# Throughput Analysis of End-to-End Measurement-Based Admission Control in IP

Giuseppe Bianchi (*), Antonio Capone (+), Chiara Petrioli (+)

(*) Università di Palermo, Dipartimento di Ingegneria Elettrica
Viale Delle Scienze, Parco D'Orleans, 90128 Palermo, Italy

(+) Politecnico di Milano, Dipartimento di Elettronica e Informazione
Piazza L. Da Vinci 33, 20133 Milano, Italy

email: {bianchi,capone,petrioli}@elet.polimi.it

*Abstract*—**This paper introduces approximate analytical models to evaluate the performance of End-to-end Measurement Based connection Admission Control (EMBAC) mechanisms, devised for the setup of real time flows over the Internet. These mechanisms rely on users probing the current congestion status of their required network path using a succession of probing packets. If the probing rate measured at the end receivers is greater than a certain threshold, users are allowed to switch to a phase of data exchange; otherwise they abort the call setup attempts. In conformance with the differentiated services framework, routers are oblivious to individual flows, and only need to give higher priority to data packets than to probing traffic. Despite the approximations introduced to make the analysis tractable, our model appears to be extremely accurate for a scenario of constant rate connections. Much less accurate, but useful as a possible starting point for future work, is the extension of the model to a scenario of variable rate connections. Simulation results are also presented in the paper to gain additional quantitative insights on the effectiveness of EMBAC to provide support for tight QoS requirements.**

## I. Introduction

It is widely accepted that the best effort model of the today Internet is not able to satisfactorily support the emerging market demand for real-time audio and video services. These services may require very stringent Quality of Service requirements (e.g. 150 ms mouth to ear delay for toll quality IP telephony), that have to be maintained for all the call holding time. The goal of achieving such a tight QoS control on the Internet, meaningwhile leaving untouched its fundamental architectural principles, is an open research issue. The question is whether an Internet stateless and scalable architecture, such as that envisioned by the Differentiated Services framework [1], [2], has the intrinsic capability to provide performance comparable to that achievable by heavyweight per-flow resource management approaches, such as Integrated Services/RSVP [3], [4].

Several proposals have appeared in the literature, either targeted to simplify and make more scalable the Integrated Services approach [5], [6], or to enable some advanced form of QoS control in a stateless DiffServ architecture. For example, [7] envisions a stateless Internet where packets carry control information from the edge to the core, and routers can take advantage of this information. The opposite approach is taken in [8], [9], where packet marking mechanisms within the core routers allow to extract congestion information from the network. Such an information can be suitably used at the network edge to implement fully distributed control mechanisms (e.g. pricing-based [8]).

Recent literature [10], [11], [12] (as well as, in part, [8]) has envisioned the possibility to introduce in IP networks a purely distributed call admission control scheme, whose operation is based on measurements taken by each connection end points over probing packets injected in the network during the connection setup phase. The goal of the in-band measure is to estimate whether there are enough resources in the network to accept the new connection without degrading accepted calls QoS. In what follows, we will refer to solutions based on the described operation with the descriptive name End-to-end Measurement Based Admission Control (EMBAC) mechanisms. The fundamental differences with classical MBAC mechanisms proposed in the literature [13], [14], [6] are two. First, the decision on whether to accept or reject a connection is not taken by the internal network routers, which are oblivious to individual flows, but it is independently taken by the edge nodes of each specific connection. Second, the measures are taken end-to-end by each connection instead of being centrally taken by core routers.

Scope of this paper is to develop analytical models aimed at providing a quantitative throughput performance evaluation of a reference EMBAC scheme, described in section II, based on bandwidth measurements. In section III a model is presented for a scenario of constant rate connections. In this case, the notable agreement with simulation results allows us to validate the approximations adopted to make the analysis tractable. The performance evaluation presented allows to draw some important comments on the behavior of EMBAC schemes. Next, in section IV, the analysis is preliminary extended to the case of homogeneous ON/OFF traffic pattern modeled as a Markov chain. Though the model only partially captures the phenomena occurring in this scenario, we feel that it is an effective starting point for more complex and accurate enhancements. Finally, section V provides additional simulation results to prove the capability of EMBAC to support connections with tight QoS requirements.

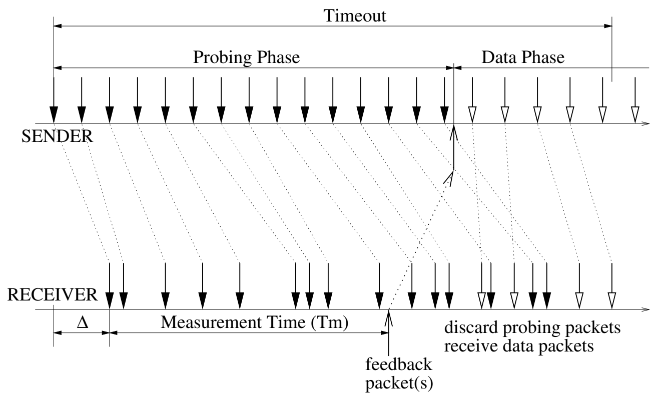


Fig. 1. *EMBAC connection setup scheme*

## II. EMBAC BASIC OPERATION

The basic and characterizing feature of EMBAC schemes is to rely on end-to-end in-band measures to determine whether there are enough resources in the network to accept a new connection. As shown in figure 1, a connection is composed of two phases: a probing phase, eventually followed, in the case the connection is accepted, by a data phase. A traffic source that wants to set-up a real time connection starts transmitting a constant rate signaling flow, composed of low priority packets tagged as *Probing* in the IP header. Upon reception of the first probing packet (i.e. after a time interval $\Delta$ as shown in the figure), the destination node starts measuring the probing packets arrival statistics over a fixed length measurement period $T_m$. The measured statistics can be as simple as the average rate received during the time interval $T_m$ (e.g. computed by counting the number of received packet within the considered interval), or more complex including for example delay and delay jitter statistics [10], [11].

At the end of the measurement period, on the basis of the statistics collected, the receiver estimates whether there are enough resources available along the connection path to meet a predetermined QoS requirement. This decision is notified back to the sender (by means of one - or more - *feedback packets*), which either switches from probing to data phase, and starts transmitting high priority *Data Packets*, or aborts the call setup. As an additional possibility, disconnection occurs when a timeout expires. This last feature is important in practice, when network congestion does not allow any of the probing packets to reach destination. Unless otherwise specified, in what follows we will assume a basic EMBAC scheme based on bandwidth measurements only. In such a case, the receiver measures the probing packets arrival rate $B_r$ (averaged over the measurement time $T_m$), and compares it with a predetermined threshold $B_t$. Clearly, $B_p > B_t \geq B$, being $B_p$ the probing rate and $B$ the data rate that must be established (e.g. the peak rate of a constant rate connection, or the average rate of a variable rate connection).

The role of the probing phase is to stress the network as much as if the connection were already offered to the network in its data phase. This implies that the probing packet rate must be as large as the additional load generated by the call when accepted. This is evident for the case of constant rate flows, in which a probing flow transmitted at a rate lower than that requested for the data flow may result in a call acceptance when not enough resources are available, and therefore in QoS degradation when the connection switches to data phase.

In EMBAC schemes, core routers are stateless. In full agreement with the DiffServ paradigms, they need only to discriminate between classes of packets. In particular, probing traffic should not contend the use of bandwidth against established traffic, to prevent QoS degradation below the requirements in case of insufficient bandwidth. To protect the already accepted traffic, each router handles packets offered to a given output link with two distinct queues, one for (high priority) data packets, and the other for (low priority) probing packets, and serves the latter packets only when the data packets queue is empty. This forwarding mechanism has the fundamental effect that probing packets are forwarded only when resources unused by data traffic are available. This implies that, when the available bandwidth is not enough to accommodate the new calls, the probing flows will suffer severe QoS degradation which will be detected by the measurements running at the end nodes.

We now observe that, while the data packets queue saturation is prevented by EMBAC operation, this is not the case for the probing packet queue, as the probing offered load can be eventually far greater than the probing bandwidth available. It is therefore necessary to control congestion of the probing packets queue, e.g. by limiting the probing packets queue size (as in [12]). In the simulation runs presented in this paper, we have taken a slightly more complex approach, by setting a limit on the probing packet lifetime (PLT), i.e. a probing packet is discarded if it is not served within a given time interval from the instant of its arrival to the queue. The reason for the choice of this more complex solution will be explained in section V.

We conclude this section with some supplementary comments. First, we note that, as long as just two priority levels (probing/data) are used within the network routers, EMBAC schemes are limited to provide a single QoS requirement in the whole network, i.e. all accepted connections share the same loss/delay performance.

Second, an important positive feature of EMBAC schemes is their intrinsic stability and robustness. Occasionally, an increase in the accepted traffic above the QoS limits may occur either because of the unreliability of the measurement process at the edge nodes (a probing connection misrated as admissible), or because of rerouting of already accepted connections. However, due to the forwarding mechanism employed, the resulting increase in the data traffic load throttles first the probing traffic. This in turn implies a lower acceptance probability of probing connections. This effect prevents the congestion period from lasting long.

Third, the EMBAC throughput performance are necessarily lower than that of an admission control scheme based on a centralized decision, which rejects a connection only when the bandwidth available on the link is explicitly lower than that requested. In fact, probing packets of different flows compete against each other to access the bandwidth left available by accepted traffic. The net result is a decrease in realized throughput since all contending flows are likely to deduce congestion even though some of them could be admitted.

Finally, we remark that the basic distributed admission control technique of EMBAC schemes does not provide any physical mean to reserve and guarantee resources within the network. Instead, it can be envisioned as a stable and adaptive congestion control mechanism which regulates the QoS provided by the network by adaptively increasing and decreasing, during the time and according to the network status, the probability of acceptance.

## III. THROUGHPUT ANALYSIS: CONSTANT RATE SOURCES

In this section we consider the case of homogeneous constant rate (CBR) connections. In such a scenario, throughput is the only figure to account for, as delay is only due to the alignment of the different regular streams, and it is negligible even for a load up to 100% of the channel capacity. As usual in the case of CAC analysis, we concentrate over a single link, since the end-to-end bandwidth measuring process is dominated by the bandwidth available on the bottleneck link - i.e. the most loaded link in the source to destination path.

Assume that new calls (i.e. calls entering the probing phase) are offered to the considered link according to a Poisson process of rate $\lambda$ calls/s. Let $T_m$ be the constant duration of the probing phase, and let $1/\mu$ be the average connection duration. Each connection is identified by the data rate $B$, by the probing rate $B_p > B$, and by the decision threshold $B_t$ (with $B_p > B_t \geq B$).

Let us now focus on a "tagged" probing connection, whose probing phase is in progress. Being $t$ a generic time instant, let $m(t)$ be the number of *additional* probing connections and let $K(t)$ be the number of already accepted connections. Since probing connections fairly share the bandwidth left available by the accepted ones, the rate $B_r(t)$ instantaneously perceived at a probing connection end point is closely approximated (in the assumption of short probing packets queue) by:

$$B_r(t) = \min\left(B_p, \frac{\max(C - K(t)B, 0)}{m(t)+1}\right) \qquad (1)$$

being $C$ the link capacity.

Let $t_d$ be the time at which an accept/reject decision is made for the tagged probing connection. The connection is accepted if $B_r \geq B_t$, where $B_r$ is the received probing bandwidth, averaged over the whole probing interval $T_m$. In order to make the analysis tractable, a very reasonable approximation is (*Approximation A1*) to assume that $B_r \approx B_r(t_d)$, i.e. that the instantaneous received bandwidth sampled at the time of decision is representative for the average bandwidth received in the measurement time $T_m$. This in turns implies that we are assuming the processes $m(t)$ and $K(t)$ to be in statistical equilibrium over the period $T_m$. Assumption A1 is motivated by the fact that, in the time interval $T_m$ period, link status changes are triggered by opposite events. First, all the $m(t_d - T_m)$ connections found in the probing phase at the beginning of the measurement period are going to end their probing phase during this time, but are replaced by the new $m(t_d)$ connections found in the probing phase at the moment of the decision. Second, the number of accepted connections among the $m(t_d - T_m)$ probing ones that will decide during $T_m$ are balanced by the accepted connections that terminate in the considered period.

Owing to approximation A1, the acceptance condition $B_r \geq B_t$ translates into the following constraint on $m(t_d)$:

$$m(t_d) \leq \left(\frac{C}{B} - K(t_d)\right)\frac{B}{B_t} - 1 = m_{\max} \qquad (2)$$

Let us now study the two dimensional stochastic process $\{K(t), m(t)\}$. For this process, the following state transitions occur. At each new (probing) connection arrival, $m(t)$ is increased, while it is decreased at each instant of decision. After a decision a probing connection either turns into a data connection in case of acceptance (and thus $K(t)$ is increased), or terminates. Finally, $K(t)$ is decreased whenever an accepted connection terminates.

In the general case of non exponential probing interval length and non exponential call duration, this process is non Markovian[1]. However, we may dramatically simplify the analysis by assuming that $K(t)$ and $m(t)$ are independent random processes (*Approximation A2*). In fact, for general distribution of the length of the probing phase (thus including constant duration $T_m$), the number of probing connections has a Poisson distribution, i.e. the probability that $i$ connections are simultaneously probing the channel is given by:

$$p_m(i) = e^{-\lambda T_m}\frac{(\lambda T_m)^i}{i!} \qquad (3)$$

In this expression, we have assumed that both the delay $\Delta$ encountered by the first probing packet, and the feedback packet delay (see figure 1) are negligible with respect to the measurement interval $T_m$), i.e. that the probing phase duration is approximately equal to $T_m$.

Owing to the independence assumption, we can separately study the process $K(t)$. Condition (2) can be transformed into the probability $P_a(K)$ that a probing connection is accepted,

[1] Note that under the simplificative hypothesis of exponential probing times with average $T_m$, and exponential call duration with average $1/\mu$, the process is a two dimensional Markov chain. Since decisions, taken according to condition (2), ultimately depend on the state $(K, m)$ of the system at the instant of decision, the only non null entries of the infinitesimal generator $Q$ of the chain are readily found ($m_{max}$ is defined in equation (2)) as:

$$\begin{array}{lll} Q[<K,m>, <K,m+1>] & = \lambda & \forall m, \forall K \in (0, \lfloor C/B \rfloor) \\ Q[<K,m>, <K-1,m>] & = K\mu & \forall m, \forall K \in (1, \lfloor C/B \rfloor) \\ Q[<K,m>, <K+1,m-1>] & = m/T_m & \forall m \in (1, m_{\max}+1) \\ Q[<K,m>, <K,m-1>] & = m/T_m & \forall m > m_{\max}+1 \end{array}$$

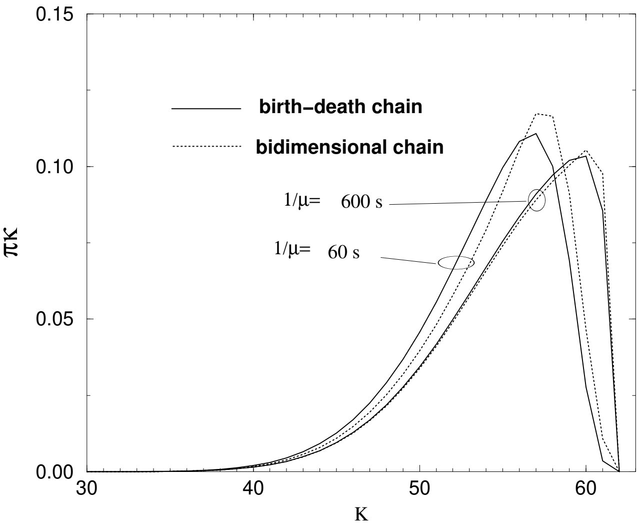


Fig. 2. *Bidimensional model versus proposed approximate model: steady state probability comparison*

conditioned to the number $K$ of already accepted connections. In fact, due to (3) and (2), and recalling from elementary queueing theory that the stationary distribution (3), representing the distribution of the number of probing connections, is equal to that seen by a probing connection departing, it is:

$$P_a(K) = \sum_{i=0}^{\lfloor m_{\max} \rfloor} p_m(i) = \sum_{i=0}^{\lfloor m_{\max} \rfloor} e^{-\lambda T_m} \frac{(\lambda T_m)^i}{i!} \quad (4)$$

where $m_{\max}$, defined in (2), gives the maximum number of additional connection in the probing phase found upon decision, in order to still receive an accept decision (in the case $m_{\max}$ is negative, the sum is assumed to yield 0).

The instant of time that a probing connection ends its probing phase is interpreted as the instant of an arrival of a "real" connection to the system. This new arrival process is still Poisson owing to Burke's well known result [15]. Since each "real" connection arriving to the system, and seeing $K$ already accepted connections, is admitted with probability $P_a(K)$, assuming independent successive decisions (*Approximation A3*), we are finally in the condition to model the process $K(t)$ as a birth-death Markov chain, with birth/death coefficients (using standard notation [15]):

$$\begin{cases} \lambda_K &= \lambda P_a(K) \\ \mu_K &= K\mu \end{cases} \quad (5)$$

This system resembles that of an Erlang loss system [15] with the difference that the arrival rate depends on the system status. Let $\pi_K(k)$ be the steady state probability of finding $k$ calls in the system, obtained by numerical solution of (5). According to known results in queueing theory [16], $\pi_K(k)$ depends on the call duration distribution only through its mean value $1/\mu$. Hence, our model applies to general distribution of the call duration.

Elementary computation allows us to finally express the normalized throughput $\eta$ as the average fraction of link occupied by accepted calls, i.e.:

$$\eta = \sum_k \frac{k \cdot B}{C} \pi_K(k)$$

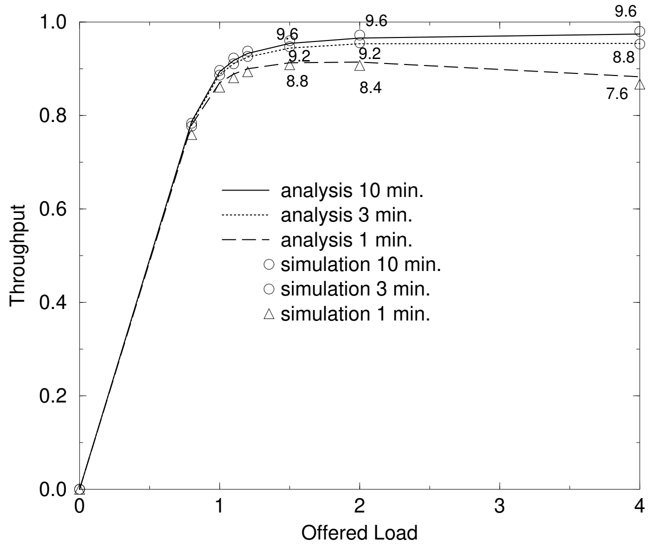


Fig. 3. *Analysis versus simulation: CBR, 2 Mbps link*

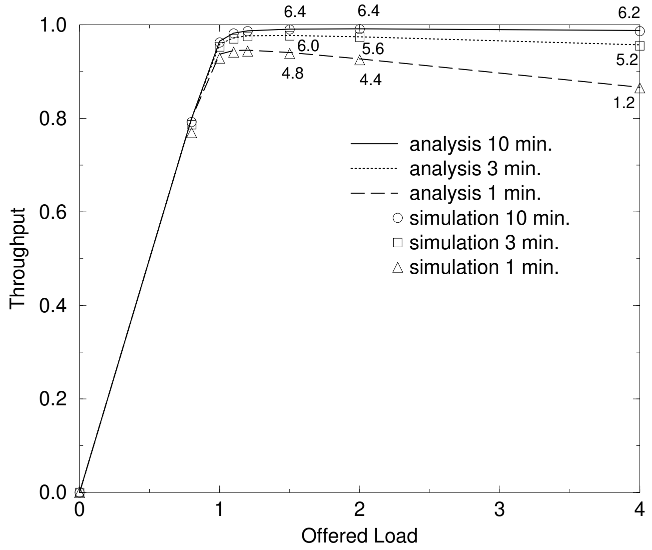


Fig. 4. *Analysis versus simulation: CBR, 20 Mbps link*

### *A. Model Validation and Performance Evaluation*

In this section, unless otherwise specified, we assume a channel capacity equal to 2 Mbps and fixed packet size equal to 1000 bits. The measurement time $T_m$ is set equal to 2 seconds. Constant rate connections are modeled with $B = 32$ packets/s (i.e. Kbps); $B_p$ is set to 38.4 Kbps (connection rate plus 20 %), while $B_t$ is set to 35.2 Kbps (connection rate plus 10 %).

Of the three approximations, the most critical are A2 and A3. To quantify the effect of these approximations, figure 2 compares the steady state probability distribution $\pi_K(.)$ obtained from (5) with that obtained by solving the two dimensional

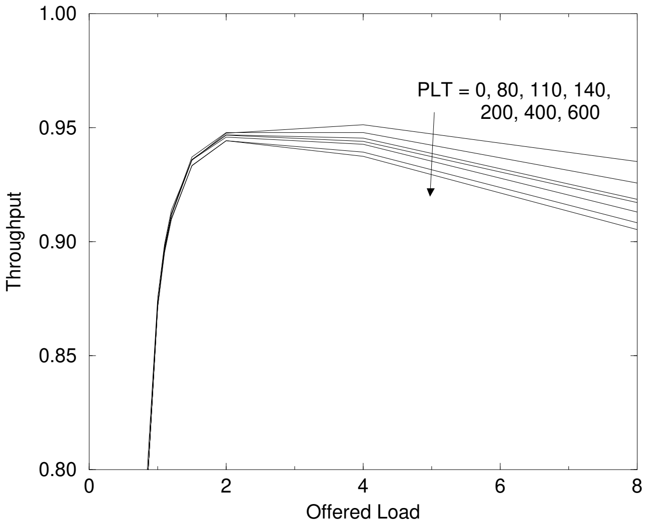


Fig. 5. *Throughput versus different packet lifetimes - simulation*

Markov chain and taking the marginal distribution. The figure reports two different cases of average connection duration equal to 1 minute and 10 minutes. We observe that the difference between the corresponding exact and approximate distributions is relatively small, especially when the average connection duration increases. Extensive throughput comparisons between the values obtained from the exact (two dimensional) and approximate model has furthermore shown differences below $1.5\%$ in all considered scenarios.

To validate the analysis presented above, we have developed a simulation program written in the C++ programming language. The program realizes an event-driven simulation that accounts for packet-level events (e.g. packet generation and queueing, packet expirations in the buffer) as well as all call-level protocol details (e.g. measurement time counted from the instant of arrival of the first packet at the receivers, and timeout).

Figures 3 and 4 compare the analytical results with that obtained by means of simulation. The figures report the throughput versus the offered load, normalized with respect of the channel capacity, respectively for 2 Mbps and 20 Mbps channel capacity. Three different average connection durations are considered: 1, 3 and 10 minutes. In the simulation runs, the probing packet lifetime has been set to 80 ms. Note that this value is far to be considered negligible (as our analysis implicitly assumes, being a fluidic approach), since it implies that approximatively up to 3 probing packets per probing connection can be buffered. Despite this significant difference in the analysis and simulation models, we observe a remarkable agreement among the simulative (symbols - the numerical values above the symbols report the corresponding $99\%$ percentile delay) and the analytical results (lines).

To investigate the effect of the probing packets lifetime, figure 5 shows the throughput performance (simulation results) versus the offered load for different PLT values. The upper line reports the analytical results (i.e. a PLT value theoretically zero). The figure shows that the throughput slightly decreases with an increasing PLT value, especially in significant overload conditions. This is easily explained by considering that, in overload and with high packet lifetime values, the probing packet queue may buffer "obsolete" packets, i.e. useless at the probing connection end points, and thus probing bandwidth is wasted. The role of the PLT will be much more critical and important in the variable rate case, as thoroughly explained in section V.

The figures presented up to now show that the throughput, in general, decreases as the offered load increases. This is especially true for high values of $T_m$, as shown in figure 6, which reports the throughput values obtained by means of the analytical model for the cases of average connection duration 1 and 3 minutes, and for $T_m$ equal to 0.5, 1 and 2 seconds. This effect is due to the increase in the probing traffic load, which in turns decreases the probability that a connections is accepted (keeping constant the number of accepted connections, as long as the value $m$ increases at the denominator of (1), the received bandwidth reduces). From the figure, it is evident that the throughput decrease at high loads is more critical for a shorter average connection duration. This is easily explained by noting that the probing load is given by:

$$\rho_p = \lambda T_m = \frac{\lambda}{\mu} \cdot \frac{T_m}{1/\mu}$$

being $\lambda/\mu$ proportional to the normalized offered load. Hence, at a given offered load the probing load is greater for a shorter connection duration, and thus the throughput decrease is larger.

While, for the constant rate case, a short measurement time implies only a marginal degradation of the quality of the measure, due to bandwidth granularity effects (i.e. received bandwidth measured as number of packets received in the time $T_m$), we will see instead, for the variable rate case, that short values of $T_m$ imply a significantly lower reliability of the measurement process (see quantitative considerations in the recent work on MBAC [13]).

Although the numerical results show that the throughput presents only a marginal decrement at high loads (up to $800\%$ of the channel capacity), the fact that a decrement occurs is a symptom of instability, and this can be a potential practical problem. In fact, in practice, blocked calls are not "cleared" (as implicitly assumed by the Poisson arrival hypothesis), but they are usually re-attempted. It is thus conceivable that congestion may become self-sustaining, i.e. the number of calls in the probing phase may become very large at the expense of data traffic capacity. This remark suggests that call re-attempts should be controlled (e.g. by means of an exponential backoff algorithm).

Finally, the parameters $B_p$ and $B_t$ only marginally affect the resulting throughput performance. In particular, for a same threshold $B_t$, different values for the probing bandwidth $B_p$ result in the same throughput (obviously, as inequality 2 does not depend at all on $B_p$). The small effect of the parameter $B_t$ is instead shown in figure 7 for $T_m$ equal to 1 and 2 s, and for three different average connection durations.

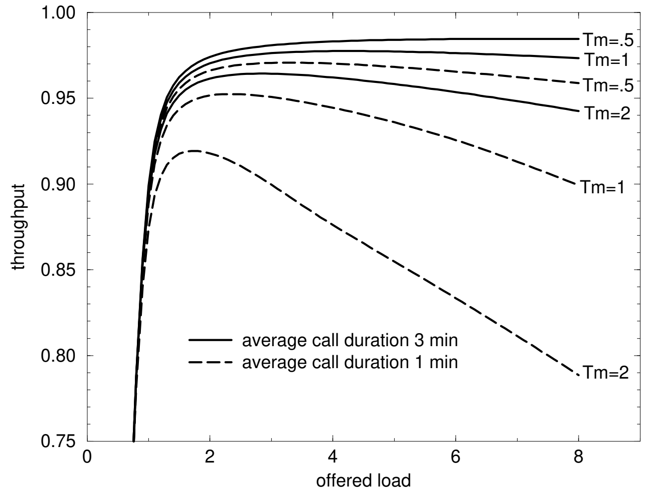


Fig. 6. *Throughput vs. offered load for several $T_m$ values.*

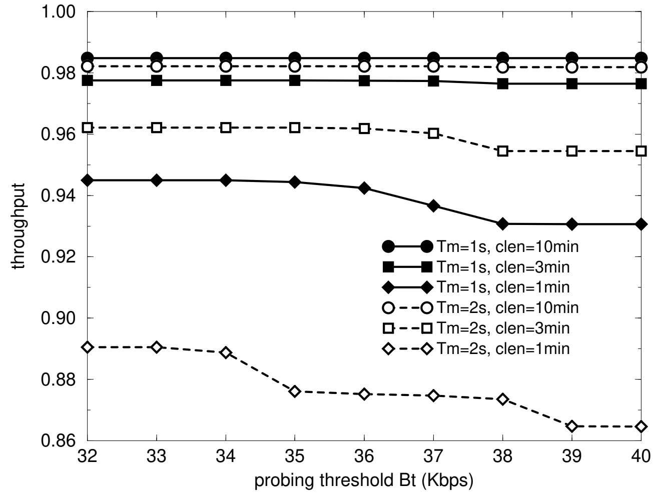


Fig. 7. *Throughput vs. acceptance threshold $B_t$*

## IV. Throughput Analysis: Variable Rate Sources

The birth-death model summarized in (5) can be adopted also in the variable rate case. The fundamental difference is that the conditional acceptance probabilities $P_a(K)$ need to be obtained via a completely different approach. The major problem in the VBR case is that the number of accepted connections $K(t)$, at time $t$, does not uniquely identify the amount of bandwidth allocated over the link. In what follows we concentrate on the case of variable rate traffic modeled as the superposition of independent homogeneous on/off sources with exponential distribution of the on/off periods.

### A. *Traffic model and preliminary results*

Owing to approximation A1, let us assume that during a measurement period $T_m$, the number of accepted connections is constant and equal to $K$. The number of *active* connections $N_K(t)$ depends on $K$ and varies with time during the measurement period. Let $B$ be the rate of packets generation in the ON state, and let $T_{\text{on}}$ and $T_{\text{off}}$ the average amount of time spent in the corresponding state. The source average rate $B_a$ is expressed as

$$B_a = B\frac{T_{\text{on}}}{T_{\text{on}} + T_{\text{off}}} = B \cdot B_u$$

where $B_u$ is the activity factor of each source. Unless otherwise specified, in the numerical results we will consider voice on/off sources, modeled according to the Brady model [17]. For such a model, $T_{\text{off}} = 1.35$ s and $T_{\text{on}} = 1$ s. Hence, the average activity factor for each source results to be $B_u = 0.4255$.

In the usual assumption of independent evolution of each traffic source, a number of straightforward results hold for the process $N_K(t)$. The probability that $i$ sources over $K$ are active at a generic time $t$ is given by the Bernoullian distribution

$$B(K,i) = \binom{K}{i} B_u^i (1 - B_u)^{K-i} \tag{6}$$

Consequently, the average number of active sources is $\gamma_K = K \cdot B_u$, while its variance is $\sigma_K^2 = KB_u(1-B_u)$. For convenience, let us denote with $\alpha = 1/T_{\text{off}}$ and $\beta = 1/T_{\text{on}}$ the transition rates from state OFF to state ON and conversely. It is simple to prove that

$$R(t) = \sigma_K^2 e^{-(\alpha+\beta)t} \tag{7}$$

is the autocovariance of the process $N_K(t)$. In fact, the transient probabilities $P_{\text{on}|\text{on}}(t)$ and $P_{\text{on}|\text{off}}(t)$ that a source is in the ON state at time $t$, given it is initially in, respectively, the ON and OFF states, are well known[2] as:

$$\begin{aligned} P_{\text{on}|\text{on}}(t) &= \tfrac{\alpha}{\alpha+\beta} + \tfrac{\beta}{\alpha+\beta} e^{-(\alpha+\beta)t} \\ P_{\text{on}|\text{off}}(t) &= \tfrac{\alpha}{\alpha+\beta}\left(1 - e^{-(\alpha+\beta)t}\right) \end{aligned} \tag{8}$$

Owing to eqs. (8), the conditional average number of active sources (over $K$ accepted ones) at time $t$, assuming $i$ active ones at time 0, is given by:

$$\begin{aligned} E[N_K(t)|N_K(0) = i] &= iP_{\text{on}|\text{on}}(t) + (K-i)P_{\text{on}|\text{off}}(t) = \\ &= \gamma_K \left(1 - e^{-(\alpha+\beta)t}\right) + ie^{-(\alpha+\beta)t} \end{aligned} \tag{9}$$

where we have applied the fact that $\gamma_K = K\alpha/(\alpha+\beta)$. This in turns yields, for $t \geq 0$ (and symmetrically for $t < 0$), to the following expression for the autocovariance:

$$\begin{aligned} R(t) &= E[N_K(0)N_K(t)] - \gamma_K^2 = \\ &= \sum_{i=0}^{K} E[N_K(t)|N_K(0) = i] \cdot i \cdot B(K,i) - \gamma_K^2 \end{aligned} \tag{10}$$

The expression for $R(t)$ given in (7) is now finally obtained after algebraic simplifications by substituting in (10) the conditional average (9).

[2] E.g. by direct solution of the Chapman-Kolmogorov differential equations:

$$\begin{cases} P'_{\text{off}}(t) = -\alpha P_{\text{off}}(t) + \beta P_{\text{on}}(t) \\ P'_{\text{on}}(t) = \alpha P_{\text{off}}(t) - \beta P_{\text{on}}(t) \end{cases}$$

with initial conditions ($P_{\text{on}}(0) = 0, P_{\text{off}}(0) = 1$) for the case $P_{\text{on}|\text{off}}(t)$ and ($P_{\text{on}}(0) = 1, P_{\text{off}}(0) = 0$) for the case $P_{\text{on}|\text{on}}(t)$.

### B. Throughput Analysis

Let us now consider a measurement time $T_m$, during which a constant number $K$ of connections are offered to the link (assumption A1). Let us now define the following stochastic process:

$$x(T_m, t) = \frac{1}{T_m} \int_{t-T_m}^{t} N_K(\tau) d\tau \tag{11}$$

$x(T_m, t)$ represents the *average* number of active sources found during a time interval $(t - T_m, t)$. It is a matter of routine (standard signal theory) to prove that the variance of the process $x(T_m, t)$ depends on the measurement time $T_m$ as:

$$\sigma_x^2(T_m) = \frac{2}{T_m} \int_0^{T_m} (1 - \frac{\tau}{T_m}) R(\tau) d\tau =$$
$$= \frac{2\sigma_K^2}{T_m^2(\alpha+\beta)^2} \left( e^{-(\alpha+\beta)T_m} + (\alpha+\beta)T_m - 1 \right) \tag{12}$$

As expected, for $T_m \to 0$, $\sigma_x^2(T_m) \to \sigma_K^2$. As $T_m$ increases, $\sigma_x^2(T_m)$ decreases (initially exponentially, then linearly), with asymptotic behavior

$$\sigma_x^2(T_m) \approx \frac{2\sigma_K^2}{T_m(\alpha+\beta)}$$

This matches the expected fact that the variance is lower as long as the measurement time is longer. Finally, since the process $N_K(t)$ is Bernoullian distributed (which in turns is closely approximated with a Gaussian distribution), we can safely assume that the distribution of the process $x(T_m, t)$ is Gaussian.

To proceed further, we need to determine, for the variable rate case, a rule like (1), to express the bandwidth received by each probing station in terms of the bandwidth allocated over the link. To keep the model's complexity under control, we appear to be forced (at least in the preliminary model presented in this paper) to make a major approximation, referred to as *Approximation A4*. This approximation consists of assuming that, at the end of the probing phase for a tagged connection, the received bandwidth is related to the bandwidth offered to the link by the relation:

$$B_r = \min\left( B_p, \frac{\int_{t-T_m}^{t} C - BN_K(\tau) d\tau}{T_m \cdot (m+1)} \right) =$$
$$= \min\left( B_p, \frac{C - x(T_m, t)B}{m+1} \right) \tag{13}$$

where $m$ is the (constant, according to approximation A1) number of additional probing connections during the measurement time. The reason and the limits of this approximation will be discussed into details in the next subsection.

Approximation A4 is the key to simply extend the constant rate model to the variable rate case. In fact, we only need to replace, in eq. (2), the term $K(t_d)$ with $x(T_m, t_d)$, thus obtaining:

$$m(t_d) \le \left( \frac{C}{B} - x(T_m, t_d) \right) \frac{B}{B_t} - 1 \tag{14}$$

or, equivalently,

$$x(T_m, t_d) \le \frac{C - B_t[m(t_d)+1]}{B} \tag{15}$$

Since inequality (15) represents the rule for accepting a probing connection, the coefficients $P_a(K)$ to use in the Markov chain (5) are obtained by computing the probability that such inequality is met:

$$P_a(K) = \sum_{m=0}^{\frac{C}{B_t}-1} P\{m(t_d) = m\} \cdot$$
$$\cdot P\left\{ x(T_m, t_d) \le \frac{C - B_t(m+1)}{B} \right\} \tag{16}$$

where the upper limit of the sum is obtained by setting $x(T_m, t_d) = 0$ in the inequality (14). Since $x(T_m, t_d)$ is a Gaussian r.v. with mean value $\gamma_K$ and variance $\sigma_x^2(T_m)$, the second term of eq. (16) is readily computed by means of:

$$P\{x(T_m, t_d) \le \xi\} = \frac{1}{2}\left( 1 + \text{erf}\left( \frac{\xi - \gamma_K}{\sqrt{2\sigma_x^2(T_m)}} \right) \right) =$$
$$= \frac{1}{2}\text{erfc}\left( \frac{(\gamma_K - \xi)T_m(\alpha+\beta)}{2\sigma_K\sqrt{e^{-(\alpha+\beta)T_m} + (\alpha+\beta)T_m - 1}} \right) \tag{17}$$

To conclude the analysis, it remains to compute the probability $P\{m(t_d) = m\}$. The simplest possibility is to adopt (3), i.e.

$$P\{m(t_d) = m\} = p_m(m) = e^{-\lambda T_m} \frac{(\lambda T_m)^m}{m!} \tag{18}$$

However, this expression may result inaccurate for very high accepted load. In fact, in the CBR case, the time interval $\Delta$ elapsing from the first probing packet generation, and its arrival to the receiver (i.e. the instant of time at which the measure begins) can be considered negligible even at high load. This is no longer the case in the VBR scenario, where the first probing packet of a connection needs to wait at least until the end of a busy period of the data packet queue before being served. To grossly account for this effect, we have added to $T_m$, in eq. (18), a corrective term

$$\Delta = \frac{C/B}{1 - \rho_K} \tag{19}$$

i.e. the average busy period of an M/M/1 queue, where $\rho_K = \min(KB_u/C, 1-\epsilon)$, being $\epsilon$ a small constant to avoid an infinite (or negative) value for $\Delta$ in the case the average accepted load reaches or exceeds 1. The effect of this heuristic correction is not particularly relevant for throughput values below 0.95, and it is explicitly shown in figure 9.

### C. Model validation

Let us first discuss into details approximation A4. This approximation is illustrated in figure 8. The figure reports the data traffic link occupation during a measurement time, and the probing bandwidth received by a single probing connection offered to the link. In case of data traffic overload, i.e. $N_K(t) \cdot B > C$

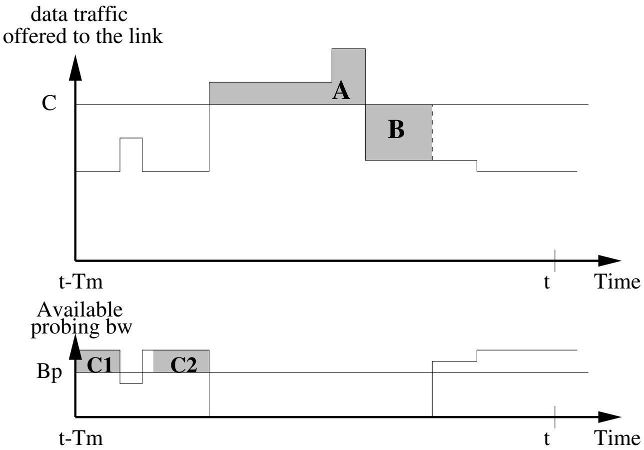

Fig. 8. *Bandwidth variation during a measurement time, VBR scenario*

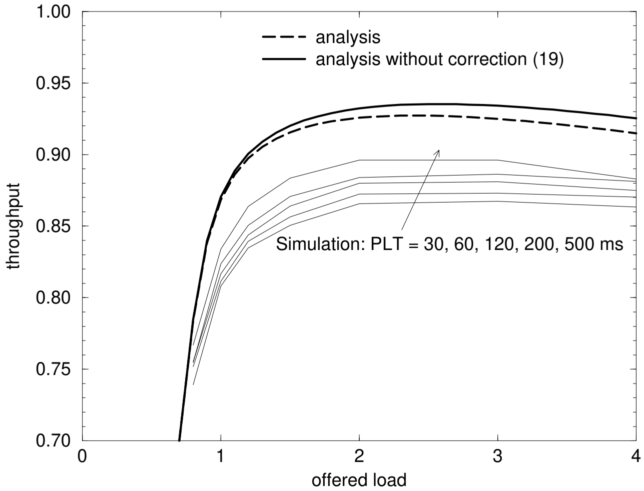

Fig. 9. *VBR: analysis versus simulation, for several PLT values*

(region A in the figure), the forwarding mechanism implies that the probing connection does not receive service. After the overload period, the probing connection starts receiving service only after the data packet queue has emptied (region B - note that the areas A and B are equal). The problem is that, during underload periods, it may happen that not all the available bandwidth may be used by the probing station (e.g. areas C1 and C2 in the figure), which is bounded to receive at most the bandwidth $B_p$ transmitted during the probing phase, plus the probing packets eventually accumulated in the probing queue. This implies that, in practice, the bandwidth received by the probing stations is lower than that assumed by formula 13 (i.e. approximation A4), and thus that our model is an upper bound of the throughput performance.

The negative effects of the above considerations are evident in figure 9. This figure compares the analytical throughput results (the analysis without correction 19 is also reported in the figure) with that obtained by means of simulation. The considered scenario is that of a 2 Mbps channel, and Brady ON/OFF connections with peak rate $B = 32$ Kbps, probing rate $B_p = 38.4$ Kbps, and threshold $B_t = 35.2$ Kbps. The average connection duration has been set to 3 minutes. The simulation results are obtained with different values for the maximum packet lifetime (PLT). From the figure, we note that the PLT value has an opposite effect on the throughput than that shown in the CBR case (refer to figure 5): in the VBR case, larger PLT values correspond to an higher throughput. In fact, a larger PLT value seems to imply a better accuracy of approximation A4. Since a greater number of probing packets are buffered at the beginning of a measurement period, it occurs less frequently for the available bandwidth to be left unused by the probing connections (i.e. the effect depicted by the regions C1 and C2 in figure 8 is less likely to occur).

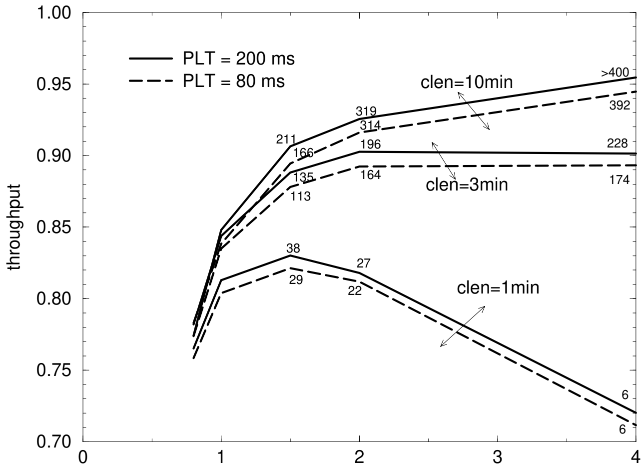

Fig. 10. *Throughput and delay performance for VBR*

¿From figure 9 we see that the analysis overestimates the simulation throughput by almost 5%. As shown before, the error is lower for high PLT values. However, in such case (as thoroughly discussed for the CBR case) the fluidic model adopted in the analysis, i.e. no probing packets buffering, results to be different from the simulation model which accounts for buffered probing packets. In conclusion, due to the intrinsic errors induced by both cases of small and large PLT values, our analytical model, in its present form, appears inherently bounded to produce not very accurate results.

## V. EMBAC DELAY PERFORMANCE AND TUNING

Besides the approximations involved, the main drawback of the previous VBR model is that it is only able to capture the throughput performance. If this is sufficient in a CBR scenario, where delays are negligible, this is not the case in a VBR environment. In fact, the primary goal of an admission control is to guarantee some minimal quality of service, for example expressed in terms of delay performance, while high link utilization can be considered a secondary goal.

As a matter of fact, it has not yet been clarified whether the delay performance of EMBAC can be controlled in the VBR scenario. Figure 10 reports the simulation throughput and delay

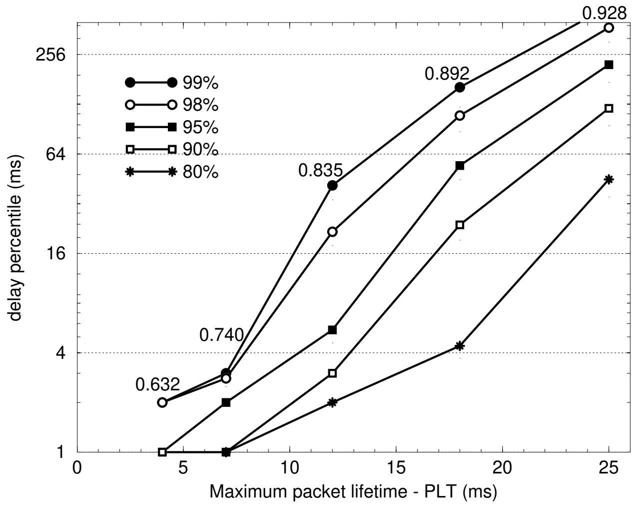


Fig. 11. *VBR: delay percentiles versus packet lifetime*

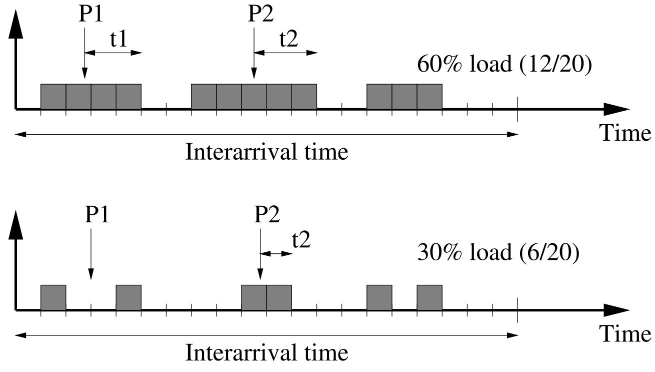


Fig. 12. *Graphical interpretation of the effect of a short PLT value*

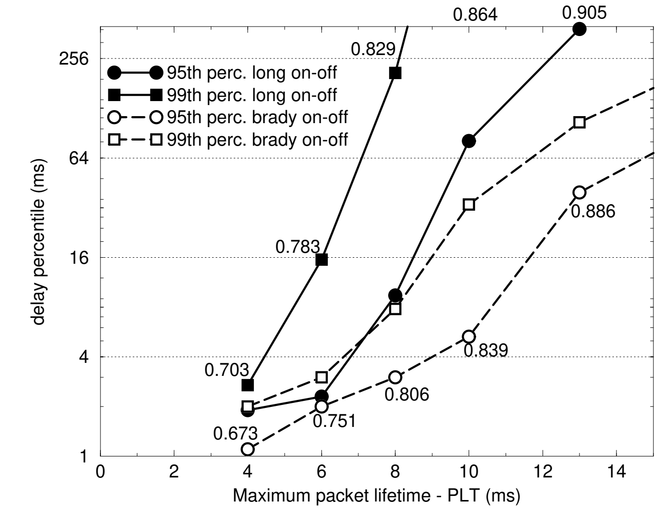


Fig. 13. *Delay performance for very short $T_m$ (500 ms) and for two different traffic models*

performance for three different cases of 32 Kbps Brady voice sources with average connection duration equal to 1 min, 3 min and 10 min, offered to a 2 Mpbs link. As before, $T_m = 2$ s, $B_p = 38.4$ Kbps, and $B_t = 35.2$. Two Packet lifetime values have been assumed: 80 and 200 ms. The numerical values (in ms) of the 99-th packet delay percentile of the accepted connections are reported on top of the throughput plot. We see, as expected, that these values basically depend on the accepted load. Hence, in the scenarios resulting in high throughput (namely, average connection duration equal to 3 and 10 minutes), these values are well above toll quality requirements.

This leads us to question whether an EMBAC scheme can be made able to control the QoS of admitted flows. *Indeed EMBAC schemes can be tuned to ensure a certain quality level*: we are going to show that a key parameter is the maximum packet lifetime within the routers (we recall from section II that the PLT represents the maximum amount of time that can elapse between the arrival of a packet in the router and its service).

Figure 11 plots various delay percentiles versus the PLT. The corresponding throughput values are numerically reported in the figure. The considered scenario is a normalized offered load equal to 4, and average connection duration equal to 3 minutes. Instead of 2 seconds, a shorter value $T_m = 750$ ms has been considered in this figure, as this value starts to be reasonable as toll quality setup time. A probing connection is accepted only if no packets are lost (i.e. $B_t = B_p = 38.4$ Kbps). Figure 11 shows that small PLT values allow an extremely tight control of the delay performance (e.g. the 99th delay percentile is as low as 2 ms for a PLT equal to 4 ms), clearly at the expense of the achieved throughput.

To understand why the PLT is such an effective parameter, it is necessary to consider what happens at the packet level[3]. For simplicity of explanation, let us focus on the graphical illustration given in figure 12. This figure reports two example patterns of link utilization by data packets. Assuming that an interarrival data packet period is divided in 20 slots (i.e. that the link capacity is equal to 20 active connections), the figure illustrates the case of link utilization equal to 60%, i.e. 12 active connections, and 30%, i.e. 6 packets to be transmitted into an interarrival period. Consider now two probing connections, P1 and P2, offered to the link. If the PLT is greater or equal than the interarrival time, the probing packets will encounter an idle slot for transmission, regardless of the link utilization pattern and of the data traffic load. In other words, if the PLT value is large enough, our measurement scheme is not able to detect *how much load* is offered to the link, but only to detect if there is enough spare capacity available.

Now consider a PLT value much shorter than the data packets interarrival time. Now, even if there is enough bandwidth to transmit the probing packets, these packets can expire, because of a short PLT value, *before an idle slot becomes available*. Moreover, as graphically shown in the figure, the expected time before receiving service strongly depends on the data traffic load (the values $t_1$ and $t_2$ are greater in the 60% case of the figure than in the 30% case). Thus, the probability that a packet is lost because of PLT expiration grows with the data traffic load.

[3]This implies that this effect cannot be captured by a fluidic model (analytical or simulative).

This remark has an important consequence for EMBAC schemes: the shorter the PLT value employed in each network node, the earlier congestion is detected by the end to end measurement process. We also conclude that when short PLTs are employed in the network routers, short measurement times may be effectively employed. To prove this point, figure 13 reports the delay percentile comparison for two different traffic models. With the label "Brady on-off" we have considered the usual Brady model, i.e. $T_{\rm on} = 1$ s and $T_{\rm off} = 1.35$ s. With the label "long on-off" we have considered the superposition of on/off exponential sources with on and off periods 10 times longer than the Brady case (i.e. $T_{\rm on} = 10$ s and $T_{\rm off} = 13.5$ s). The normalized offered load in the figure is 4, the connection duration has been set to 10 minutes (to minimize the positive effect of the probing traffic load), and the measurement time $T_m$ has been set to a value as low as 0.5 s. The resulting throughput values for the two traffic models are also reported in the figure (top = long on-off, bottom = Brady on-off). The results are impressive and show that a suitable setting of the PLT value appears to allow us to control traffic with a dynamic much longer than the measurements time.

## VI. Conclusions

In this paper we have discussed the performance of End-To-End Measurement Based Admission Control (EMBAC) schemes for IP networks. These schemes, recently proposed, allow to realize a fully distributed and scalable admission control scheme over a stateless Internet architecture, such as that considered in the Differentiated Services Framework. The EMBAC implementation allows core network routers to be oblivious to each individual flow, and only requires routers able to differentiate between two classes of packets (probing and data packets) and handle them with different forwarding priorities.

In particular, we have outlined an analytical model to evaluate the throughput performance of EMBAC schemes based on bandwidth measurements only. Our model is extremely accurate in the case of network loaded with constant rate traffic, while it is only a loose upper bound in the case of variable rate traffic. Despite the inaccuracies of the model in this latter case, we feel that the proposed approach represents a possible starting point for future research work.

In the last section of the paper, we have shown that, with a tight setting of the packet lifetime timer (PLT) in the routers, EMBAC is able to exert an impressive control on the QoS of the accepted connections, even with extremely short measurement times (necessary to provide toll-quality setup times). A more accurate study of this effect, via simulation and analysis, is object of current research.

Finally, we remark that the short PLT approach presented in this paper is one of several possible implementation alternatives. A perhaps more effective scheme (actually under investigation) is represented by a per class scheduling policy at each node that limits the bandwidth utilization of the probing traffic to a fraction of the channel capacity. Note that in this case, as well as in the PLT approach, the relevant parameter (e.g. the scheduling threshold, or the PLT value) can be adaptively tuned according to the traffic conditions, by means of traffic measurements within each router, and running for a time scale much longer than the probing measurement time (see [13] for the underlying theory).

## References


[1] K. Nichols, S. Blake, F. Baker, D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and Ipv6 Headers", RFC2474, December 1998.

[2] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, "An Architecture for Differentiated Services", RFC2475, December 1998.

[3] R. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin, "Resource Re SerVation Protocol (RSVP) - Version 1 Functional Specification", RFC2205, September 1997.

[4] J. Wroclawsky, "The use of RSVP with IETF Integrated Services", RFC2210, September 1997.

[5] W. Almesberger, T. Ferrari, J. Le Boudec, "Scalable Resource Reservation for the Internet", IEEE IWQOS'98, Napa, CA, USA.

[6] S. Jamin, P. Danzig, S. Shenker, L. Zhang, "A measurement based Admission Control algorithm for Integrated Services Packet Networks", IEEE/ACM Transactions on Networking, 5(1):56-70. February 1997.

[7] I. Stoica, H. Zhang, "Providing Guaranteed Services without Per Flow Management", ACM SIGCOMM99, Boston, August 1999.

[8] R. J. Gibbens, F. P. Kelly, "Distributed Connection Acceptance Control for a Connectionless Network", Proc. of 16th International Teletraffic Conference, Edimburgh, June 1999.

[9] Z. R. Turanyi, L. Westberg, "Load control of real-time traffic", Internet Draft draft-westberg-loadcntr-02.txt, October 1999.

[10] F. Borgonovo, A. Capone, L. Fratta, M. Marchese, C. Petrioli, "PCP: A Bandwidth Guaranteed Transport Service for IP networks", IEEE ICC'99, Vancouver, Canada, June 1999.

[11] F. Borgonovo, A. Capone, L. Fratta, C. Petrioli, "VBR bandwidth guaranteed services over DiffServ Networks", IEEE RTAS Workshop, Vancouver, Canada, June 1999.

[12] G. Karlsson, "Providing Quality for Internet Video Services", CNIT/IEEE 10th International Tyrrhenian Workshop on Digital Communications , Ischia, Italy, September 1998.

[13] M.Grossglauser, D. Tse, "A Framework for Robust Measurement-Based Admission Control", IEEE Trans. on Networking, vol. 7, no. 3, June 1999, pp. 293-309.

[14] R.J. Gibbens, F.P. Kelly, "Measurement-Based Connection Admission Control", Proc. of 15th International Teletraffic Congress, June 1997.

[15] L. Kleinrock, *Queueing Systems*, Voll. 1 & 2, John Wiley & Sons Inc., 1975.

[16] S. S. Lam, "Queueing networks with population size constraints", IBM J. Res Develop. July 1977.

[17] P. T. Brady, "A Model for Generating On-Off Speech Patterns in Two-Way Conversation", The Bell System Technical Journal, September 1969, pp.2445-2471.