Statistical Analysis of Delay Bound Violations at an Earliest Deadline First (EDF) Scheduler

Vijay Sivaraman

Department of Computer Science, 3820 Boelter Hall, UCLA, Los Angeles, CA 90095, U.S.A. (Email: vijay@cs.ucla.edu)

Fabio M. Chiussi

Bell Laboratories, Lucent Technologies, 101 Crawfords Corner Road, Room 4D-521, Holmdel, NJ 07733, U.S.A. (Email: fabio@bell-labs.com)

Abstract

In this paper, we develop an analytical framework for providing statistical delay guarantees in an Earliest Deadline First (EDF) scheduler which multiplexes traffic from multiple markovian sources with heterogeneous delay requirements. Our framework permits the computation of steady-state delay bound violation probabilities (i.e., the fraction of traffic that does not meet its delay bounds) at the EDF scheduler, and can be therefore used to characterize the schedulable region of EDF in a statistical setting. Our method employs results from the theory of large deviations and the theory of effective bandwidths, and demonstrates that effective bandwidths at both infinite and finite time scales have to be considered in the analysis of delays at the EDF scheduler (this is in contrast to the analysis of packet losses at a multiplexor, where only the effective bandwidth at infinite time scales is relevant). Our framework is of general use, and suitable to handle a broad range of markovian sources. As illustrating examples, we apply our method to two simple models, poisson and markovian on-off fluid traffic, and compare the analytical results with simulations, showing that the analysis is quite accurate. The framework presented in this paper can serve as the basis for the design of a Call Admission Control (CAC) mechanism which provides statistical guarantees on traffic transfer delays. Such a statistical CAC approach can offer dramatic advantages in network utilization over CAC frameworks based on deterministic delay bounds.

Key words: Scheduling, Quality-of-Service, Earliest Deadline First, Statistical Delay Bounds.

1 Introduction

Emerging broadband packet-switched networks are expected to provide a notion of Quality of Service (QoS) to real-time applications with widely different characteristics. Applications such as voice and video typically require QoS guarantees in terms of data transfer delays. The heterogeneity in the delay requirements of these applications, together with the need of having real-time traffic coexist with best-effort traffic without compromising network utilization, necessitates the use of packet scheduling mechanisms more sophisticated than First-In-First-Out (FIFO) service or simple priority schemes in the switches. Generalized Processor Sharing (GPS) and Earliest Deadline First (EDF) constitute the two most popular examples of such scheduling mechanisms.

In the last few years, an important body of research has focused on developing Call Admission Control (CAC) frameworks which employ these scheduling schemes to guarantee *deterministic* delay bounds to connections with constrained traffic (see for example [20] for a framework centered on GPS, and [13] for a treatment of EDF). However, these deterministic frameworks to provide guaranteed delay bounds are intrinsically conservative, since they have to account for the worst-case scenarios that can be encountered in the schedulers, even though such worst cases may occur with exceedingly low probabilities. As a result, the allocation of bandwidth to the connections is over-engineered, and the achieved network utilization may be far from optimal. Deterministic bounds are also "excessive", since real-time applications are typically resilient to infrequent violations in their delay bounds (i.e., are not unduly hindered if a small fraction, say 10^{-4}, of their packets get excessively delayed or dropped within the network). The known inefficiencies of the deterministic approach, together with the fact that most real-time applications actually require only "less than perfect" QoS guarantees, have recently generated a compelling need for *statistical* frameworks where the delay bounds are guaranteed probabilistically. Such statistical frameworks are expected to allow the links to operate at a much higher utilization, and still meet the QoS requirements of real-time traffic.

Schedulable regions and CAC schemes in the statistical setting have been studied extensively in literature, but mainly in the context of packet losses rather than delays. Beginning with the seminal work of Anick, Mitra and Sondhi in [1], an elegant theory of "effective bandwidths" has emerged in literature (see for example [3,6,15,16,23] and the references therein), and the related results allow efficient computation of aggregate packet losses when multiple connections are multiplexed into a shared buffer. These frameworks can also produce a distribution of the packet delays, but only under FIFO scheduling. Buffer occupancy processes have also been more recently studied in the presence of sophisticated schedulers like GPS [5,27] and priority scheduling [8,2].

In this paper, we are interested in formulating a statistical framework which employs a scheduling scheme more sophisticated than FIFO to guarantee heterogeneous delays to connections. A framework which provides statistical delays has been developed in [17] for priority scheduling; two frameworks which use GPS have been developed in [28,9], although the one in [28] assumes a rather specific source model, while the one in [9] is limited to two classes of traffic. In our case, we have chosen the EDF scheduling scheme, since it is known to provide the *optimal* delay performance [11,18] in the deterministic environment (as reviewed in greater detail in Section 2). We therefore consider a switch operating at a constant service rate and employing the EDF scheduling scheme, at which a multiplicity of markovian traffic sources are multiplexed. Given the markovian description of the sources and their required delay bounds at the multiplexor, we develop a framework which makes it possible to compute the steady-state probability of delay violations (i.e., fraction of *aggregate* traffic at the multiplexor which does not meet its delay requirement), which in turn allows to characterize the schedulable region of EDF in the statistical setting.

The theoretical basis for our statistical framework, drawing inspiration from the EDF schedulability constraints corresponding to the deterministic setting, relates the delay violation probability to the queue length in a hypothetical system derived from the real one. By making appropriate assumptions and employing results from the theory of large deviations and effective bandwidths, this probability is computed with relative ease. The framework is very general, and can handle arbitrary markovian descriptions of the traffic sources. As specific examples, we present simple expressions for poisson and fluid on-off markovian sources, and validate them with simulation results. To arrive at a framework that is simple and general, we have chosen to focus on computing the delay violation probability for the *aggregate* traffic at the switch, rather than on an individual connection or class basis; in fact, not only would the individual metrics be harder to evaluate due to the strong traffic interactions at the EDF scheduler, but they also are, at least to an extent, controllable by an appropriate choice of the packet discard policy, which is beyond the scope of this work. In this paper, we consider the case of an EDF scheduler in isolation (single node); the extension of our method to the multi-node case is possible, using reshaping at each node (in a way inspired by the corresponding approach for deterministic end-to-end delay developed in [26,13,4]) and is the topic of a forthcoming publication.

The rest of the paper is organized as follows. In Section 2, we provide a brief review of EDF scheduling and its CAC framework in the deterministic setting. In Section 3, after introducing the system model, we present the mathematical framework which allows the computation of the delay violation probabilities. In Section 4, we apply the results to two specific source models (namely poisson and fluid on-off markovian), and compare the analysis with corresponding results from simulations. We present concluding remarks and directions for future work in Section 5.

2 Background on EDF

We briefly review the basic concepts of EDF scheduling and some relevant results of the related deterministic analysis that has been developed in literature. The EDF scheduling discipline [19,10,22] works as follows: each connection i at the switch is associated with a *local* delay deadline d_i; then, an incoming packet of connection i arriving to the scheduler at time t is stamped with a deadline $t + d_i$, and packets in the scheduler are served by increasing order of their deadline.

In the deterministic setting, EDF is known to be the optimal scheduling policy at a single switch [11]. Optimality is defined in terms of the *schedulable region* associated with the scheduling policy. Given N connections with traffic envelopes $A_i(t)$ [1] $(i = 1, 2, \ldots, N)$ sharing an output link, and given a vector of delay bounds $\vec{d} = (d_1, d_2, \ldots d_N)$, where d_i is an upper bound on the scheduling delay that packets of connection i can tolerate, the schedulable region of a scheduling discipline π is defined as the set of all vectors $\vec{d}$ that are schedulable under π. The authors in [11,18] have shown that EDF has the largest schedulable region of all scheduling disciplines, and its Non-Preemptive version (NPEDF) has the largest schedulable region of all the non-preemptive policies. The schedulable region of the NPEDF policy consists of those vectors which satisfy the following constraints [11,18]:

$$\frac{L}{C} \le d_1 \tag{1}$$

$$L + \sum_{i=1}^{N} A_i(t - d_i) \le Ct, \qquad \frac{L}{C} \le t \le d_N \tag{2}$$

$$\sum_{i=1}^{N} A_i(t - d_i) \le Ct, \qquad t > d_N \tag{3}$$

[1] The traffic envelope $A_i(t)$ is such that the amount of traffic from connection i entering the network in any interval of length t is bounded by $A_i(t)$. A typical traffic envelope specification could be in terms of the *leaky bucket* parameters (p_i, σ_i, ρ_i) which denote the envelope $A_i(t) = \min\{p_i t, \sigma_i + \rho_i t\}$.

where $d_1 \leq d_2 \leq \ldots \leq d_N$, L is the packet size (if the packet size is variable, L is the maximum packet size), C the link rate, and $A_i(t) = 0$ for $t < 0$.

Given the traffic envelopes and the delay requirements of each connection, equations (1)-(3) can directly be used to devise a single-node CAC mechanism. However, this deterministic framework has to account for the worst-case scenarios that can be encountered by the scheduler, regardless of the likelihood of those scenarios, and is therefore overly conservative in admitting connections into the network. This motivates us to focus instead on a statistical framework where the delay bounds are "softer", i.e., guaranteed with a reasonably high probability. This allows the network to operate at a higher efficiency, while ensuring that the resulting degree of QoS can be quantified and managed at a desirable level.

3 Analytic Framework

We first present the system model and its associated assumptions, and then proceed to establish the mathematical method which allows the computation of the delay violation probabilities at the EDF scheduler.

3.1 System Model

Consider a single EDF scheduler which multiplexes connections[2] onto a transmission link operating at a constant rate C. The connections are categorized into J classes, with class j ($j = 1, \ldots, J$) comprising of k_j stochastically identical sources, each of which requires a delay bound d_j at the scheduler. Without loss of generality, we assume $d_1 \leq d_2 \leq \ldots \leq d_J$. We use the two-tuple (j, i) to refer to the i-th source belonging to the j-th class. Further, $A_{ji}[0, t]$ denotes the amount of work arriving from connection (j, i) in the interval $[0, t)$. We make the following assumptions:

Assumption 1 *Traffic is modeled as a fluid; hence packetization issues are ignored*[3] *.*

Assumption 2 *The connection (j, i) arrival traffic $A_{ji}[0, t]$ has stationary increments.*

[2] We use the words *connection* and *source* interchangeably throughput the paper
[3] However, we loosely use the term "packet" to refer to an infinitesimal quantity of the traffic.

Assumption 3 *Each connection generates traffic independent of all other connections.*

Assumption 4 *The fraction of traffic that does not meet its delay bound at the scheduler has a steady-state value; equivalently, the probability of delay bound violations P_{vio} exists and has a stationary value.*

The above assumptions are quite similar to the ones routinely adopted in literature for analyzing buffer occupancies. Our objective is to estimate, under these assumptions, the probability of delay bound violations at the scheduler. In our context, delay bound violations arise purely due to the scheduler's inability to meet every packet's deadline; the issue of packet losses due to lack of buffer space at the switch is orthogonal [4] and not addressed here. We thus make the additional assumption:

Assumption 5 *Buffer space at the switch is unlimited; hence there are no packet losses due to buffer overflow.*

The infinite buffer assumption not only makes the analysis tractable, but also provides an upper bound on the probability of delay bound violations in a finite buffered system; thus, it is a *conservative* approximation. Finally, we assume:

Assumption 6 *Packets are not discarded at the scheduler, even if their deadline has expired.*

Thus, packets (if any are present) are transmitted in order of their deadline, regardless of whether the deadline of the packet chosen for transmission has expired or not [5]. This simplifies the analysis and circumvents the issue of choosing an appropriate discard policy, which could range from something as simple as eliminating the packet chosen for transmission if its deadline has already expired, to a highly sophisticated one (and clearly non-trivial to be implemented in practice) where at each packet arrival instant it is determined if the deadlines of all packets in the system can be met, and if not, the packet chosen for discarding is one belonging to a connection with the least (weighted) number of deadline violations in the past (see [25,24] for an argument that such a discard policy is optimal). Including a specific discard scheme in the analysis would make our task more difficult (or in some cases even intractable), and the resulting framework less general. Again, this assumption of no discards is a *conservative* approximation which provides an upper bound on the fraction of packets violating their delay bounds.

[4] Indeed, if buffers are sufficiently large and packet losses are rare enough, the delay violation probability and the packet loss probability can be analyzed independently.
[5] Again, this is similar to the infinite buffer assumption commonly employed when analyzing packet losses.

3.2 Mathematical Model

We now present the framework which permits the computation of the delay violation probability, i.e., the fraction of traffic which does not meet its delay requirement. In order to formulate a framework that is conceptually simple and of general use, our objective is to compute this quantity for the *aggregate* traffic at the switch, not on an individual connection or class basis. The individual metrics are not only hard to evaluate due to the strong traffic interactions at the EDF scheduler, but are also controllable to an extent by an appropriate choice of the packet discard policy, which is beyond the scope of this work.

Under the assumptions stated above, we establish the following theorem which provides a basis for computing the stationary probability of delay violations:

Theorem 1 *The stationary probability P_{vio} of delay bound violations at a server employing EDF scheduling equals the probability that, at a random time t, the queue would be non-empty if, for every source (j, i), all arrivals in $[t - d_j, t)$ were to be discarded.*

PROOF. We show that there is a deadline violation at the EDF scheduler at time t if and only if the queue size $Q^H(t)$ at time t is non-zero in a hypothetical system H which discards all traffic arriving from every connection (j, i) in interval $[t - d_j, t)$. The proof of this is as follows.

We partition the traffic queued at the server into two (logical) queues Q^1 and Q^2 such that all traffic from connection (j, i) arriving in $[0, t-d_j)$ enters queue Q^1 while all traffic arriving in $[t - d_j, t)$ enters queue Q^2. Every packet, upon arrival, is assigned a *timestamp* representing the time by which the packet has to be served in order to satisfy its delay bound. By the definition of timestamp, all traffic queued in Q^1 has timestamp less than t, while no traffic queued in Q^2 has timestamp less than t. Since the EDF scheduler serves traffic in strict order of timestamp, traffic queued in Q^1 is given pure priority in service over traffic queued in Q^2. Thus, the evolution of Q^1 is as if the server is always available to serve it, i.e., as if Q^2 is completely ignored. This means that Q^1 behaves in the exact identical way as the queue length Q^H in the hypothetical system H which discards all arrivals from every source (j, i) in the interval $[t - d_j, t)$. Having established that $Q^H \equiv Q^1$, we also note that if Q^1 is non-empty at time t, then the traffic being served at time t has an expired deadline (since all timestamps in Q^1 are less than t), while if Q^1 is empty at time t, there is no delay violation at time t. This proves the result. $\triangle$

We make two observations about the above theorem. First, for $P_{\text{vio}} = 0$, the above theorem is compatible with the deterministic setting constraints of equations (1)-(3) (which collapse into the single equation $\forall t \ : \ \sum_{i=1}^{N} A_i(t - d_i) < Ct$ for the case of fluid traffic). Second, the above theorem holds *only* for EDF scheduling, since it is the only scheduling discipline which guarantees strict priority of Q^1 over Q^2. For a scheduler other than EDF, even though $Q^H(t)$ may be zero, the traffic served in the actual scheduler at time t could have an expired deadline, since Q^1 may not be zero because packets from Q^2 may have been served.

With the assumption that P_{vio} has a stationary value, the above theorem can be used, at least in principle, to compute the exact probability of delay violations. Let the fluid assumption be relaxed (i.e., have the traffic discretized into small units of constant size, to enable continuous-time discrete-space analysis). Now, let Q be the random variable denoting the stationary queue length of the system and $\vec{\pi}$ its distribution vector. Further, let $Q^H(t)$ be the random variable denoting the queue length of the hypothetical system H described in the proof of theorem 1 above, and let $\mathbf{P_j}(j = 1, \ldots, J-1)$ be the transition rate matrix which ignores all connections with a class number higher than j.

From theorem 1 above, the probability of delay violations is given by

$$P_{\text{vio}} = P[Q^H(t) > 0] = P[Q^H(t - d_1) > Cd_1] \tag{4}$$

(C is the link capacity), since the system H discards all arriving traffic in the interval $[t - d_1, t)$. Now let $\vec{\pi'}$ denote the distribution vector of $Q^H(t - d_1)$. The stationary probability of delay violations is thus

$$P_{\text{vio}} = \sum_{i > Cd_1} \vec{\pi'}(i) \tag{5}$$

Further, $\vec{\pi'}$ can be computed by

$$\vec{\pi'} = \vec{\pi} \exp \left[\int_0^{d_J - d_{J-1}} \mathbf{P_{J-1}} du + \int_0^{d_{J-1} - d_{J-2}} \mathbf{P_{J-2}} du + \ldots + \int_0^{d_2 - d_1} \mathbf{P_1} du \right] \tag{6}$$

In relation to the proof of theorem 1, the above expression for the distribution of $Q^H(t - d_1)$ corresponds to starting with the distribution of $Q^H(t - d_J)$ (given by the vector $\vec{\pi}$, since the real system and the hypothetical system H behave identically in $[0, t - d_J)$), and for successive $j = J-1, \ldots, 1$ doing a transient analysis to get the distribution of the queue length $Q^H(u)$ at $u = t - d_j$ by ignoring all connections with class number higher than j. This allows $\vec{\pi'}$, and hence P_{vio} to be determined.

Clearly, the above method is computationally complex, even for the simplest source models. For most source models of interest, computing the stationary queue length distribution vector $\vec{\pi}$ and the transition rate matrices $\mathbf{P_j}$ is a complex enough task; the need for computing the transients (in interval $[t - d_J, t - d_1]$) makes the task even more challenging. Therefore, in what follows, we develop approximations which are easily computed without much sacrifice in accuracy. Our first assumption to move in this direction is as follows:

Assumption 7 *The delay bounds $d_1, \ldots, d_J$ are "reasonably large", and the spread $(d_J - d_1)$ of the delay bounds is "reasonably small".*

We require the delay bounds to be "reasonably" large to allow large deviations results on queue length distribution tail probabilities to be applicable (this will become more evident as the discussion unfolds). The rationale behind requiring the delay bounds to be "reasonably" close together (say of about the same order of magnitude) is as follows: The queue length $Q^H(t - d_1)$ of the system H is described in terms of the workload process as the maximum of two terms:

$$Q^H(t - d_1) = \max\{Q^H(t - d_J) + A^H[t - d_J, t - d_1] - C(d_J - d_1), \\ \max_{0 \le T < d_J - d_1} (A^H[t - d_1 - T, t - d_1] - CT)\} \quad (7)$$

where $A^H[.,.]$ denotes total arrivals to the system H in the specified interval. At large values of $d_J - d_1$, the "initial" queue length $Q^H(t - d_J)$ has no impact on $Q^H(t - d_1)$, and hence the second term in the maximization on the right hand side dominates. However, as $d_J - d_1$ gets smaller, the first term within the maximization above becomes more significant (this is all the more so if the delay bounds $d_1, \ldots, d_J$ are large, in which case the contribution of the term $Q^H(t - d_J)$ is required to be larger for delay bound violations to occur anyway). Thus, for small values of $d_J - d_1$, it is reasonable to use the approximation $Q^H(t - d_1) \approx Q^H(t - d_J) + A^H[t - d_J, t - d_1] - C(d_J - d_1)$. We point out that this approximation is in general an optimistic one, i.e., could result in the queue lengths and hence the probability of delay violations being underestimated.

With assumption 7, therefore, the delay violation probability $P[Q^H(t - d_1) > Cd_1]$ is approximated by $P[Q^H(t - d_J) + A[t - d_J, t - d_1] > Cd_J]$. Noting that $Q^H(t - d_J)$ has the stationary distribution of the queue length Q, and that the traffic from each connection has stationary increments, we conclude that:

Theorem 2 *Under the approximations listed above, the stationary probability of delay violations is given by*

$$P_{\text{vio}} \approx P[Q + A > Cd_J] \quad (8)$$

where Q has the stationary queue length distribution, and $A = \sum_{j,i} A_{ji}[0, d_J - d_j]$.

Computing P_{vio} as given by equation (8) above requires knowledge of the distribution of the queue length Q. Determining the exact queue length distribution is in general complex, so we resort to large deviations estimates which have been developed in the literature for a wide variety of source models. For this purpose we find the following definition (see Kelly [16]) useful:

Definition 1 *The "effective bandwidth" or "logarithmic moment generating function" $\alpha_j(s,t)$ of source (j,i) is given by*

$$\alpha_j(s,t) = \frac{1}{st} \log E[e^{sA_{ji}[0,t]}] \qquad 0 \leq s,t \leq \infty \tag{9}$$

Here s and t denote the "space" and "time" scales respectively. Also, its long-term effective bandwidth is given by $\alpha_j(s) = \lim_{t\to\infty} \alpha_j(s,t)$.

Effective bandwidths have been computed in literature for a wide variety of source models, and large deviations estimates of the queue length tail probabilities have been studied [21]. For our purposes, we use the following form for the queue length tail probabilities [7] which is found to work well for markovian source models :

$$P\{Q > q\} \approx e^{-\delta q} \tag{10}$$

where δ is the queue length decay rate computed as follows:

$$\delta = max\{s \ : \ \sum_{j=1}^{J} k_j \alpha_j(s) \leq C\} \tag{11}$$

where C is the link capacity. It is important to point out that the queue length tail probability estimate of the form (10) is but one example of such an estimate, chosen here simply for computational convenience; our analysis does not preclude the choice of alternate approximations.

The probability measure in equation (8) can now be bounded in two ways:

- We can rewrite $P[Q + A > Cd_J]$ as $\sum_k P[A = k]P[Q > (Cd_J - k) \mid A = k]$, which, since Q and A are independent quantities, equals $\sum_k P[A = k]P[Q > (Cd_J - k)^+]$. Using (10) as an estimate of the queue lengths, this yields $\sum_k P[A = k]e^{-\delta(Cd_J - k)^+} = e^{-\delta Cd_J} \sum_k P[A = k]e^{\delta \min(k, Cd_J)}$. Since $\min(k, Cd_J) \leq k$, we have

$$P_{\text{vio}} \leq e^{-\delta Cd_J} E[e^{\delta A}] \tag{12}$$

- Applying Chernoff's theorem we have $P[Q + A > Cd_J] \leq \min_s \frac{E[e^{s(Q+A)}]}{e^{sCd_J}}$.

Again, Q and A are independent quantities. Moreover, from (10), $E[e^{sQ}] \approx \frac{\delta}{\delta - s}$ for $0 < s < \delta$. Thus

$$P_{\text{vio}} \leq \min_{0<s<\delta} \frac{\delta}{\delta - s} e^{-sCd_J} E[e^{sA}] \tag{13}$$

The right hand side expression in s is convex on $(0, \delta)$; the minimum exists and is unique, and hence can be determined numerically.

Computation of the bounds at various loads leads to the observation that the first bound is tighter in general at higher loads while the second dominates at low loads. Lastly, we note that

$$\begin{aligned} E[e^{sA}] &= \exp(\sum_{j=1}^{J} k_j \log(E[e^{sA_{ji}[0,d_J-d_j]}])) \\ &= \exp(\sum_{j=1}^{J} k_j (d_J - d_j)\alpha_j(s, d_J - d_j)) \end{aligned}$$

which leads to the following estimate of the probability of delay violations:

$$\begin{aligned} P_{\text{vio}} \approx \min \Bigg(& \exp\left[-\delta C d_J + \delta \sum_{j=1}^{J} k_j(d_J - d_j)\alpha_j(\delta, d_J - d_j)\right] , \\ & \min_{0<s<\delta} \frac{\delta}{\delta - s} \exp\left[-sCd_J + s\sum_{j=1}^{J} k_j(d_J - d_j)\alpha_j(s, d_J - d_j)\right]\Bigg) \end{aligned} \tag{14}$$

where δ is the queue length decay rate as given by equation (11). It should be pointed out from the above expression that effective bandwidths at *both* infinite and finite time scales become significant in the delay analysis of the EDF scheduler; this is in contrast to most of the existing analyses on losses at the multiplexor, in which only the effective bandwidths at infinite time scales play a role.

The framework developed above is very general and applicable to a broad range of markovian sources. In the next section, we apply this general model to two specific source models, the poisson and the fluid on-off markovian, and validate the analytic results against those obtained from simulations.

4 Analytic and Simulation Results

The poisson and fluid markovian on-off models are chosen purely because of their simplicity which leads to succinct explanation and simple expressions;

type	k	p(Mbps)	σ(Kbits)	ρ(Mbps)
type-0 (video conference)	varied	10	80	0.5
type-1 (stored video)	15	10	800	3
type-2 (audio)	200	0.064	10	0.064

Table 1
Connection parameters

our framework is suitable to handle more complex source models which are believed to adhere to real-world traffic more accurately (as, for example, the multistate markov model for video conferencing traffic developed in [7]), although, of course, such more complex models would introduce higher computational complexity.

The simulation setup is as follows: connections of three types, 0, 1, and 2, are multiplexed at an EDF server operating at a link of capacity 100Mbps and with unlimited buffer capacity. For simplicity, packet sizes are chosen to be fixed at 10Kbits, close to the Ethernet packet size (variable packet sizes would be treated similarly, since the analysis assumes fluid traffic, and hence variations in packet sizes are not significant). Table 1 shows the number of connections of each of the three types and their leaky bucket characterization in terms of the peak rate p, the bucket size σ, and the mean rate ρ. The values shown for type-0, type-1 and type-2 traffic in the table are consistent with the corresponding values for video conferencing, stored video, and audio connections reported in [12]. For both the poisson and fluid markovian on-off source models, the parameters are derived from the leaky bucket description of table 1 (this is described in more detail below). Type-1 and type-2 traffic are considered background traffic and kept constant. As the number of type-0 connections is varied, we plot the probability of delay violations as obtained from the analysis, along with the fraction of traffic violating its delay bounds as observed from the simulation of EDF scheduling. For comparison purposes, we also plot the fraction of traffic which violates its delay requirements under FIFO scheduling.

4.1 Poisson Sources

Let λ_j denote the traffic arrival rate corresponding to source (j, i). Its effective bandwidth is given by

$$\alpha_j(s, t) = \lambda_j(e^s - 1) \tag{15}$$

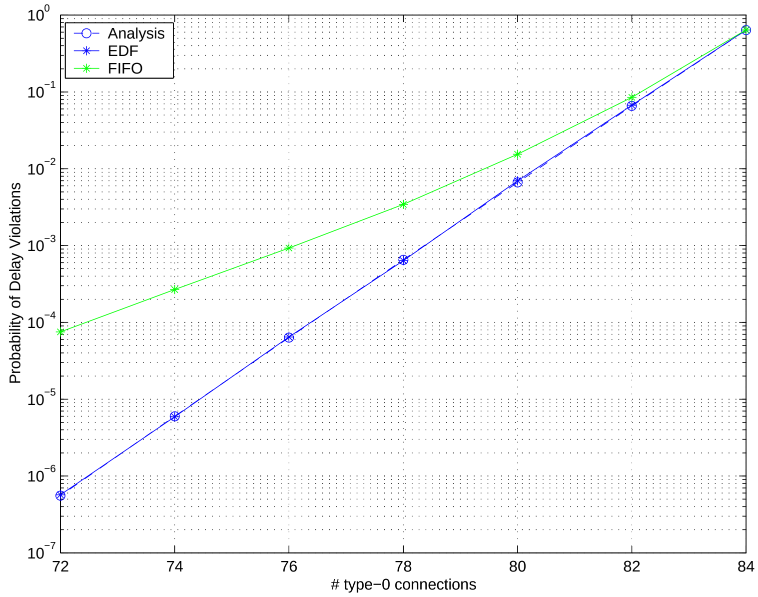

Fig. 1. Poisson Sources: Load varies from 93.8% to 99.8%

Of course, since a poisson source has independent increments, $\alpha_j(s,t)$ does not depend on t. The queue length decay rate δ is computed from (11) as

$$\delta = \max\{s \ : \ \sum_{j=1}^{J} k_j \lambda_j (e^s - 1) \leq s\} \tag{16}$$

Equations (15) and (16) can be directly used to compute P_{vio} for poisson traffic from equation (14).

In order to compare the analytic and simulation values, the poisson arrival rate λ_j for connection (j,i) is chosen to be identical to its average rate ρ_j from Table 1 above (the burstiness of the sources is therefore not captured by the poisson model). We have set the delay requirement of the video conferencing traffic at 10ms, of the stored video at 14ms, and of the audio at 6ms (since poisson traffic has a very low burstiness, we had to choose high utilizations and rather tight delay bounds to get delay bound violation probabilities high enough to be observable in the simulations).

Figure 1 plots on logarithmic scale the probability of delay violations as the number of type-0 connections is varied from 84 down to 72 (thus varying the utilization from 99.8% to 93.8%), as obtained from the analysis, and from simulations of both EDF and FIFO scheduling (the simulations were run sufficiently long; the confidence intervals are not plotted). As expected, EDF scheduling drastically reduces the fraction of traffic not meeting its delay requirement at the scheduler as compared to FIFO (the reduction can be as much as a couple of orders of magnitude), thus validating once more that EDF offers dramatic advantages over FIFO for supporting real-time traffic.

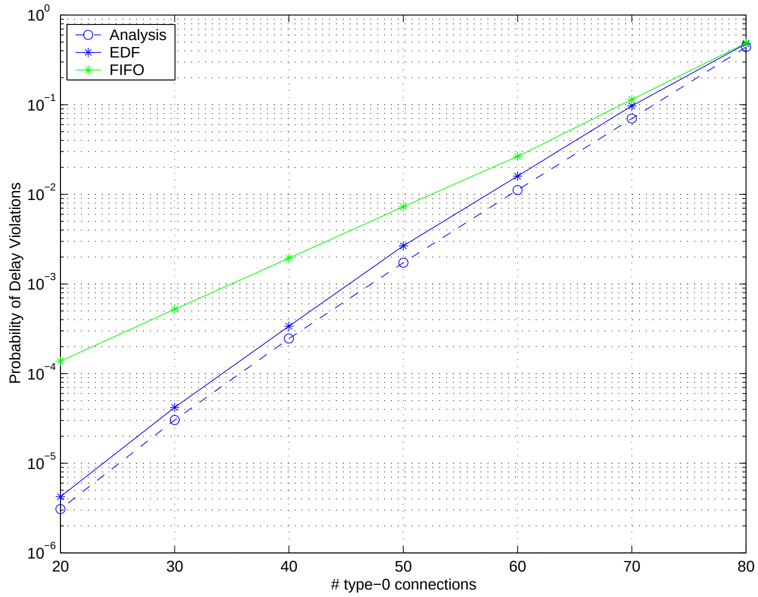

Fig. 2. On-Off Sources: Load varies from 67.8% to 97.8%

The plot shows that the analytic model matches the values obtained from simulations very closely (in fact the two curves are almost indistinguishable), thus demonstrating that the analytic framework for poisson traffic characterizes the schedulable region of EDF very accurately.

4.2 On-Off Markovian Fluid Sources

Let each source (j, i) be described by a two-state Markov chain. The transition rate from state 2 (the Off state) to state 1 (the On state) is λ_j, and the transition rate from state 1 to state 2 is μ_j. While the Markov chain is in state 1, workload (fluid) is produced at a constant rate h_j; while it is in state 2, no workload is produced. The effective bandwidth of such a source is given by [14–16]:

$$\alpha_j(s,t) = \frac{1}{t}\log\left\{\left(\frac{\lambda_j}{\lambda_j+\mu_j}, \frac{\mu_j}{\lambda_j+\mu_j}\right)\exp\begin{pmatrix} -\mu_j t + h_j s t & \mu_j t \\ \lambda_j t & -\lambda_j t \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\} \tag{17}$$

and $\alpha_j(s) = \lim_{t\to\infty} \alpha_j(s,t)$ is given by

$$\alpha_j(s) = \frac{1}{2}\left(h_j s - \mu_j - \lambda_j + \sqrt{(h_j s - \mu_j + \lambda_j)^2 + 4\lambda_j\mu_j}\right) \tag{18}$$

The above expressions for $\alpha_j(s,t)$ and $\alpha_j(s)$ can be used directly in equations (11) and (14) to determine the probability of delay violations at the EDF scheduler for the given markovian on-off fluid sources.

For comparison of analytic and simulation values, the markov parameters of source (j,i) are derived from the leaky bucket parameters of table 1 as follows: the on-rate h_j is set equal to the peak rate p_j, the mean holding time $1/\mu_j$ in the on-state is chosen as $\frac{1}{\log 5}\frac{\sigma_j}{p_j-\rho_j}$ (this corresponds to fixing the probability of getting a burst larger than the maximum burst admissible by the leaky bucket at 20%), and the mean holding time $1/\lambda_j$ in the off-state is chosen as $\frac{1}{\log 5}\frac{\sigma_j}{\rho_j}$, thus ensuring that the average rate equals to ρ_j. The parameters are chosen in an attempt to model as closely as possible the periodic worst-case on-off behaviour of a source consistent with its leaky bucket description. Unlike the poisson model, the burstiness of the sources is captured by the on-off markovian model. In this case, we have set the delay requirement of the video conferencing traffic at 40ms, of stored video at 60ms, and of audio at 20ms (these are reasonably realistic numbers consistent with [12]).

Figure 2 plots on logarithmic scale the probability of delay violations as the number of type-0 connections is varied from 80 down to 20 (thus varying the utilization from 97.8% to 67.8%) as obtained from the analysis, and from simulations of both EDF and FIFO scheduling. As in the poisson case, the fraction of traffic traffic violating its delay bounds under EDF is lower by up to a couple of orders of magnitude as compared to FIFO. Also in this case, the analysis is found to match the EDF simulation values quite well. The small discrepancies between the two can be attributed to the rather wide spread in the delay requirements of the connections (thereby not being in strict accordance with assumption 7) which results in the probability of delay violations being underestimated. Nevertheless, the match is quite close, and the analytical model hence provides a reasonably good characterization of the schedulable region of EDF.

To give a feeling of the dramatic benefits that a statistical framework can offer over a deterministic framework in terms of network utilizations, it is worth noting that the deterministic framework for EDF scheduling could not accommodate even a *single* type-0 connection under this traffic scenario.

5 Conclusions and Future Work

In this paper, we have developed an analytical framework for evaluating the probability of delay bound violations at an EDF scheduler which multiplexes traffic from markovian sources with reasonably heterogeneous delay bounds. The delay violation probability is expressed in terms of the effective bandwidths of the sources at both infinite and finite time scales, and is easily computed given that effective bandwidths for a wide variety of markovian sources are available in the literature. As illustrative examples, we have shown results obtained for poisson and markovian on-off fluid sources, which match quite closely the values obtained from simulations. In this paper, we have focused on the single node in isolation; the extension to the multi-node case, using traffic reshaping at each node (again inspired by the corresponding approach for deterministic end-to-end delays, see [26,13,4]) is the topic of a forthcoming publication. Such a statistical framework is potentially of great significance in the design of CAC schemes for networks which provide QoS in the form of statistical guarantees on data transfer delays, and can offer very substantial advantages in terms of network utilization over CAC schemes based on deterministic delay guarantees.

One direction for future work involves assumption 7, which limits the spread in the delay requirements of the various connections; a more general framework which can handle greater heterogeneity in the delay requirements is highly desirable.

References

[1] D. Anick, D. Mitra, M. M. Sondhi, Stochastic Theory of a Data-Handling System with Multiple Sources, *Bell Systems Technical Journal*, 61(8):1871–1894, 1982.

[2] A. W. Berger, W. Whitt, Effective Bandwidths with Priorities, *IEEE/ACM Transactions on Networking*, 6(4):447–460, August 1998.

[3] C.-S. Chang, J. A. Thomas, Effective Bandwidth in High-Speed Digital Networks, *IEEE Journal on Selected Areas in Communications*, 13(6):1091–1100, August 1995.

[4] F. M. Chiussi, V. Sivaraman, Achieving High Utilization in Guaranteed Services Networks using Early-Deadline-First Scheduling, In *International Workshop on Quality of Service*, pages 209–217, Napa, California, May 1998.

[5] G. de Veciana, G. Kesidis, Bandwidth Allocation for Multiple Qualities of Service using Generalized Processor Sharing, *IEEE Transactions on Information Theory*, 42:268–272, 1996.

[6] G. de Veciana, G. Kesidis, J. Walrand, Resource Management in Wide-area ATM Networks using Effective Bandwidths, *IEEE Journal on Selected Areas in Communications*, 13(6):1081–1090, August 1995.

[7] A. Elwalid, D. Heyman, T. V. Lakshman, D. Mitra, A. Weiss, Fundamental Bounds and Approximations for ATM Multiplexors with Applications to Video Teleconferencing, *IEEE Journal on Selected Areas in Communications*, 13(6):1004–1016, August 1995.

[8] A. I. Elwalid, D. Mitra, Analysis, Approximations and Admission Control of a Multiservice Multiplexing System with Priorities, In *Proceedings of INFOCOM '95*, pages 463–472, Boston, MA, April 1995.

[9] A. I. Elwalid, D. Mitra, Design of Generalized Processor Sharing Schedulers which Statistically Multiplex Heterogeneous QoS Classes, In *Proceedings of INFOCOM '99*, New York, NY, March 1999.

[10] D. Ferrari, D. Verma A Scheme for Real-time Channel Establishment in Wide-area Networks, *IEEE Journal on Selected Areas in Communications*, 8(3):368–379, April 1990.

[11] L. Georgiadis, R. Guérin, A. Parekh, Optimal Multiplexing on a Single Link: Delay and Buffer Requirements, *IEEE Transactions on Information Theory*, 43(5):1518–1535, September 1997.

[12] L. Georgiadis, R. Guérin, V. Peris, R. Rajan, Efficient Support of Delay and Rate Guarantees in an Internet, In *Proceedings of ACM SIGCOMM*, Stanford University, CA, August 1996.

[13] L. Georgiadis, R. Guérin, V. Peris, K. Sivarajan, Efficient Network QoS Provisioning Based on per node Traffic Shaping, *IEEE/ACM Transactions on Networking*, 4(4):482–501, August 1996.

[14] R. J. Gibbens, P. J. Hunt, Effective Bandwidths for the Multi-type uas Channel, *Queueing Systems*, 9:17–27, 1991.

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

[16] F. Kelly, Notes on Effective Bandwidth, *Stochastic Networks: Theory and Applications*, 4:141–168, 1996.

[17] E. D. Knightly, Enforceable Quality of Service Guarantees for Bursty Traffic Streams, In *Proceedings of INFOCOM '98*, pages 635–642, San Francisco, CA, March 1998.

[18] J. Liebeherr, D. Wrege, D. Ferrari, Exact Admission Control for Networks with a Bounded Delay Service, *IEEE/ACM Transactions on Networking*, 4(6):885–901, December 1996.

[19] C. L. Liu, J. W. Wayland, Scheduling Algorithms for Multiprogramming in a Hard Real Time Environment, *Journal of ACM*, 20(1):46–61, January 1973.

[20] A. Parekh, R. Gallager, A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Multiple Node Case, *IEEE/ACM Transactions on Networking*, 2(2):137–150, April 1994.

[21] A. Shwartz, A. Weiss, *Large Deviations for Performance Analysis*, Chapman & Hall, 1995.

[22] D. Verma, H. Zhang, D. Ferrari, Guaranteeing Delay Jitter Bounds in Packet Switching Networks, In *Proceedings of TRICOMM*, pages 35–46, Chapel Hill, NC, April 1991.

[23] W. Whitt, Tail Probabilities with Statistical Multiplexing and Effective Bandwidths in Multi-class Queues, *Telecommunication Systems*, pages 71–107, 1993.

[24] Y. Xie, T. Yang, Cell Discarding Policies Supporting Multiple Delay and Loss Requirements in ATM Networks, In *Proceedings of GLOBECOM '97*, pages 1075–1080, Phoenix, AZ, September 1997.

[25] Y. Xie, T. Yang, Efficient Admission Control for EDF Scheduler with Statistical QoS Guarantee, In *Proceedings of ICCCN '97*, pages 242–247, Los Alamitos, CA, September 1997.

[26] H. Zhang, D. Ferrari, Rate-controlled Service Disciplines, *Journal of High Speed Networks*, 3(4):389–412, 1994.

[27] Z.-L. Zhang, Large Deviations and the General Processor Sharing Scheduling for a Multiple-Queue System, *Queueing Systems*, 28(4):349–376, 1998.

[28] Z.-L. Zhang, D. Towsley, J. Kurose, Statistical Analysis of the Generalized Processor Sharing Scheduling Discipline, *IEEE Journal on Selected Areas in Communications*, 13(6):1071–1080, August 1995.

Vitae

Vijay Sivaraman received his B. Tech in Comp. Sc. and Eng. from the Indian Institute of Technology at Delhi, India, in 1994, and his M.S. in Comp. Sc. from NCSU in 1996. He is currently pursuing his Ph.D. at UCLA; his research interests include packet scheduling issues and QoS support in High-Speed Networks.

Fabio M. Chiussi received the Ph. D. in Electrical Engineering from Stanford University in 1993. Since 1993, he has been with Bell Laboratories, Lucent Technologies. He is currently a distinguished member of technical staff in the High-Speed Networks Research Department, where he has been conducting fundamental research in the area of broadband networks, and has been leading the architectural design and development of three generations of silicon for ATM and IP switches. Dr. Chiussi has written more than 60 technical papers and holds 6 patents, with 20 more pending. He was named the 1997 Eta Kappa Nu Outstanding Young Electrical Engineer. His current research interests include ATM switching, IP switching and routing, traffic management and scheduling, congestion control, and VLSI design.