Submitted to IEEE Journal on Selected Areas in Communications.
An earlier version to appear in IEEE Infocom 2000.

Comments on the Performance of Measurement-Based Admission Control Algorithms

Lee Breslau
ATT Labs - Research
75 Willow Road
Menlo Park, CA 94025
breslau@research.att.com

Sugih Jamin
EECS Department
University of Michigan
Ann Arbor, MI 48109-2122
jamin@eecs.umich.edu

Scott Shenker
International Computer Science Institute
1947 Center Street
Berkeley, CA 94704-1198
shenker@icsi.berkeley.edu

Abstract

Relaxed real-time services that do not provide guaranteed loss rates or delay bounds are of considerable interest in the Internet, since these services can achieve higher utilization than hard real-time services while still providing adequate service to adaptive real-time applications. Achieving this higher level of utilization depends on an admission control algorithm that does not rely on worst-case bounds to guide its admission decisions. Measurement-based admission control is one such approach, and several measurement-based admission control algorithms have been proposed in the literature. In this paper, building on earlier work in [20], we use simulation to compare the performance of several of these algorithms. We find that all of them achieve nearly the same utilization for a given packet loss rate, and that none of them are capable of accurately meeting loss targets.

Keywords

Admission control, measurement, quality of service

I. Introduction

In an effort to better support applications with real-time constraints, several new per-flow packet delivery services have been proposed for the Internet (e.g., [24], [26]).[1] Lying between the extremes of *hard* real-time services (which provide worst-case guarantees) and the vagaries of the current best-effort service are *soft* real-time services that provide an enhanced quality of service without making hard guarantees. Specifications for these services might provide a delay *target*, rather than a bound, and permit periodic excursions above this target [6], or they might specify that the service provides low delay and low loss without quantifying actual performance [26].

One key difference between hard and soft real-time services is the nature of their admission control algorithms. Hard real-time services necessarily use *parameter-based* admission control algorithms that are based on worst case bounds derived from the parameters describing the flow; these algorithms typically result in low network utilization in the face of bursty network traffic. Soft real-time services can use less stringent admission control algorithms. It has long been

This work was begun while Lee Breslau and Scott Shenker were with the Xerox Palo Alto Research Center. At Xerox, this work was supported in part by the Defense Advanced Research Projects Agency, monitored by Fort Huachuca under contract DABT63-96-C-0105. Sugih Jamin's research is supported in part by the NSF CAREER Award ANI-9734145 and the Presidential Early Career Award for Scientists and Engineers (PECASE) 1998. Additional funding is provided by MCI WorldCom, Lucent Bell-Labs, and Fujitsu Laboratories America, and by equipment grants from Sun Microsystems Inc. and Compaq Corp.

[1] Here we are restricting our attention to services that make per-flow assurances; we are not addressing services that only give aggregate service assurances, such as the recent Differentiated Services proposals [5], [22], since they do not rely on per-flow admission control.

recognized that *measurement-based* admission control algorithms (MBACs) are more appropriate for these soft real-time services [6], [18]. Because they base admission control decisions on measurements of existing traffic rather than on worst-case bounds about traffic behavior, MBACs can achieve much higher network utilization than parameter-based algorithms while still providing acceptable service [19]. Of course, traffic measurements are not always good predictors of future behavior, and so the measurement-based approach to admission control can lead to occasional packet losses or delays that exceed desired levels. However, such occasional service failures are acceptable given the relaxed nature of the service commitment provided by soft real-time services.

In designing a measurement-based admission control algorithm, one can conceivably have two goals. One is to provide a parameter that accurately estimates *a priori* the level of service failures that will result. The other is to achieve the highest possible utilization for a given level of service failures. Several measurement-based admission control algorithms have been proposed in the literature (see, for example, [17], [10], [14], [15], [16], [19], [20], [11], [7], [13], [21]) and they implicitly or explicitly seek to achieve one or both of these design goals.

The proposed algorithms, although embracing similar goals, differ in four important ways. First, some algorithms are *principled*, based on solid mathematical foundations such as Large Deviation theory, and others are *ad-hoc*, in that they lack a theoretical underpinning. Second, the specific equations used in making admission decisions are quite different. Third, while all algorithms have a parameter that varies the level of achieved performance and utilization (by making the algorithm more or less aggressive), some algorithms attempt to calibrate this parameter and have it serve as an accurate estimate of the resulting performance, while others leave the parameter uncalibrated; in the latter case it is assumed the network operator will learn appropriate parameter settings over time. Fourth, the measurement processes used to produce an estimate of network load are very different; they range from a simple point sample estimate, to an exponentially weighted average, to estimates based on both the mean and variance of measured load. Thus, the space of measurement-based admission control algorithms is both heavily and broadly populated.

Somewhat surprisingly, given the number of papers on the subject, no comprehensive comparison of these algorithms exists. Previous comparisons (including our own previous work on the subject) look only at a few test cases, and then only for a few of the algorithms [20], [21]. In this paper we extend this previous work by considering more (although by no means all) of the proposed algorithms, and by subjecting them to more extensive tests. In all of these tests we use packet losses as the definition of a service failure.[2] We evaluate the algorithms according to how well they are able to meet the two goals of MBACs. First, we compare the *performance frontier* or *loss-load curve* (we will use these terms interchangeably) achieved by each algorithm; the loss-load curve depicts the rate of losses that occur at a given level of utilization. Second, for those algorithms that attempt to predict the resulting level of losses, how close is the resulting performance to the target?

On the first goal, we find that even though the algorithms are derived from diverse motivations and theories, they all produce essentially the same performance frontier. The particular theory upon which they are based and the specific admission equations they use seem to be of little consequence. Regarding the second goal, we find that none of the algorithms achieve the specified performance targets consistently. However, some algorithms do somewhat better than others; whether these differences are important, and whether future algorithms can do better, remains an open question.

The remainder of this paper is organized as follows. In the next section we describe the algorithms we include in our study and briefly review previous performance comparisons of the algorithms. In Section III we describe our simulation methodology and present experimental results comparing the performance frontiers of the various algorithms. In Section IV we study the extent to which algorithms can accurately predict the resulting loss level. We summarize our findings in Section V.

II. Measurement-Based Admission Control Algorithms

To give the context necessary for discussing our results, in this section we very briefly describe the six admission control algorithms whose performance we study. These algorithms represent a broad, though not complete, sample of existing MBACs. Each algorithm has two key components: a measurement process that produces an estimate of network load, and a decision algorithm that uses this load estimate to make admission control decisions. After presenting each of the six algorithms, we elaborate on some common features of the algorithms.

[2] Violations of a delay target may also be a relevant characteristic. However, for the fixed buffer regime we study, this is sufficiently similar to loss and so we do not treat it separately.

For the purposes of this study, we assume that applications use a signalling protocol, such as RSVP [3], to make their requests for service to the network. These service requests contain a traffic descriptor describing the worst case behavior of the application traffic. The traffic descriptor takes the form of a token bucket with parameters r and b denoting the token rate and bucket depth, respectively.[3] We measure the quality of the service delivered in terms of packet drops. Soft real-time services are typically intended to be scalable, therefore we only consider MBACs that require no per-flow state; that is, the measurements are taken on the aggregate traffic, not on individual flows. Since measurement is done on the aggregate and admission control decisions are made on a per flow, rather than a per packet basis, implementation overhead is not critical [20] and is not explored in this paper.

Some admission control algorithms do not fit within the framework we consider and are excluded from our study. For example, we do not include one of the MBACs described in [13] because it depends on per-flow (rather than aggregate) measurements. In addition to excluding algorithms that require per-flow measurements, we also do not consider algorithms that make any assumptions, either implicitly or explicitly, about the average behavior of flows. For example, we do not include the MBAC presented in [16] because it computes a per-flow average estimate and assumes that all arriving and departing flows conform to that average. We only consider algorithms that make no assumption about what a flow's contribution will be to aggregate load beyond the worst case parameters supplied by the flow. Similarly, when a flow departs the network, its prior contribution to aggregate load can only be determined by measuring subsequent aggregate load.

Following are brief sketches of the six admission control algorithms we compare:

- **Measured Sum (MS).** The Measured Sum algorithm [20] admits a new flow if the sum of the token rate of the new flow and the estimated rate of existing flows is less than a utilization target times the link bandwidth. A time window estimator is used to derive the estimated rate of existing flows.
- **Hoeffding Bounds (HB).** The admission control algorithm described in [11] computes the equivalent bandwidth for a set of flows using the Hoeffding bounds. A new flow is admitted if the sum of the peak rate of the new flow and the measured equivalent bandwidth is less than the link utilization. An exponential averaging measurement mechanism is used to produce the load estimate.
- **Tangent at Peak (TP).** Four measurement-based admission control algorithms are presented in [13]. The first algorithm, based on the tangent at the peak of an equivalent bandwidth curve computed from the Chernoff Bounds, admits a new flow if the following condition is met:

$$np(1 - e^{-sp}) + e^{-sp}\widehat{\nu} \leq \mu, \tag{1}$$

where n is the number of admitted flows, p is the peak rate of the flows, s is the space parameter of the Chernoff Bound, $\widehat{\nu}$ is the estimate of current load, and μ is the link bandwidth. This algorithm uses a point sample measurement process.

- **Tangent at Origin (TO).** A second algorithm presented in [13] uses a tangent to the equivalent bandwidth curve at the origin. Here, a new flow is admitted if the following equation is satisfied:

$$e^{sp}\widehat{\nu} \leq \mu. \tag{2}$$

This admission control algorithm also uses the point sample measurement process.[4]

- **Measure CAC (MC).** The *Measure* admission control algorithm [7], which is based on large deviation theory, admits a new flow if the sum of the peak rate of the flow and the estimated bandwidth of existing flows is less than the link bandwidth. The estimated bandwidth takes as input a target loss rate and makes use of the scaled cumulant generating function of the arrival process.
- **Aggregate Traffic Envelopes (TE).** The admission control algorithm in [21] uses measurements of the maximal traffic envelopes of the aggregate traffic, capturing variability on different time scales. Both the average and variance of these traffic envelopes, as well as a target loss rate, are used as input into the admission algorithm.

The brief descriptions presented above ignore the details of the individual algorithms, but the key point is that the algorithms differ both in their underlying theory and in the specific measurement and admission control equations

[3] Some of the admission control algorithms require a peak rate p. Following [11], the peak rate is computed from the token bucket parameters as $p = r + b/T$, where T is the basic measurement interval used by the algorithm.

[4] A third algorithm presented in [13] is equivalent to the HB algorithm. As described above, the fourth algorithm is excluded because it depends on per-flow measurements.

they use. While these differences are what we seek to understand in this paper, certain similarities are worth noting. For instance, each of these algorithms has one component that derives a load estimate based on measured traffic and another component that makes an admission decision using this load estimate. Rather than treating each algorithm as a monolithic block, it is possible in some cases to pair the estimation process of one algorithm with the decision process of another. This allows us to ask whether differences in performance derive from the estimation process, the decision process, or both. We undertake this "mix and match" analysis in Section III.

In addition to the equations that form the basis of the algorithms described above, there are also certain MBAC features that address specific practical concerns. For instance, when a new flow is admitted to the network, the existing load estimates will not immediately reflect the presence of the new flow. In such a case, the network runs the risk of admitting too many flows before recognizing that load has increased. To prevent this situation, some of the algorithms (MS, HB, MC) artificially increase the load estimate to account for a newly admitted flow. This feature, while included in the specifications of three algorithms, can be seen as an independent mechanism that can be applied to any of them. We eliminate this feature as a source of performance differences between algorithms by including it in all of the algorithms in our performance comparison.[5]

A final observation is that each of the admission control equations has one or more parameters that control their operation. For example, the MS algorithm has a *utilization target* that affects how many flows will be admitted, the MC and TE algorithms use a *target loss rate*, and the HB algorithm has a parameter that indicates the probability that the actual bandwidth requirement exceeds the estimates. While these parameters were not all intended as tuning parameters by the designers of the algorithms, adjusting these parameters will make the algorithms either more conservative or more aggressive with regard to the number of flows they admit. Hence, instead of providing a single level of performance, each algorithm enables a range of loss rates and utilizations depending on the values of these parameters. Thus, we describe the utilization performance of these algorithms by their loss-load curves or performance frontiers.

This paper is an extension of our earlier work [20]. In that paper we compared three different measurement-based admission control algorithms (MS, HB, and an acceptance region based MBAC from [14] which was later generalized in [13]) and one simple parameter-based admission control algorithm. These algorithms were compared for several different traffic loads (similar to those we use here, to be described in Section III-A) and on single link and multiple link network topologies (as we discuss in Section III-A, we only use a single link network topology in this paper). The simulation results in the earlier paper were deficient in several respects. The algorithms were only tested at one parameter value setting. Such *point comparisons* cannot describe the entire performance frontier provided by an admission control algorithm, and so do not adequately characterize the performance of an MBAC. Moreover, for the particular parameter values and traffic models used in [20], the admission control algorithms recorded no losses, so only the utilization figures could be compared. Also, there was no attempt to compare the target loss rate with the actual loss rates, so there are no results analogous to those in Section IV. Thus, this previous work did not adequately answer the relevant question: how well do the various MBACs satisfy the two goals of measurement-based admission control?

There have been few other attempts to systematically compare the performance of measurement-based admission control algorithms. The closest work is [21], in which the performance of the TE algorithm is compared to that of HB and the algorithm specified in [19].[6] The authors of [21] compare utilization achieved for particular quality of service targets, and do not compare the performance frontiers of the algorithms; however, the main thrust of [21] is on achieving accurate loss estimates, and to evaluate success along that dimension it is not necessary to investigate the entire performance frontier.

In one other related piece of work, in a short (three page) discussion paper [4] we briefly review some of the research presented here and then use that to argue that the research agenda in measurement-based admission control should address certain policy issues (such as how to allocate admission between large and small flows, and between flows travelling many hops and those travelling fewer hops).

[5]Results of simulations not included in this paper show the importance of this feature. Under highly dynamic conditions, performance can degrade if estimation algorithms do not account for the presence of newly admitted flows.

[6]Based on communication with the author of [11], we do not interpret the parameter in HB as a performance target; however, one could easily make that interpretation, and that is what is done in [21].

III. Performance Frontiers

In this section we evaluate how well each of the six algorithms performs with respect to the first goal: achieving high network utilization and low packet loss. We first describe our simulation methodology and present our basic results for the MBACs with several different source models. We then focus on three specific issues: the impact of heterogeneous traffic, a comparison between MBACs and an ideal parameter-based algorithm, and implications of long range dependent traffic on measurement-based admission control. Throughout the discussion and accompanying figures, we refer to the algorithms by the abbreviations introduced in the previous section: MS, HB, TO, TP, MC, TE.

A. Simulation Methodology

We use discrete event simulation to generate performance frontiers for each algorithm. Simulations were carried out using the *ns* network simulator.[7] In order to understand the behavior of the algorithms in the most simple case, we used a simple topology in which admission control was employed on a single bottleneck link. While interesting issues may arise when studying admission control in a multi-link scenario, the basic performance aspects of these algorithms are most easily revealed in this simpler one-link configuration, particularly since the admission control decisions for each of the algorithms are made on a link-by-link basis. Further, we expect that issues arising in a multi-link scenario (e.g., discrimination against larger flows and flows traversing longer paths [4], [19]) are independent of the particular algorithms and are, therefore, orthogonal to the questions we ask here.[8]

A simulation experiment consists of a random process of flow arrivals. Each flow requests service from the network using a simple resource reservation protocol, and it is admitted or rejected according to the specifics of the algorithm in question. A rejected flow departs the network without sending any data packets and does not retry its service request again. A flow that is accepted sends data packets for a flow lifetime chosen from a random distribution. Packets are generated according to a source model selected for the flow when it is created.

We use two kinds of source models in our experiments. The first is an ON/OFF source, in which the source transmits at a constant rate during a randomly chosen ON period, and then remains idle for a randomly chosen OFF time. The second kind of source model uses a trace of video traffic to drive the simulation. The specific parameters are described below. Packets generated by a source are subject to policing by a token bucket filter. The token bucket parameters (rate and bucket depth) are included in the reservation request that is handed to the admission control module.

For each simulation, the average utilization and packet loss rate are measured. Data collected during an initial warmup period are discarded. All simulations were repeated using different random number seeds. The number of repetitions and the length of each simulation were varied depending on the underlying variability of the source model and offered load used in each experiment. The averages across all repetitions are reported in our results.

In all experiments, the bottleneck link bandwidth is 10 Mbps. Unless otherwise noted, packets are 128 bytes long, and there is buffering for 160 packets at the bottleneck link. In most of our experiments, the total offered load (in terms of the number of flows requesting service) is high, leading to a high call rejection rate. While the actual rejection rates may be unrealistically high, it is in the regime of overload that the behavior of the admission control algorithms is most interesting. Below we also present results from simulations with more moderate demand.

Each of the algorithms has several parameters that control how much history is maintained by the estimation algorithm. We tried, when possible, to use parameter settings suggested in the original references. However, in some cases we found that changing these values yielded better performance. We suspect that this is due to differences between our source models and offered load and those used by other researchers. In all cases, we used those parameter values that yielded the best performance in our experiments.

B. Results

Our first experiments use homogeneous on/off sources with exponentially distributed on and off times (325ms average). The transmission rate during on periods is 64kbps, making the average rate 32 kbps. The token rate and bucket depth are set to 64 kbps and 1 packet, respectively (assuring no loss at the token bucket filter). These parameters are

[7] http://www-mash.cs.berkeley.edu/ns/.

[8] This is not to say we don't think these issues are interesting. In fact, given the results we present here, we make the case in [4] that these issues of discrimination mentioned above should be considered more seriously by MBAC researchers.

consistent with PCM coded voice that might be produced by an IP telephony application. On average each source consumes about .3% of the link bandwidth. Flow inter-arrival times are exponentially distributed with a mean of 400 ms. Flow lifetimes, which are also exponentially distributed, have a mean of 300 seconds. We refer to this traffic model as the EXP1 source. Simulations were run for 6000 simulation seconds; data collected during the first 1500 seconds was discarded. Each simulation was repeated 5 times with different seeds to the random number generator.

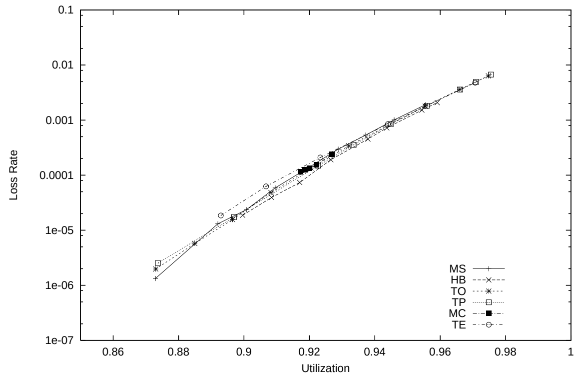


Fig. 1. Performance frontiers of measurement-based admission control algorithms with EXP1 traffic

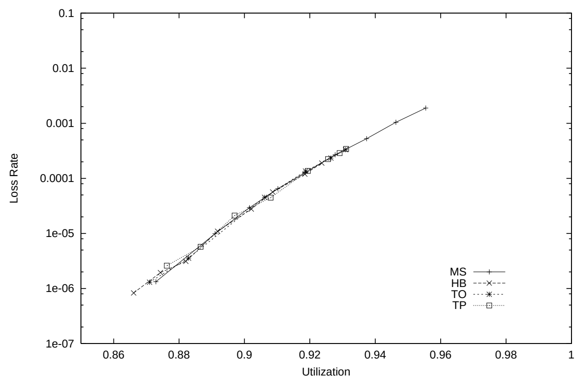
Results for this experiment are shown in Figure 1. This graph plots the packet loss rate on a log scale as a function of link utilization. A performance frontier is shown for each of the six algorithms.[9] It is difficult to distinguish between the performance frontiers in the graph, indicating that all of the algorithms yield very similar performance. That is, they all permit essentially the same choices in the tradeoff between loss rate and utilization. Further, the very slight differences in performance are not of practical importance, because even if one algorithm yields a marginally higher loss rate than another at a given level of utilization, the loss rates can be made equivalent with extremely small changes in utilization. Because there is variance in both the x and y values in the figure (i.e., a given MBAC input parameter determines both the utilization and packet loss rate), and these variations are highly correlated and not normally distributed, we do not depict these variations as error bars in our graphs. Table II shows the means and standard deviations for points plotted in Figure 1. As is evident, the variance across simulation runs is small.

In some cases, the interfaces between the estimation and decision components of each algorithm are such that the estimation process of one can be used with the decision process of another. When this was possible, we "mixed and matched" the various components. Specifically, the MS, HB, TO and TP decision algorithms were run with the Time Window, Exponential Averaging, and Point Sample estimators in order to understand the degree to which each component impacts the results. Figure 2 shows the results for these three estimators with the four different decision algorithms. Relative to Figure 1, the slight variations across algorithms have been reduced. This result demonstrates two things. First, the conclusion above that each algorithm has nearly the same performance frontier does not depend on any particular coupling between estimation and decision processes. Second, the reduced variance indicates that it is the estimation process, and not the decision algorithm that is responsible for the slight variations in Figure 1.

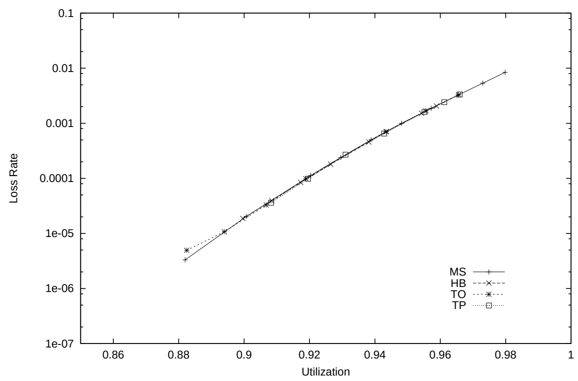
We performed additional experiments using the following source models:

- EXP2 – in this source model, the peak rate is increased by a factor of 10 (640 kbps versus 64 kbps) relative to the EXP1 source while the average rate is held constant, leading to a burstier source model. As can be seen in Figure 3, the increased burstiness is reflected in the fact that equivalent loss rates are achieved at much lower link utilization compared to our earlier results.
- POO1 – this is an on/off source with the same averages as the EXP1 source. However, they are taken from a Pareto distribution. Flow lifetimes are taken from a lognormal distribution with a median of 300 seconds following [2], [9]. The aggregation of these sources produces traffic that is long range dependent [8], [25]. Because of increased variability with these traffic sources, simulations were run for 20,000 seconds, with data for the first 10,000 seconds discarded. The resulting performance frontiers are shown in Figure 4.

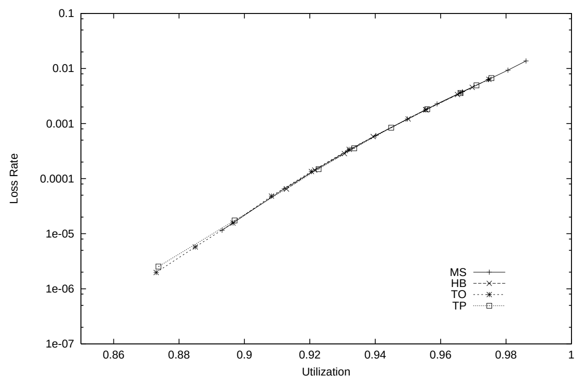
[9] Because utilization is not an independent variable in these experiments, data points are not plotted for the same x values for each algorithm. The actual number of points plotted varies across algorithms, but we have covered an overlapping range on the x axis for each curve.



(a) Time Window



(b) Exponential Averaging



(c) Point Sample

Fig. 2. Decision algorithms paired with different estimators for the EXP1 source model

• STARWARS – this source model is taken from a trace file produced by an MPEG encoding of the Star Wars motion picture [12]. Each source starts from a random place within the trace file in order to avoid correlation among the sources. This source model differs from the previous ones in that it has a higher average rate (350kbps vs 32kbps) resulting in a lower degree of multiplexing, and it is characteristic of traffic produced by a video source rather than an on/off model. With this source model, packets are 200 bytes long and there are 500 packet buffers at the bottleneck link. Figure 5 exhibits the performance frontiers for the various MBACs with this source model.
• HET – this experiment consists of a mix of six different on/off sources, with varying average rates, idle times and burst times. Each arriving flow chooses from among these source models at random. All flows have the same leaky bucket parameters, so they appear identical to the admission control algorithm. Figure 6 shows that the performance frontiers remain quite similar under this heterogeneous traffic load.

The results from these experiments reveal that our basic result holds across different traffic models. That is, in the presence of burstier sources, long range dependent traffic, lower multiplexing, traffic derived from a video trace, and heterogeneous traffic (with identical token bucket parameters), all of the algorithms achieve roughly the same performance frontier.

The experiments presented thus far used a very heavy offered load, yielding a high flow rejection rate. We now present results from experiments with the EXP1 traffic source with more moderate load (approximately 10% above capacity.) The results from this experiment indicated slightly more variation across algorithms than was evident previously. However, these variations again disappeared when the “mix and match” methodology was employed. That is, different admission control algorithms achieved the same performance frontiers when using the same estimator. This reinforces the point above that it is the estimation algorithm that is responsible for the small variations in performance across algorithms. Figure 7 shows the performance frontier for the four algorithms included in the “mix and match” experiments with EXP1 traffic source and lower offered load.

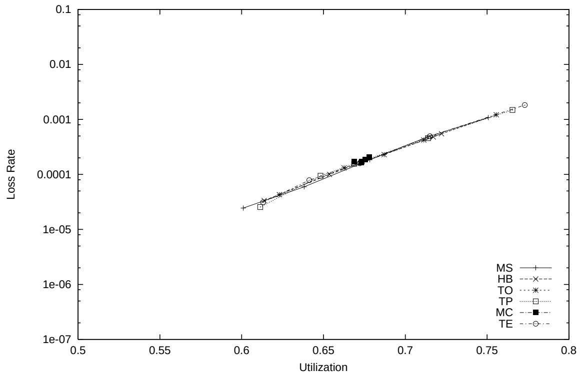


Fig. 3. Performance frontiers for measurement-based admission control with EXP2 traffic

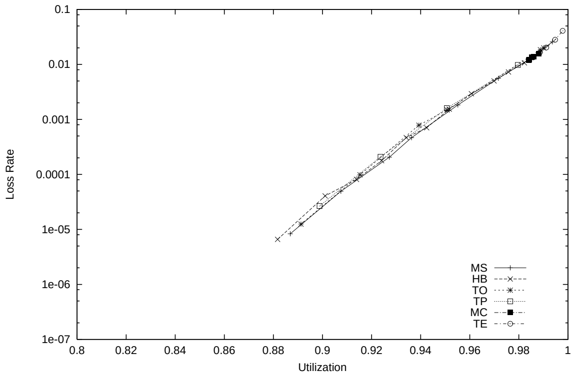


Fig. 4. Performance frontiers for six measurement-based admission control algorithms with long range dependent traffic

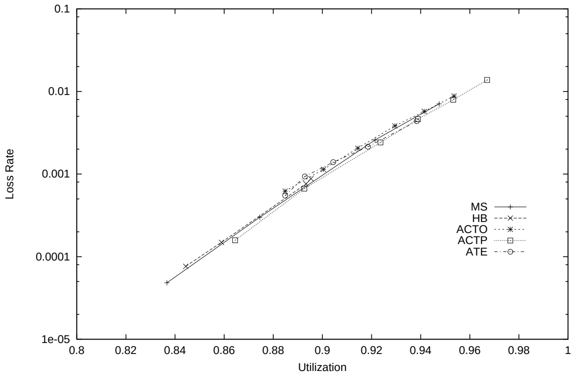


Fig. 5. Performance frontiers for measurement-based admission control algorithms with the *Star Wars* trace

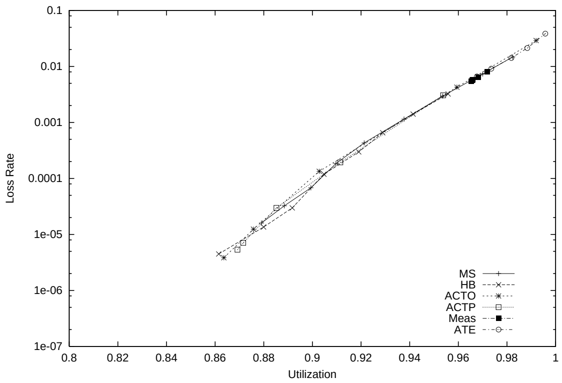


Fig. 6. Performance frontiers for measurement-based admission control algorithms with heterogeneous traffic

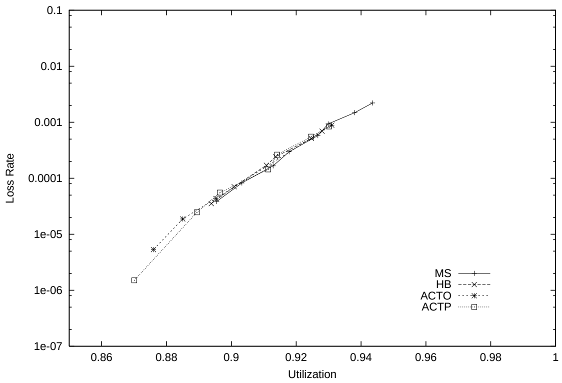


Fig. 7. Performance frontiers for measurement-based admission control algorithms using Point Sample estimators, EXP1 traffic and moderate offered load

C. *More on Heterogeneous Traffic*

We now briefly return to the issue of heterogeneous traffic. In the simulation with heterogeneous traffic described above, all flows had identical token bucket parameters, and so were indistinguishable to the admission control algorithms. The results in this case were consistent with the homogeneous experiments. We now ask what happens when the token bucket parameters are no longer identical, allowing the admission control algorithms to admit them differentially. Not surprisingly, when the flows are distinguishable, different admission control algorithms lead to different mixtures of traffic, and hence to different performance frontiers. To illustrate this, we consider an experiment where each arriving flow used one of the following two source models, chosen with equal probability. The first source model was the Star Wars trace introduced above. This trace had an average rate of approximately 350 kbps. In order to accommodate its burstiness, the token bucket parameters are $r = 800kbps$ and $b = 200kb$. The second source model was a Constant Bit Rate (CBR) source sending at 800 kbps. The token bucket parameters for this source are $r = 800kbps$ and $b = 1.6kb$ (to hold a single packet).

Figure 8 shows results for this heterogeneous traffic mix with 2 admission control algorithms. The first is the Measured Sum (MS) algorithm, which uses the token rate of the new flow. The second algorithm is a variant of Measured Sum using the peak rate (computed as $p = r + b/T$, with $T = 500ms$ in our experiments) of the incoming flow, rather than its token rate, in the admission control equation. The first version of the MS algorithm does not discriminate between the two kinds of flows because they have the same token rates. This leads to a traffic mix that is made up of roughly equivalent numbers of the two kinds of flows. The peak rate algorithm, on the other hand, discriminates against the trace driven flows, as they have a higher peak rate ($1200kbps$ vs. $800kbps$). This leads to a traffic mix in which the CBR sources outnumber the video sources by a ratio of approximately 3:1. Consequently, the peak rate algorithm has a better performance frontier than the token rate algorithm. We introduced the peak rate version of the MS admission control algorithm to accentuate the extent of discrimination. One finds similar, but less extreme, results when comparing the six admission control algorithms we have discussed in this paper under heterogeneous traffic loads with distinguishable flows.

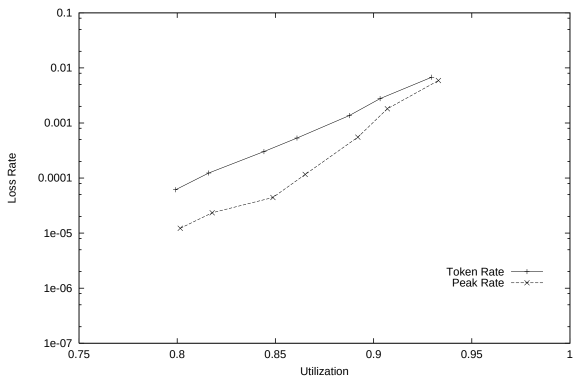


Fig. 8. Peak rate versus token rate versions of the MS admission control algorithm with heterogeneous traffic

Note that the traffic mix admitted by the peak-rate algorithm is, in the aggregate, less bursty than the one admitted by the token rate algorithm; thus, the loss rate experienced at an equivalent utilization is lower than is experienced with the token rate admission control algorithm. In general, when admission control algorithms admit different mixtures of flows, the aggregate traffic will have different degrees of burstiness, and so the performance frontiers will no longer be the same. Thus, in the face of heterogeneous and distinguishable flows, MBACs don't necessarily produce the same performance frontier.

One might think that this would undercut our observation about the equivalence between various MBACs. However, we think that the question of which traffic mixture should be admitted is one of policy, not efficiency. Clearly one could minimize the loss rates by admitting only CBR-like flows, but such a limitation would be unwise as it would preclude bursty sources from obtaining reasonable service. Admission control algorithms that happen to pick less bursty flows to admit, while providing superior performance frontiers (in the presence of heterogeneous and distinguishable traffic) are not necessarily more desireable and in fact have only made one particular policy choice out of a broad range of possible choices.

D. Comparison with an Ideal Algorithm

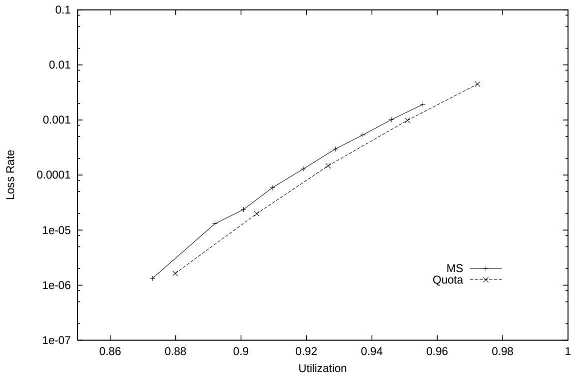
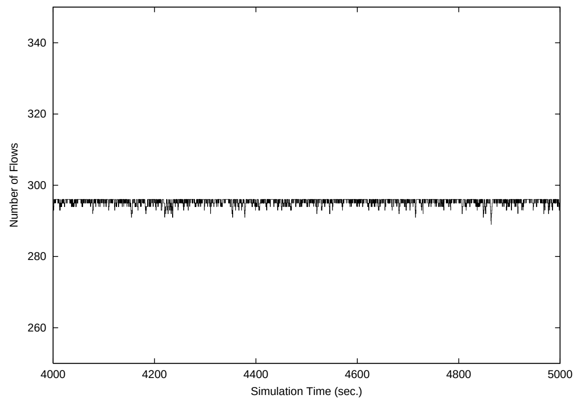
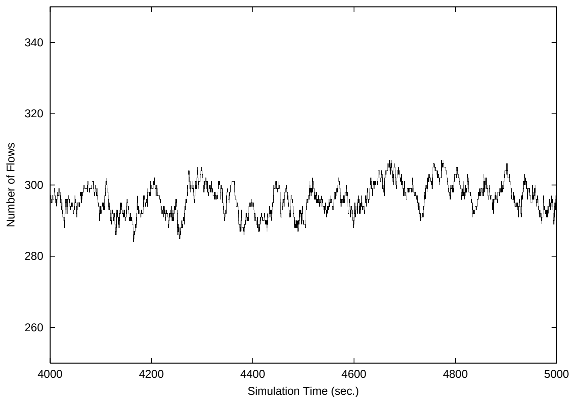


Fig. 9. Performance frontiers for MS and Quota algorithms with EXP1 traffic

We now elaborate on our result that all of the algorithms have similar performance frontiers. With so much effort going into the design of measurement-based admission control algorithms, one might have assumed that the effort would lead to improved performance. Our simulations suggest quite the opposite, that even very simple *ad hoc* algorithms achieve the same performance frontier as more complicated and more principled ones. Given this, we ask the following two questions: why are the differences in performance between the algorithms so small? Are there untapped advantages not yet realized by any of these algorithms or are they in fact all performing at or near some optimal level? To answer these questions, we construct an "ideal" algorithm.



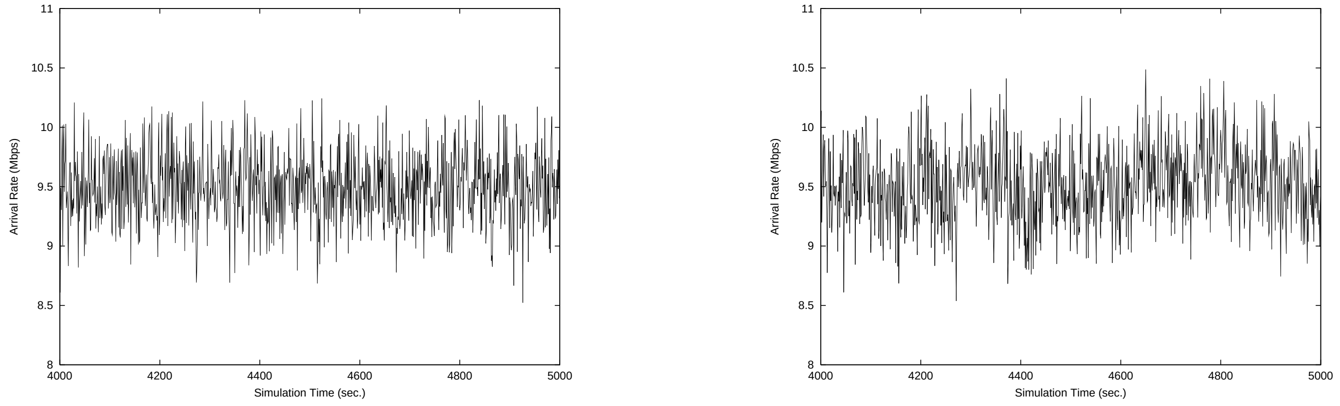
(a) Quota



(b) MS

Fig. 10. Number of admitted flows as a function of time for (a) Quota and (b) MS algorithms with EXP1 traffic source

Consider our initial experiment with the EXP1 traffic source. In this simulation, all flows in the network were homogeneous exponential on/off sources. The aggregate traffic generated by these sources has no long term correlation. Further, the time scale at which individual sources change between the idle state and the active state (100s ms) is shorter than the time scale at which new flows are admitted to the network (seconds). Thus, it is impractical for the admission control algorithm to attempt to adjust to short term fluctuations in traffic (i.e., on the time scale of bursts). Given that there are no long term correlations in the aggregate traffic, the ideal strategy for admission control is to keep long term average load constant. While this might present a challenge in reality, it is trivial in our simulation environment when we have homogeneous flows with no long term correlations. Hence, for present purposes we define the *Quota* algorithm, which does not depend on measurements. This simple algorithm admits a newly arriving flow if there are less than n flows currently receiving service, and rejects the flow otherwise. The parameter n controls how conservative



(a) Quota

(b) MS

Fig. 11. Arrival rate as a function of time for (a) Quota and (b) MS algorithms with EXP1 traffic source

or aggressive the algorithm is. While this algorithm is helpful in better understanding the limits of the performance of MBACs, it is impractical in any real setting since it requires homogeneous flows.

Figure 9 plots the performance frontiers for the Quota algorithm and for one of the measurement-based algorithms (MS) with the EXP1 traffic. As the figure shows, the Quota algorithm outperforms the measurement-based algorithm; across the load levels tested, the loss rate for the measurement-based algorithm is between 50% and 250% higher than that of the Quota algorithm.

The differences between the aggregate performance of the Quota and measurement-based algorithms can be understood by looking at their behavior on shorter time scales. Figure 10 shows the number of admitted flows as a function of time for a portion of a single simulation for each of the 2 algorithms. As expected, the Quota algorithm yields essentially a straight line with very small deviations (representing the random delay between the departure of a flow and the arrival of a new one.) The MS algorithm mimics the Quota algorithm fairly well, but there is significant variation in the number of admitted flows. Similar variations in load occur with the other MBACs we evaluated. The differences in the number of flows admitted by the Quota and MS algorithms are reflected in number of packet arrivals at the bottleneck link. Figure 11 plots the arrival rate (in Mbps) at the bottleneck link over 1 second intervals as a function of time for the 2 algorithms.[10] The MS algorithm exhibits higher variance in utilization than the Quota algorithm. The ratios of the standard deviation to the mean are .034 and .031 with the MS and Quota algorithms, respectively. Increased variance in load leads to higher loss rates for a given average load. Thus, with the EXP1 traffic source, it is precisely the variation in admitted load that leads to the worse performance frontier for the measurement-based admission control algorithm. Is this variation inevitable, or can MBACs eventually match the performance of the Quota algorithm?

There are two distinct causes for this variation leading to the performance degradation relative to the Quota algorithm. The first is the way that the measurement-based algorithms must deal with the arrival and departure of flows. Because the measurement-based algorithms we consider use aggregate rather than per-flow measurements, they do not know how much a departing flow was contributing to the previous estimate of load. Measurement-based algorithms must therefore wait before admitting a new flow until new measurements reflect the departure of the previous flow. During this time, additional flows may depart, and the number of flows in the system may drop. The Quota algorithm on the other hand, with its perfect but unrealistic knowledge of the departing flow, can immediately admit a new flow. Similarly, when a new flow is admitted to the system, measurement-based algorithms must assume worst case behavior about the new flow until new measurements reflect its presence. In contrast, the Quota algorithm can admit flows based on their average behavior and need not delay further admissions.

The second factor leading to variation in the number of admitted flows is that measurement-based admission control algorithms, by their reliance on measurements of current traffic, must necessarily respond to significant fluctuations in the load even when the number of flows has not changed. That is, the MBAC cannot distinguish between having too many flows admitted and a long fluctuation to a higher level of aggregate traffic by a fixed set of flows; not being able

[10]Note that in any interval, the offered load (measured by the bit arrival rate) on the link can exceed the link bandwidth. Excess packets are buffered or dropped.

to detect the difference, the MBAC is forced to turn away flows during such a fluctuation even when there are too few flows present and similarly, if the current flows fluctuate to a lower level of traffic, the MBAC is forced to admit flows even when too many are already present.

The impact of these issues on the performance of MBACs is illustrated by examining an enhanced, if unrealistic, MBAC that is designed to avoid these problems. Our new algorithm, which we call MS+, operates as follows. In each measurement interval, the average per-flow bandwidth is computed (i.e., total load divided by the number of flows.) Exponential averaging is used to obtain a long-term per-flow average. A new flow is admitted if the following condition is satisfied:

$$\widehat{r} * (n+1) < \upsilon\mu, \tag{3}$$

where $\widehat{r}$ is an estimate of per-flow bandwidth based on measurements, n is the number of admitted flows, μ is the link bandwidth, and υ is a utilization target intended to limit the maximum link load. The aggregate estimate of load in this algorithm is $\widehat{r} * n$, so when a flow is admitted or departs the estimate is immediately updated to reflect the change. This algorithm uses knowledge about the average contribution of each flow to aggregate load. The algorithm is unrealistic as an actual MBAC because it depends on flow homogeneity (i.e., a newly arriving flow is assumed to have the same characteristics as existing flows). In addition, the algorithm requires explicit *teardown* messages, which may not be provided by all signalling protocols. Nonetheless, its performance shows what an MBAC that has knowledge about average flow behavior can do. Results for this algorithm are shown along with the MS and Quota algorithms in Figure 12. As is evident, by taking advantage of knowledge about average behavior in an environment with homogeneous flows, the MS+ algorithm performs much better than the MS algorithm.

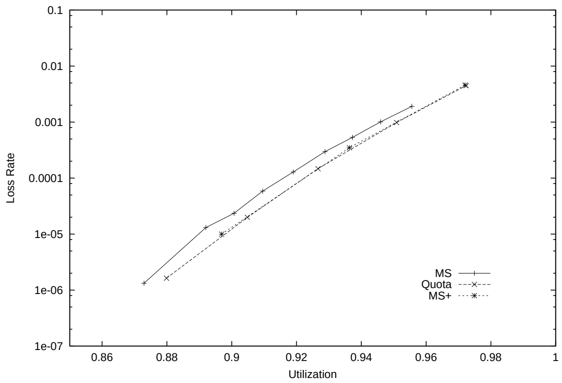


Fig. 12. Performance frontiers for MS, Quota and MS+ algorithms with EXP1 traffic.

Note that there is an inherent tension between the two factors that cause MBAC performance to degrade relative to the Quota algorithm. To avoid adapting to short term fluctuations in load, longer measurement intervals are suggested [15], [19]. Longer measurement intervals, on the other hand, will only slow down the reaction of the measurement-based algorithms to the departure and arrival of flows. Therefore, it is likely that these two factors will prevent any measurement-based algorithm from ever performing as well as the Quota algorithm.

If MBACs could emulate the Quota algorithm, then they would all have the same performance frontier, and our results in Section III-B would be rendered obvious. However, the discussion above shows that MBACs cannot accurately emulate the Quota algorithm. The surprise in our results in Section III-B is that the set of MBACs we tested all had such similar deviations from the *ideal* behavior of the Quota algorithm. One might have thought (indeed, we did think) that different admission control equations and different measurement procedures would make a difference in how well this ideal was followed; our results suggest that this is not the case.[11]

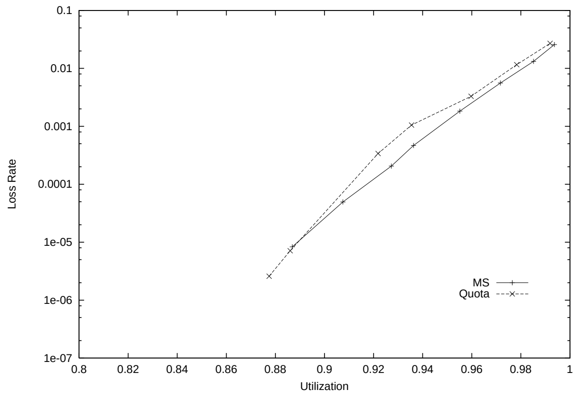
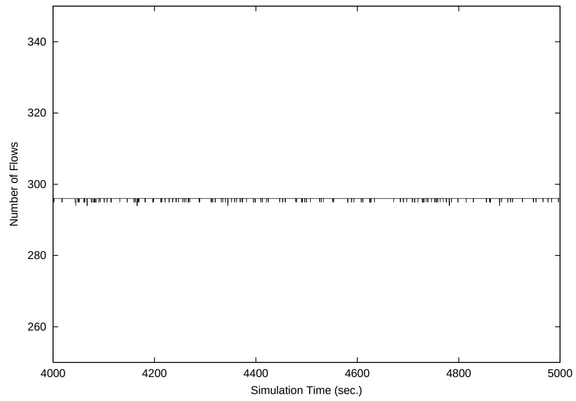
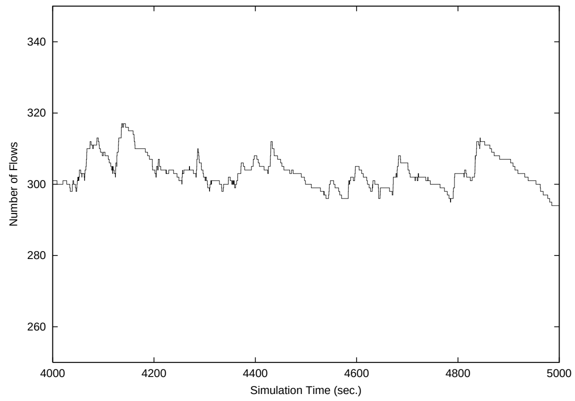


Fig. 13. Measurement-based algorithm vs. Quota Algorithm for LRD traffic

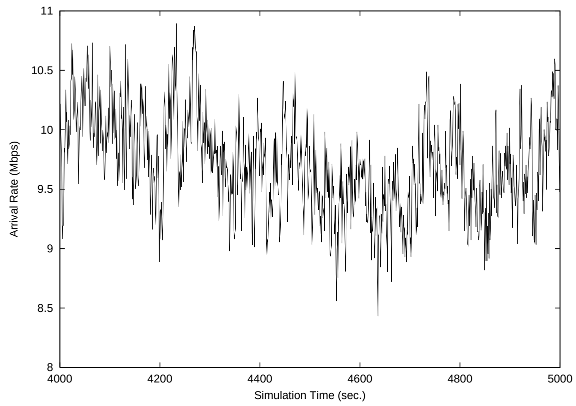


(a) Quota

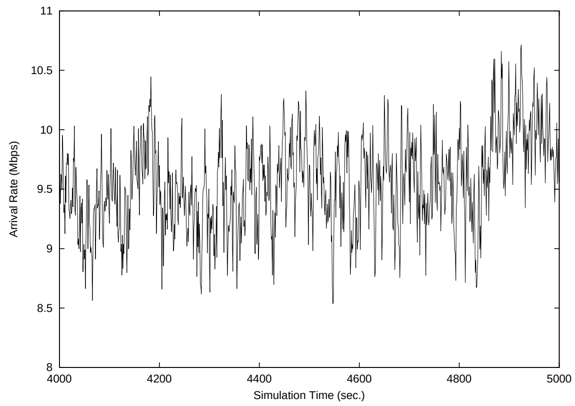


(b) MS

Fig. 14. Number of admitted flows as a function of time for (a) Quota and (b) MS algorithms with POO1 traffic source



(a) Quota



(b) MS

Fig. 15. Arrival rate as a function of time for (a) Quota and (b) MS algorithms with POO1 traffic source

E. Long Range Dependence

Before turning to the second goal of MBACS (performance targets) we briefly discuss long range dependence and its effect on admission control. Long range dependence has been observed in video traffic [1], [12] and may also arise from the aggregation of audio traffic [11], two traffic classes that may be subject to admission control. Results for the POO1 traffic source showed that our basic result, that the various MBACs have similar performance frontiers, remains unchanged in the face of long range dependent traffic. However, the relative performance of the Quota algorithm, held up as an *ideal* algorithm in the previous Section, is quite altered by the presence of long range dependent traffic.

Above we showed that the Quota algorithm, which admits a fixed number of flows, performs better than the measurement-based algorithms with the EXP1 traffic source (which does not give rise to long range dependent aggregate traffic). We repeated these experiments using the POO1 traffic source. Results are shown in Figure 13, and for clarity we again show only the Measured Sum algorithm and the Quota algorithm. With long range dependent traffic, we see the opposite results. In this case, the measurement-based algorithm performs *better* than the Quota algorithm (which one can think of as a simple, if unrealistic, parameter-based algorithm). The explanation for this is straightforward. The long range dependent traffic exhibits variations over long time scales. By keeping the number of flows fixed, the Quota algorithm does nothing to smooth these variations. The measurement-based algorithm, on the other hand, is able to adjust the number of flows admitted in response to the variations. As the aggregate load increases, departing flows need not be replaced by new ones, and when load decreases additional flows can be admitted. This behavior is shown in Figures 14 and 15, which plot the number of admitted flows and the bit arrival rate as a function of time for the Quota and MS algorithms. While increased variance in the number of flows and offered load were positively correlated with the EXP1 traffic, the opposite is true with the POO1 traffic source. The MS algorithm varies the number of flows admitted over time (Figure 14), but it actually smooths variations in load over time relative to the Quota algorithm (Figure 15). The ratios of the standard deviation to the average load are .037 and .042 for the MS and Quota algorithms, respectively.

We believe the implications of this are important. The original arguments for using measurement-based admission control claimed that the worst-case behavior of bursty traffic is far worse than the average case, and that it is hard *a priori* to know the average behavior of a bursty traffic flow. Since the average behavior is unknown, any parameter-based algorithm must be based on worst-case parameters, leading to low network utilization. However, our results here indicate that that argument should be taken one step further. Even if the *average* behavior of traffic flows were known, the existence of long range dependent traffic would still mandate the need for measurement-based admission control in order to adapt to these long time scale fluctuations. Thus, while it has previously been suggested that long range dependence may present certain challenges for measurement-based admission control [11], [23] (and we do not disagree with those arguments), we believe that long range dependence also provides additional motivation for the use of measurement-based admission control. When the time scale of flow arrivals and departures is shorter than that of the ebb and rise of traffic, measurement-based admission control enables the network to react to these traffic fluctuations.

These results on long range dependence also shed light on another issue. Some have argued that our basic result—that the performance frontiers of MBACs are very similar—follows quite directly from the observation that all algorithms seek to mimic the Quota algorithm. In Section III-B we found that there are inherent limitations to how closely any MBAC can mimic the Quota algorithm. Our results about long range dependence further show that mimicking the Quota algorithm is not always the optimal behavior.

IV. Performance Targets

Results in the previous section showed that all the measurement-based algorithms are capable of making the same tradeoff between utilization and loss. However, network operators who will deploy these algorithms may be interested in more than just knowing that the algorithms achieve the same tradeoff. Rather, it may be important for a network operator to know *how* to end up at a particular point on the performance frontier, so that a desired loss rate can be achieved. When comparing algorithms, it is important to ask to what extent their input parameters are useful in predicting actual performance. An algorithm that allows an operator to control resulting performance will be preferred over one that does not.

[11] Our fuller set of simulations (not presented here) suggest that the length of the averaging periods, and the way in which new flows are treated, are much more important than the equations themselves in determining how close MBACs come to the performance frontier of the Quota algorithm. This is consistent with the observations above about the two causes of the variations, since they both relate to measurement intervals and the treatment of new flows, and are orthogonal to the specific equations used in the admission decision.

We note that not all of the designers of the algorithms we study intended their algorithms to be tunable, nor did they all make claims about how well the algorithms were able to meet a particular performance target. Hence, we undertake this evaluation not to judge whether a particular algorithm meets its design objectives. Rather, we begin with the observation that each algorithm has one or more parameters that can be adjusted to control performance. We ask whether these parameters are able to provide functionality that network operators may find useful.

The tuning parameter in the TP and TO algorithms represents the space parameter of the Chernoff Bound used to compute the equivalent bandwidth curve upon which the algorithms are based. As such, this parameter does not represent a meaningful performance target. One may then ask whether this parameter can be mapped into a useful performance value in a deterministic way. For instance, if a particular parameter value in the TO algorithm always yields the same loss rate, then the parameter can be useful in predicting actual performance. However, a review of our simulation results shows that this is not the case. As an example, with the TP algorithm, a parameter value of $4.0e^{-7}$ yields loss rates of .0098, .0018 and less than 10^{-7} with the POO1, EXP1 and Star Wars sources, respectively. These kinds of inconsistencies were also observed with the TO algorithm. Thus, the tuning parameter in the TO and TP algorithms can not be used to predict actual performance.

The MS algorithm has a parameter, υ, which represents a cap on the fraction of the link bandwidth that can be used by traffic subject to admission control. As such, its semantics are easily understood, and we can ask whether it is useful as a utilization target. Simulation results indicate that it is not. For example, with the EXP1 traffic source, when $\upsilon = 1.0$, average utilization is 94% of the link bandwidth. With the EXP2 traffic source, utilization is only 75% of the link bandwidth with the same value of υ. Further, even if the utilization target was consistently met, we question the value of this parameter as a performance target. We expect loss rate to be a more relevant parameter, since loss rate directly affects user performance.

The HB algorithm uses a paramter, ϵ, to represent the probability that the stationary bandwidth requirement of a set of flows exceeds the computed equivalent bandwidth of the flows. In practice, this does not turn out to be a useful predictor of loss. For example, in the simulations shown previously, we typically use values of ϵ above .9. Further, these values do not map into actual loss in any consistent manner. For example, with $\epsilon = .9$, the loss rates are .00045, .005 and less than 10^{-7} with the EXP1, POO1, and Star Wars source models, respectively.

Algorithm	Source Model	Target Loss Rate	Actual Loss Rate
TE	EXP1	10^{-6}	1.9×10^{-5}
TE	EXP1	10^{-2}	4.8×10^{-2}
TE	Star Wars	10^{-6}	5.5×10^{-4}
TE	Star Wars	10^{-2}	4.4×10^{-3}
TE	EXP2	10^{-6}	3.1×10^{-5}
TE	EXP2	10^{-2}	1.8×10^{-3}
TE	POO1	10^{-6}	1.3×10^{-2}
TE	POO1	10^{-2}	4.1×10^{-2}
MC	EXP1	10^{-6}	1.1×10^{-4}
MC	EXP1	10^{-2}	2.4×10^{-4}
MC	Star Wars	10^{-6}	3.0×10^{-3}
MC	Star Wars	10^{-2}	4.5×10^{-3}
MC	EXP2	10^{-6}	1.7×10^{-4}
MC	EXP2	10^{-2}	2.0×10^{-4}
MC	POO1	10^{-6}	1.2×10^{-2}
MC	POO1	10^{-2}	1.6×10^{-2}

TABLE I

TARGETED VERSUS ACTUAL LOSS RATES FOR THE TE AND MC ALGORITHMS

The final two algorithms, TE and MC, use target loss rate as a tuning parameter. Table I shows both the target and actual loss rates for both algorithms and several traffic sources. These data show that the algorithms are unable to achieve performance close to their targeted performance in a consistent manner. Indeed, for each algorithm the table shows examples in which the actual loss rate is both higher and lower than the target, sometimes by 2 or 3 orders of magnitude. While the TE algorithm comes closer to its targets in general, it still misses by a couple of orders of magnitude in some cases. As such, even though the targets are achieved under certain scenarios, they do not predict performance reliably.

In sum, none of the algorithms provide tuning parameters that are useful as performance targets. At best, these parameters can be seen as largely uncalibrated knobs that can increase or decrease utilization and loss.

V. Conclusions

In this paper we compared several different measurement-based admission control algorithms. We evaluated the algorithms according to two criteria. First, what tradeoff of loss and load do they each achieve? This criterion shows how well the algorithms are able to balance the conflicting goals of providing good quality of service to individual users and achieving high network utilization (i.e., satisfying many users). Here our results were unambiguous. Across a range of traffic sources, all the algorithms, whether *ad hoc* or principled, achieved nearly identical performance. This result argues that there is no particular performance benefit of one over the others. Our study also yielded several additional insights about measurement-based admission control. First, we showed that for many algorithms, the measurement estimation and admission decision processes can be decoupled. Second, differences in performance caused by flow heterogeneity are a matter to be addressed by policy, rather than by algorithmic differences. Third, simulation results showed that measurement-based admission control algorithms not only cope well with long range dependence in traffic, in some circumstances they are more adept at handling it than are parameter-based algorithms.

The second criterion we used to evaluate the algorithms was the extent to which they provided performance tuning knobs that allow network operators to set a target performance level for the network. Such a knob would allow the network operator to decide where on the performance frontier the network should operate. Here the results were less impressive. None of the algorithms was able to reliably match actual performance to targeted performance levels. Thus, we believe that for any of these algorithms, network operators will need to monitor actual performance in order to learn appropriate parameter settings. On the other hand, some algorithms did better than others in this regard in the sense that they tended to get closer to targets on average than others. While the magnitude of the errors was in all cases large enough to call into question the value of the knobs as performance targets, whether or not this difference is important is a subject of debate. The ability of future algorithms to improve in this regard is an open question.

Acknowledgements

We had many helpful conversations with Nick Duffield, Sally Floyd, Richard Gibbens, Matt Grossglauser, Frank Kelly, Ed Knightly, Andrew Moore and Jingyu Qui about the measurement-based admission control algorithms discussed in this paper. We would like to thank Sandeep Bajaj for his assistance with the implementation of some of the admission control algorithms in the *ns* simulator. Patrick McDaniel provided valuable feedback on an earlier version of this paper.

References

[1] Beran, J., Sherman, R., Taqqu, M. S., and Willinger, W. Long-range dependence in variable-bit-rate video traffic. *IEEE Transactions on Communications 43*, 2 (Feb. 1995), 1566–1579.

[2] Bolotin, V. "Modeling Call Holding Time Distributions for CCS Network Design and Performance Analysis". *IEEE Journal on Selected Areas in Communications 12*, 3 (Apr. 1994), 433–438.

[3] Braden, R., Ed., Zhang, L., Berson, S., Herzog, S., and Jamin, S. Resource ReSerVation protocol (RSVP) – version 1 functional specification. Tech. Rep. RFC 2205, Internet Engineering Task Force, Sept. 1997.

[4] Breslau, L., Jamin, S., and Shenker, S. Measurement-based admission control: What is the research agenda? In *Proc. of IEEE/IFIP Seventh International Workshop on Quality of Service (IWQOS '97)* (London, England, 1999).

[5] Clark, D., and Wroclawski, J. An approach to service allocation in the Internet. Internet draft, Internet Engineering Task Force, July 1997.

[6] Clark, D. D., Shenker, S., and Zhang, L. Supporting real-time applications in an integrated services packet network: Architecture and mechanism. In *Proceedings of ACM Sigcomm* (Aug. 1992), pp. 14–26.

[7] Crosby, S., Leslie, I., McGurk, B., Lewis, J. T., Russell, R., and Toomey, F. Statistical properties of a near-optimal measurement-based cac algorithm. In *Proceedings IEEE ATM '97* (June 1997).

[8] CROVELLA, M., AND BESTAVROS, A. Self-similarity in world wide web traffic: Evidence and possible causes. *IEEE/ACM Transactions on Networking 5*, 6 (Dec. 1997), 835–846.

[9] DUFFY, D., MCINTOSH, A., ROSENSTEIN, M., AND WILLINGER, W. "Statistical Analysis of CCSN/SS7 Traffic Data from Working CCS Subnetworks". *IEEE Journal on Selected Areas in Communications 12*, 3 (Apr. 1994), 544–551.

[10] DZIONG, Z., JUDA, M., AND MASON, L. A framework for bandwidth management in ATM networks – aggregate equivalent bandwidth estimation approach. *IEEE/ACM Transactions on Networking 5*, 1 (Feb. 1997), 134–147.

[11] FLOYD, S. Comments on measurement-based admissions control for controlled-load services. Technical report, Lawrence Berkeley Laboratory, July 1996.

[12] GARRETT, M. W., AND WILLINGER, W. Analysis, modeling and generation of self-similar VBR video traffic. *Computer Communications Review 24*, 4 (Oct. 1994). SIGCOMM '94 Symposium.

[13] GIBBENS, R., AND KELLY, F. "Measurement-Based Connection Admission Control". *15th International Teletraffic Congress* (Jun. 1997).

[14] GIBBENS, R. J., KELLY, F. P., AND KEY, P. B. A decision-theoretic approach to call admission control in ATM networks. *IEEE Journal on Selected Areas in Communications SAC-13*, 6 (1995), 1101–1113.

[15] GROSSGLAUSER, M., AND TSE, D. A framework for robust measurement-based admission control. *Computer Communications Review 27*, 4 (Oct. 1997), 237–248. ACM SIGCOMM'97, Sept. 1997.

[16] GROSSGLAUSER, M., AND TSE, D. N. C. A time-scale decomposition approach to measurement-based admission control. In *Proceedings of the Conference on Computer Communications (IEEE Infocom)* (New York, Mar. 1999).

[17] GROSSGLAUSER, M., TSE, D. N. C. C., KUROSE, J., AND TOWSLEY, D. A new algorithm for measurement-based admission control in integrated services packet networks. In *Proceedings of the Fifth International Workshop on Protocols for High-Speed Networks* (Antipolis, France, Oct. 1996).

[18] JAMIN, S., DANZIG, P., SHENKER, S., AND ZHANG, L. A measurement-based admission control algorithm for integrated services packet networks. In *Proceedings of ACM Sigcomm* (Sept. 1995).

[19] JAMIN, S., DANZIG, P. B., SHENKER, S. J., AND ZHANG, L. A measurement-based admission control algorithm for integrated services packet networks. *IEEE/ACM Transactions on Networking 5*, 1 (Feb. 1997), 56–70.

[20] JAMIN, S., SHENKER, S., AND DANZIG, P. "Comparison of Measurement-based Admission Control Algorithms for Controlled-Load Service". *Proceedings of the Conference on Computer Communications (IEEE Infocom)'97* (Apr. 1997).

[21] KNIGHTLY, E. W., AND QIU, J. Measurement-based admission control with aggregate traffic envelopes. In *IEEE ITWDC '98* (Ischa, Italy, September 1998).

[22] NICHOLS, K., JACOBSON, V., AND ZHANG, L. A two-bit differentiated services architecture for the Internet. Internet Draft, Internet Engineering Task Force, May 1999. Work in progress.

[23] PAXSON, V., AND FLOYD, S. Wide area traffic: The failure of poisson modeling. *IEEE/ACM Transactions on Networking 3*, 3 (June 1995).

[24] SHENKER, S., PARTRIDGE, C., AND GUERIN, R. Specification of guaranteed quality of service. RFC 2212, Internet Engineering Task Force, Sept. 1997.

[25] WILLINGER, W., TAQQU, M., SHERMAN, R., AND WILSON, D. "Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level". *Proceedings of ACM Sigcomm'95* (Aug. 1995), 100–113.

[26] WROCLAWSKI, J. Specification of the controlled-load network element service. RFC 2211, Internet Engineering Task Force, Sept. 1997.

Algorithm	Parameter	$\bar{x}$	σ_x	$\bar{y}$	σ_y
	υ				
MS	0.93	0.873	5.97e-04	1.33e-06	1.72e-06
MS	0.95	0.892	8.59e-04	1.31e-05	4.39e-06
MS	0.97	0.91	1.23e-03	5.87e-05	1.47e-05
MS	0.99	0.929	1.18e-03	2.97e-04	2.97e-05
MS	1	0.937	6.34e-04	5.34e-04	2.95e-05
MS	1.02	0.956	7.73e-04	1.91e-03	3.56e-05
	ϵ				
HB	0.5	0.9	4.89e-04	1.87e-05	6.06e-06
HB	0.7	0.917	2.59e-04	7.35e-05	9.21e-06
HB	0.9	0.938	4.36e-04	4.52e-04	3.24e-05
HB	0.9417	0.944	2.82e-04	7.23e-04	3.94e-05
HB	0.99	0.954	1.57e-04	1.52e-03	5.69e-05
HB	0.999	0.959	3.32e-04	2.08e-03	1.15e-04
	s				
TO	1.8e-06	0.873	2.71e-04	1.98e-06	2.65e-06
TO	1.4e-06	0.897	7.60e-04	1.56e-05	5.31e-06
TO	1.2e-06	0.908	3.45e-04	4.82e-05	9.45e-06
TO	8e-07	0.932	3.27e-04	3.40e-04	1.58e-05
TO	4e-07	0.955	2.83e-04	1.79e-03	8.81e-05
TO	2e-08	0.975	4.06e-04	6.35e-03	2.44e-04
	s				
TP	2e-06	0.874	2.50e-04	2.53e-06	3.57e-06
TP	1.5e-06	0.897	2.52e-04	1.74e-05	8.43e-06
TP	8e-07	0.934	4.31e-04	3.55e-04	2.71e-05
TP	4e-07	0.956	1.02e-04	1.82e-03	8.93e-05
TP	1e-07	0.971	3.12e-04	4.95e-03	6.97e-05
TP	1e-08	0.976	2.56e-04	6.70e-03	1.92e-04
	clr				
MC	1e-06	0.917	6.65e-04	1.15e-04	1.77e-05
MC	1e-05	0.919	1.37e-03	1.26e-04	2.56e-05
MC	0.0001	0.92	1.61e-03	1.35e-04	2.40e-05
MC	0.001	0.922	7.10e-04	1.53e-04	1.54e-05
MC	0.01	0.927	8.47e-04	2.42e-04	3.88e-05
	clr				
TE	1e-06	0.893	2.45e-03	1.85e-05	5.87e-06
TE	1e-05	0.907	9.19e-04	6.25e-05	5.66e-06
TE	0.0001	0.923	9.19e-04	2.09e-04	1.21e-05
TE	0.001	0.944	1.24e-03	8.46e-04	7.87e-05
TE	0.01	0.971	5.29e-04	4.78e-03	1.98e-04

TABLE II

STANDARD DEVIATIONS FOR EXP1 DATA IN FIGURE 1