Initial background traffic based on a real traffic trace

The traffic trace was gathered by the University of Stuttgart [16] on Sunday October 31^{st} 2004, between 6pm and 10pm, on a 100 Mb/s link in the dormitory network "Selfnet". In our case, we used only the first 15 minutes of this trace and we adjust it to a 10 Mb/s link by scaling it down such that its average rate of transmitted packets is equal to 4.5 Mb/s.

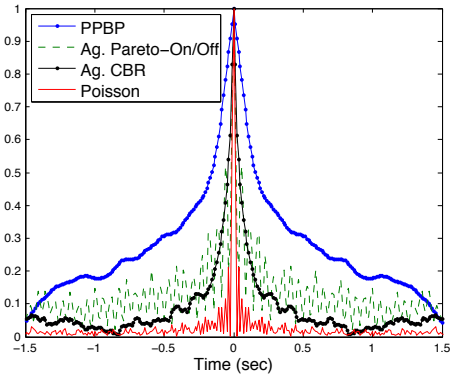


Figure 2: Autocorrelation function for a Poisson process, a PPBP process, a superposition of 100 independent CBR flows and a superposition of 20 independent Pareto On/Off flows

Although these three processes have equal values with regard to their average packet sending rate, other statistical properties (*e.g.*, variance, degree of autocorrelation) may significantly differ. As shown in Figure 2, the degree of autocorrelation for a Poisson process is null, moderate for an aggregation of 100 independent CBR flows as well as for an aggregation of 20 independent Pareto-On/Off flows, and large for a PPBP process. Note that we did not represent the degree of autocorrelation for the real trace since its value fluctuates greatly over time. Using various types of models for the background traffic allows us to investigate in a broader framework than previous studies how each admission control solution performs.

4.2 Why admission control is required?

As said above, the initial background traffic has an average rate of 4.5 Mb/s. On the other hand, as shown in Equation (10), if no admission control were to be performed, the cumulated sending rate of VBR flows would be of 5.5 Mb/s. In our scenario, the goal of admission control is then to limit the number of VBR flows so as to keep the total rate of all combined traffics at the "right" level (below the link capacity of 10 Mb/s), and thus preventing packets from experiencing excessive queueing time in the buffer and exceedingly high levels of loss.

4.3 Calibration of the admission control algorithms according to a target loss rate, Pr

We now detail the configuration of the investigated admission controls. As said before, we calibrate their parameters such that all of them have an identical target in terms of maximum tolerable packet loss rate. We let Pr denote this target loss rate.

4.3.1 Measured Sum

The authors of the Measured Sum algorithm did not provide specific guidelines for selecting the value of ν. Obeying to the analysis principle proposed in [12], we choose the value of ν as equal to the ratio of the average packets arrival rate to the average transmission (service) rate so that the link modeled by a corresponding $M/M/1/K$ queue, with K set to 131 packets, leads to a packet loss rate equal to Pr.

4.3.2 *Equivalent Capacity*

The authors denote by ϵ the probability that the instantaneous arrival rate of the background traffic, modeled by a Normal distribution, exceeds the *Equivalent Capacity* of the communication link, $C(\epsilon)$. To link the value of ϵ to the value of Pr, we proceed as follows. Assuming that the probability ϵ also represents the steady-state probability of having the buffer full (which would be the case for a buffer length of 1), and assuming that the steady-state probabilities are the same as the probability of the state seen by an arriving packet (which would be the case if the incoming flows were Poisson, see the PASTA property [1]), then, ϵ would also be the probability for an incoming packet to be rejected, namely Pr. Based on this rationale, we select ϵ equal to Pr, and thereby computing the value of α.

4.3.3 *Aggregate Traffic Envelopes*

The selected value for the confidence level, α_E, determines the expected probability that on-going flows do not experience any packet loss, $\phi(\alpha_E)$. To choose the value of α_E, we simply associate the value of $\phi(\alpha_E)$ to the target packet loss rate for accepted flows, namely Pr.

We report in Table 1 the numerical values selected for the tested MBAC algorithms.

4.4 Calibration of the admission control algorithms according to a target queueing delay, D

We describe here how we parameterize the admission controls according to a target queueing delay. Note that *Equivalent Capacity* is obviously out of this section as it does not provide a control on the packet delay. Recall that D denote the target maximum tolerable queueing delay over a single communication link (see formulas (2) and (6)).

Table 1: Summary of the parameters involved in the each MBAC solutions

		Measured Sum	**Equivalent Capacity**	**Aggregate Traffic Envelopes**
Measured quantities	**Aggregated rate**	R	$\widehat{r}$	$\overline{R_k}$ $(k = 1, \dots t)$
	History	Single measurement window	Exponential moving average, $\omega = 0.1$	Last 20 measurement windows
	Standard-deviation	-	σ	σ_k $(k = 1, \dots t)$
	History	-	Last 20 measurement windows	Last 20 measurement windows
	Estimated delay	$\widehat{D}$	-	-
	History	Single measurement window	-	-
	Measurement window	$T = 4$s	$T = 200$ms	$T = 200$ms
		200ms *Sampling periods*		10ms *Sampling periods* $(t = 20)$
Calibrated parameters	**Target loss rate**			
	$Pr : 10^{-2}$	$\nu = 0.9543$	$\alpha = 2.7152$	$\alpha_E = 2.325$
	Target queueing delay			
	$D : 10$ms	$\lambda = 2$	*unavailable*	$\alpha_E = 3.6$

4.4.1 Measured Sum

The value of λ aims at tuning the stringency level of the admission control. The greater λ, the more conservative the admission control is, and the less accepted flows. As no specific guidelines are given by the authors of *Measured Sum* for setting the value of λ, we let λ be equal to the most favorable value that we observed in several experiments, namely $\lambda = 2$.

4.4.2 *Aggregate traffic envelopes*

There is no clear recommendation from the authors on the choice for α_E. Therefore, we set it to the most favorable value among couple of experimented values, namely $\alpha_E = 3.6$.

Table 1 relates the parameter values selected for the MBAC algorithms tested in the case of a target queueing delay.

4.5 Estimating the peak rate of incoming flows

As said previously, we focus our studies in the context of *semantic networks*. In such networks, the network acquires knowledge on flows by itself via an analysis of the on-going traffic. With such an approach, the *a priori* knowledge on the flow peak rate is not necessary. This *a priori* knowledge can be obtained via signaling and/or the use of a tocken bucket. Tocken buckets are difficult to parameterize and may induce conservative results for the admission control (since the decision algorithm uses a conservative value for r). Our approach is more simple and does not need any signaling as it is only based on data packets. In this section, we detail the procedure we implement to let the network estimate the peak rate of a new flow requesting admission.

To estimate the peak rate of a new incoming flow, we track the first n packets of this flow. We use a sliding window of length equal to k packets. For every possible window on the first n packets, we compute the average rate. Finally, the peak rate corresponds to the highest value among the $(n - k + 1)$ windows. With this approach, flows consisting of less than n packets (often called *mice* flows) are always admitted and only flows with more than n packets (often called *elephant* flows) can be rejected.

In this work, the estimated peak rate of an injected flow is computed based on the 20 first packets ($n = 20$) with a sliding window of length equal to 5 packets ($k = 5$). Note that in our experiments, the VBR flows may achieve a maximal rate of 150 kb/s.

4.6 Ideal admission control

At this point the *ideal admission control*, which will serve as a benchmark to compare the outcome of each MBAC solution, can be clearly defined. It should accept the maximum number of flows, thus achieving the maximum utilization rate, while successfully meeting the QoS target (i.e., no false positives and no false negatives).

Given the huge number of flows coming into the link during the simulation time (more than 630), any exhaustive approach that will consider every feasible combinations of accepted / rejected flows will lead to approximately $2^{630} \simeq 10^{190}$ possible sequences, and thus would be intractable. We rather rely on an iterative method to determine the sequence of flows accepted by the ideal admission control under the policy *First come, First served* (if the flow does not violate the QoS target). At iteration (i), k flows have been accepted (some of them may still be going on) and j have been refused. As soon as a new flow will arrive, we will accept it, and then we will keep the simulation running until this flow ends but, meanwhile, any subsequent VBR flow will be refused. Once the flow is done, we check whether the QoS target was preserved for this flow as well as for any previously accepted flow. If this is the case, then we grant this flow as acceptable by the ideal admission control and the value of k is incremented. Otherwise, the flow will not be part of the sought sequence of flows and j is incremented.

Table 2: Summary of solutions performance with Poisson process for the initial background traffic

	Measured Sum	**Equivalent Capacity**	**Aggregate Traffic Envelopes**	**Ideal Admission Control**
Target loss rate				
Number of accepted flows	236	203	210	247
Percentage of accepted flows	37.3%	32.1%	32.2%	39.1%
Mean loss rate over the simulation	4×10^{-4}	$< 10^{-8}$	$< 10^{-8}$	4.1×10^{-3}
Percentage of violation	1.3%	0%	0%	0%
Target queueing delay				
Number of accepted flows	228	-	207	245
Percentage of accepted flows	36.1%	-	32.7%	38.8%
Mean queueing delay over the simulation	3.01ms	-	0.59ms	5.15ms
Percentage of violation	4.9%	-	0%	0%

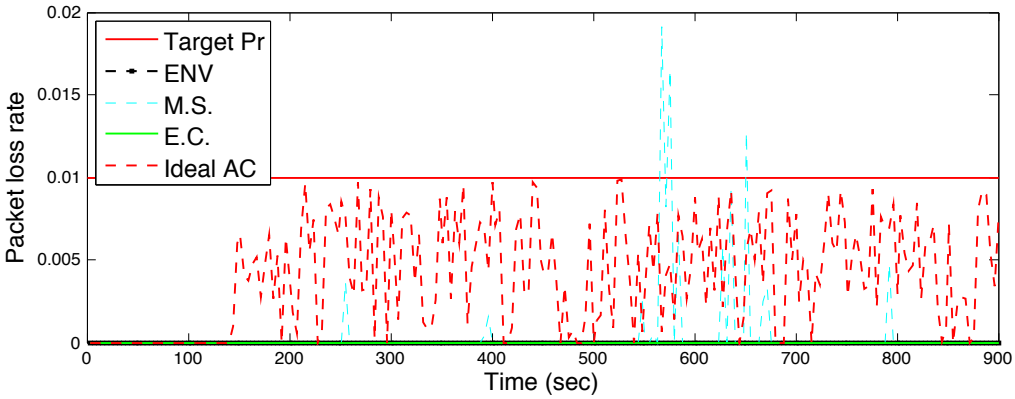
5 Performance comparison

In this section, we evaluate the performance of the three MBAC solutions using ns-3 simulations. Each simulation is run for a period of 15 minutes.

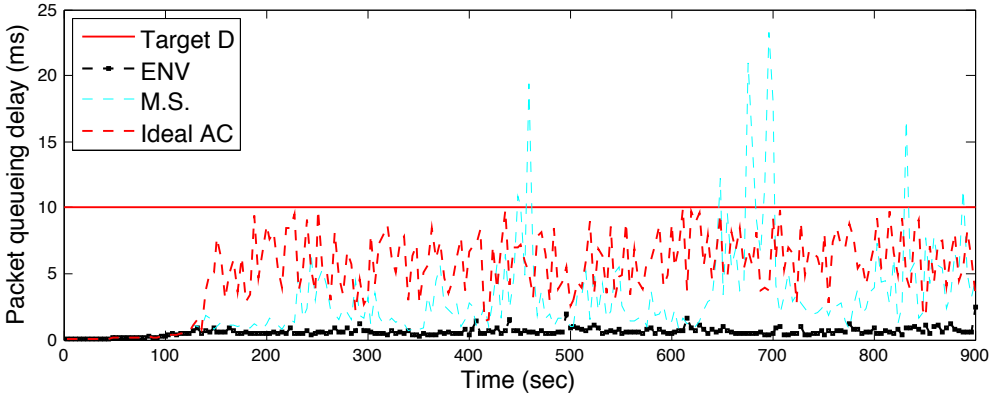
As explained before, new VBR flows will randomly request admission to the link. Then it is up to the MBAC solution to decide whether the flow is accepted or not. If a flow is accepted, it keeps transmitting packets until it ends. Incoming flows are rejected when the MBAC solution assumes that the target loss rate Pr (the target queueing delay D, respectively) will be violated if the flow workload would be summed up to the on going traffic. Let us remind that in our experiments the target loss rate Pr is set to 10^{-2} and that the maximum tolerable queueing delay D is set to 10ms.

In our results we consider two time scales: (i) A short time scale that reflects the “instantaneous"

values of the measured performance parameters. These values are computed on a sliding window of length equal to 4s; (ii) A long time scale used to relate the percentage of accepted flows, the percentage of violation that represent the ratio of time during which the QoS target is violated and the overall values of the performance parameters. These latter values are then computed over the entire duration of the simulation, viz. 15 minutes. We purposely consider these two time scales to highlight the fact that a QoS fulfillment on a large time scale does not necessarily imply its fulfillment on a shorter time scale.



(a) Target packet loss rate $Pr = 10^{-2}$



(b) Target queueing delay $D = 10$ms

Figure 3: Poisson process for the initial background traffic

5.1 Initial background traffic based on a Poisson process

Figure 3 and Table 2 relate the experimental results obtained for a Poisson initial background traffic over a short time scale and a long time scale, respectively.

Figure 3(a) represents the instantaneous packet loss rate obtained by each MBAC solution. Both the *Aggregate Traffic Envelopes* and the *Equivalent Capacity* solutions lead to steadidly and excessively low levels of loss rates, several orders of magnitude below Pr. On the other hand, the *Measured Sum* solution appears to be less conservative and overall emerged as the best solution even if it may, in some cases, temporarily violate the target loss rate (less than $1,5\%$ of time).

It is also worth noting that Figure 3(a) clearly confirms that our implanted ideal admission control keeps the instantaneous loss rates below the target loss rate while approaching it very nearly.

Figure 3(b) shows the instantaneous packet queueing delay with regards to the target delay D. Overall, these results are in line with the previous results. The *Aggregate Traffic Envelopes* solution exceedingly fulfills the target queueing delay, and appears to be overly conservative. On the contrary, the *Measured Sum* solution comes closer to the ideal admission control.

Table 2 indicates the overall performance for each admission control solution. First, it states that the ideal admission control can accept up to 247 flows in case of target loss rate and up to 245 for the target queueing delay. Second, it states that the *Measured Sum* solution accepts around 10 flows less than the ideal admission control, while both the *Aggregate Traffic Envelopes* and the *Equivalent Capacity* solutions respectively accept around 35 and 40 flows less.

Overall, in the case of a Poisson process for the initial background traffic, the *Measured Sum* solution tends to significantly outperform other investigated solutions (even if in some very few cases, the QoS target is violated).

5.2 Initial background traffic based on a PPBP process

Table 3: Summary of solutions performance with PPBP process for the initial background traffic

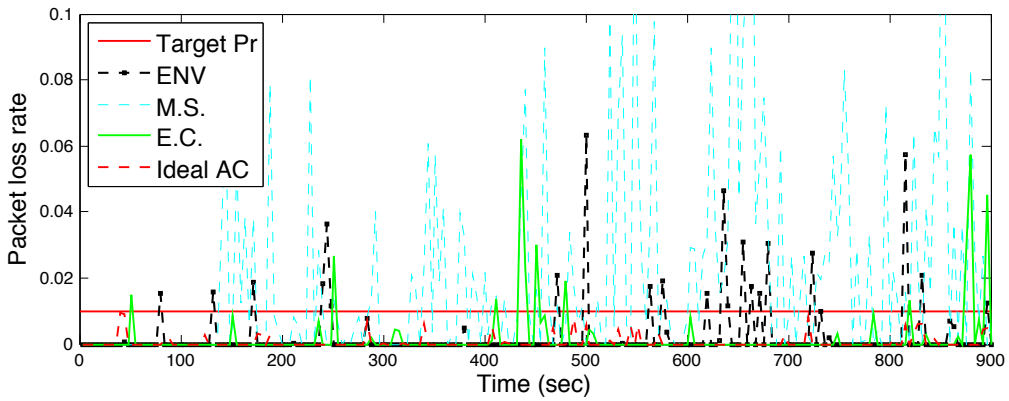
	Measured Sum	**Equivalent Capacity**	**Aggregate Traffic Envelopes**	**Ideal Admission Control**
Target loss rate				
Number of accepted flows	209	116	138	117
Percentage of accepted flows	33.1%	18.3%	21.8%	18.5%
Mean loss rate over the simulation	2.2×10^{-2}	1.8×10^{-3}	2.5×10^{-3}	6×10^{-4}
Percentage of violation	43.6%	4.9%	9.3%	0%
Target queueing delay				
Number of accepted flows	152	-	62	125
Percentage of accepted flows	24.1%	-	9.8%	19.8%
Mean queueing delay over the simulation	7.59ms	-	0.96ms	0.98ms
Percentage of violation	29.3%	-	0.9%	0%

We now consider that the initial background traffic is made of a PPBP process. As said before, the PPBP process has been recognized to share fundamental statistical characteristics with Internet traffic such as a potential high degree of autocorrelation.

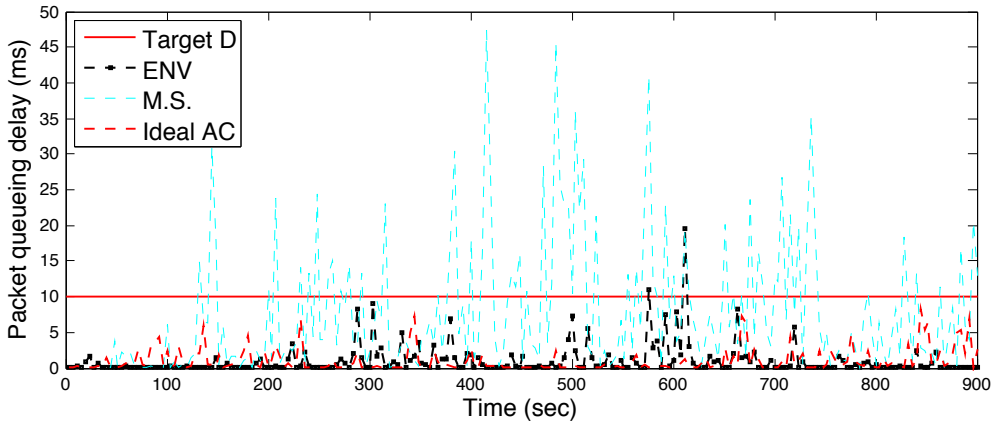
Figure 4 and Table 3 report the corresponding results yielded by simulations.

For the target loss rate, Figure 4(a) clearly states a clear advantage for the *Equivalent Capacity* and the *Aggregate Traffic Envelopes* solutions over the *Measured Sum* solution, even though their performance can be viewed as relatively moderate. When considering the target queueing delay, this gap comes even wider (see Figure 4(b)). The performance of the *Aggregate Traffic Envelopes* solution are excellent (almost constantly meeting the target delay and exhibiting a behavior close to the ideal admission control). On the other hand, the *Measured Sum* solution tends to deviate very frequently beyond the target threshold.

We now turn to Table 3. First, with regard to the target loss rate, Table 3 indicates that, apart from the *Measured Sum* solution, others MBAC solutions lead to a number of accepted VBR flows close to the one delivered by the ideal admission control. Second, when dealing with a target queueing delay, these results show that that the *Measured Sum* solution tend to accept too many flows while the *Aggregate Traffic Envelopes* solution can be considered as overly conservative. These latter results are in line with those previously described for Figure 4(a). Finally, it is worthwhile noting that, when the



(a) Target packet loss rate $Pr = 10^{-2}$



(b) Target queueing delay $D = 10$ms

Figure 4: PPBP process for the initial background traffic

performance parameters are averaged over the long time scale (i.e., the whole simulation time) as this is the case in Table 3, they almost all fall below the target thresholds. This could lead to misleading interpretations. Conversely, Figure 4 which reports instantaneous values (computed on shorter time scale), clearly exhibits frequent target violations for some investigated solutions, as the percentage of violation in Table 3 also suggests.

5.3 Initial background traffic based on a real traffic trace

Finally we consider the case where the initial background traffic corresponds to a trace collected on a real network. Figure 5 and Table 4 relate the results. Interestingly, Figure 5(a) shows that, when dealing with a target loss rate, both the *Aggregate Traffic Envelopes* and the *Equivalent Capacity* solutions fairly approximate the behavior exhibited by the ideal admission control. On the contrary, the *Measured Sum* solution leads to much poorer results as it frequently leads to instantaneous loss rates much above the target threshold (around 80% of time). In fact, this latter solution accepts a too large number of flows (see Table 4).

If the target is expressed as a queueing delay limit, the considered solutions perform almost identically, and both can be viewed as overly conservative. As shown by Figure 5(b), the delay experienced by packets in the queue remains at levels significantly lower than the ideal admission control. In fact, Table 4 shows

Table 4: Summary of solutions performance with real traffic traces for the initial background traffic

	Measured Sum	**Equivalent Capacity**	**Aggregate Traffic Envelopes**	**Ideal Admission Control**
Target loss rate				
Number of accepted flows	256	163	178	191
Percentage of accepted flows	40.5%	25.8%	28.2%	30.2%
Mean loss rate over the simulation	2×10^{-2}	2.6×10^{-3}	3.6×10^{-3}	4.1×10^{-3}
Percentage of violation	80%	1.3%	4%	0%
Target queueing delay				
Number of accepted flows	10	-	0	197
Percentage of accepted flows	1.58%	-	0%	31.2%
Mean queueing delay over the simulation	1.31ms	-	1.25ms	6.01ms
Percentage of violation	0%	-	0%	0%

that the *Aggregate Traffic Envelopes* refused every incoming VBR flows, while the *Measured Sum* merely accepted 10 flows. These results are in clear contrast with the ideal admission control that accepts up to around 200 flows during the simulation time.

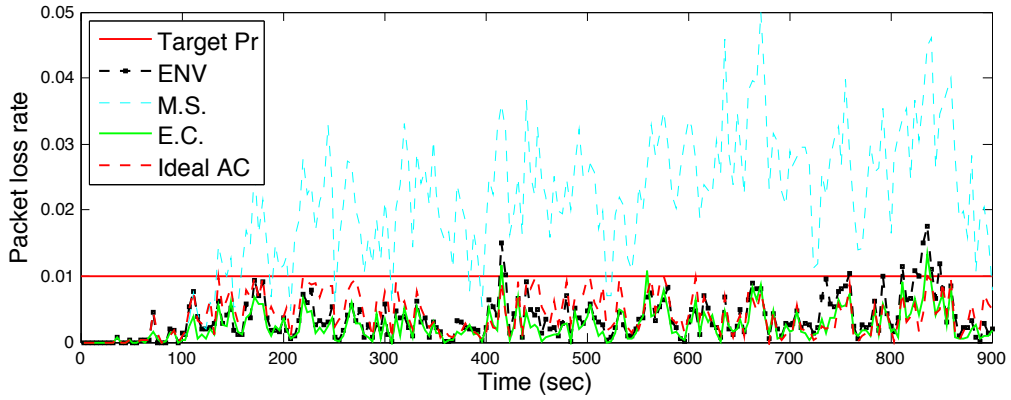
6 Conclusion

In this paper we compared the performance of three existing MBAC solutions in the context of semantic networking. Each MBAC solution was carefully parameterized to have identical performance target either expressed in terms of maximum tolerable loss rate or in terms of maximum acceptable queueing delay. For the sake of generality, we considered various assumptions for the background traffic. The simulations were carried out using alternately a Poisson process, a PPBP process or a real traffic trace. We collected results at two different time scales since flows performance have to be considered both instantaneously and over longer periods. We also included the results that would be obtainable by an ideal admission control so as to benchmark the performance of existing MBAC.

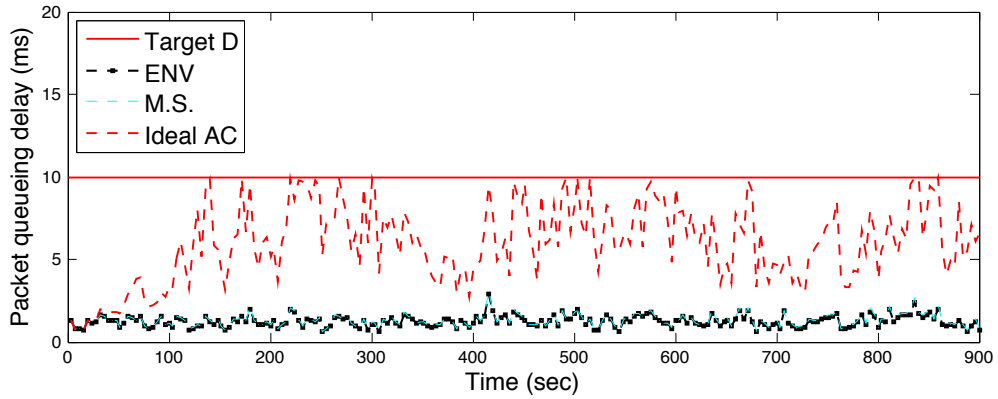
Our results tend to show that, in case of a target loss rate, when the characteristics of background traffic is “regular" enough (e.g., a Poisson process), the *Measured Sum* solutions arises as the best solution. However, when the characteristics of background traffic deviate to more variable sources (e.g., a PPBP process, a real trace), the *Aggregate Traffic Envelopes* and *Equivalent Capacity* solutions brings satisfactory results, much better than those obtained by the *Measured Sum* solution.

When dealing with a queuing delay target, the outcomes of MBAC solutions are generally less successful. The solutions typically tend to deviate further from the ideal admission control, often in an overly conservative way.

To conclude, it is the authors point of view that the major difference between any existing MBAC solutions lies in their easiness in parameter tuning and their ability to meet the criterion chosen as the target (typically a maximum tolerable loss rate or queueing delay). On the other hand, the quantitative performance of MBAC solutions result mainly from the “physical" laws that shares the network resources among a set of competing flows. Indeed, no admission control can attain a lower packet loss rate than another, while accepting at the same time a greater number of (statistically identical) flows.



(a) Target packet loss rate $Pr = 10^{-2}$



(b) Target queueing delay $D = 10$ms

Figure 5: Real traffic trace for the initial background traffic

Acknowledgment

This work has been partly supported by the project *Semantic Networking* within the common laboratory INRIA - Alcatel Lucent-Bell Labs.

References

[1] Arnold O. Allen. Probablity, Statistics and Queueing Theory with Computer Science Applications, Second Edition. In *Int. CMG Conference*, 1990.

[2] Doreid Ammar, Thomas Begin, Isabelle Guérin-Lassous, and Ludovic Noirie. Evaluation and comparison of MBAC solutions. In *Proceedings of the 36th Conference on Local Computer Networks, IEEE LCN 2011*, Germany, Oct. 2011.

[3] Thomas Begin, Bruno Baynat, Alexandre Brandwajn, and Francis Sourd. A DFO technique to calibrate queueing models. *Computers & Operations Research*, 37(2):273 – 281, February 2009.

[4] Thomas Begin, Alexandre Brandwajn, Bruno Baynat, Bernd Wolfinger, and Serge Fdida. High-level approach to modeling of observed system behavior. *Performance Evaluation*, 67(5):386 – 405, May 2010.

[5] Lee Breslau, Sugih Jamin, and Scott Shenker. Comments on the Performance of Measurement-Based Admission Control Algorithms. In *Infocom*, 2000.

[6] Lee Breslau, Edward W. Knightly, Scott Shenker, Ion Stoica, and Hui Zhang. Endpoint admission control: architectural issues and performance. *SIGCOMM Comput. Commun. Rev.*, 30, August 2000.

[7] Sally Floyd. Comments on Measurement-based Admissions Control for Controlled-Load Services. Technical report, 1996.

[8] S. Georgoulas, P. Trimintzios, G. Pavlou, and K. Ho. An integrated bandwidth allocation and admission control framework for the support of heterogeneous real-time traffic in class-based IP networks. *Computer Communications*, 31:129–152, 2008.

[9] R. Guerin, H. Ahmadi, and M. Naghshineh. Equivalent capacity and its application to bandwidth allocation in high-speed networks. *IEEE JSAC*, 9(7):968–981, Sep 1991.

[10] Trace II. Brescia university. `http://www.ing.unibs.it/ntw/tools/traces/`.

[11] Sugih Jamin and Peter B. Danzig. A measurement-based admission control algorithm for integrated services packet networks. *IEEE/ACM Transactions on Networking*, 5(1):56–70, Feb 1997.

[12] Sugih Jamin, Scott Shenker, and Peter B. Danzig. Comparison of Measurement-based Admission Control Algorithms for Controlled-Load Service. In *Infocom*, 1997.

[13] Andrew W. Moore. An implementation-based comparison of Measurement-Based Admission Control algorithms. *J. High Speed Netw.*, 13, April 2004.

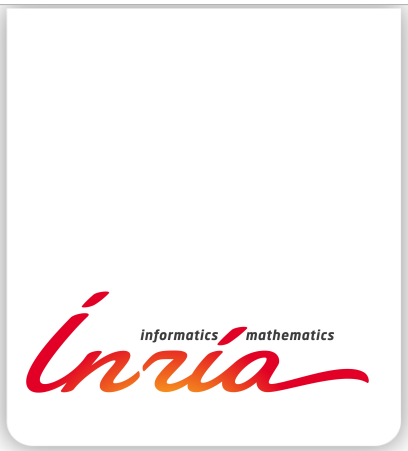
[14] A. Nevin, Y. Jiang, and P. J. Emstad. Robustness Study of MBAC Algorithms. In *ISCC*, 2008.

[15] Jingyu Qiu and Edward W. Knightly. Measurement-Based Admission Control with Aggregate Traffic Envelopes. *IEEE/ACM Transactions on Networking*, 9(2):199–210, Apr 2001.

[16] D. Sass. Internet traces of the“Selfnet" university dormitory network, Trace UST2, University of Stuttgart, Trace UST2, 2004.

[17] George Arthur Frederick Seber. *Multivariate observations*. Wiley series in probability and mathematical statistics. Wiley, New York, NY [u.a.], 1984.

Contents



RESEARCH CENTRE
GRENOBLE – RHÔNE-ALPES

Inovallée
655 avenue de l'Europe Montbonnot
38334 Saint Ismier Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399