Comments on Measurement-based Admissions Control for Controlled-Load Services

Sally Floyd*
Lawrence Berkeley National Laboratory
One Cyclotron Road, Berkeley, CA 94720
floyd@ee.lbl.gov
DRAFT DRAFT DRAFT
to be submitted to CCR

July 9, 1996

Abstract

This paper considers measurement-based admissions control procedures for Controlled-Load services [Wro96]. First, we discuss some of the factors that reduce the requirements on a measurement-based admissions control procedure in an Internet environment with a strong mix of realtime and data traffic. Next, we discuss measurement-based admissions control procedures for Controlled-Load services that are based on the approach of equivalent capacity. Finally, we show some measurements of audio traffic for small audioconferences, and discuss why such traffic could be among the most problematic for measurement-based admissions control procedures.

1 Introduction

This paper makes several observations about measurement-based admissions control procedures for Controlled-Load services [Wro96].

We begin with a general discussion of the factors that would allow a measurement-based admissions control procedure to succeed even in the presence of traffic that is difficult to predict.

Next, we discuss a specific measurement-based admissions control procedure that is a modest variant of previously-proposed admissions control procedures based on the notion of equivalent capacity.

Finally, we describe traffic that we believe will be among the most problematic for measurement-based admissions control procedures. This traffic consists of audio traffic for a k-person audioconference, for moderate sizes for k. From the perspective of a specific link carrying such traffic, the audio flow would be at the negotiated peak rate for a sustained period of time, and the flow could also be idle for sustained periods of time.

This paper does not discuss mechanisms for scheduling and policing for Controlled-Load services. The current Internet-draft for Controlled-Load service specifies that packets for a Controlled-Load flow that exceed the token-bucket specifications for that flow should be forwarded on a best-effort basis if resources permit, and that routers should implement mechanisms in the best-effort queueing path to protect elastic best-effort traffic from inelastic best-effort flows. These concerns are orthogonal to the admissions-control issues discussed in this paper.

Section 2 gives the assumptions underlying the paper, and discusses some of the factors that reduce the demands on a measurement-based admissions control procedure. Section 3 describes a measurement-based admissions control procedure based on the approach of equivalent capacity. Section 4 discusses the source characteristics of audio traffic from small audioconferences.

2 Assumptions

This section explains some of the assumptions underlying this paper.

Our assumptions include the following:
(1) Some Internet applications will require explicit admissions control procedures;
(2) Many realtime applications that use explicit admissions control procedures will have adaptive playback times;

*Supported by the Director, Office of Energy Research, Scientific Computing Staff, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

(3) Controlled-Load service will meet the needs of many of the realtime applications that use explicit admissions control procedures;
(4) Much of the traffic that uses Controlled-Load service cannot be tightly and accurately characterized in advance;
(5) Measurement-based admissions control procedures show the most potential for most effectively utilizing the bandwidth available for Controlled-Load service; and
(6) It is likely that a large fraction of Internet traffic will continue to use the traditional best-effort service.

While each of these assumptions is accepted in some subset of the research community, each of these assumptions also has its challengers. We have not tried to collect the many citations to both sides for each of these assumptions.

Our first assumption is that for the short-term, at the very least, some applications will require explicit admissions control procedures to see adequate performance over the Internet. (We do not attempt to make predictions about the scarcity of bandwidth in the long-term future of the Internet.)

Opponents of explicit admissions control procedures [SCH$^+$95] argue that the first response to inadequate capacity should be the use of variable-bandwidth applications (e.g., source-based or receiver-based variable-bandwidth audio and video applications). While clearly the Internet architecture should encourage the use of variable-bandwidth applications whenever possible, there are limits to this approach, and some applications effectively have minimum bandwidth requirements [She95].

Opponents of explicit admissions control also have argued that the use of admissions control procedures assumes answers to a wealth of social, political, and economic questions about who does and doesn't get to use scarce bandwidth in times of congestion. While we agree that these are problematic issues, we don't think that the difficulties of these questions should prevent us from making the controlled decisions about the use of bandwidth that are implicit in an admissions control procedure.

Our second assumption is that much real-time traffic will be both adaptive and somewhat drop-tolerant [CSZ92]. That is, we assume that much real-time traffic will have adaptive playback times, and therefore will be able to adjust, within limits, to varying delays in the network. We further assume that such adaptive applications are likely to be tolerant of an occasional packet drop.

This does not mean that we believe that *all* realtime applications will be adaptive, and therefore tolerant of a measurement-based admissions control. Applications that require stricter guarantees can use more specialized services such as guaranteed service where they are offered.

Our third assumption is of the context of a Controlled-Load service [Wro96] amenable to a bandwidth-based admissions control procedure. The Integrated Services Working Group [PWC] of the IETF has discussed four separate qualities of service for realtime traffic in the Internet: Guaranteed [SP95], Predictive [SPD95], Controlled Delay [SPW95], and Controlled Load [Wro96] Service. On these, the Guaranteed and Controlled Load services are to become proposed standards, for further deployment and experimentation in the Internet.

Guaranteed Service provides an assured level of bandwidth that, when used by a policed flow, produces a delay bounded service with no dropped packets. The Predictive and Controlled Delay Services, which are not currently planned for deployment, were both intended as realtime services that provide low packet loss. Predictive service was intended for applications that require an upper bound on end-to-end delay. The measurement-based admissions control procedures designed for predictive service [JDSZ95, CSZ92] are based on measurements of packet delay as well as of bandwidth, and use measurements of the maximal packet delay seen by a predictive service packet over a recent time interval.

Controlled Delay Service was intended for applications that can dynamically change the level of packet delay they request from the network when their current level of delay is not adequate.

In contrast, Controlled-Load Service is intentionally minimal, offering only a single class of service, and requires an admissions control procedure whose goal is simply to ensure that "adequate bandwidth and packet processing resources are available to handle the requested level of traffic". This paper is only concerned with admissions control procedures for Controlled-Load Service or similar bandwidth-based services.

Our fourth assumption is that tightly characterizing source traffic is neither necessary nor plausible, given the diversity of Internet applications that might use real-time services. Clearly, some traffic, including some video traffic, lends itself to tight characterization. Nevertheless, for much real-time traffic, such as the audio traffic described in Section 4, peak-rate characterization is the tightest appropriate characterization.

Our fifth assumption is that, given the need for some admissions controls, the difficulty in tightly characterizing the traffic, and the pervasiveness of adaptive applications, then at locations with sufficient levels of aggregation, measurement-based admissions control procedures show the most potential for allowing as many flows as possible to share the available bandwidth [YKTH93].

[CSZ92] was one of the first papers to argue for the importance of measurement-based admissions control procedures for realtime services in the Internet. In the absence of tight characterizations of the source traffic, an admissions control procedure based on the traffic characterizations provided at admit-time will use the available bandwidth conservatively, and will end up in unnecessarily rejecting some requests for admissions. This might be unavoidable at a location with only a small number of admitted connections. However, for a location with a larger number of admitted connections, the admissions control procedure can be guided not only by the traffic characterizations of the accepted flow and of the flow applying for admissions, but also by measurements of the past bandwidth in fact used by the other admitted connections.

Our sixth assumption is that a dominant fraction of Internet traffic will continue to use best-effort service, and will not use explicit reservations. One of the main dangers with a measurement-based admissions control procedure is that the measured average arrival rate could be unrepresentatively low for a period of time, resulting in a succession of admissions and more problems when the average arrival rate returns to its normal higher level. For this reason, for a link with a low level of aggregation, the admissions control procedure will probably be simply based on the peak rates of the admitted flows. For links with a higher level of aggregation, where measurement-based admissions control procedures can be of substantial benefit, the presence of a large fraction of best-effort traffic significantly reduces the demands on the accuracy of the measurement-based admissions control procedure.

For example, although the admissions control procedure might have the goal of restricting the admitted realtime traffic to a fraction p of the link bandwidth, the scheduling algorithm might allow the admitted realtime traffic to use up to p_1 of the link bandwidth, for p_1 significantly larger than p. This still guarantees the best-effort traffic a fraction $1 - p_1$ of the link bandwidth, while ensuring that the realtime traffic won't be negatively affected by the failure of the admissions control procedure until the actual arrival rate for the realtime traffic exceeds p_1.

We also assume that an increasing fraction of realtime traffic to be able to vary the bandwidth used in times of congestion. Variable-bandwidth techniques include both source-based rate adaption and receiver-based solutions [BTW94, MJV96]. Although this paper does not address the potential requirements of realtime services for rate-adaptive realtime traffic, we assume that this will be a key development for realtime traffic over the Internet.

3 Admissions control procedures based on the equivalent capacity

In this section we use the approach of equivalent capacity for a class of aggregated traffic [GAN91] to give a measurement-based admissions control procedure. Following previous work [GAN91], we define the *equivalent capacity* of a class of traffic as that value $C(\epsilon)$ such that the stationary arrival rate for the class exceeds $C(\epsilon)$ with probability at most ϵ. We propose an admissions control scheme that estimates the *equivalent capacity* of a class. A connection is admitted to a class if, with the addition of the new connection, the equivalent capacity for the class would be less than the allocated bandwidth for that class. This allows a simple, fairly easily computed admissions control procedure.

For the measurement-based admissions control procedure described in this paper, we assume that the traffic parameters for new connections include token bucket parameters with the token rate and bucket size, as in Controlled-Load service, and that these token bucket parameters are policed at the router. Traffic from admitted connections that exceeds these token bucket parameters is forwarded on a best-effort basis if resources are available, but does not have to be considered by the admissions control procedure. The admissions control procedure outlined in this paper uses the policed token bucket parameters to infer a peak rate for each flow over some fixed time interval. For such an admissions control procedure, a request for admissions is most likely to be accepted if the token rate is set close to the peak rate of the flow, and the bucket size is set to some small value to accommodate the accumulated jitter as packets are jostled in the network.

We assume that the admissions control procedure can determine when a connection has ended (that is, when a reservation has been torn down), even if the reservation is not explicitly torn down by the session participants.

The simpliest measurement-based admissions control procedures rely simply on the measured average arrival rate over some previous interval. The approach presented in this section uses measurement-based estimates of the average aggregate arrival rate, along with the peak rates of the individual flows (as derived from the token bucket parameters).

The proposal from [GAN91] suggested that the class arrival rate be modeled as a random variable with a normal distribution. This approach is likely to be most appropriate for a large realtime class with hundreds of admitted connections with similar peak rates. In this section we present a formulation of equivalent capacity that does not assume a normal approximation of the arrival rate distribution, and that is suitable for classes with either a moderate number of admitted connections or a wide range in the peak rates of the admitted connections.

Notation: Let $X_{i,T}$ be the instantaneous arrival rate used by connection i at time T. We assume that for $i \neq j$, $X_{i,T}$ and $X_{j,T}$ are independent; we don't necessarily assume that $X_{i,T}$ and $X_{i,T+\delta}$ are independent. Let $S_T = \sum_{i=1}^{n} X_{i,T}$ be the instantaneous arrival rate used by the class at time T. Let $\hat{\mu_S}$ be the estimated average arrival rate, based on measurements of the aggregate class arrival rate.

3.1 Estimating the equivalent capacity using the normal distribution

The formulation for the equivalent capacity of aggregated traffic given by [GAN91] models the instantaneous aggregate arrival rate S_T as having a normal distribution. Assume that upper bounds for the average μ_S and variance σ^2 of the class arrival rate S are derived either computed from the token bucket parameters of the individual connections or estimated directly from measurements. Then $\hat{C_N}$, the equivalent capacity based on the normal distribution [GAN91], is given by

$$\hat{C_N}(\mu_S, \sigma^2, \epsilon) = \mu_S + \alpha\sigma,$$

$$\text{for } \ \alpha = \sqrt{2 \ln \frac{1}{\epsilon} + \ln \frac{1}{2\pi}}.$$

That is, if the arrival rate S has a normal distribution, then at time t in the future, the arrival rate S_t should exceed the estimated equivalent capacity $\hat{C_N}$ with probability at most ϵ.

[GAN91] also presents an alternate formulation of the equivalent capacity for a single flow, using a queueing theory approach that takes into account the maximum buffer size and mean burst period for the connection as well as other token bucket parameters. Because we are concerned in this paper with measurement-based admissions control procedures, we do not pursue this approach.

For admissions controls based on equivalent capacity, the parameter ϵ gives an estimated upper bound on the probability that the instantaneous arrival rate exceeds the allocated throughput for a class. The smaller the value for ϵ, the more conservative the admissions control algorithm for that class. A new flow's admission request would be accepted if the peak rate of the new flow added to the equivalent capacity is less than the allocated bandwidth for that class.

As noted in [GAN91], because the formula for $\hat{C_N}(\mu_S, \sigma^2, \epsilon)$ assumes a normal distribution for the class arrival rate, this approach is not suitable for small or moderate-sized classes. As Figure 1 shows later in the paper, for smaller classes the estimated equivalent capacity based on the normal distribution can underestimate the actual equivalent capacity. This paper gives an alternate formula for the equivalent capacity that is valid even for moderate-sized classes.

3.2 Estimating the equivalent capacity using Hoeffding bounds

This section gives an upper bound for the equivalent capacity that uses an upper bound on the peak rate for each admitted connection along with the measured aggregate arrival rate. Appendix C describes how an upper bound on the peak rate can be derived from the token bucket parameters. For an admitted connection i with instantaneous arrival rate X_i, the minimum rate is zero, and an upper bound on the peak rate p_i over a fixed time interval can be derived from the token bucket parameters.

To efficiently estimate an upper bound on the equivalent capacity, we use a result derived from Hoeffding.

Theorem 1 *[Hoe63, p.16], [Hof87, p.104]: Let X_1, X_2, ..,X_n be independent, and let $0 \leq X_i \leq p_i$. Then for $t > 0$,*

$$Prob[S \geq \mu_S + nt]$$
$$\leq e^{-2n^2t^2/\sum_{i=1}^{n}(p_i)^2}.$$

□

This theorem is one of a family of Chernoff-style bounds that considers a sum of random variables, and gives an upper bound on the tail of the distribution. In particular, Theorem 1 quantifies the probability that the sum of the n random variables will be greater than the average of the sum by $n\ t$ or more.

We want to find t such that

$$Pr\left[S_T \geq \mu_S + n * t\right] \leq \epsilon.$$

This is satisfied if

$$e^{-2n^2t^2/\sum_{i=1}^{n}(p_i)^2} \leq \epsilon,$$

for ϵ a parameter that controls the degree of risk of the admissions control procedure. This is satisfied for

$$t \leq \frac{1}{n}\sqrt{\frac{\ln(1/\epsilon)\sum_{i=1}^{n}(p_i)^2}{2}}$$

Thus the equivalent capacity $\hat{C_H}$ based on peak rate policing is

$$\hat{C_H}(\mu_S, \{p_i\}_{1 \leq i \leq n}, \epsilon) =$$

$$\mu_S + \sqrt{\frac{\ln(1/\epsilon)\sum_{i=1}^{n}(p_i)^2}{2}}. \qquad (1)$$

As described before, a new flow's admission request would be accepted if the sum of the peak rate of the new flow and the equivalent capacity was less than the allocated bandwidth for that class. In this case, the estimated equivalent capacity is based on the measured average arrival rate used by the class and on the peak rates on the individual flows derived from the policed token bucket parameters.

The estimated equivalent capacity given by $\hat{C_H}$ uses the assumption that the arrival rates for the different flows are independent. The estimated equivalent capacity also assumes that the admissions control procedure knows μ_S, the average of the aggregate arrival rate.

Note that the admissions control procedure has no information about how $X_{i,T}$ is correlated with $X_{i,T+\delta}$, and no knowledge of how S_T is correlated with $S_{T+\delta}$. In the rigorous sense, this approach to equivalent capacity is based on the assumption that, for δ sufficiently large, $X_{i,T}$ and $X_{i,T+\delta}$ are not correlated, and therefore S_T and $S_{T+\delta}$ are not correlated. It is obvious, from the discussions of audio traffic earlier in the paper, that these assumptions are not necessarily correct.

The Hoeffding bound is *guaranteed* to give an upper bound on the equivalent capacity only if the assumptions above are true, and if we have an upper bound for the average arrival rate for the aggregate admitted connections. Thus, the measurement-based admissions control procedure is only reliable to the extent that our measurements of the aggregate arrival rate are good predictors for the future average arrival rate. The second part of the admissions control procedure, namely the measurement of the average arrival rate, is discussed briefly in a later section.

3.3 Evaluating the Hoeffding bounds

This section considers a class with n admitted ON/OFF connections, and numerically compares the exact equivalent capacity, the estimated equivalent capacity $\hat{C_N}$ based on the normal distribution, and the estimated equivalent capacity $\hat{C_H}$ based on the Hoeffding bounds. The results show that, while $\hat{C_N}$ is generally a closer approximation to the exact equivalent capacity than is $\hat{C_H}$, $\hat{C_N}$ sometimes underestimates the exact equivalent capacity.

For each scenario we randomly create n ON/OFF connections, each with a different peak rate p_i and ON-probability o_i. From this we compute the exact equivalent capacity and the estimated equivalent capacities $\hat{C_H}$ and $\hat{C_N}$. The computations of $\hat{C_H}$ and $\hat{C_N}$ assume that the admissions control procedure has exact knowledge of the average arrival rate

$$\sum_{i=1}^{n} p_i o_i$$

(and, for $\hat{C_N}$, exact knowledge of the variance

$$\sum_{i=1}^{n} p_i^2 o_i(1-o_i)$$

of the arrival rate).

For Figure 1, for each numerical experiment we randomly created n ON/OFF connections as follows: for the i-th connection, the peak rate p_i was chosen from the uniform distribution on [0,1], and the probability o_i of being in the ON state was chosen from the uniform distribution on [0, M]. The top row shows scenarios with M=0.5, and the bottom row uses M=0.1. For a scenario with n connections, the ratio between the largest peak rate and the smallest peak rate is likely to be on the order of n to 1.

The four graphs show results for both values of M, and for ϵ set to 0.01 and to 0.000001. For each graph, for selected values of n starting with $n = 10$, we randomly created n ON/OFF connections, computed the exact equivalent capacity as described in Appendix D, and computed the two estimates of the equivalent capacity.

For each graph in Figure 1, the x-axis shows n, the number of connections. The solid line shows the exact equivalent capacity, expressed as a fraction of the sum of the peak rates, and the two dashed lines show the two approximations to the equivalent capacity. The dotted line shows the sum of the average rates for the n connections, again expressed as a fraction of the sum of the peak rates.

Figure 1 shows that statistical multiplexing has benefit even for moderate-sized classes. The y-axis is normalized so that an admissions control procedure that reserved the sum of the peak rates would reserve a capacity of 1. The figures suggest that for classes with only 10 connections, an admissions control procedure based on equivalent capacity would be essentially the same as a peak-rate admissions control procedure. However, for classes with as few as 50 connections, an admissions control procedure based on equivalent capacity could be of benefit, providing the required service while reducing the number of flows whose admissions requests are refused. As n increases, the exact and the estimated equivalent capacity both gets closer to the average rate, illustrating the increased benefits of statistical multiplexing as classes grow larger.

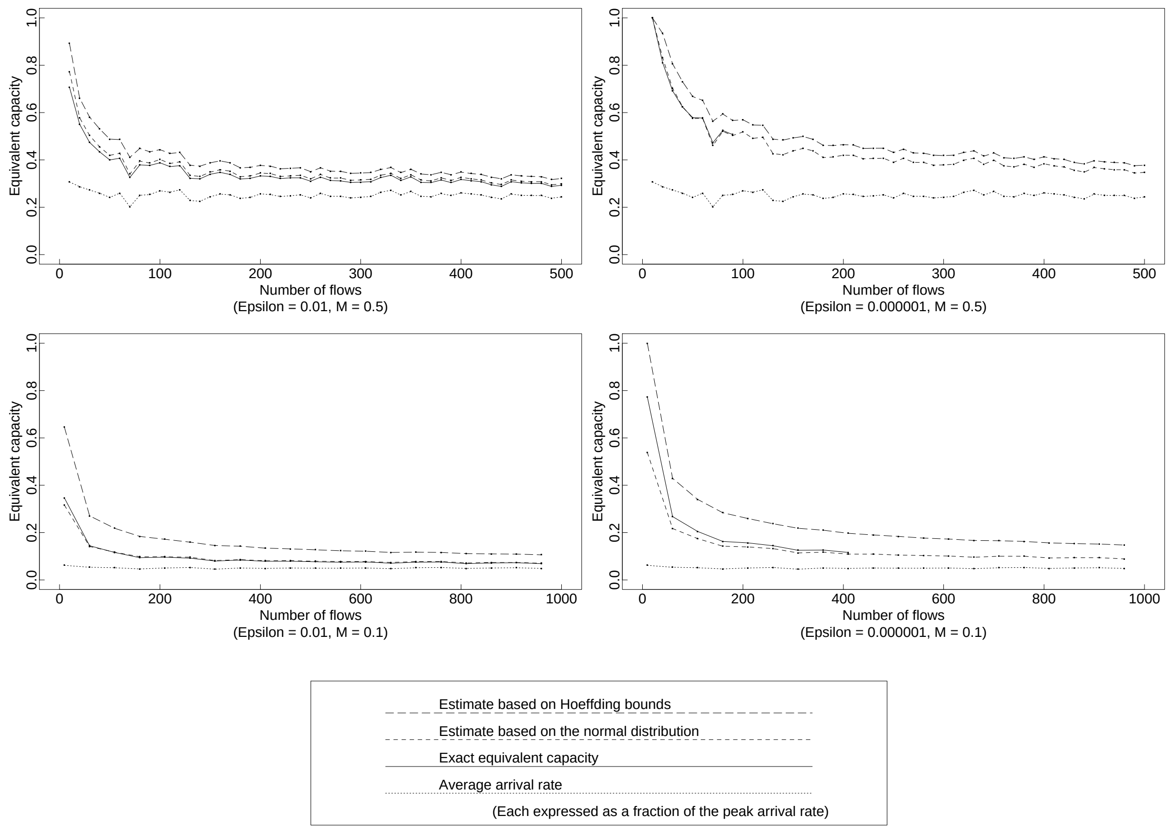

Figure 1: A class with n connections, with different peak and average rates. The ratio between the average rate and the peak rate for an individual connection ranges from 0 to M.

3.4 Description of the admissions control procedure

The Hoeffding bound leads to the following measurement-based admissions control procedure. Given n admitted flows, and measured aggregate arrival rate $\hat{\mu_S}$, admit the new flow with peak rate p_{n+1} if

$$\hat{C_H} + p_{n+1} \leq C,$$

for $\hat{C_H}$ the estimated equivalent capacity and C the target bandwidth for the class.

After the new flow is admitted, the peak rate p_{n+1} for that flow is added to the estimated class mean $\hat{\mu_S}$. If the aggregate arrival rate is estimated using an exponential weighted moving average, then the new estimate will gradually begin to reflect the actual added traffic from the new flow, and not the peak rate p_{n+1}. Note, however, that this constrains the time constants for estimating the average arrival rate; the estimator must not "forget" the peak rate p_{n+1} added to the class mean before the router is likely to begin seeing traffic from that new flow.

Following [GKK95], when a request for admissions is refused, then the admissions control procedure should not consider other requests of the same magnitude until one of the admitted flows has terminated.

When a reservation terminates, the admissions control procedure removes that flow from the list of admitted flows used in computing the equivalent capacity. However, the admissions control procedure does not decrement the measured aggregate arrival rate; the network has no knowledge of how much the terminated flow was contributing to the aggregate arrival rate. The new measurements of the aggregate arrival rate will begin to reflect the aggregate arrival rate without the recently-terminated flow.

3.5 Choosing a value for ϵ

For an admissions control procedure based on equivalent capacity, a key parameter is the value chosen for ϵ. Given an environment with a large fraction of best-effort traffic, the best strategy is probably for network administrators simply to choose a value of ϵ that works well in practice. (that is, a value of ϵ that minimizes both the number of refused requests for admissions and the time for which the aggregate Controlled Load arrival rate exceeds the desired fraction of the link bandwidth). Given the current lack of information about the likely traffic characteristics of Controlled Load traffic, the best approach in probably to make choices for ϵ based on an accumulation of experience with Controlled Load traffic. It would probably be useful to start with a small value for ϵ, representing a fairly conservative admissions control procedure, and to increase ϵ if experience indicates that a less conservation admissions control procedure would be adequate.

3.6 Estimating the average arrival rate

A key concern in the measurement-based admissions control procedure is how to estimate the average arrival rate (or the average bandwidth) for the class.

There are several difficulties in estimating the average arrival rate from measurements. First, even with the number of admitted connections fixed, the assumption of the existence of a "steady-state" average is not necessarily accurate.

Second, even if there was a "steady-state" average arrival rate, this average arrival rate could be difficult to determine from measurements because of the possible long-range dependence of the aggregate traffic. This gives a second reason why the measured average arrival rate over some time interval might not be a good predictor for the average arrival rate for a future time interval. This seems to us an unavoidable risk for a measurement-based admissions control procedure. For this reason, using a measurement-based admissions control procedure in an environment where the goal is to use all of the link bandwidth as efficiently as possible strikes us as highly problematic. However, because realtime services are likely to be placed in an Internet where a large fraction of the link bandwidth is used by rate-adaptive best-effort traffic, measurement-based admissions control procedures should be able to accomodate some lack of precision in the measured average arrival rate.

A simple procedure for estimating the average arrival rate would be to measure the arrival rate r_i once every A seconds, for that A-second interval. The average arrival rate could then be calculated using an exponential-weighted moving average with a weight w: $avg = (1 - w) * avg + w * r_i$. The time constant for this is

$$t = -1/\ln(1 - w) * A$$

seconds. [That is, assume that the class arrival rate changes abruptly from 0 to 1, and that the arrival rate remains at the value 1. After t seconds, the estimated average arrival rate will have reached 63% of the new arrival rate value of 1.]

This time constant t should be as long as the interval from when a connection is admitted, until the measurements are likely to actually reflect traffic from that connection. If the time constant is too short, then the potential traffic from newly-admitted connections will not be taken into account in successive requests for admissions. If the time constant is too large, then the estimate for the average arrival rate will reflect the contribution of old connections long after they have in fact been terminated. Similarly, the estimated average arrival rate would reflect the peak rates of new connections long after the admissions time.

For a time constant of at most t, that is, for

$$-1/\ln(1 - w) * A \leq t,$$

we require

$$w \geq 1 - e^{-A/t}.$$

In this paper, we have referred to an estimate of the average *arrival rate* of the aggregate admitted traffic. In practice, an estimate of the average *bandwidth* actually used by the admitted traffic would be sufficient. This is particularly true if the class is allowed to use slightly more bandwidth that the target bandwidth of the class.

If the aggregate arrival rate is not estimated using an exponential weighted moving average, then the admissions control procedure needs some procedure (e.g., after waiting a simple fixed time interval) for determining when the peak rate p_{n+1} for the recently-admitted flow can be deleted as a component of the estimated equivalent capacity.

3.7 Previous work on measurement-based admissions control for bandwidth-based admissions control procedures

In this section we briefly discuss some of the related work on measurement-based admissions control for bandwidth-based admissions control procedures.

[SS91] proposed a dynamic admissions control procedure based on measurements of the number of cells arriving during a fixed interval. [SS91] pointed out that an admission procedure based on actual measurements of the arrival rates at the gateway tolerates possible errors in the policing at the edge of the network. In [SS91], after a new connection is admitted, the traffic parameters of the new connection are used to update the estimate of the distribution of cell arrivals. This estimated distribution is then updated using exponential weighted moving averages.

There is a growing literature on admissions control procedures based on equivalent capacity [GAN91]. We discuss one paper in particular [GKK95] that considers an approach based on equivalent capacity in the context of measurement-based admissions control procedures.

[GKK95] describes a measurement-based admissions control procedure based on measurements of the current admitted load. The assumptions outlined in [GKK95] are that statistical multiplexing has benefits over peak rate allocations, that it is not viable to assume that each flow has an efficiently-policed mean rate, and that the admissions control procedure should not depend critically on traffic assumptions.

[GKK95] describes a measurement-based admissions control procedure based on Chernoff bounds that is similar in spirit to the admissions control procedure described in this paper. A new call is accepted when the measured arrival rate is below a certain threshold, and rejected otherwise. After a call is rejected, no further requests are considered until some accepted call ends.

Unlike the Hoeffding bound given in this paper, which gives a measurement-based admissions control procedure based on the policed peak rates of the admitted connections as well as measurements of the aggregate arrival rate, the Chernoff bound in [GKK95] gives a measurement-based admissions control procedure based on measurements of the aggregate arrival rate and on a single burstiness parameter p for all of the admitted connections, where p is the probability that a flow is active at a particular point in time.

[GKK95] goes on to develop a decision-theoretic measurement-based admissions control procedure that does not require previous knowledge of the value of the burstiness parameter p, and only assumes that p is relatively stable. Instead of using a parameter ϵ, and admitting a new connection when the probability of exceeding the target bandwidth is estimated to be at most ϵ, the decision-theoretic approach balances the "reward" for each packet that is successfully delivered against a "penalty" for each packet that is dropped.

The experiments of the admissions control procedure in [GKK95] are quite encouraging. Nevertheless, the experiments use a relatively tractable model that assumes that the instantaneous load at time t and at time $t + \delta$ are independent, given the burstiness parameter p. Unfortunately, little can be known at this time about the likely mix of realtime traffic for future realtime services in the Internet. [GKK95] acknowledges the problem that, given successive highly correlated load measurements, a sequence of calls could be admitted in quick succession that later lead to problems.

4 Characterizing the source

In this section we examine some of the difficulties of trying to tightly characterize the sources (other than by peak rate). In particular, we believe that the following approaches to traffic characterization are fraught with danger: (1) characterizations based on unicast traffic models; (2) characterizations based on exponential ON/OFF models, or other models that are not compatible with long-range dependence; (3) characterizations based on restrictive assumptions (e.g., a particular video coding algorithm, a particular dominant traffic type such as video, or a particular likely size for multicast groups for multicast traffic). Along with many other researchers [CSZ92, GKK95, JDSZ95], we have considered measurement-based admissions control procedures that make as few assumptions as possible about traffic characterizations.

4.1 Lessons from audio traffic

This section considers audio traffic from small audioconferences, traffic that we believe will be among the more challenging for measurement-based admissions control procedures. In particular, such audio traffic seems more problematic than video traffic, audio traffic from two-way conversations, or one-to-many audio traffic. We consider this traffic not because we claim that this will be a dominant traffic load for future realtime services in the Internet, but to emphasize the danger of making assumptions that future realtime traffic will be well-behaved.

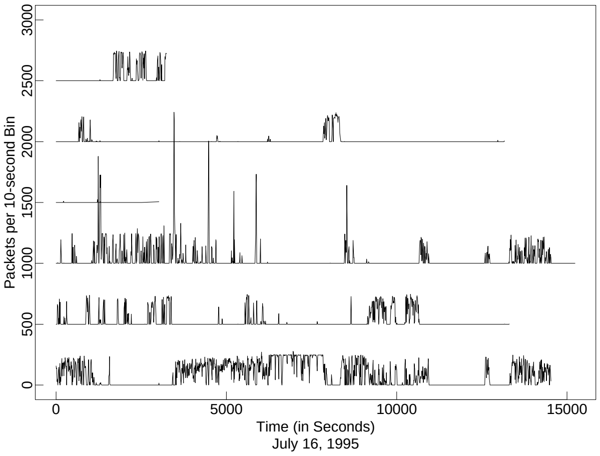

Figure 2: Audio patterns from a four-hour trace of a July 1995 audioconference.

Figures 2 and 3 show traffic patterns for two very different small audioconferences. For each audioconference we used tcpdump to record the number of data packets sent by each IP source address over 0.1-second intervals. (The tcpdump command is given in Appendix B.) We filtered out the smaller session packets; thus, these traces do not show session members who are listen-only. We also filtered out the actual IP addresses. While these traces only show the traffic pattern as received at LBL, this should be adequate for the time scales of interest for this discussion.

In these graphs, each row of the graph represents a different IP source. The x-axis shows time, and the y-axis shows the number of audio packets sent by each

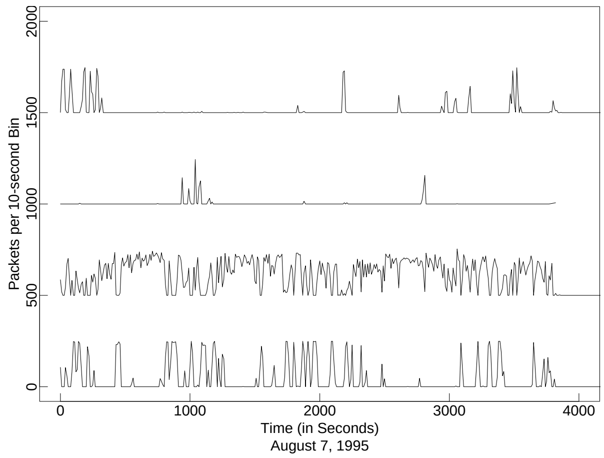

Figure 3: Audio patterns from a one-hour trace of a August 1995 audioconference.

source over successive 10-second bins. Each line shows the number of audio packets from one of the conference participants for successive 10-second bins, and the y-axis for each participant should be read as ranging from 0 to 500 packets per 10-second bin. At the peak rate sources were sending 64 kbps (or 248 328-byte packets over 10-second intervals).

Figure 2 shows the audio traffic sent to a particular multicast address over a four-hour time period. There were speakers from six sites (including LBL). The occasional spikes for source #3 are due to problems with an echo-canceller, resulting in peak rates that exceed the the default peak rate of 64 kbps.

The audio pattern in Figure 2 could be problematic for measurement-based admissions control procedures, depending on the level of aggregation and mix of real-time traffic on the link. This pattern is also quite different from the standard audio model in the literature of a single speaker in a two-way conversation, of independent ON and OFF time intervals with exponential distributions [NKT91]. In Figure 2 a single speaker can have the floor for an extended period (15 minutes, say, for speaker #1), then be completely silent for half an hour, and then speak at peak rate for much of the following hour. Figure 2 suggests that there are long-range correlations in the data. For example, there can be very long OFF intervals even for the most active speakers.

Note that in Figure 2 the entire multicast group is idle for a half-hour (from 10940 seconds until 12600 seconds) and then starts up again. Whether long group-wide idle periods will persist for traffic subject to a measurement-based admissions-controlled procedure, or whether reservations will be instead torn down and explicitly re-established a half-hour later, will depend on such factors as the pricing structures and the perceived possibility of a future rejection of an admissions request.

In contrast to Figure 2, Figure 3 shows a small audioconference that would be more tractable for a measurement-based admissions control procedure. The floor control patterns remain fairly consistent for each speaker for the duration of the audioconference.

One possibility for modeling audio traffic would be to collect data from a number of audioconferences, and to try to characterize them in terms of distributions of ON and OFF intervals (e.g., is the distribution of ON and/or OFF times heavy-tailed?). As an example of this approach, Figure 4 shows the complementary distribution plots [WTSW95] for the bottom three speakers in Figure 2. The bottom row in Figure 4 shows the ON and OFF times for the bottom speaker in Figure 2. (We do not explain complementary distribution plots in this paper. We refer the interested reader to [WTSW95] for a discussion of the use of complementary distribution plots to investigate the distribution of ON and OFF times in traffic generated by a particular source. A distribution with hyperbolic tails is indicated by an approximately straight line on the complementary distribution plot with a slope between -1 and -2.) Figure 2 suggests that while the ON times for two of the speakers might fit an exponential distribution, the ON and OFF times for the third speaker are more suggestive of a Pareto distribution. A statistically more rigorous examination of the traces would use Hill's estimate, as described in [WTSW95].

However, there are a number of difficulties with an attempt to systematically characterize the ON and OFF times of Mbone audio traffic. First, it is unclear what type of audio traffic is likely to be typical of future internet audio traffic: two-way conversations, small group audioconferences, and large audioconferences (e.g., IETF, rock concerts) are likely to have quite different traffic patterns.

Second, even for small audio conferences, it is unclear what would be the most appropriate model. For example, a model that assumes that successive ON and OFF times are uncorrelated might not be adequate.

A third difficulty is that the admissions control procedure is concerned not with the traffic pattern of a particular source, but with the traffic pattern of the k sources in the multicast group whose multicast traffic happens to be carried on that link.

Rather than attempting to construct source models for audio traffic at this stage, we emphasize measurement-based admissions control procedures that make as few assumptions as possible about the traffic patterns for individual sources or flows. We do, however, rely on the assumption that traffic patterns of separate multicast groups are uncorrelated.

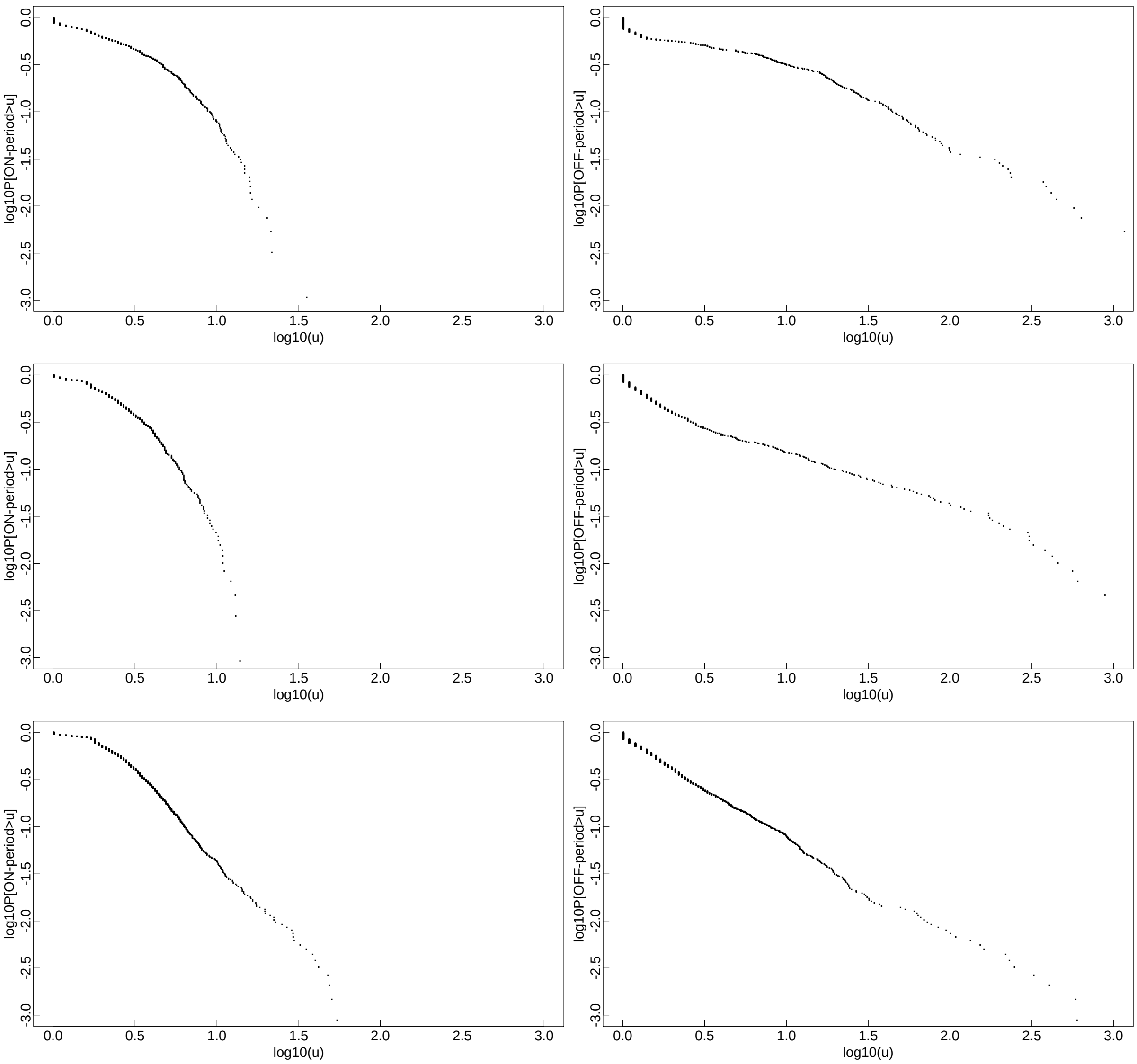

Figure 4: Complementary distribution plots for the ON-periods (left) and OFF-periods (right) for three speakers in the July 1995 audioconference.

Figures 2 and 3 underscore the potential advantages for measurement-based admissions control procedures for audio traffic in particular. For an n-person audioconference, any particular link carries audio traffic from some subset k of those participants, where the exact subset depends on the underlying multicast trees. Unlike video traffic, where the peak-to-mean ratio for a particular video flow is likely to be a small constant (e.g., two or three), the audio flow from different speakers in an audioconference can have quite different peak-to-mean ratios. For example, for Figure 2, considering each audio flow as starting at time 0 and ending with the last data packet generated by that flow, and assuming peak rates of 64 kbps, the peak-to-mean ratios range from 2.9 for speaker 1 to 33 for speaker 5 (and 1250 for speaker 4). For Figure 3, the peak-to-mean ratios range from 3 for speaker 2 to 100 for speaker 3. Thus, some links in the underlying multicast trees will be carrying audio traffic for a particular audioconference with a very large peak-to-mean ratio. In this case the benefits of measurement-based admissions control procedures can be significant.

At the same time, two speakers with similar peak-to-mean ratios for their generated audio traffic (e.g., speaker 1 in Figure 2 and speaker 2 in Figure 3) could have quite different ON/OFF patterns, in terms of the degree of burstiness over longer time scales. This argues that a measurement-based admissions control procedure should make as few assumptions about the underlying traffic as possible.

Some researchers have suggested that audio traffic for a particular audioconference on the Internet is likely to be accompanied by video traffic, and that the aggregate audio and video traffic will be dominated by the characteristics of the video traffic. In an environment with cheap and plentiful bandwidth, this might well be true; in such an environment, however, admission control procedures themselves are of minor interest. In an environment with limited bandwidth (relative to the current demand), we think it likely that audio traffic will often be accompanied either by low-bandwidth video, variable-bandwidth video, or no video at all. In such circumstances, the nature of the audio traffic has a significant impact on the characteristics of the aggregate traffic.

One conjecture is that audio traffic from a particular flow can be modeled by an ON/OFF model with heavy-tailed ON and OFF times, and therefore that aggregate audio traffic is long-range dependent. It has already been shown by Garrett that video traffic even for a single flow can be self-similar. Of course, it is possible that aggregate video traffic could be self-similar and at the same time have a low variance; in this case the aggregate video traffic would present few difficulties for a measurement-based admissions control procedure. Because Internet-based audio traffic has a wider range of peak-to-mean ratios than does video traffic, aggregate audio traffic could be more of a challenge for measurement-based admissions control procedures.

4.2 Other measurement studies of realtime traffic

This section gives a non-systematic overview of some of the measurement studies of realtime traffic.

[LTWW94] shows that Ethernet traffic can be modeled by self-similar (i.e., fractal) models. Extending this approach to realtime traffic, [BSTW94, GW94] has shown that variable-bit-rate (VBR) video traffic exhibits self-similar characteristics. Other studies of video traffic [HDLK95, BSTW94, ENW96] show that VBR video has strong short-term correlations, and that models for VBR video must reflect both short-range dependence (SRD) and long-range dependence (LRD).

[Mah95] describes measurements of IP multicast traffic taken at UC Berkeley. This includes measurements of audio and video traffic, along with traces of the still image distribution program `imm`. The traces of `imm` show essentially ON/OFF traffic, with half-hour idle periods between images not uncommon. One day-long trace of a session has a one-second peak of 8.8 Kbps, with half-hour idle periods between images not uncommon, resulting in an average throughput of 0.27 Kbps. Even though the peak-to-mean ratio is large, the peak rate is still quite small, compared to voice and video traffic.

5 Conclusions and open problems

There are many open problems in designing measurement-based admissions control procedures, and there is much to be learned from the incremental deployment of realtime services in the Internet. Work needs to be done on characterizing the traffic load, on developing procedures for accurately estimating the average arrival rate, and so on. Because we are unlikely to be successful in predicting the characteristics on future realtime traffic in the Internet, it is important to rely on admissions control procedures that make as few assumptions as possible about traffic characteristics, other than the per-flow token bucket parameters policed at the router.

While equivalent-capacity-based admissions control procedures seem a perfect fit to Controlled Load services, there is clearly much to be learned about the most appropriate admissions control procedures, and about the nature of the relevant measurements of aggregate traffic for measurement-based admissions control procedures. In particular, the estimated equivalent capacity based on the Hoeffding bound described in this paper is not necessarily the best upper bound on the equivalent capacity for moderate sized classes. For example, the work of Bennett [Ben62] could lead to another approach to equivalent capacity. Bennett explores the distribution of the sum of independent random variables when we are given the variance of the aggregate and an upper bound for the variables (e.g., the largest per-flow peak rate.) These upper bounds are applicable when the number of component random variables is small and/or have different distributions.

Given a network in which realtime traffic is mixed with a sizeable fraction of non-reservation-based data traffic, the constraints are sufficiently loose that much can and should be learned not only from analytical research, but also from understanding the behavior of measurement-based admissions control procedures as they actually begin to be deployed.

6 Acknowledgements

This paper is an outgrowth of earlier work [Flo92] on admissions control procedures using the Hoeffding bound. I have benefited from conversations with a large number of people over the years about measurement-based admissions control procedures, most principally to Dave Clark, Van Jacobson, Sugih Jamin, Frank Kelly, Vern Paxson, Scott Shenker, John Wroclawski, and Lixia Zhang. It should not, however, be assumed that

any of these people agree with any or all of the assumptions discussed in this paper.

References

[Ben62] G. Bennett. Probability inequalities for the sum of independent random variables. *American Statistical Association Journal*, pages 33–45, Mar. 1962.

[BSTW94] J. Beran, R. Sherman, M. Taqqu, and W. Willinger. Long-range dependence in variable-bit-rate video traffic. *IEEE Transactions on Communications*, 1994.

[BTW94] J.-C. Bolot, T. Turletti, and I. Wakeman. Scalable feedback control for multicast video distribution in the internet. *SIGCOMM Symposium on Communications Architectures and Protocols*, pages 58–67, Aug. 1994.

[Cox84] D. R. Cox. Long-range dependence: A review. *Statistics: An Appraisal, Proceedings 50th Anniversary Conference, Iowa State Statistical Library*, pages 55–74, 1984.

[CSZ92] D.D. Clark, S. Shenker, and L. Zhang. Supporting real-time applications in an integrated services packet network: Architecture and mechanism. *SIGCOMM Symposium on Communications Architectures and Protocols*, pages 14–26, 1992.

[DO93] N. Duffield and N. O' Connell. Large deviation and overflow probabilities for the general single-server queue, with applications. Technical report, 1993. Technical Report DIAS-STP-93-30, Dublin Institute for Advanced Studies.

[ENW96] A. Erramilli, O. Narayan, and W. Willinger. Experimental studies on the performance impacts of long range dependence. *IEEE/ACM Transactions on Networking*, 1996.

[Flo92] S. Floyd. One proposal for admission control (viewgraphs). Technical report, Jun. 1992. Presentation to the Internet End-to-End Research Group.

[GAN91] R. Guerin, 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, Sep. 1991.

[GKK95] R. Gibbens, F. Kelly, and P. 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.

[GW94] M. Garrett and W. Willinger. Analysis, modeling and generation of self-similar vbr video traffic. *SIGCOMM Symposium on Communications Architectures and Protocols*, 1994.

[HDLK95] C. Huang, M. Devetsikiotis, I. Lambardaris, and A. Kaye. Modeling and simulation of self-similar variable bit rate compressed video: A unified approach. *SIGCOMM Symposium on Communications Architectures and Protocols*, pages 114–125, 1995.

[Hoe63] W. Hoeffding. Probability inequalities for sums of bounded random variables. *American Statistical Association Journal*, 58:13–30, Mar. 1963.

[Hof87] M. Hofri. *Probabilistic Analysis of Algorithms*. Springer-Verlag, 1987.

[JDSZ95] S. Jamin, P. Danzig, S. Shenker, and L. Zhang. A measurement-based admission control algorithm for integrated services packet networks. *SIGCOMM Symposium on Communications Architectures and Protocols*, pages 2–13, 1995.

[LCH95] S.-Q. Li, S. Chong, and C.-L. Hwang. Link capacity allocation and network control by filtered input rate in high-speed networks. *IEEE/ACM Transactions on Networking*, pages 10–25, Feb. 1995.

[LH93a] S.-Q. Li and C.-L. Hwang. Queue response to input correlation functions: Continuous spectral analysis. *IEEE/ACM Transactions on Networking*, pages 678–692, Dec. 1993.

[LH93b] S.-Q. Li and C.-L. Hwang. Queue response to input correlation functions: Discrete spectral analysis. *IEEE/ACM Transactions on Networking*, pages 522–533, Oct. 1993.

[LTG95] N. Likhanov, B. Tsybakov, and N. Georganas. Analysis of an atm buffer with self-similar(fractal) input traffic. *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, pages 985–992, 1995.

[LTWW94] W. Leland, M. Taqqu, W. Willinger, and D. Wilson. On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, pages 1–15, Feb. 1994.

[Mah95] B. Mah. Measurements and observations of ip multicast traffic. Technical report, 1995.

[MJV96] Steven McCanne, V. Jacobson, and M. Vetterli. Receiver-driven layered multicast. *SIGCOMM Symposium on Communications Architectures and Protocols*, Aug. 1996. to appear.

[NKT91] R. Nagarajan, J. Kurose, and D. Towsley. Approximation techniques for computing packet loss in finite-buffered voice multiplexers. *IEEE Journal on Selected Areas in Communications*, 9(3):368–377, Apr. 1991.

[Nor94] I. Norros. A storage model with self-similar input. *Queueing Systems 16*, pages 387–396, 1994.

[PWC] C. Partridge, J. Wroclawski, and D. Clark. Integrated services working group. Technical report. URL http://www.ietf.cnri.reston.va.us/html.charters/intserv-charter.html.

[SCH+95] S. Shenker, D. Clark, C. Huitema, S. Deering, and D. Ferrari. Reservations or no reservations. *Infocom '95 Panel*, Apr. 1995. URL ftp://ftp.parc.xerox.com/pub/net-research/infocom95.html.

[She95] S. Shenker. Is admission control necessary (viewgraphs). *Infocom '95 Panel: Reservations or No Reservations*, Apr. 1995. URL ftp://ftp.parc.xerox.com/pub/net-research/infocom95.html.

[SP95] S. Shenker and C. Partridge. Specification of guaranteed quality of service. (Internet draft, work in progress), Dec. 1995.

[SPD95] S. Shenker, C. Partridge, and B. Davie. Specification of predictive quality of service. (Internet draft, work in progress), 1995.

[SPW95] S. Shenker, C. Partridge, and J. Wroclawski. Specification of the controlled delay quality of service. (Internet draft, work in progress), 1995.

[SS91] H. Saito and K. Shiomoto. Dynamic call admission control in atm networks. *IEEE Journal on Selected Areas in Communications*, 9(7):982–989, Sep. 1991.

[Wro96] J. Wroclawski. Specification of the controlled-load network element service. (Internet draft, work in progress), 1996.

[WTSW95] W. Willinger, M. Taqqu, R. Sherman, and D. Wilson. *SIGCOMM Symposium on Communications Architectures and Protocols*, pages 100–113, Aug. 1995.

[YKTH93] D. Yates, J. Kurose, D. Towsley, and M. Hluchyj. On per-session end-to-end delay distributions and the call admission problem for real-time applications with qos requirements. *SIGCOMM Symposium on Communications Architectures and Protocols*, pages 2–12, Sep. 1993.

A Questions of long-range dependence

In the appendix we summarize some of the research on the relationship between long-range dependent traffic and the resulting queue length distribution when that traffic is fed into a queue. We then give definitions for the *low-frequency*, *mid-frequency*, and *high-frequency* components of arriving traffic for a realtime class, and argue that it is the *mid-frequency* components that will be of concern for a measurement-based admissions control procedure for controlled-load traffic.

Traffic that is *self-similar* does not have a characteristic frequency or burst length, but looks similar over all time scales. The larger, more inclusive category of *long-range dependent* traffic exhibits correlations over a wide range of time scales. More precisely, traffic is *long-range dependent* if its autocorrelation function $r(k)$ is nonsummable (i.e., $\sum_k r(k) = \infty$) [Cox84]). Thus, while the technical definition of long-range dependence applies only to infinite time series, it does not require that the time series look the same over different time scales.

There are a number of analytical studies investigating the impact of long-range dependent traffic on queueing delay. [LTG95] and others show that while with traditional Markovian traffic models the queue length distribution decreases exponentially fast, for self-similar traffic the decrease is much slower. From [LTG95], the queue length distribution for self-similar traffic decreases algebraically with buffer size. [Nor94] has shown that for Fractional Brownian Motion, one form of self-similar traffic, the probability $P(V > x)$ that the queue-length V exceeds x is at least $exp[-cx^{2-2H}]$, for Hurst parameter H, and for c a function of both the Hurst parameter H and the utilization ρ. [DO93] has shown that this bound is asymptotically tight.

[ENW96] is an experimental study investigating the impact of long-range dependent traffic on queueing delay. [ENW96] looks at queueing delay from traces of long-range dependent Ethernet traffic, and compares the results to the delay curve predicted by GI/G/1 approximations based on two-moment characterizations of the input traffic. The delay-vs.-utilization curves of the two scenarios are dramatically different. For the LRD traffic, delay increases sharply at low utilization factors; large increases in the average delay can occur with very small increases in the utilization. Further experiments verify that the large delays are due to long-range dependence, and not to the short-term or one-step correlations in the traffic. These experiments show that correlations in the traffic over extremely long time scales have measurable and practical consequences in terms of queueing delay. The experimental studies are also based on data traffic, and do not necessarily lead to conclusions about aggregate real-time traffic.

The measurements of audio traffic and the possibility of long-range dependence in aggregate realtime traffic lead to a discussion of measurement-based admissions control procedures in the language of the frequency domain [LH93b, LH93a]. While our discussion of frequency components borrows from the approach of [LCH95], our definitions of the low, high, and mid-frequency components are specific to the context of aggregate controlled-load traffic. For this discussion, we assume a scheduling algorithm where the service received by a class is limited largely by the characteristics of the arriving traffic in the class, and not by the characteristics of possibly higher-priority classes.

In this context, we can define the *high-frequency* components of the aggregate arriving traffic for a class at a gateway as those frequencies that simply result in a certain acceptable queueing delay. That is, the "high-frequency" components refer to the short bursts in the arriving traffic that lead to a transient queue. For the scheduling algorithms discussed above, this frequency range is determined by the pre-defined "acceptable delay" for that class at that gateway. These frequencies in the aggregate arriving traffic are not a problem for a measurement-based admissions control procedure. The router's estimated average arrival rate can easily be designed to filter out these high-frequency components of the aggregate arrival rate.

We define the *low-frequency* components of the aggregate arriving traffic as those frequencies on a similar or greater time scale than the time scale of arrivals and departures of flows for that class. The time scale of arrivals and departures of flows will be quite different at different gateways that have different levels of multiplexing for the realtime traffic. Very slow changes in the aggregate bandwidth at these low frequencies can be corrected for largely by waiting for old connections to depart, and refusing to admit new connections, and again are not a problem for a measurement-based admissions control procedure.

The difficulties for the measurement-based admissions control come from the *mid-frequency* components of the aggregate arriving traffic, defined simply as that (possibly very large) frequency range that is neither high nor low. The *mid-frequency* components of the aggregate arriving traffic can translate to unacceptable local delays, to occasional packet drops, or to both. (Unacceptable delay can only result if the buffer is sufficiently large or the scheduling algorithm sufficiently unfriendly to allow transmitted packets with unacceptable delays.)

Note that if the average arrival rate were close to the available bandwidth for the class, then these *mid-frequency* components could be problematic even if the variance of the aggregate traffic was fairly small. That is, if the aggregate arrival rate is just a little higher than the available bandwidth, and maintains this "just a little higher" rate for a sufficiently long time, then the result will ultimately still be a buffer overflow and a packet drop rate determined by the excess arrival rate of the class. Thus, the frequencies in the *mid-frequency* class are those that can cause most difficulties for a measurement-based admissions control procedure, and could lead to either unacceptable delays or to packet drops if there is not sufficient lee-way in the bandwidth available to the class. We would conjecture that these mid-range frequencies are determined not by the jostling of packets as they progress in the network, but by the ON/OFF patterns of underlying sources with some long ON and OFF times.

One of the dangers of strong mid-frequency components in the aggregate arrival rate would be that the measurement-based admissions control would underestimate the actual average arrival rate, and would therefore admit more new flows than it should. One of the open questions for measurement-based admissions control procedures concerns the optimal algorithm for estimating an upper bound on the aggregate arrival rate.

B Procedures for tracing Mbone traffic

To use tcpdump to trace encapsulated IP data packets, for example as sent to multicast address 224.2.0.1:

```
tcpdump 'ip[9]==4 && ip[36]=224 &&
 ip[37]=2 && ip[38]=0 && ip[39]=1 &&
 greater 200' > filename
```

To trace unencapsulated UDP/IP data packets, for example as sent to multicast address 224.0.1.11:

```
tcpdump host 224.0.1.11 and
 greater 200 > filename
```

The "greater 200" phrase excludes small packets, which generally are session packets.

To make a data file that, for each connection, simply prints the connection number and the number of data packets sent in each 0.1-second time interval:

```
tcpdump -tt -n -r filename |
 awk -f vat1.awk > datafile
```

for vat1.awk as follows:

```
# vat1.awk
# output:  time, source-no., count
# optional parameters:
# start_time:  Start time in seconds
{
if ($6>=100) {
  split($1, time, ".")
  if (start_time == 0)
    start_time = time[1]
  seconds = time[1] - start_time
  tenths = substr(time[2], 1, 1)
  split($2, ip, ".")
  ipnum = ip[1]*1000000000
    + ip[2]*1000000 + ip[3]*1000
    + ip[4]
  if (num[ipnum] == 0) {
    ++count;
    num[ipnum] = count }
  if (last_sec[ipnum] != seconds ||
    last_ten[ipnum] != tenths) {
    if (sum[ipnum] > 0)
      print last_sec[ipnum]"."
      last_ten[ipnum],
      num[ipnum], sum[ipnum];
    sum[ipnum] = 1
    last_sec[ipnum] = seconds
    last_ten[ipnum] = tenths
  }
  else ++sum[ipnum]
} }
```

C Token-bucket policing - deriving the peak rate.

Given a token bucket of b bytes and a token fill rate of r bytes/sec., the source can send at most $b + rU$ bytes in U seconds. This gives an upper bound of $r + b/U$ bytes per second for the peak rate over a U-second interval. The interval U of interest is not necessarily a global parameter, but might be specific to a particular router.

This implies that, for an admissions control procedure based on peak rates of flows applying for admissions, a request for admissions is most likely to be accepted if the token bucket parameters are set with a small token bucket and the token fill rate set to the peak rate of the connection. The small token bucket is needed to accommodate small variations in packet delay that accumulate in the network. However, the admissions control procedure does not *require* a token bucket with a small bucket size; an upper bound on the peak rate over time interval U can be derived from any setting of the token bucket parameters.

D Computing the exact equivalent capacity

Given n ON/OFF connections, the exact equivalent capacity could be computed in the following straightforward but time-consuming manner: Let the j-th of the 2^n possibilities for the set of ON and OFF assignments to the n connections be represented by a set of n boolean variables $b_{1,j}$ to $b_{n,j}$, where connection i is ON if $b_{i,j} = 1$, and OFF otherwise. The j-th possibility has probability P_j, for

$$P_j = \prod_{i=1}^{n} (b_{i,j} o_i + (1 - b_{i,j})(1 - o_i)),$$

and an aggregate arrival rate A_j, for

$$A_j = \sum_{i=1}^{n} b_{i,j} o_i.$$

The arrival rate A_j is the sum of the peak rates of the ON connections. From these, the exact equivalent capacity is given by the following formula:

$$\mathrm{Max}\{A_j : \sum_{k=1}^{2^n}(A_k > A_j)P_k \leq \epsilon\},$$

This is the largest aggregate rate for which a total probability at most ϵ is assigned to the set of possibilities with larger aggregate rates.

Even for offline simulations this exact equivalent capacity is feasible to compute only for small values of N. For larger values of N, we estimated the exact equivalent capacity numerically. For each trial, we randomly assigned ON and OFF assignments to the n connections, according to the appropriate probabilities, and used repeated trials to construct an empirical distribution function for the aggregate arrival rate. From this empirical distribution function, we chose that aggregate rate for which a fraction ϵ of the trials resulted in larger aggregate rates.