Measurement-based Admission Control Algorithms for Controlled-load Service: A Structural Examination

Sugih Jamin*
Computer Science and Engineering Division
Department of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, Michigan 48109-2122
jamin@eecs.umich.edu

Scott Shenker
Palo Alto Research Center
Xerox Corporation
Palo Alto, California 94304-1314
shenker@parc.xerox.com

CSE-TR-333-97
April 1997

Abstract

The Internet Engineering Task Force (IETF) is considering the adoption of the controlled-load service, a real-time service with very relaxed service guarantees [Wro95]. Measurement-based admission control algorithms (MBAC's), as opposed to the more conservative *worst-case* parameter-based methods, were expressly designed to achieve high levels of network utilization for such relaxed real-time services. Most researchers studying MBAC's, including ourselves [JDSZ97], have focused primarily on the design of the admission control equations using a variety of principled and *ad hoc* motivations. In this paper we show that, much to our own surprise, the admission control equations developed so far in the MBAC literature give essentially the same performance. Furthermore, we observe that the equations, even though they are derived and motivated in quite different ways, have the same structural form.

1 Introduction

The Internet Engineering Task Force (IETF) is considering the adoption of the controlled-load service, a real-time service with very relaxed service guarantees [Wro95]. Measurement-based admission control algorithms (MBAC's), as opposed to the more conservative *worst-case* parameter-based methods, were expressly designed to achieve high levels of network utilization for such relaxed real-time services. An MBAC consists of a measurement mechanism, and a set of admission control equations that take these measurements as inputs. Most researchers studying MBAC's, including ourselves [JDSZ97], have focused primarily on the design of the admission control equations. In this paper we show that, much to our own surprise, several of the admission control equations developed so far in the MBAC literature give essentially the same

*This research is supported in part by equipment grant from Sun Microsystems Inc.

performance. We therefore realize that our initial focus on the design of the admission control equations was mistaken, and that research on MBAC's should instead focus on the setting of the measurement parameters and on the global issues involved in admission control (see Section 7.2).

There is a rapidly growing literature on MBAC's; before proceeding, we briefly review some of the more relevant work. The controlled-load service specification uses a very parsimonious source characterization; sources are described only by a peak rate and a token-bucket filter. Hence in this paper we only consider MBAC's that do not require precise source characterization. Aside from our own ([JDSZ97]), several MBAC's requiring only a token-bucket filter description of sources have been proposed in the literature [GKK95, Flo96, GK97]. A number of studies have also been conducted to explore alternative measurement processes and their implementations in hardware [DJM97, $C^{+}91$, WCKG94]. The more recent works on MBAC's ([JSD97, GK97]) consisted of comparative studies of several different MBAC schemes. In [JSD97] we looked at the utilizations achieved by three MBAC's given a particular performance bound, e.g. 1e-6 packet loss rate. Two of the MBAC's we studied derived their admission decisions by evaluating principled equations that take this performance bound as input.

In the current paper, we study these MBAC's more closely. This paper is different from our previous work in two crucial aspects. First, we do not rely on the predicted or calculated performance of these principled methods, but instead measure the actual performance through simulation. We find that the observed behavior departs substantially from the calculated predictions. Thus, to achieve a particular performance goal, one must treat the input performance target as merely an arbitrary and tunable parameter. Second, rather than focusing on a single point, we study the full range of achievable performance of each MBAC.[1] We call the range of achievable utilizations and observed loss rates the *loss-load curve*.

The study reported in [GK97] is similar in spirit to ours. Through formal analysis, the authors showed that the MBAC equations in [Flo96] and [GKK95] belong to the same class: tangents on the exact equivalent bandwidth curve computed using large-deviation analysis. The authors further conducted numerical investigations on the trade off between loss rate and network utilization when different MBAC's, reflecting tangents at different points of the equivalent bandwidth curve, were employed. Whereas [GK97] took an analytical and numerical approach, we run simulations to discover the performance characteristics of each MBAC. As we shall see later in Section 3, the predictions of these numerical calculations are far from the actual performance.

[1] This idea arose in private discussions with Frank Kelly.

The main results of this paper may be summarized as follows:

1. In the operating region where losses occur under all MBAC's, they can all be induced to give the same loss-load curve by tuning their measurement parameters.
2. All the MBAC's we study perform similarly because they are all based on admission equations of the same form:

$$\widehat{\nu} < f(\cdot)\mu - g(\cdot), \tag{1}$$

   where $\widehat{\nu}$ is the measured load, μ is the link bandwidth, and $f(\cdot)$ and $g(\cdot)$ are functions of the source's reserved rate and the number of admitted sources. (See Section 4 for an elaboration of $f(\cdot)$ and $g(\cdot)$.)
3. For immediate implementation of MBAC for controlled-load service, we recommend the following algorithm:

$$\widehat{\nu} < \upsilon\mu - \kappa r, \tag{2}$$

   where υ is a utilization factor to be set based on historical load pattern and desired performance, μ the link bandwdith, $\kappa > 0$ a constant to be determined, again, from historical data, and r the reserved rate of an incoming flow. Any of the measurement processes mentioned in the Section 2.2 may be used in conjunction with this algorithm.
4. The performance of these algorithms, while somewhat insensitive to the form of the admission control equations, appears rather sensitive to changes in the parameters controlling the measurement process.

2 Description of the MBAC's

The MBAC's we study in this paper were originally published in [JSD97, Flo96, GK97]; we summarize them here for the sake of completeness.

2.1 The Admission Control Algorithms

Measured Sum (MS). Our MBAC for controlled-load service first published in [JSD97] uses measurement to estimate existing traffic load and admits a new flow of rate r if the following test succeeds:

$$\widehat{\nu} < \upsilon\mu - r, \tag{3}$$

where $\widehat{\nu}$ is the measured load of existing traffic, and υ is a user-defined utilization target intended to limit the maximum link load. We describe the measurement process in the next section. Upon admission of the new flow the load estimate is increased by $\widehat{\nu}' = \widehat{\nu} + r$.

Hoeffding Bounds (HB). The MBAC described in [Flo96] computes the equivalent bandwidth for a set of flows using the Hoeffding bounds. The equivalent bandwidth of a set of flows is defined in references [Flo96, GAN91] as the bandwidth $C(\epsilon)$ such that the stationary bandwidth requirement of the set of flows exceeds this value with probability at most ϵ. The measured equivalent bandwidth based on Hoeffding bounds ($\widehat{C}_H$) of n flows, assuming peak rate (p) policing, is:

$$\widehat{C}_H(\widehat{\nu}, \{p_i\}_{1 \leq i \leq n}, \epsilon) = \widehat{\nu} + \sqrt{\frac{ln(1/\epsilon) \sum_{i=1}^{n} (p_i)^2}{2}}, \tag{4}$$

The admission control check when a new flow requests admission is:

$$\widehat{C}_H + p \leq \mu. \tag{5}$$

Upon admission of a new flow, the load estimate is increased by $\widehat{\nu}' = \widehat{\nu} + p$. If a flow's peak rate is unknown, it is derived from its token-bucket filter parameters (r, b) using the following equation:

$$\widehat{p} = r + b/U, \tag{6}$$

where U is a user-defined averaging period. Similar to the algorithm in [GKK95], if a flow is denied admission, no other flow of a similar type will be admitted until an existing one departs.

Acceptance Region. Let a and p be the average and peak rates of an ON/OFF source, the equivalent bandwidth (C) of the source can be computed using the following equation [Kel96c, GK97]:

$$C(s) = \frac{1}{s} log \left[1 + \frac{a}{p}(e^{sp} - 1) \right], \tag{7}$$

where $s > 0$. One can then draw an equivalent bandwidth curve varying the average rate on the *x-axis* and with the resulting equivalent bandwidth on the *y-axis*. A linear bound at different points of this curve can be computed as tangent at that point [Kel96c, Kel96a]:

$$c + \alpha\widehat{\nu} \leq \mu, \tag{8}$$

where c determines the location and α the slope of the tangent. This linear bound can then be used as an MBAC [Kel96a]. Four MBAC's, each based on a different tangent of the equivalent bandwidth curve were presented in [GK97]:

1. **Tangent at peak (TP)**:
$$np(1 - e^{-sp}) + e^{-sp}\widehat{\nu} \leq \mu, \tag{9}$$
where n is the number of admitted flows.

2. Tangent at arbitrary point:
$$\frac{pe^{sp}}{(p + \bar{a}(e^{sp} - 1))^2}(n\bar{a}^2(e^{sp} - 1) + p\widehat{\nu}) \leq \mu, \tag{10}$$
where $\bar{a}$ is the measured average rate of a source.

3. Tangent of slope one:
$$\widehat{\nu} + \sqrt{\frac{\gamma}{2}p^2 n} \leq \mu, \tag{11}$$
where γ is a constant.

4. **Tangent at the origin (TO)**:
$$e^{sp}\widehat{\nu} \leq \mu. \tag{12}$$

As pointed out in [GK97], Eqn. 11 is the same as Eqn. 4, and Eqn. 12 is similar to the MBAC described in [GKK95]. We do not consider Eqn. 10 in this study because it requires measurement of per-source average rate, which we do not believe can be accurately and cost effectively measured. The measured load used in Eqns. 9 and 12 is not artificially adjusted upon admittance of a new flow. For flows described by a token-bucket filter (r, b) but not peak rate, we use Eqn. 6 to derive its peak rate. If a flow is rejected, the admission control algorithm does not admit another flow until an existing one departs.

2.2 The Measurement Processes

We now describe the measurement mechanisms used with each of the MBAC's.

Time-window. Following [JDSZ97], we use a simple time-window measurement mechanism to measure network load as input to the MS algorithm. We compute an average load every S sampling period. At the end of a measurement window T, we use the highest average from the just ended T as the load estimate for the next T window. When a new flow is admitted to the network, the estimate is increased by the rate of the new flow, as described in the previous section. If a newly computed average is above the estimate, the estimate is immediately raised to the new average. At the end of every T, the estimate is adjusted to the actual load measured in the previous T.

Point Samples. The measurement mechanism used with both of the acceptance region MBAC's takes an average load sample every S' period [Kel96b]. Or, equivalently, the measurement mechanism is the time-window mechanism with a T/S ratio of 1.

Exponential Averaging. An exponential average load is used as input to the HB MBAC. An average load ($\widehat{\nu}^S$) is measured once every S sampling period. The exponential average load is computed using an infinite impulse response (IIR) function with weight w:

$$\widehat{\nu}' = (1 - w) * \widehat{\nu} + w * \widehat{\nu}^S. \tag{13}$$

Recall that HB MBAC requires peak rate policing and, when the peak rate is not explicitly specified, derives a flow's peak rate from its token-bucket parameters using Eqn. 6. To be safe, the averaging period U in Eqn. 6 should be smaller than or equal to S, the measurement sampling period. We set $U = S$ to reflect the peak rate seen by the measurement mechanism, with the caveat that too small a U could result in practically infinite peak rate when the bucket depth is large.

3 Loss-Load Behavior of the MBAC's

As we mentioned in the Introduction, the main goal of this paper is to map out the full range of achievable performance of the MBAC's under study. We measure the performance of an MBAC along two axes: the

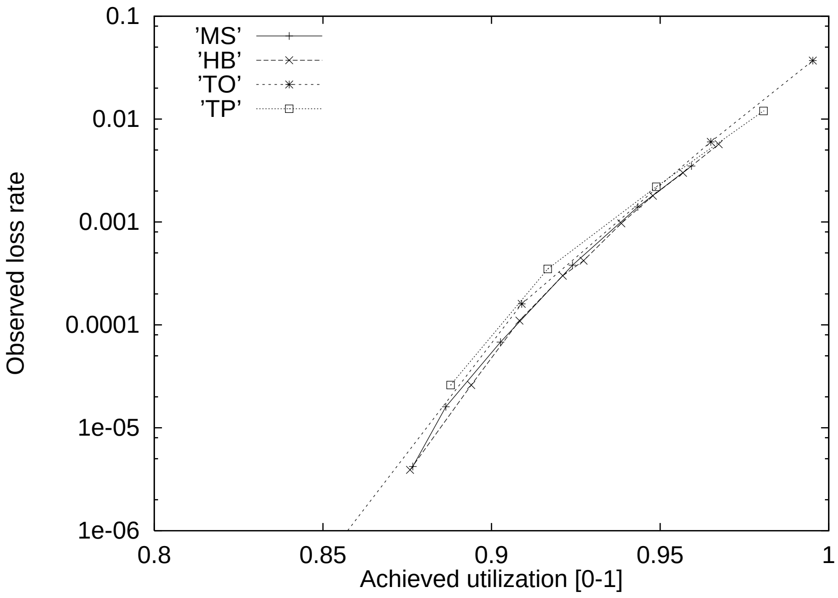

Figure 1: The loss-load curves for POO1 source.

achievable link utilization and the observed loss rate. For a given source model, a given set of parameter settings of an MBAC results in a certain link utilization at a certain loss rate. We run several simulations for each source model under the same MBAC, varying the MBAC's parameter settings to collect several such utilization-vs.-loss-rate data points. The curve connecting these data points we call the *loss-load curve*. In this paper we only show portions of loss-load curves where the loss rate is non-zero with adequate statistical significance, considering that our maximum simulation run can only serve 1e8 packets. The simulation scenarios from which we draw the loss-load curves, and the source models we use to run these simulations, are the same ones we have been using for our previous work on MBAC's ([JDSZ97, JSD97]). For completeness sake, we describe them both again in the Appendix of this paper.

Figure 1 shows the loss-load curve of the POO1 source under the MS, HB, TO, and TP MBAC's. Where they overlap, the curves are practically on top of each other.[2] (Note that the *y-axis* is on a log scale.) Thus even though the three MBAC's are derived in very ! different ways, they essentially give the same performance. From left to right, the MS data points are from simulations with the υ parameter of the equation set to 0.91, 0.92, 0.94, 0.96, 0.98, and 1.0 respectively. In all cases, the sampling period, S, is set to 5e3 packet transmission times and the window size, T, to 10S. From left to right, the HB data points are from simulations with the ϵ parameter of the equation set to 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9 respec-

[2]We have repeated some of the simulations presented in this paper with different random seeds and will show the confidence interval of each data point in the final version of the paper.

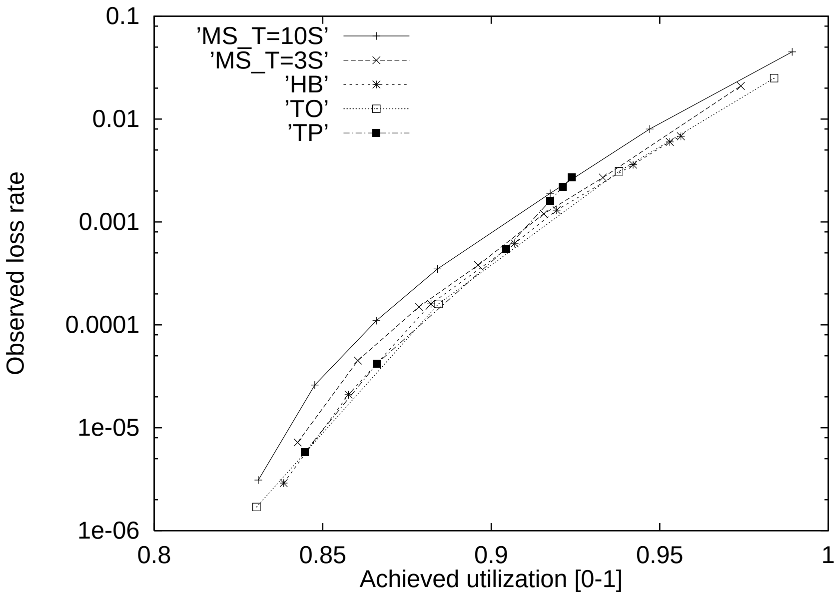

Figure 2: The loss-load curves for POO2 source.

tively. The sampling period, S, is set to 5e3 packet transmission times, and the weight of the IIR function, w, is set to 2e-3. From left to right, the parameter s for the TO curve is set to 3e-3, 2e-3, 1e-3, and 1e-4 respectively, and that for the TP curve is set to 2e-3, 1.5e-3, 1e-3, and 4e-4 respectively. The sampling period S' is set to 2.5e4 packet transmission times for both TO and TP cases. To summarize, we map out the loss-load curves of the four MBAC's by tuning v, ϵ, or s, respectively. Observe that the resulting loss rates and utilization levels from these simulations is very different from the input performance bounds. For instance, when ϵ =0.4, the actual loss rate under the HB MBAC is 3e-4. The two primary conclusions from this example are: (1) the three MBAC's give equivalent performance, and (2) the performance of the principled MBAC's do not match the input parameters. As we shall see these conclusions also apply to the other data sets to which we now turn.

Figure 2 shows the loss-load curve of the POO2 source under the four MBAC's as before, but with two data sets for the MS MBAC. The measurement parameters for the HB, TO, and TP MBAC's are as before. From left to right, the HB data points are from simulations with the ϵ parameter set to 0.6, 0.7, 0.8, 0.9, 0.947, 0.99, 0.999, and 0.9999 respectively. The data points for the TO MBAC are from simulations with the parameter s set to 3e-3, 2e-3, 1e-3, and 1e-4 going from left to right. The setting of s for the TP MBAC are 3e-4, 2.5e-4, 2e-4, 1e-4, 6e-5, 4e-5, and 3e-5. Data points for the MS MBAC with $T = 10S$ are from simulations with v set to 0.94, 0.96, 0.98, 1.0, 1.02, 1.04, and 1.06 going from left to right. The distance between this and both the HB and TO curves is somewhat more pronounced than in the POO1 case. However when we adjust the time window size to $T = 3S$, we see that the MS curve can be brought down to the vicinity of the curves from the other two MBAC's. This illustrates our claim that where losses occur, all MBAC's under study can be induced to give the same loss-load curve by tuning their measurement parameters. In effect, the measurement parameters determine the offset of the loss-load curve along the *y-axis*—a point we will explore further in Section 7.1. From left to right, the data points for the MS with $T = 3S$ case are from simulations with the υ parameter set to 0.9, 0.92, 0.94, 0.96, 0.98, 1.0, and 1.02 respectively. In both MS cases, S =5e3 packet transmission times.

Figure 3 shows the loss-load curves of the EXP1 source under the same four MBAC's. The curves are somewhat more distributed than in the previous two scenarios. In the MS case, the time window measurement parameter is set to $T = 3S$, with S =5e3 packet transmission times. The υ parameter, for the data points from left to right, is set to 0.95, 0.96, 0.97, 0.98, 0.99, 1.0, 1.01, and 1.02 respectively. Data points for the HB MBAC are from simulations with ϵ set to 0.7, 0.8, 0.9, 0.9417, 0.99, and 0.999 going from left to right. The weight w of the measurement process is set to 2e-4. For the TO case, going from left to right, the s parameter is set to 2e-3, 1.8e-3, 1.6e-3, 1.4e-3, 1.2e-3, 1e-3, 8e-4, and 8e-5; and for the TP case, s is set to 2e-3, 1e-3, 8e-4, 6e-4, 4e-4, 2e-4, and 1e-4 going from left to right. The measurement parameters for the TO and TP MBAC's are set to S' =2.5e4 packet transmission times; we do not experiment with other settings of this parameter for this source model.

4 On the Structural Similarity of the MBAC's

While the fact that all four MBAC's give rise to practically the same loss-load curves may be surprising at first, further studies of the equations used in the MBAC's show that they can all be expressed as:

$$\widehat{\nu} < f(\cdot)\mu - g(\cdot), \tag{14}$$

where $\widehat{\nu}$ is the measured load, μ is the link bandwidth, and $f(\cdot)$ and $g(\cdot)$ are functions of the source's reserved rate and the number of admitted sources. For the MS algorithm $f_{MS}(\cdot)$ is the tunable parameter υ and $g_{MS}(\cdot)$ is the source's reserved rate, r. Function $f_{TO}(\cdot)$ in the TO case is e^{-sp}, where p is the source's peak rate and $s > 0$ is a tunable parameter; $g_{TO}(\cdot)$ in this case is 0. In the TP case, $f_{TP}(\cdot)$ is e^{sp} and

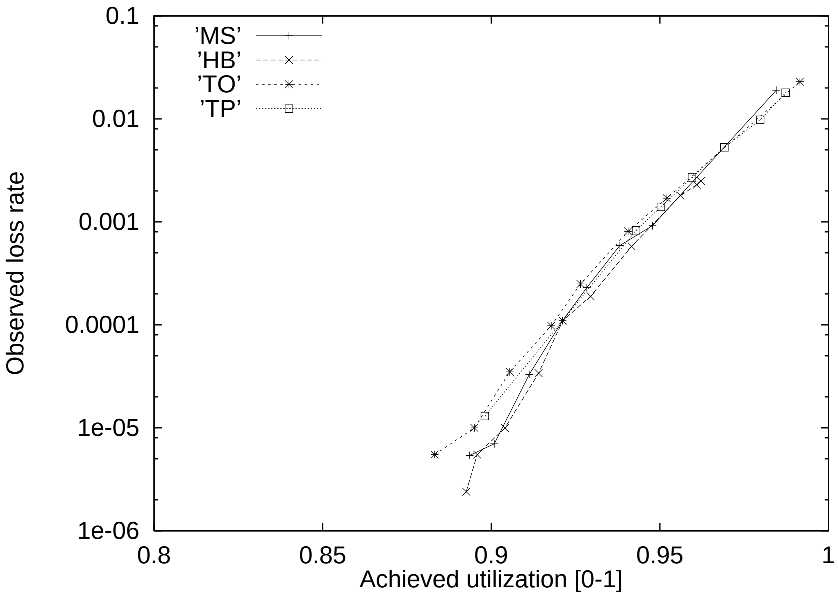

Figure 3: The loss-load curves for EXP1 source.

$g_{TP}(\cdot) = (e^{sp} - 1)np$, where n is the number of admitted flows and p the peak rate of the source model. In the HB MBAC, $f_{HB}(\cdot) = 1$ and

$$g_{HB}(\cdot) = \sqrt{\frac{ln(1/\epsilon)\sum_{i=1}^{n}(p_i)^2}{2}} + p, \tag{15}$$

where n is the number of admitted sources and the p and p_i's are either the sources' declared peak rates or peak rates derived from the sources' token-bucket filter parameters.

While we initially thought that the details of the admission control equation (i.e., the form of $f(\cdot)$ and $g(\cdot)$) were the critical factors in the performance of an MBAC, the data suggests that, at least for the three choices of $f(\cdot)$ and $g(\cdot)$ explored here, the form of $f(\cdot)$ and $g(\cdot)$ have little effect on the performance of the MBAC. It turns out that the tuning of the parameters has a much greater effect on the performance.

We identify two degrees of freedom in which to operate an MBAC: (1) by tuning the measurement parameters such that $\hat{\nu}$ is more stable or more adaptive to actual utilization, and (2) by tuning the parameters in $f(\cdot)$ and/or $g(\cdot)$ to control the amount of slack bandwidth set aside to accommodate traffic fluctuations. In Section 3 we showed that tuning the parameters in the $f(\cdot)$ and $g(\cdot)$ functions of an MBAC allow us to map out the MBAC's entire loss-load curve. That is, these variations tended to take us *along* the loss-load curve of an MBAC.

Table 1: Experiments with the Hoeffding bounds based MBAC

Model	HB($\widehat{p}$)		HB(r)		HB(r, ϵ=1e-1)	
Name	%Util	#Actv	%Util	#Actv	%Util	#Actv
EXP2	13	13	43	42	68	67
EXP3	0.3	3	6	45	23	171
POO3	2	31	13	177	37	504

In contrast, as shown in Figure 3, tuning the measurement parameters of an MBAC shifts the loss-load curve of the MS MBAC along the *y-axis*. That is, this change in measurement parameters moved us *above* the loss-load curve of the MBAC, giving us worse performance (i.e., higher loss at the same utilization). Later we will see that this is not the only effect varying the measurement parameters has. It also effects the extent of an MBAC's loss-load curve.

5 Experiments Tuning MBAC's

In this section we experiment with the tunable parameters identified in the previous section to explore their contributions to the admission decisions an MBAC makes. Table 1 shows results from simulations of HB MBAC. The "%Util" columns show the average link utilization achieved for various settings of the HB MBAC's parameters. The "#Actv" columns show the average number of concurrently active connections. The columns under "HB($\widehat{p}$)" contain results from simulations of the HB MBAC with ϵ =1e-6 and sources characterized by peak rates derived from their token-bucket filter parameters using Eqn. 6. To lead off our experiments with the parameters of the HB MBAC, we ask the question, "Can the HB MBAC allow more flows into the network, and achieve higher link utilization, than the numbers under HB($\widehat{p}$)?"

Studying Eqn. 15, we note two tunable variables: ϵ and the sources' peak rate (p_i's). In the case of EXP2, EXP3 and POO3 sources, the flows' peak rates are derived from their token-bucket parameters using Eqn. 6. In the case of POO3, for S=5e3 packet transmission times, the derived peak rate is 363 Kbps, higher than the actual peak rate. In [Flo96] the author suggests that token rate of a source's token-bucket filter be set to its peak rate and the bucket depth to a small number that will "accommodate small variations in packet delay that accumulate in the network." To see how a less conservative peak rate effects the performance of HB MBAC in the EXP2, EXP3, and POO3 cases, we simulate them using the token-bucket rate of each as its peak rate, ignoring the token-bucket depths. The two columns of Table 1 under the HB(r) heading show

Table 2: Further experiments with the Hoeffding bounds based MBAC

Model	%Util		
Name	ϵ=0.9	ϵ=0.99	ϵ=0.9999
EXP2($\widehat{p}$)	73	82	88
EXP2(r)	89	93	95
EXP3($\widehat{p}$)	18	39	56
EXP3(r)	63	76	83
POO3($\widehat{p}$)	63	75	85
POO3(r)	73	82	87

results from these simulations. Next we experiment with tuning ϵ. Relaxing ϵ from 1e-6 to 1e-1, we see further increase in the number of concurrent flows and link utilization (as the numbers under HB($r, \epsilon = 1e-1$) show). The sampling period S of the measurement process used with EXP2 source is 1e3, with EXP3 source 5e2, and with POO2 source 5e3 packet transmission times. The same S is used for all of the simulations involving each of the sources. The exponential averaging weight is 2e-3 in all cases.

Table 2 shows the achieved utilization for the various source models when ϵ is further relaxed (the measurement parameters are not modified from the above cases). Rows marked with ($\widehat{p}$) contain results from cases when sources' peak rates are derived from their token-bucket filter parameters and the rows marked with (r) are results from using the token-bucket rates as the sources' peak rates. Except for the EXP2(r) cases when ϵ =0.99 and ϵ =0.9999 and the observed loss rates are 1e-5 and 1e-4 respectively, the other simulations do not result in packet losses, hence we cannot present their loss-load curves. Note that Eqn. 15 requires $\epsilon \leq 1$ for real-valued solutions. When $\epsilon = 1$, the HB MBAC reduces to the MS MBAC with $\upsilon = 1$. Therefore the HB MBAC cannot access points on the loss-load curve that MS MBAC achieves with $\upsilon > 1$.

From Eqn. 4, we note that aside from ϵ and the source's peak rate, the load estimate ($\widehat{\nu}$) also determines the equivalent bandwidth computation. A load estimate that tracks actual utilization more closely may result in a smaller estimated equivalent bandwidth and a higher flow admittance rate. The exponential averaging process used in the HB MBAC to measure load consists of two tunable parameters: the weight w and the sampling period, S. As we pointed out in Section 2.2, we also use S as the averaging period, U, in deriving a source's peak rate from its token-bucket filter parameters using Eqn. 6. Unfortunately, this makes it hard for us to isolate the effect of tuning S in cases where the source's peak rate must be derived from its token-bucket filter parameters: increasing S not only smooths out our samples, it gives a smaller derived peak rates. Hence we will limit our experiments tuning the measurement parameters to sources that are already described by their actual peak rates, namely the EXP1, POO1, and POO2 sources.

In scenarios with homogeneous sources such as the ones we are investigating here, knowing the peak (p) and average (a) rates of the sources allows one to deterministically compute the number of flows that may be admitted under the HB MBAC by solving for n in the quadratic equation:

$$C_H = na + \sqrt{\frac{ln(1/\epsilon)np^2}{2}}. \tag{16}$$

with the solution being the n that also satisfies $C_H - na > 0$. Achievable utilization is then na/μ. For the EXP1 source, ϵ =1e-6, and C_H upper bounded by the bottleneck link bandwidth, 10Mbps, Eqn. 16 gives $n = 286$, with achievable utilization upper bounded at 89%. With the exponential averaging parameters S =5e3 and w =2e-3, our simulation results show an achieved average link utilization of 72.5%, serving an average of 228 concurrently active flows. A smaller sampling period increases the burstiness of our samples but also increases the sampling frequency, allowing us to track actual utilization closer. Using S =1e3, keeping ϵ at 1e-6, we can concurrently admit 232 EXP1 flows, achieving 73.8% utilization.

One could admit more flows into the network if actual aggregate utilization of n flows is lower than na, where a is the sources' declared average rate. Recall that the POO1 model has the same *declared* peak and average rates as the EXP1 model but has heavy-tailed ON and OFF times distributions, which lead to burstier aggregate traffic. And indeed, in our simulations of POO1 sources, with ϵ =1e-6, w =2e-3, and S =5e3, we see an average of 277 concurrently active flows, 21% more flows than EXP1 case under similar circumstances, including achieved average link utilization of 72.4%.

A closer investigation of Eqns. 15 and 16 further reveals that achievable link utilization $\widehat{\nu}$ can be approximated by na when n is small and $g_{HB}(\cdot)$ is not a limiting factor. When n is large, however, $\widehat{\nu}$ is bounded by $\mu - g_{HB}(\cdot)$. In effect, $g_{HB}(\cdot)$ acts as a utilization threshold to ensure sufficient slack bandwidth to accommodate traffic fluctuations. For the EXP1 source, when ϵ =1e-12, the intersection of the two lines $\widehat{\nu} = na$ and $\widehat{\nu} = v\mu - g_{HB}(\cdot)$ is at n =187 and $\widehat{\nu}$ =59.8%, regardless of the sources' actual burstiness or the accuracy of the measurement mechanism.

Next we turn our attention to the TP MBAC and note that if the s parameter in Eqn. 9 is set to 0, the admission control check simply says that measured load must be less than link bandwidth. For s sufficiently large and $e^{-sp} \to 0$, the admission control check becomes $np < \mu$ and achieved utilization is determined by the sources' peak to average rates ratio. Hence for EXP1 and POO1 sources, for $s \geq 0.4$, link utilization is 51%. To limit the link bandwidth share of a source below this "natural" lower bound, e.g. to limit a CBR source to 50% of link bandwidth, would require introducing another tuning parameter to Eqn. 9 or a separate link-sharing mechanism.

One other interesting observation we made with our simulation results is with regard to source EXP2 under the MS MBAC. For any value of v between 1.03 and 1.03125 the achieved link utilization is 87.5%, with no packet loss. Then for v between 1.03126 and 1.036, the achieved link utilization is 99.9%, with a uniform loss rate of 4.8e-1. The measurement parameters are $S =$1e3 and $T = 10S$. We observe no similar behavior with the TO MBAC, which gives link utilization of 86.5, 94.5, and 95.9%, with loss rates of 4.6e-4 and 1.2e-3 in the later two cases, when the s parameter is varied from 5e-4, 2e-4, to 1e-4. The measurement parameter S' is set to 5e3 throughout. Close examination of the MS admission process reveals that at $v =$1.03126, enough extra bandwidth is alloted that one more EXP2 source may pass the $g_{MS}(\cdot) = r$ constraint; whereas in the TO MBAC, a new flow's reserved rate effects the admission process only fractionally, as multiple of s. This motivates our modification of the MS MBAC to the form recommended for use with controlled-load service in the next section.

6 An MBAC for Controlled-load Service

Given that all the MBAC's studied in this paper produce very similar loss-load curves, we advocate for controlled-load service a very simple and flexible MBAC:

$$\hat{\nu} < v\mu - \kappa r, \tag{17}$$

where v is a utilization factor to be set based on historical load pattern and desired performance, μ the link bandwdith, $\kappa > 0$ a constant to be determined from historical data, and r the reserved rate of an incoming flow. The utilization factor v determines the MBAC's operating region, and κ determines the degree of conservativeness when admitting new flows. Any of the measurement processes mentioned in Section 2.2 may be used in conjunction with this algorithm. For performance reasons, one may implement the load sampling mechanism in hardware, allowing users to set the sampling period. As results from the

Table 3: Parameters settings for five sets of simulations with the EXP1 source model under the MS MBAC. S is in packet transmission time unit.

S	T/S	υ values
5e3	13	1.02, 1.03, 1.04, 1.05, 1.06, 1.07, 1.08
5e3	10	1.0, 1.01, 1.02, 1.03, 1.04, 1.05, 1.06, 1.07
5e3	7	1.0, 1.01, 1.02, 1.03, 1.04, 1.05
6.5e3	7	1.0, 1.01, 1.02, 1.03, 1.04, 1.05, 1.06, 1.07, 1.08
5e3	3	0.95, 0.96, 0.97, 0.98, 0.99, 1.0, 1,01, 1.02

DT MBAC demonstrate, when the sampling period is properly set, a simple point sample measurement process may be sufficient. To be more conservative, one may choose to implement the time-window or exponential averaging measurement process (in software, with $T/S > 1$ or $w < 1$(!)). These processes keep history and provide users with a tuning knob that governs the length of history to keep. Intuitively, processes that keep history give more stable estimates. However, in the presence of long-range dependent traffic, the efficacy of measurement history in providing more stable estimates is a topic for further research [Kel96b].

7 A New Research Agenda for MBAC's

Instead of designing another set of principled MBAC equation, or a measurement process that may be expensive to implement or operate, we suggest that the research agenda for MBAC's should be directed towards the determination of their operating region and the tuning of their measurement parameters, both based on historical data. Another direction MBAC research should take is to solve the structural problems that arise from interactions between reservations that we first pointed out in [JDSZ95] and describe again in Section 7.2 below.

7.1 Tuning the Parameters

There are two items to this research agenda: First, tuning the parameters of an MBAC can either move one along the loss-load frontier or away from it. We do not yet understand which way of tuning the parameters accomplishes which effect. Second, now that we have shown the complete lack of relationship between the input performance bounds of an MBAC to its actual performance, we cannot tell a network operator desiring some loss rate on its network where to set the parameters of its MBAC. In this section

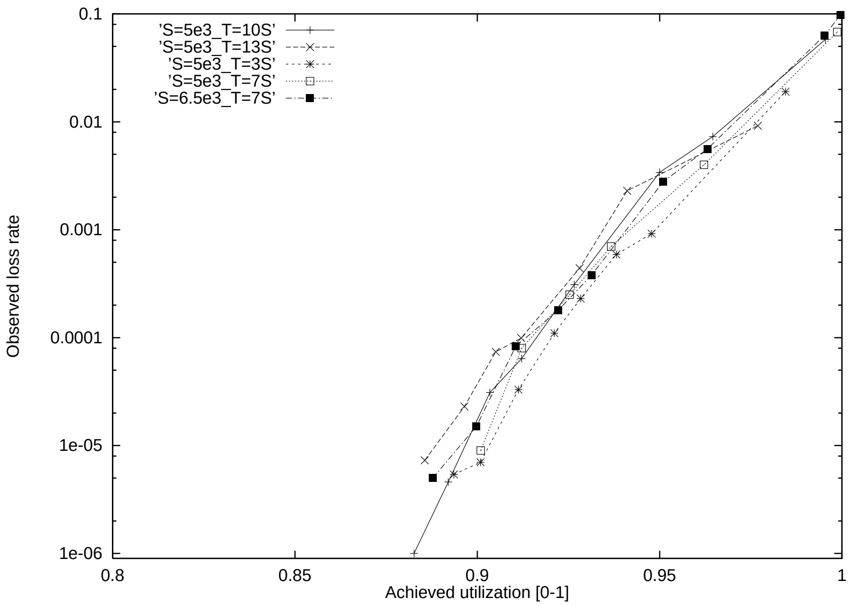

Figure 4: Five loss-load curves of the EXP1 source under the MS MBAC, varying the parameter settings

we show results from several experiments which illustrate our lack of understanding on how to set the parameters of an MBAC. We start with five sets of simulations with the EXP1 source model under the MS MBAC. Each set of simulations uses a specific measurement parameter setting, varying the v parameter. Figure 4 shows the loss-load curves resulting from these five sets of simulations. Table 3 summarizes the parameter settings of the five sets of simulations. The first two columns list the sampling period, S, and the time window size, T, expressed as T/S ratio, used with each set of the simulations. The last column lists the v values corresponding to the data points that make up the loss-load curve of that set, from left to right, in Figure 4. As we mentioned earlier, tuning the measurement parameters seem to change the offset of an MBAC loss-load curve along the *y-axis*. Figure 4 certainly bears this out, especially in the region where achieved utilization is low. However, we do not claim to have understood, much less be able to predict, how any given set of parameter setting influences the outcome of a run.

Figure 5 shows two loss-load curves each from simulations of POO1 and POO2 sources. Again we note the loss-load curve offset along the *y-axis* for the POO2 source when the measurement parameters are changed, but not for the POO1 source. Since the only difference between EXP1, POO1, and POO2 sources lies in their degree of burstiness, with POO1 source being burstier than EXP1 source, and POO2 source being the burstiest of the three, we cannot explain the lack of any signifant gap between the two POO1 loss-load curves when we change the T/S ratio of the measurement parameters from 10 to 3 when the same

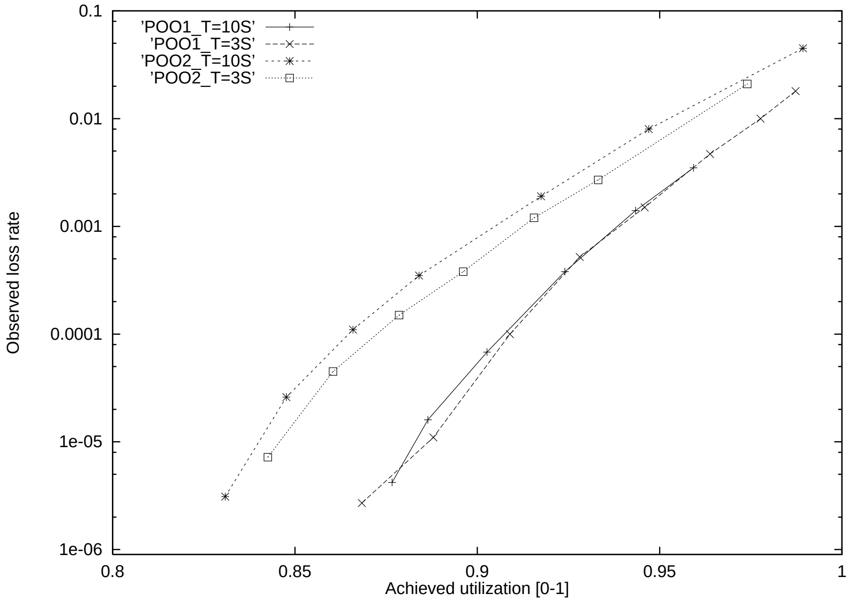

Figure 5: Two loss-load curves each from POO1 and POO2 sources under the MS MBAC, varying the parameter settings

Table 4: Parameters settings for two sets of simulations each with the POO1 and POO2 source models under the MS MBAC. S is in packet transmission time unit.

Model	S	T/S	υ values
POO1	5e3	10	0.91, 0.92, 0.94, 0.96, 0.98, 1.0
	5e3	3	0.86, 0.88, 0.9, 0.92, 0.94, 0.96, 0.98, 1.0
POO2	5e3	10	0.96, 0.98, 1.0, 1.02, 1.04, 1.06, 1.08
	5e3	3	0.9, 0.92, 0.94, 0.96, 0.98, 1.0, 1.02

modification of the ratio produces noticable gap in both the EXP1 and POO2 cases. Table 4 summarizes the parameter settings used to obtain the four loss-load curves of Figure 5.

Finally in Figure 6 we present three loss-load curves of the EXP1 sources under the HB MBAC, varying its parameters. Table 5 gives the parameter settings used in these simulations. The first two columns list the averaging period, S, and exponential averaging weight, w, used in each set of simulations, while the last row lists the ϵ values corresponding to each data point from left to right, on the loss-load curve of that set of simulations. We iterate here that we do not yet understand how any given set of parameter setting influences the outcome of a run. For the same ϵ value of 0.9999, two of the loss-load curves in Figure 6 do not manage to reach utilization beyond 94.1%, while the other reaches 96.2%, attesting to our earlier observation that settings of the measurement parameters also effect the extent of the loss-load curves.

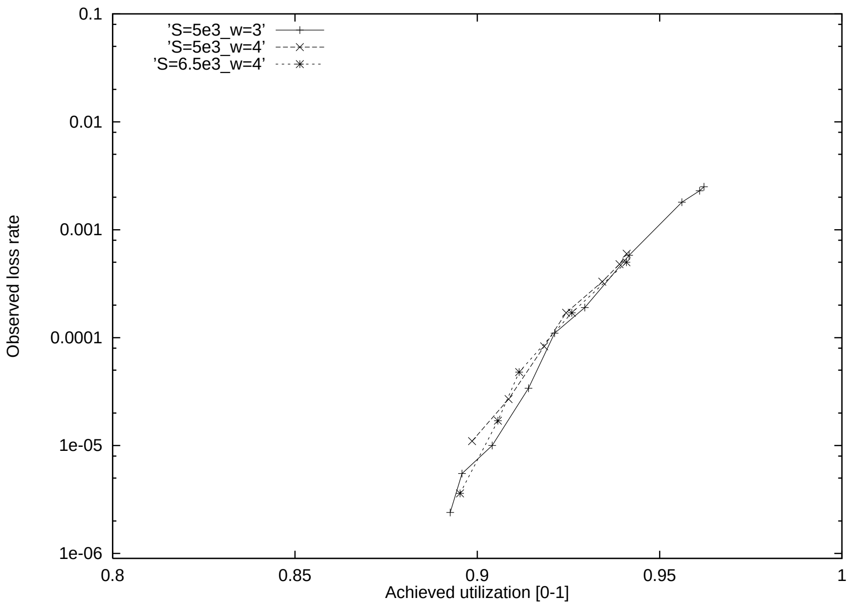

Figure 6: Three loss-load curves of the EXP1 source under the HB MBAC, varying the parameter settings

Table 5: Parameters settings for three sets of simulations with the EXP1 source model under the HB MBAC. S is in packet transmission time unit.

S	w	ϵ values
5e3	3	0.35, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.99, 0.999, 0.9999
5e3	4	0.7, 0.8, 0.9, 0.9417, 0.99, 0.999, 0.9999
6.5e3	4	0.8, 0.9, 0.9417, 0.999, 0.9999

Extending our work in [JDSZ95], the authors of [CKT96] suggest a feedback control mechanism to automatically tune the measurement parameters based on observed loss rate. The appropriate time scale over which these and parameters in the $f(\cdot)$ and $g(\cdot)$ functions are tuned is still a matter of future research. Too long a time scale may lead to long convergent time for these tuning parameters. On the other hand, one must allow enough time for traffic to stabilize after a change in parameter settings to tell the true effect of that change. In other words, the reaction time of the feedback control mechanism itself requires further tuning. Note however, tuning the reaction time of the feedback control mechanism will be done at a time-scale much larger than that of tuning the MBAC parameters. Whereas one might be tempted to tune the MBAC parameters every few days or weeks, the feedback mechanism should require attention only when the underlying traffic pattern changes.

7.2 Global Structural Limitations

As we saw in Section 4, the effectiveness of the tuning parameters may be constrained by the source's characteristics. In particular we saw that tuning υ in the MB MBAC case must take into consideration the source's reserved rate. Alloting an extra amount of bandwidth that is less than the source's reserved rate will not result in increase admittance rate. In our other simulations reported in [JDSZ97], we note the possibility that on highly congested links with infinite offered load, when a large grain flow, i.e. a flow with a high ratio of requested rate to link bandwidth, leaves the network, before the measurement mechanism of an MBAC has completely deflated its load estimate following this flow's departure, a small grain flow can immediately gain admittance at the expense of future larger grain flows. Upon departure of a small grain flow, if the link is congested, only another small grain flow can be admitted. Thus over time it is probable that only small grain flows are present on the link and large grain flows are completely shut off. We call this the *under-representation* problem, and consider it a structural limitation of all admission control algorithms.

We further identify two other structural limitations: One, if the only flows that depart from the network are short-lived flows, the link population will become dominated by flows with long call holding times. This is a structural limitation only if it prevents a network service provider from implementing certain traffic composition policies. Two, a flow that traverses a longer path runs a higher chance of being rejected by one of the switches along the path, and hence rejected admission to the network, than one that traverses a shorter path. As the flow collects admissions from the switches, it makes reservations for resources and holds on to them. If the flow is denied admission to some switches, this can lead to a deadlock situation. While the first two forms of structural limitations are results of decisions local to a switch, higher rejection rates experienced by many-hop flows results from the dynamics of admission decisions across multiple switches and affects both the flow requesting admission and the state of the network.

Our initial results in [JDSZ97, JSD97] show that these structural limitations are present in all admission control algorithms that base their admission decision solely on the *violation prevention* paradigm, i.e. when the sole criterion of admission is that some service commitments are no violated, and not limited to measurement-based admission control algorithms.

8 Summary

In this paper we have presented simulation results which show that all MBAC's in the form of Eqn. 14 can be tuned to provide the same loss-load curve. We consequently recommended for use with controlled-load service an MBAC that allows for the greatest degree of flexibility in operation by relying on historical data for the tuning of its parameters. This MBAC is cost effective to implement and operate, relying only on the simplest measurement process. We closed this paper by suggesting two areas of research that should form the new MBAC research agenda, namely: the tuning of the MBAC parameters and solutions to the global structural limitations found on all admission control algorithms.

Acknowledgment

We thank Richard Gibbens and Frank Kelly for assiting us in understanding their approach.

References

[Bol94] V.A. Bolotin. "Modeling Call Holding Time Distributions for CCS Network Design and Performance Analysis". *IEEE Journal on Selected Areas in Communications*, 12(3):433–438, Apr. 1994.

[C+91] M. Conti et al. "Interconnection of Dual Bus MANs: Architecture and Algorithms for Bandwidth Allocation". *Journal of Internetworking: Research and Experience*, 2(1):1–22, March 1991.

[CKT96] C. Casetti, J. Kurose, and D. Towsley. "An Adaptive Algorithm for Measurement-based Admission Control in Integrated Services Packet Networks". *Proc. of the Protocols for High Speed Networks Workshop*, Oct. 1996.

[DJM97] Z. Dziong, M. Juda, and L.G. Mason. "A Framework for Bandwidth Management in ATM Networks — Aggregate Equivalent Bandwidth Estimation Approach". *ACM/IEEE Transactions on Networking*, Feb. 1997.

[DMRW94] D.E. Duffy, A.A. McIntosh, M. Rosenstein, and W. Willinger. "Statistical Analysis of CCSN/SS7 Traffic Data from Working CCS Subnetworks". *IEEE Journal on Selected Areas in Communications*, 12(3):544–551, Apr. 1994.

[Flo96] S. Floyd. "Comments on Measurement-based Admissions Control for Controlled-Load Service". Submitted to *Computer Communication Review*, 1996. **URL** ftp://ftp.ee.lbl.gov/papers/admit.ps.Z.

[GAN91] R. Guérin, H. Ahmadi, and M. Naghshineh. "Equivalent Capacity and Its Application to Bandwidth Allocation in High-Speed Networks". *IEEE Journal on Selected Areas in Communications*, 9(7):968–981, Sept. 1991.

[GK97] R.J. Gibbens and F.P. Kelly. "Measurement-Based Connection Admission Control". *International Teletraffic Congress15*, Jun. 1997.

[GKK95] R.J. Gibbens, F.P. Kelly, and P.B. Key. "A Decision-Theoretic Approach to Call Admission Control in ATM Networks.". *IEEE Journal on Selected Areas in Communications*, 13(6):1101–1114, Aug. 1995.

[JDSZ95] S. Jamin, P. B. Danzig, S. J. Shenker, and L. Zhang. "A Measurement-based Admission Control Algorithm for Integrated Services Packet Networks". *Proc. of ACM SIGCOMM '95*, pages 2–13, 1995. **URL** http://netweb.usc.edu/jamin/admctl/sigcomm95.ps.Z.

[JDSZ97] S. Jamin, P. B. Danzig, S. J. Shenker, and L. Zhang. "A Measurement-based Admission Control Algorithm for Integrated Services Packet Networks (Extended Version)". *ACM/IEEE Transactions on Networking*, Feb. 1997. **URL** http://netweb.usc.edu/jamin/admctl/ton96.ps.Z.

[JSD97] S. Jamin, S. J. Shenker, and P. B. Danzig. "Comparison of Measurement-based Admission Control Algorithms for Controlled-Load Service". *Proc. of IEEE INFOCOM '97*, Apr. 1997. **URL** http://netweb.usc.edu/jamin/admctl/info97.ps.gz.

[Kel96a] F. Kelly. Effective bandwiths, pricing, and admission control. Workshop on System and Control Issues in Communication Networks, Airlie, VA, Aug. 10, 1996.

[Kel96b] F. Kelly. Personal communication. Workshop on System and Control Issues in Communication Networks, Airlie, VA, Aug. 10, 1996.

[Kel96c] F.P. Kelly. *Notes on Effective Bandwidths*. Technical Report Research Report 1996-11, University of Cambridge, Statistical Laboratory, Feb. 1996.

[Mol27] E.C Molina. "Application of the Theory of Probability to Telephone Trunking Problems". *The Bell System Technical Journal*, 6:461–494, 1927.

[Par92] A.K. Parekh. *A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks*. PhD thesis, MIT, Lab. for Information and Decision Systems, Tech. Report LIDS-TR-2089 1992. Parts of this thesis were also published with R.G. Gallager in the *ACM/IEEE Transactions on Networking*, 1(3):344-357 and 2(2):137-150.

[PF94] V. Paxson and S. Floyd. "Wide-Area Traffic: The Failure of Poisson Modeling". *Proc. of ACM SIGCOMM '94*, pages 257–268, Aug, 1994. An extended version of this paper is available as **URL** ftp://ftp.ee.lbl.gov/papers/poisson.ps.Z.

[WCKG94] R. Warfield, S. Chan, A. Konheim, and A. Guillaume. "Real-Time Traffic Esitmation in ATM Networks". *International Teletraffic Congress*, June 1994.

[Wro95] J. Wroclawski. *Specification of the Controlled-Load Network Element Service*. Internet-Draft, Nov. 1995.

[WTSW95] W. Willinger, M.S. Taqqu, R. Sherman, and D.V. Wilson. "Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level". *Proc. of ACM SIGCOMM '95*, pages 100–113, Aug. 1995.

Appendix: Simulation Scenarios

For this paper, we run our simulations on a network topology where two switches are connected by a bottleneck link of 10 Mbps. Each switch is connected to a host with an infinite bandwidth link. Buffer space of the switch connected to the bottleneck links are shared by all admitted flows and are sized differently for each simulation, as we explain below. Traffic flows in one direction from one host to the other.

We use two kinds of source model in our simulations. The first is an ON/OFF model with exponentially distributed ON and OFF times. During each ON period, an exponentially distributed random number of packets, with average N, are generated at fixed rate p packet/sec. Let I milliseconds be the average of the exponentially distributed OFF times, then the average packet generation rate a is given by

Table 6: Six Instantiations of the Two Source Models

Model Name	Model Parameters				TB Filter			Switch Parameters	
	p pkt/ sec	I msec	N pkts	p/a	r tkn/ sec	b tkns	max qlen	D^* msec	D msec
EXP1	64	325	20	2	64	1	0	16	16
EXP2	1024	90	10	10	320	50	17	160	160
EXP3	∞	684	9	∞	512	80	1	160	160
				β					
POO1	64	325	20	1.2	64	1	0	16	16
POO2	64	2925	20	1.2	64	1	0	16	16
POO3	256	360	10	1.9	240	60	220	256	160

$1/a = I/N + 1/p$. The second model is an ON/OFF process with Pareto distributed ON and OFF times (for ease of reference, we call this the *Pareto*-ON/OFF (POO) model). Pareto distribution is a heavy-tailed distribution that can be described by two parameters: its location and shape. A Pareto shape parameter less than 1 gives data with infinite mean; shape parameter less than 2 results in data with infinite variance. The Pareto location parameter can be computed from the formula: $mean * (shape - 1)/shape$. During each ON period of the Pareto-ON/OFF model, a Pareto distributed number of packets, with mean N and Pareto shape parameter β, are generated at some peak rate p packet/sec. The OFF times are also Pareto distributed with mean I milliseconds and shape parameter γ. Each Pareto-ON/OFF source by itself does not generate LRD series; the aggregation of them does.

We use six instantiations of the above source models as summarized in Table 6. The shape parameter of the Pareto distributed ON time (β) of the Pareto-ON/OFF sources are selected following the observations in [WTSW95]. According to the same reference, the shape parameter of the Pareto distributed OFF time (γ) stays mostly below 1.5; in this paper we use $\gamma = 1.1$ for all Pareto-ON/OFF sources. For the POO1 and POO2 models, we use a token-bucket rate equal to the source's peak rate so that the token-bucket filter does not reshape the traffic. For the POO3 model, some of the generated packets were queued; this means during some of the source's alleged "OFF" times, it may actually still be draining its data queue onto the network. In the table, $p = \infty$ means that after each OFF time, packets for the next ON period are transmitted back to back. (On real networks, packets are sent back to back when the applications generate traffic faster than the network can transmit it.)

In the same table, we also list the settings of the token-bucket parameters assigned to each source. In this study, we assign each flow a data queue with infinite length (i.e. packets that arrive at an empty token-bucket are queued, and the queue never overflows). Column 7 of the table, labeled *max qlen*, shows the maximum data queue length a flow can expect to see. Our packets are of fixed size (1 Kbits) and each of our token is worth 1 packet of data.

When a flow with token-bucket parameters (r, b) is served with WFQ, the maximal queueing delay (ignoring terms proportional to a single packet time) is given by b/r [Par92]. This is also the "burst time" queueing delay acceptable under the definition of controlled-load service. Column 8 of the table, labeled D^*, lists the maximal delay for each source given its assigned token-bucket filter. Column 9, labeled D, lists the delay bound we assigned to each source. We have chosen the token-bucket parameters such that, in most cases, the delay bounds given to a flow will be the same as its "burst time" queueing delay. This facilitates analyzing the performance of the algorithms under controlled-load service. In the few cases where the delays are not the same, such as in the POO3 case, the analysis is less applicable. For each simulation with measurement-based admission control algorithm, we size the buffer at the switches with enough space to accommodate the delay bound (D). For example, simulations with EXP1 sources, given a link speed of 10 Mbps, use a buffer size of 160 packets.

In addition to each source's burstiness, network traffic dynamics is also effected by the arrival pattern and duration of flows. We use exponentially distributed lifetimes with the Markov-ON/OFF sources, following [Mol27]. The duration of Pareto-ON/OFF sources, however, are taken from a lognormal distribution, following [Bol94, DMRW94]. The interarrival times of all flows are exponentially distributed [PF94]. We choose an average flow interarrival time of 400 milliseconds. The average holding time of all Markov-ON/OFF sources is 300 seconds. The Pareto-ON/OFF sources have lognormal distributed holding times with median of 300 seconds and shape parameter 2.5. We run simulations with Markov-ON/OFF sources for 3e3 seconds simulated time, serving 1e7 packets. The data presented are obtained from the later half of each simulation. By visual inspection, we determined that 1500 simulated seconds is sufficient time for the simulations involving Markov-ON/OFF sources to warm up. Simulations involving Pareto-ON/OFF sources, however, require a longer warmup period and a longer simulation time for the LRD effect to be seen, thus we run them for 5.5 hours simulation time, serving close to 1e8 packets, with reported data taken from the later 1e4 seconds.