Queueing at large resources driven by long-tailed $M/G/\infty$-modulated processes

N. G. Duffield
AT&T Laboratories
Room 2C-323, 600 Mountain Avenue, Murray Hill, NJ 07974, USA
duffield@research.att.com

December 30, 1996

Abstract

We analyze the queue at a buffer with input comprising sessions whose arrival is Poissonian, whose duration is long-tailed, and for which individual session detail is modeled as a stochastic fluid process. We obtain a large deviation result for the buffer occupation in an asymptotic regime in which the arrival rate nr, service rate sr, and buffer level nb are scaled to infinity with a parameter n. This can be used to approximate resources which multiplex many sources, each of which only uses a small proportion of the whole capacity, albeit for long-tailed durations.

We show that the probability of overflow in such systems is exponentially small in n, although the decay in b is slower, reflecting the long tailed session durations. The requirements on the session detail process are, roughly speaking, that it self-averages faster than the cumulative session duration. This does not preclude the possibility that the session detail itself has long-range dependent behavior, such as fractional Brownian motion, or another long-tailed $M/G/\infty$ process.

We show how the method can be used to determine the multiplexing gain available under the constraint of small delays (and hence short buffers) for multiplexers of large aggregates, and to compare the differential performance impact of increased buffering as opposed to load reduction.

1 Introduction.

Description of the Model. Consider a buffer whose input is a fluid process, possibly stochastic, driven by a discrete-time $M/G/\infty$ process, with G having a long-tailed distribution. By this we mean a resource at which sessions arrive according to a discrete time Poisson process of intensity r, i.e. the number M_t of sessions arriving at time $t \in \mathbf{Z}$ has a Poisson distribution with mean r and is independent of the $M_{t'}$ for $t \neq t'$; each session remains at the resource for an independent duration with distribution G. At the end of its duration the session stops, and can be considered to leave the resource. The stochastic fluid process enters as follows. Immediately after arriving, and for the duration of its stay at the resource, each session empties fluid into a common infinite buffer according to (an excerpt of) an independent copy of a stationary random process Z. The buffer is drained at a rate s per unit time.

In this paper we shall use large deviation methods to investigate the stationary distribution of the amount of fluid in the buffer in an asymptotic regime that the amount of resources used by each session, as a proportion of the whole, goes to zero. We do this by scaling the parameters of the system by an increasingly large integer n: setting the Poissonian arrival rate of sessions to be rn, the service rate of the fluid buffer to be sn and calculating the stationary probability that the amount of fluid in the buffer exceeds bn. n can be thought of as an aggregation or intensity level for arrivals, and a pooling level for resources. The principal application we have in mind in the statistical multiplexing on a resource which is a shared communication link of the aggregation over many sources of fluid sessions which have moderate bandwidth requirements, but long-tailed durations. The $M/G/\infty$ process models the number of active sessions, the stochastic fluid process the pattern of arrivals within each session. It should be noted that the theory also includes the special case that the fluid arrival rate of an individual session being constant over the duration of the session.

The distribution G of the session duration will have the following property. Let G_e be the stationary excess distribution corresponding to G, i.e. G_e is the distribution with mass $G^c(r)/\gamma$ at $r \in \mathbf{N}$, where γ is the mean from G and $G^c = 1 - G$ is the complementary cdf to G. Then we will require that $v(t) := -\log G_e^c(t)$ is $o(t)$ as $t \to \infty$. This class of G includes, for example, Pareto distributions with power $\alpha > 1$, in which case $v(t)$ is proportional to $\log t$. A session arriving at time t, and of duration d empties quantities of fluid $\{z_t, z_{t+1}, \ldots, z_{t+d-1}\}$ into the buffer at times $\{t, t+1, \ldots, t+d-1\}$, where the $\{z_t, z_{t+1}, \ldots, z_{t+d-1}\}$ have the distribution of some stationary process $(Z_{t'} : t' \in \mathbf{Z})$, from $t' = 1$ up to the random time d. Z may be simply be constant for all time, or a Markov process, or even long-range dependent, subject to certain conditions governing its decay of correlation. We call Z, or an instance of it, the session detail process.

Summary of the Results and Some Implications. Let Q^n be the fluid buffer level in the stationary regime. We summarize the main theoretical results of the paper concerning the tails of the distribution of Q^n.

Large aggregation asymptotics. Under hypotheses stated in Theorem 3 below,

$$\lim_{n\to\infty} n^{-1} \log \mathbf{P}[Q^n > nb] = -I(b), \tag{1.1}$$

for some *shape function* I depending on the detailed statistical properties of the arrival process.

Asymptotic behavior of the shape function. Under hypotheses on the session detail process stated in Theorem 4 below, then for some extension h of v (from $\mathbf{N}$ to $\mathbf{R}_+$)

$$\lim_{t\to\infty} \frac{I(b)}{h(b)} = \delta \tag{1.2}$$

and δ is a constant which depends on the session detail process only through its long term average.

Lower bound on the shape function. There exists a real η such that for all b

$$I(b) \le \delta h(b) + \eta. \tag{1.3}$$

Combining (1.1) and (1.2) we make the successive approximations

$$\mathbf{P}[Q^n > nb] \approx e^{-nI(b)} \quad \text{for large } n \tag{1.4}$$

$$\approx e^{-n\delta v(b)} \quad \text{for large } b. \tag{1.5}$$

The first approximation in (1.4) says the tail probabilities decay exponentially fast in n. This indicates *economies of scale*: that good multiplexing gain is available for many sessions on a pooled resource. The reason this occurs is that our hypotheses on the session detail process imply that the distribution of the amount of fluid generated over any fixed time interval is not long tailed. Consequently, if the buffer level exceeds nb, it typically requires a number of sessions proportional to n to be active. The probability that this happens is exponentially small in n. Such behavior has been recently established for classes of queueing models: see [6, 11, 15, 34].

The second approximation uses only the mean of the session detail process (though δ) and the asymptotic properties of the session duration (through v). For example, when G is Pareto of power $(\alpha - 1)$, $v(t) \sim (\alpha - 1)\log t$ and so the second approximation in (1.4) becomes

$$\mathbf{P}[Q^n \geq nb] \approx b^{-(\alpha-1)\delta n}. \tag{1.6}$$

For G Pareto, it is shown Theorem 4 that

$$\delta = s\bar{z}^{-1} - r\gamma. \tag{1.7}$$

The tail probability decays only comparatively slowly in b. This slow decay is a direct consequence of that of G^c since large buffer occupation occur as a result of sessions being active over the (large) duration required to achieve this level. But our hypotheses on the session detail process will imply that its time-averages settle down at a faster rate than those of the cumulative session duration. (We note that the model class does not exclude the possibility that the session detail does itself possess dependence which is long-range, albeit not so long as that of the session duration process). Over the time scales of the long bursts required to achieve large buffer occupation nb, the time-averaged activity of the session detail process has already settled down to its expectation. This provides a heuristic explanation for the fact that the session detail process only enters through its mean into the the large b approximation.

Eq. (1.3) gives a (conservative) upper bound on large n asymptotic tail probabilities, while retaining a form which is, like (1.5), comparatively simple compared with the full shape function approximation (1.4).

Relation to Other Results. In contrast to the logarithmic large aggregation tail-asymptotics considered in the present paper, Parulekar and Makowski [29, 30, 31] have investigated logarithmic large buffer tail-asymptotics of the queue for the $M/G/\infty$ buffer process, while Jelenkovic and Lazar [22] have obtained non-logarithmic large buffer tail-asymptotics for the queue length at the start of the arrival activity periods. In a result closer in emphasis to the present paper, Heath, Resnick and Samorodnitsky [18] have recently found that time to overflow in buffers of superposed long-tailed on-off sources benefits from the pooling of resources. Tail probabilities for fluid queue lengths for these models has been investigated by Choudhury and Whitt [9].

The $M/G/\cdot$ arrival process has itself appeared as a limit as $L \to \infty$ of an L-fold superposition of on-off processes (as modeled by alternating renewal processes): see Likhanov, Tsybakov and Georganas [25], and earlier Cohen [10] who considered exponentially distributed off times (with

more recent generalizations by Boxma [7]). In these cases, the on periods have distribution G, while the off periods have distribution function $D_L(\cdot) = D(\cdot/L)$ for some distribution D. In the limit, D only enters into the $M/G/\cdot$ arrivals through its mean, which determines the mean arrival rate. However, the service discipline and considered in [25] was $M/G/1$: the resource could be consumed by a single session. Anantharam [3] has considered a multiple-server generalization of this model, and has discussed the different manners in which queues arise for long and short tailed service times [2].

$M/G/1$ would also describe a variation from our model in which the fluid supplied by a session, instead of being spread out over the duration of a session, is instead delivered instantaneously at the start of it. Suppose, for example, that G has Pareto distribution with index $\alpha > 1$. Then Q^n has the same distribution as Q^1 and so it follows from Pakes [28] that $\mathbf{P}[Q^n > nb] \propto (nb)^{\alpha-1}$: decay of the tail probabilities is only algebraic in n.

Finally, we comment that potential limitation of the model here is that we consider an infinite buffer rather than a finite one. However, this results in a conservative approximation as far as estimating overflow probabilities are concerned: it is not difficult to show that the probability of the queue exceeding a level b in an infinite buffer is an upper bound for the probability of overflow from a finite buffer of size b. Another limitation of the model is that the sessions are never blocked. In some circumstances a better model might be $M(rn)/G/nk/nk$ so that new sessions are blocked if nk sessions are active. However, numerical calculations on the model have demonstrated that loss ratios of occupying sessions (as opposed to overflows from the fluid buffer) are increasingly close to those for $M(nr)/G/\infty$ as n becomes large; see Abate and Whitt [1].

Outline of the Paper. The main limit theorem behind (1.1) is stated and proved in Section 2, along with some technical background. The asymptotics and bounds for the shape function (1.2) and (1.3) are established in Section 3. In Section 4 we give some examples of processes satisfying the conditions of the theorems, including examples where both the session duration and session detail process possess long range dependence. In Section 5 we use simulations to investigate the nature of the convergence as $n \to \infty$ in (1.1). We also use (1.1) to examine determine the multiplexing gain available for large aggregations which can be buffered only within a delay constraint. Typically this will require the buffer allocation per unit activity to be small: use of the large b asymptotic (1.5) will be not be appropriate if the typical time scale to overflow is not over bursts of sessions. We illustrate these points some numerical examples. We show how to compare the relative performance impact of increased buffering under the delay constraint as opposed to reducing the offered load.

2 Queue tail probabilities at large resources.

2.1 Mathematical Background.

The work of this paper rests on being able to combine three stands. The first is a characterization of the asymptotics properties of the cumulative session duration process, namely the $M/G/\infty$ buffer process, which has been established recently by Parulekar and Makowski [29, 30]. The second is an asymptotic characterization of the cumulative session detail process, which can be regarded as modulated by the above $M/G/\infty$ process. The third strand is a recent result by Duffield [15] on the tail asymptotics as $n \to \infty$ of queues fed by superpositions of n sources, possibly long range-dependent. For us n will be the scaling factor on the arrival and service rates. This yields

the asymptotic form (1.1). The first two strands are need to establish technical conditions for the third, which in turn yields (1.2).

In all three strands, the key object of interest is the transient log moment generating function (MGF)

$$m_t(\theta) = \frac{1}{w(t)} \log \mathbf{E}[e^{\theta w(t) C_t/t}], \tag{2.1}$$

where w is a *scaling function*, i.e., a positive function increasing to infinity, and where C_t is some accumulated activity over the times $\{1, 2, \ldots, \}$. Depending on the context, this activity may be the total duration of all sessions present over the interval of length t; or the total fluid supplied during those sessions; or the fluid supplied during an individual session of duration t.

The role of the moment generating function m_t and scaling function w is as follows. Recall from [13] that the pair $(C_t/t, w(t))$ is said to satisfy a *large deviation principle* (LDP) with rate function J if for all Borel subsets B of $\mathbf{R}_+$

$$-\inf_{x \in B^0} J(x) \leq \liminf_{t \to \infty} \frac{1}{w(t)} \log \mathbf{P}[C_t/t \in B] \leq \limsup_{t \to \infty} \frac{1}{w(t)} \log \mathbf{P}[C_t/t \in B] \leq -\inf_{x \in \bar{B}} J(x), \tag{2.2}$$

where B^o is the interior and $\bar{B}$ the closure of B. Sufficient conditions for the LDP to hold specified in the Gärtner-Ellis Theorem (2.3.6 in [13]). The upper bound holds if $m(\theta) = \lim_{t \to \infty} m_t(\theta)$ exists as an extended real number for all θ and is finite in a neighborhood of 0. In this case J is the Legendre transform m^* of m, namely

$$m^*(x) = \sup_{\theta \in \mathbf{R}} (\theta x - m(x)). \tag{2.3}$$

The lower bound holds under additional hypotheses. In all cases we consider, C will have stationary increments and the mean activity $\bar{c} = \mathbf{E}[C_1]$ will be finite. If (2.2) holds, then $J(\bar{c}) = 0$, while under mild supplementary conditions J is strictly convex so that $J(k) > 0$ for $k \neq \bar{k}$. Thus the meaning of (2.2) is that the mass of the distribution of C_t/t concentrates to $\bar{k}$ as $t \to \infty$, at rate which is exponentially fast on a scale determined by w. Actually, for what follows, we will not require convergence of these moment generating function for all θ. However, we do require that the scaling function $v(t)$ for the cumulative session duration diverges slower than the scaling function $u(t)$ for the session detail process in the sense that $v(t) = o(u(t))$ as $t \to \infty$. In other words, we require a separation of time-scales between the two levels of description of the arrival process. We mention that the existence of m, with appropriate side conditions, be used to derive asymptotic results for the asymptotic queue length distribution for fixed n as $b \to \infty$. (See, for example, [16, 21, 23]). However, the conditions are violated for the $M/G/\infty$ buffer process for the class of G that we consider; see [14] and [29].

2.2 Asymptotics of the cumulative session duration.

Let M_t be the number of sessions arriving at time t. The M_t are independent Poissonian variables of mean nr. We adopt the convention that a session arriving at time t and duration d is present at the resource and emptying fluid into the buffer according to the session detail process at times $\{t, t+1, \ldots, t+d-1\}$. Thus if N_t is the number of sessions active at time t, including the session of number M_t which at that time, we have

$$N_{t+1} = N_t + M_{t+1} - D_t, \tag{2.4}$$

where D_t is the number of sessions whose last active slot is at time t. The duration of the arriving sessions are independent random variables, each with the distribution of an integer-valued random variable σ. The stationary distribution of the remaining duration of an active session, the stationary excess distribution G_e, is that of a random variable $\hat{\sigma}$ for which,

$$\mathbf{P}[\hat{\sigma} = r] = \frac{\mathbf{P}[\sigma \geq r]}{\mathbf{E}[\sigma]}. \tag{2.5}$$

Finally, it can be shown in this case that in the stationary version of the process N, the N_t have Poisson distribution with mean equal to the session arrival rate times $\mathbf{E}[\sigma]$; in this case, $nr\mathbf{E}[\sigma]$. (see [12]).

The cumulative session duration C_t^n in $\{1, 2, \ldots, t\}$ of all sessions present during that interval is the random variable

$$C_t^n = \sum_{t'=1}^{t} N_{t'}. \tag{2.6}$$

Let us re-express C_t^n in terms of the random variables $N_0, \sigma, \hat{\sigma}$. Each session present at $t = 0$ remains present beyond and including $t = 1$ for a duration of distribution $\hat{\sigma} - 1$. Thus the total duration between $t = 1$ and an arbitrary t is $(\hat{\sigma} - 1) \wedge t$, where $\wedge$ denotes the minimum. Similarly, sessions arriving at time $t - s + 1$ with $s \in \{1, 2, , \ldots, t\}$ are present for a duration $\sigma \wedge s$. Thus C_t^n has the same distribution as

$$\sum_{i=1}^{N_0} (\hat{\sigma}^{(i)} - 1) \wedge t + \sum_{t'=1}^{t} \sum_{i=1}^{M_{t-t'+1}} \sigma^{(i,t')} \wedge t', \tag{2.7}$$

where $\hat{\sigma}^{(i)}$ and $\sigma^{(i,t')}$ are independent copies of $\hat{\sigma}$ and σ respectively. Define the log moment generating function

$$\omega_t^n(\theta) = \frac{1}{nv(t)} \log \mathbf{E}[e^{\theta C_t^n v(t)/t}]. \tag{2.8}$$

Since the N_0 and M_s have independent Poisson distributions with mean nr and $nr\mathbf{E}[\sigma]$ respectively we have for any scaling function v

$$\omega_t^n(\theta) = \omega_t(\theta) := \frac{r\mathbf{E}[\sigma]}{v(t)} \left(\mathbf{E}[e^{\theta((\hat{\sigma}-1)\wedge t)v(t)/t}] - 1 \right) + \frac{r}{v(t)} \sum_{t'=1}^{t} \left(\mathbf{E}[e^{\theta(\sigma \wedge t')v(t)/t}] - 1 \right), \tag{2.9}$$

independent of n. Parulekar and Makowski have proved the following convergence result for $\omega_t(\theta)$ as $t \to \infty$. They make the key observation that convergence of the moment generating function obtains if the scale $v(t)$ is chosen according to the distribution of σ.

Theorem 1 *Set $v(t) = -\log \mathbf{P}[\hat{\sigma} > t] = -\log G_e^c(t)$ and suppose that $v(t) = o(t)$ with $v(t)/t$ eventually decreasing. Assume that there exists a mapping $\Gamma : \mathbf{N} \to \mathbf{N}$ such that (i) $\Gamma(t) < t$ for all t, (ii) $\lim_{t\to\infty} v(t)\Gamma(t)/t = \infty$ and (iii) $\lim_{t\to\infty} v(t)\Gamma(t)/(t\, v(\Gamma(t))) = 0$. Then the limit $\omega(\theta) = \lim_{t\to\infty} \omega_t(\theta)$ exists and is given by*

$$\omega(\theta) = \begin{cases} -\theta r \mathbf{E}[\sigma] & \text{if} \quad \theta < 1 \\ +\infty & \text{if} \quad \theta > 1 \end{cases}, \tag{2.10}$$

with $\omega(1)$ determined by further detail of the model in question.

2.3 Asymptotics of a cumulative session detail.

The setting of the previous section is sufficient to describe the simplest example of a session detail process: that for which each session empties fluid into the buffer at identical constant rates. In this section we show how to include stochastic session detail processes which, in a sense which we will make precise, average on a faster time-scale that the $M/G/\infty$ process. Let $(Y_t : t \in \mathbf{N})$ be the cumulative activity $Y_t = \sum_{t'=1}^{t} Z_{t'}$ of the prototype session detail process Z. For some scaling function u define the MGF

$$\lambda_t(\theta) = \frac{1}{u(t)} \log \mathbf{E}[e^{\theta Y_t u(t)/t}] = \frac{1}{u(t)} \mu(\theta u(t)/t). \tag{2.11}$$

Let A_t^n be the total fluid arriving during $\{1, 2, \ldots, t\}$ from all sessions active during that period and define the MGF

$$\nu_t(\theta) = \frac{1}{nv(t)} \log \mathbf{E}[e^{\theta A_t^n v(t)/t}], \tag{2.12}$$

with $v(t) = -\log G_t^c(t)$ as before.

Theorem 2 *Assume the hypotheses of Theorem 1. Let u be a scaling function as in (2.11) such (i) $v(t) = o(u(t))$ as $t \to \infty$; and (ii) $t/u(t)$ is either bounded or eventually increasing; and (iii) for some $\tilde{\theta} > 0$, the session detail process Y is such that the $\lambda_t(\tilde{\theta})$ are finite and the limit $\lambda(\tilde{\theta}) = \lim_{t\to\infty} \lambda_t(\theta)$ exists and is finite. Then*

$$\lim_{t\to\infty} \nu_t(\theta) =: \nu(\theta) = \omega(\bar{z}\theta), \tag{2.13}$$

where $\bar{z} = \mathbf{E}[Y_1] = \lambda_1'(0)$.

Proof: By repeating the argument which gave (2.7) we find that A_t^n has the same distribution as

$$\sum_{i=1}^{N_0} Y^{(i)}{}_{(\hat{\sigma}^{(i)}-1)\wedge t} + \sum_{t'=1}^{t} \sum_{i=1}^{M_{t-t'+1}} Y^{(i,t')}{}_{\sigma^{(i,t')}\wedge t'}, \tag{2.14}$$

where $\hat{\sigma}^{(i)}$ and $\sigma^{(i,t')}$ are independent copies of $\hat{\sigma}$ and σ respectively, and the $Y^{(i)}$ and $Y^{(i,t)}$ are independent copies of the process Y. Define

$$\mu_t(\theta) = \log E[e^{\theta Y_t}]. \tag{2.15}$$

(Note that μ_t, unlike ν_t, contains no scaling functions in its definition). The analog of (2.9) is

$$\nu_t^n(\theta) = \nu_t(\theta) := \frac{r\mathbf{E}[\sigma]}{v(t)} \left(\mathbf{E}[e^{\mu_{(\hat{\sigma}-1)\wedge t}(\theta v(t)/t)}] - 1 \right) + \frac{r}{v(t)} \sum_{t'=1}^{t} \left(\mathbf{E}[e^{\mu_{\sigma\wedge t'}(\theta v(t)/t)}] - 1 \right), \tag{2.16}$$

independent of n. We now show how (2.17) can be approximated by an expression of the form (2.9).

By Jensen's inequality, μ_t is convex for each t. Clearly $\mu_t(0) = 0$, so when $\mu_{t'}$ is finite and differentiable in a neighborhood of 0 we have

$$(t'\theta v(t)/t)t'^{-1}\mu_{t'}'(0) \le \mu_{t'}(\theta v(t)/t) \le (t'\theta v(t)/t)t'^{-1}\mu_{t'}'(\theta v(t)/t). \tag{2.17}$$

Since Y has stationary increments, $t^{-1}\mu_t'(0) = \mu_1'(0) = \bar{z}$ the mean activity per session per slot. So the lower bound in (2.17) is equal to $s\bar{z}\theta v(t)/t$. This means we can bound $\nu_t(\theta)$ below by an expression of the form of (2.9), but with θ replaced by $\theta\bar{z}$. So using Theorem 1, $\liminf_{t\to\infty}\nu_t(\theta)$ is bounded below by an expression of the form (2.10), but with θ replaced by $\theta\bar{z}$.

We turn to the upper bound in (2.17). By the hypothesis of Theorem 1, $v(t)/t \to 0$ as $t \to \infty$. Suppose, therefore, that $t'^{-1}\mu_{t'}'(\theta v(t)/t)$ converges to $t'\mu_{t'}'(0) = \bar{z}$ uniformly for $t' \le t$ as $t \to \infty$. Then for each $\varepsilon > 0$ we can for t sufficiently large write an upper bound to $\nu_t(\theta)$ in the form of (2.9) with θ replaced by $\theta(\bar{z}+\varepsilon)$. Using Theorem 1 for any $\theta < (\bar{z}+\varepsilon)^{-1}$, then taking $\varepsilon \to 0$ we obtain an upper bound to $\limsup_{t\to\infty}\nu_t(\theta)$ which is the same as the lower bound just found. Thus $\nu(\theta) = \lim_{t\to\infty}\nu_t(\theta)$ exists and takes the form (2.13).

It remains only to establish the differentiability of μ_t and uniformity of convergence of the derivatives. Conditions (i) and (ii) guarantee that $\varepsilon(t) = \sup_{t'\le t}(t'v(t)/(tu(t'))$ goes to 0 as $t \to \infty$. Thus

$$t'^{-1}\mu_{t'}'(\theta v(t)/t)) = \lambda_{t'}'(\theta(t'/u(t'))v(t)/t) \le \lambda_{t'}'(\varepsilon(t)) \le \lambda_{t'}'(0) + \int_0^{\varepsilon(t)} dx\ \lambda_{t'}''(x), \tag{2.18}$$

whenever these derivatives are defined. But condition (iii) means that $\lambda_{t'}(\theta)$ is uniformly bounded for all t' and all $\theta \le \tilde{\theta}$. Since the Y_t are non-negative, it is not difficult to see, by explicit calculation of the derivatives, that the same uniform boundedness property holds for λ_t' and λ_t''. Since $\varepsilon(t) \to 0$, then we have, in conjunction with the the lower bound in (2.17), the required uniform convergence property, namely,

$$\lim_{t\to\infty}\sup_{t'\le t}|t'^{-1}\mu_{t'}'(\theta v(t)/t) - \bar{z}| = 0. \tag{2.19}$$

■

2.4 Asymptotics for large arrival and service rate.

The third and last strand to be used in deriving (1.1) has as its starting point a result from [15] on the asymptotics of queue length distribution for queues whose input is, for example, a superposition of many sources. The present case has weaker hypotheses than used in [15], but also some properties that allow for some simplification. For this reason we will supply the relevant new details of the derivation of (1.1).

Define the backward cumulative arrivals $\tilde{A}_t^n$, the total fluid arrivals over times $\{-1, -2, \ldots, -t\}$ with $\tilde{A}_0^n = 0$. Our starting point is a pathwise relation between the queue length Q^n at $t = 0$ and the process A^n (see, e.g., [5]):

$$Q^n = \sup_{t\ge 0}(\tilde{A}_t^n - snt). \tag{2.20}$$

Since arrivals are stationary, $\tilde{A}_t^n \stackrel{d}{=} A_t^n$. Recall that the effective domain of an extended real-valued function is the set on which it is finite.

Theorem 3 *In addition to the hypotheses of Theorems 1 and 2, assume*

(i) $r\bar{z}\mathbf{E}[\sigma] < s$ *(Stability).*

(ii) For all $\varepsilon > 0$, $\lim_{t_0\to\infty}\limsup_{n\to\infty} n^{-1}\log\sum_{t>t_0} e^{-\varepsilon n v(t)} = -\infty$.

Then

$$\lim_{n\to\infty} n^{-1}\log \mathbf{P}[Q^n > nb] = -I(b), \quad \text{where} \tag{2.21}$$

$$I(b) = \inf_{t>0} v(t)\xi_t^*(x/t), \tag{2.22}$$

where

$$\xi_t(\theta) = \nu_t(\theta) - s\theta. \tag{2.23}$$

Proof of Theorem 3. *Lower bound:* Since the session arrivals are Poissonian, $\tilde{A}^t - snt$ has the same distribution as a superposition of n i.i.d. copies of $\tilde{A}^1 - st$. Thus, by Cramer's Theorem (see Section 2.2 in [13]), for each t, the pair $(n^{-1}\tilde{A}_t^n - st, n)$ satisfies an LDP with a rate function which is the Legendre transform of the function $\theta \mapsto \log \mathbf{E}[e^{\theta(A_t^1 - st)}] = v(t)\nu_t(\theta t/v(t)) - st\theta = v(t)\xi_t(\theta)t/v(t))$. By a rescaling one shows that $(v(t)\xi_t(\cdot t/v(t)))^*(x) = v(t)\xi_t^*(x/t)$. Now since

$$\{Q^n > nb\} = \{(\sup_{t>0} \tilde{A}_t^n - stn) > nb\} = \cup_{t>0}\{\tilde{A}_t^n - snt > nb\} \tag{2.24}$$

we have $\mathbf{P}[Q^n > nb] \geq \sup_{t>0}\mathbf{P}[A_t^n - snt > nb]$. In conjunction with the above LDP for each t we get

$$\liminf_{n\to\infty} n^{-1}\log \mathbf{P}[Q^n > nb] \geq \liminf_{n\to\infty} n^{-1}\sup_{t>0}\log \mathbf{P}[A^n - nb > nst] \tag{2.25}$$

$$\geq \sup_{t>0}\liminf_{n\to\infty} n^{-1}\log \mathbf{P}[A^n - nb > nst] \tag{2.26}$$

$$\geq -\inf_{t>0} v(t)\xi_t(b^+/t), \tag{2.27}$$

where the $^+$ denotes limit from the right. If we can show that ξ_t^* is continuous on $\mathbf{R}_+$ then we can replace $\xi_t(b^+/t)$ by $\xi(b/t)$. Combining (2.17) with (2.16) we see that for each t, $\nu_t(\theta) > k_{1,t}e^{\theta k_{2,t}}$ for some positive constants $k_{1,t}, k_{2,t}$. From this it follows that the effective domain of ν_t^*, and hence also that of ξ_t^*, includes $\mathbf{R}_+$. Since ν_t^* is, being a Legendre transform, convex, then by Theorem 10.1 of [33] it is continuous on the interior of its effective domain. This includes $(0,\infty)$, as required.

Upper bound: The proof is the same as for Theorem 1 of [15]. ■

The role of the scaling function v in the proof of Theorem 3 is twofold. Firstly, convergence of ξ_t as $t \to \infty$ is required to establish the upper bound. Secondly, as we shall see in the next section, v determines the asymptotic form of the shape function I. However, v can be removed explicitly from the variational expression for I by defining the (unscaled) MGF

$$\psi_t(\theta) = \log E[e^{\theta A_t^1}] - s\theta t = v(t)\nu_t(\theta t/v(t)) - s\theta t = v(t)\xi_t(\theta t/v(t)). \tag{2.28}$$

Then (2.22) is equivalent to

$$I(b) = \inf_{t>0}\psi_t^*(b). \tag{2.29}$$

3 Analysis of the shape function

3.1 Asymptotic behavior of $I(b)$ as $b \to \infty$

Recall from, for example, Section 1.4 of [4], that a measurable function h is said to be Baire *regularly varying* (at infinity) if for all y in a Baire subset of $(0,\infty)$, the limit $\lim_{x\to\infty} h(xy)/h(x)$ exists, in

which case

$$\lim_{x\to\infty} h(xy)/h(x) = y^r \tag{3.1}$$

for all $y > 0$, for some r called the *index* of h. For a given regularly varying function of index $\tilde{v}$ we set $\hat{h}(x) = x^{\tilde{h}}$. Examples of regularly varying functions are powers: $h(x) = \hat{h}(x) = x^{\tilde{h}}$, and the logarithm $\log x$, which is regularly varying with index zero: $\widehat{\log}(x) = 1$ for $x > 0$. When $\tilde{h} \geq 0$, the convergence in (3.1) is uniform for $y \in (y_1, y_2]$ with $0 < y_1 < y_2 < \infty$. When $\tilde{h} > 0$, this holds with $y_1 = 0$.

Theorem 4 *Assume the hypotheses of Theorems 1, 2 and 3, and furthermore that v has an extension, h, to $\mathbf{R}_+$ which is regularly varying, and hence, since $v(t) = o(t)$, has index $\tilde{h} \in [0, 1)$. Then*

$$\lim_{b\to\infty} \frac{I(b)}{h(b)} = \delta, \tag{3.2}$$

where, setting $\phi = s\bar{z}^{-1} - r\mathbf{E}[\sigma]$,

$$\delta = \begin{cases} \phi & \text{if} \quad \tilde{h} = 0 \\ \left(\tilde{h}\bar{z}\right)^{-\tilde{h}} \left(\frac{\phi}{1-\tilde{h}}\right)^{1-\tilde{h}} & \text{if} \quad \tilde{h} \in (0,1) \end{cases} \tag{3.3}$$

The proof requires the following lemma from [15].

Lemma 1 *Let χ_n be a sequence of convex functions on $\mathbf{R}$. If for some function χ, $\lim_{n\to\infty} \chi_n = \chi$ pointwise on the interior of the effective domain of χ, then $\lim_{n\to\infty} \chi_n^* = \chi^*$ pointwise on the interior of the effective domain of χ^*.*

Combining Theorem 2 with (2.10) and (2.23) we find that

$$\lim_{t\to\infty} \xi_t(\theta) = \begin{cases} \bar{z}\phi\theta & \text{if} \quad \theta < \bar{z}^{-1} \\ +\infty & \text{if} \quad \theta > \bar{z}^{-1} \end{cases} . \tag{3.4}$$

Let $\bar{\xi}$ denote any monotonic extension of this limit function to the point $\theta = \bar{z}^{-1}$. Then

$$\bar{\xi}^*(x) = \begin{cases} +\infty & \text{if} \quad x < -\bar{z}\phi \\ \phi + x\,\bar{z}^{-1} & \text{if} \quad x \geq -\bar{z}\phi \end{cases} \tag{3.5}$$

Combining with Lemma 1 we have proved the first part of the following:

Lemma 2 *(i)* $\lim_{t\to\infty} \xi_t^*(x) = \bar{\xi}^*(x), \quad x > -\bar{z}\phi.$

(ii) The above convergence is uniform on bounded subsets of $\mathbf{R}_+$.

Proof of Lemma 2(ii). Being Legendre transforms, ξ_t^* and $\bar{\xi}^*$ are convex. So since they are finite on $\mathbf{R}_+$, they are also, by Theorem 10.1 of [33], continuous there. Hence the convergence in (i) is uniform on compact sets. ■

Proof of Theorem 4. It is convenient to re-parameterize time as $t = \lfloor b/c \rfloor$ for $c > 0$, where $\lfloor \cdot \rfloor$ denotes the integer part. Setting $f_b(c) = \frac{h(b)}{v(\lfloor b/c \rfloor)}$ then

$$\frac{I(b)}{h(b)} = \inf_{c>0} \frac{\xi^*_{\lfloor b/c \rfloor}\left(b/\lfloor b/c \rfloor\right)}{f_b(c)}. \tag{3.6}$$

Upper bound: By the regularly varying property of h, $\lim_{b\to\infty} f_b(c) = \hat{h}(c)$. Using this and Lemma 2,

$$\limsup_{b\to\infty} \frac{I(b)}{h(b)} \le \limsup_{b\to\infty} \frac{\xi^*_{\lfloor b/c\rfloor}(b/\lfloor b/c\rfloor)}{f_b(c)} = \frac{\bar{\xi}^*(c)}{\hat{h}(c)}. \tag{3.7}$$

Thus

$$\limsup_{b\to\infty} \frac{I(b)}{h(b)} \le \inf_{c>0} \frac{\bar{\xi}^*(c)}{\hat{h}(c)} = \inf_{c>0} c^{-\tilde{h}}(\phi + c\bar{z}^{-1}) = \delta, \tag{3.8}$$

with δ as in (3.3), where the last equality follows by direct evaluation of the variational expression.

Lower bound: We show that δ is a lower bound on $\liminf_{b\to\infty} I(b)/h(b)$. Suppose the infimum if (3.6) is achieved at some c_b. (If the infimum is not achieved, use instead for any $\varepsilon > 0$ a point c_b where the infimum is achieved to within ε, then take $\varepsilon \to 0$ at the end). Then

$$\liminf_{b\to\infty} \frac{I(b)}{h(b)} = \lim_{b\to\infty} \frac{\xi^*_{\lfloor b/c_b\rfloor}(b/\lfloor b/c_b\rfloor)}{f_b(c_b)} \tag{3.9}$$

along some subsequence, which we also denote by (c_b). Now (c_b) must have a sub-subsequence (which we also denote by (c_b)) with one of the following properties. Either (i) there exist $0 < d_- < d_+ < \infty$ such that c_b eventually lies within $[d_-, d_+]$; or (ii) $\lim_{b\to\infty} c_b = 0$; or (iii) $\lim_{b\to\infty} c_b = \infty$. We examine the three cases in turn.

In case (i), then by the uniform convergence described above for (3.1) implies convergence of $f_b(c)$ to $\hat{h}(c)$ as $b \to \infty$, uniformly for $c \in [d_-, d_+]$. Together with Lemma 2 this implies that for all $\varepsilon > 0$ there exists b_ε such that for all $b > b_\varepsilon$.

$$\frac{\xi^*_{\lfloor b/c_b\rfloor}(b/\lfloor b/c_b\rfloor)}{f_b(c_b)} \ge (\bar{\xi}^*(c_b) - \varepsilon)\left(\frac{1}{\hat{h}(c_b)} - \varepsilon\right). \tag{3.10}$$

Since $\bar{\xi}^*$ and $\hat{h}$ are uniformly bounded on $[d_-, d_+]$ we have

$$\liminf_{b\to\infty} \frac{I(b)}{h(b)} \ge \inf_{c\in[d_-,d_+]} \frac{\xi^*_{\lfloor b/c\rfloor}(b/\lfloor b/c_b\rfloor)}{f_b(c_b)} \ge \frac{\bar{\xi}^*(c)}{\hat{h}(c)} + O(\varepsilon) \ge \delta + O(\varepsilon). \tag{3.11}$$

The required lower bound follows since ε is arbitrary. In the remainder of the proof we exclude the possibility of the existence of subsequences (c_b) satisfying (ii) or (iii).

In case (ii), $c_b \to 0$, note first that $\xi_t'(0) = \bar{\xi}'(0) < 0$ for all t. So since $\xi_t(0) = 0$ and ξ_t is convex, $\xi_t(\theta) > 0$ for all $\theta < 0$. Thus for $x > 0$, we can restrict the supremum in the definition (2.3) of $\xi_t^*(x)$ to positive θ; consequently $\xi_t^*(x) = \sup_{\theta>0}(\theta x - \xi_t(\theta))$ is increasing. Thus $\xi_t^*(x) \ge \xi_t^*(0) \to \xi^(0) > 0$ as $t \to \infty$ and so $\xi^*_{\lfloor b/c\rfloor}(b/\lfloor b/c_b\rfloor)$ remains bounded away from zero as $b \to \infty$. However, since $c_b \to 0$, $f_b(c_b) \to 0$ and so $\liminf_{b\to\infty} I(b)/h(b) = +\infty$, in contradiction with the upper bound derived above. Hence there is no such subsequence (c_b) converging to 0.

In case (iii), $c_b \to \infty$, we distinguish between two sub-cases, according to whether (c_b) has a further subsequence for which b/c_b is bounded, or if it does not, in which case b/c_b is divergent. In the first case, with b/c_b bounded, then $I(b)/h(b) \ge k_1 \min_{1\le t\ t_0} \xi_t^*(k_2 b)/h(b)$ for some positive t_0, k_1, k_2. Since each ξ_t^* is increasing and convex, the for some constant k_3, $\xi_t^*(x) > k_3 x$ uniformly for $1 \le t \le t_0$ for large x large enough. Thus for b large and some constant k_4, $I(b)/h(b) > k_4 b/h(b)$ which diverges as $b \to \infty$, in contradiction with the lower bound: there is no such subsequence.

The second case has c_b and b/c_b divergent as $b \to \infty$. The reasoning here is the same as for case (ii) above. $\xi^*_{\lfloor b/c \rfloor}(b/\lfloor b/c_b \rfloor)$ is bounded away from zero, while $1/f_b(c_b) = v(\lfloor b/c_b \rfloor)/h(b) \to \infty$ as $b \to \infty$. $\liminf_{b\to\infty} I(b)/h(b) = +\infty$ as before so there is no such subsequence. ■

3.2 A conservative bound for the shape function.

Although the approximation $\delta h(b)$ to $I(b)$ suggested by Theorem 4 is of a far simpler form that the full variational expression for b, it is only asymptotically true for large b. To what extent can one get a more accurate estimate of the shape function for smaller b, while still retaining something of the comparative simplicity of the approximation?

For $b = 0$ one can show (see [6]) that the infimum in (2.22) is attained at $t = 1$: $I(0) = \psi_1^*(0)$. This is just the large deviation result for blocking probabilities for large aggregations at a bufferless resource given by Hui [19], i.e.

$$\lim_{n\to\infty} n^{-1} \log \mathbf{P}[A_1^n > sn] = -\psi_1^*(0), \tag{3.12}$$

where as before A_1^n is the total fluid arriving in one time slot. In the following theorem we find upper and lower bounds on I which holds for all b. The lower bound is conservative in that it yields an upper bound for overflow probabilities.

Theorem 5

$$I(b) \geq \delta h(b) + \eta \quad \text{where} \quad \eta = -\sup_{t>0} (\mu_t^* \circ h^{-1})^*(\delta). \tag{3.13}$$

Proof:

$$\begin{aligned}
I(b) - \delta h(b) &\geq \inf_{b>0} I(b) - \delta h(b) && (3.14)\\
&= \inf_{b>0} \inf_{t>0} (\mu_t^*(b) - \delta h(b)) && (3.15)\\
&= -\sup_{t>0} \sup_{b>0} (\delta h(b) - \mu^*(b)) && (3.16)\\
&= -\sup_{t>0} \sup_{b>0} (\delta b - \mu^*(h^{-1}(b))) && (3.17)\\
&= -\sup_{t>0} (\mu_t^* \circ h^{-1})^*(\delta). && (3.18)
\end{aligned}$$

■

The utility of (3.13) is in numerical calculations. For a queue with offered load $r\bar{z}\mathbf{E}[\sigma]/s < 1$, one can show that μ_t^* is convex and increasing on $\mathbf{R}_+$. Furthermore, for all the examples treated in Section 4, we can take h^{-1} to be convex. Hence $\mu_t^* \circ h^{-1}$ is convex, and so the calculation of its Legendre transform is straightforward through the Euler-Lagrange equation. A corresponding upper bound does not obtain easily since mirroring the steps above by exchanging supremums and infimums over b does not give such an optimization problem.

4 Examples of Session Duration and Detail.

Session duration.

Parulekar and Makowski have announced a number of examples of session durations G which satisfy the hypotheses of Theorem 1. In each case, the given G defines the scale $v(t) = -\log G^c(t)$, so it remains to find a suitable Γ. We give three examples from [31]:

Pareto Distribution. Let G be discrete Pareto with index $\alpha > 1$, so that $G^c(x) \sim x^{-\alpha}$ as $x \to \infty$. Then $v(t) \sim -(\alpha - 1)\log t$ and we can take $\Gamma(t) = \lfloor t/(1 + \log(1 + \log t))\rfloor$.

Log-normal Distribution. Let G be the log-normal distribution with $G^c(t) = 1 - \Phi(\delta^{-1}\log(t/m))$, where Φ is the cdf for a standard Gaussian random variable. Then $v(t) \sim 2(\delta \ln t)^2$ and we can take $\Gamma(t) = \lfloor t/(1 + \log(1 + \log t))\rfloor$.

Weibull Distribution. Let G be discrete Weibull with parameters $\alpha > 0$ and $\beta \in (0, 1)$, so that $G^c(t) = e^{-\alpha t^\beta}$. One shows easily that $v(t) \sim \alpha t^\beta$ as $t \to \infty$. We require $\beta < 1$ in order that $v(t) = o(t)$, and can then take $\Gamma(t) = \lfloor t/(1 + \log t)\rfloor$.

Session detail

Detail with linear large deviation scalings. By this we mean Y_t for which we can take $u(t) = t$ and $\lim_{t\to\infty} t^{-1}\log E[\theta Y_t]$ exists, at least for some $\theta > 0$. Generally, mixing condition can be invoked to obtain convergence in the linear scale; see [13] for more details. Concretely, the following three classes can be included, the first of which is trivial from the stochastic point of view.

Constant Rates. Here $Y_t = \bar{z}t$, $\nu_t(\theta) = \nu(\theta) = \bar{z}\theta$.

Markovian Arrivals. More precisely, the additive component of a Markov Additive Process (see [20]), which satisfy a certain recurrence condition.

Processes with Associated Increments. Let $Y_{t,t'} = \sum_{t''=t+1}^{t'}$ We say the session detail process has associated increments if

$$\mathbf{E}[f(Y_{t_1,t_2})g(Y_{t_3,t_4})] \geq \mathbf{E}[f(Y_{t_1,t_2})]\mathbf{E}[g(Y_{t_3,t_4})] \tag{4.1}$$

for all positive functions f, g and all $t_1 < t_2 < t_3 < t_4$. As pointed out by Chang [8], using $f(x) = g(x) = e^{\theta x}$, an argument involving sub-additivity shows that when $u(t) = t$, $\nu(\theta)$ exists and is finite if $\nu_t(\theta)$ is finite for t sufficiently large.

Fractional Brownian Motion. If Y_t is fractional Brownian Motion with Hurst parameter $H \in [1, 1/2)$, [26] then in particular Y_t is Gaussian with variance t^{2H}. Thus with $u(t) = t^{2-2H}$, $\nu_t(\theta) = \nu(\theta) = \theta^2$. This is the first session detail process we consider which not always compatible, in the sense of Theorem 2, the the duration processes. The condition $v(t) = o(u(t))$ as $t \to \infty$ requires that the session durations, if Weibull, have parameter $\beta > 2 - 2H$. Fractional Brownian motion has been proposed as a traffic model in [24]. See also [27] for queueing analysis.

Multiscale session duration processes We can use as the session detail process another $M/G/\infty$ process with distribution $G^{(1)}$ and associated scaling function $v^{(1)}(t)$, provided $v(t) = o(v^{(1)}(t))$ and the interior duration process has its own session detail process satisfying the hypotheses of Theorem 2. One can iterate this construction producing multiple levels of inter-burst structure on different time-scales. So for example, one could, if desired, have n levels of detail described by Weibullian burst distributions $G^{(1)}, \ldots G^{(n)}$, $G^{(1)}$ being the distribution of the outermost session level, $G^{(n)}$ being that of the innermost, comprising, say, sessions of constant rate. The conditions of Theorem 2 are satisfied if the corresponding constants $\beta^{(1)}, \ldots \beta^{(n)}$ satisfy $\beta^{(1)} < \beta^{(2)} < \ldots < \beta^{(n)}$.

5 Examples and Applications

5.1 Convergence and comparison of shape function and bound with simulation.

We investigated the closeness of the approximation (1.4) though simulation. The parameters were as follows: arrival rate $r = 1$; session duration Pareto with $G^c(x) = (1 + x/X)^{-\alpha}$ for $\alpha = 1.5$ and $X = 0.2$ a scale factor; session detail constant with $z_t = \bar{z} = 1$. Simulations were conducted for $n = 10, 50, 100$. The results are displayed in the three lower curves of Figure 1. The curves are rescaled in the sense that for each n they join the points $(b/n, \log \hat{\mathbf{P}}[Q^n > b]/n), b = 0, 1, 2, \ldots, B_n$, where B_n is a cutoff chosen as described below. By Theorem 3, this curve should approach that of the shape function. This has been calculated numerically and also displayed. For each n, three runs over 10^8 time slots were conducted, and the cutoff B_n chosen in order that the plots for each run agreed at roughly the resolution used in the Figure. This was at a probability level of about 10^{-6}, although $B_{10}/10$ is greater than the largest value displayed, namely 2. The uppermost curve in Figure 1 is that of the upper bound $-\delta h(b) - \eta$ to $-I(b)$. η was calculated numerically. In this case it is negative, but we have no general result concerning its sign.

5.2 Applications to performance modeling.

General estimation through the shape function. Specify a model by a session arrival rate R, session duration distribution G, session detail Z, service rate S and a queue-level B. By the remarks following Theorem 3, the approximation from (1.4) for the queue length Q is

$$-\log \mathbf{P}[Q > B] \approx \mathcal{I}(B) := \inf_{t>0} \Psi_t^*(B), \tag{5.1}$$

where Ψ is as in (2.28) but with arrival rate R and service rate S. This formula is invariant w.r.t the choice of n in the division of R, S, B into nr, ns, nb.

In the case of unit constant session detail, the stationary distribution of fluid arriving per time instant is Poisson with mean $R\mathbf{E}[\sigma]$, hence

$$\Psi_1(\theta) = R\mathbf{E}[\sigma](e^\theta - 1) - S\theta. \tag{5.2}$$

Using the fact again from [6] that the infimum in (5.1) for $B = 0$ is attained at $t = 0$, or equivalently Hui's result,

$$\mathcal{I}(0) = \Psi_1^*(0) = S \log(S/R\mathbf{E}[\sigma]) + R\mathbf{E}[\sigma] - S. \tag{5.3}$$

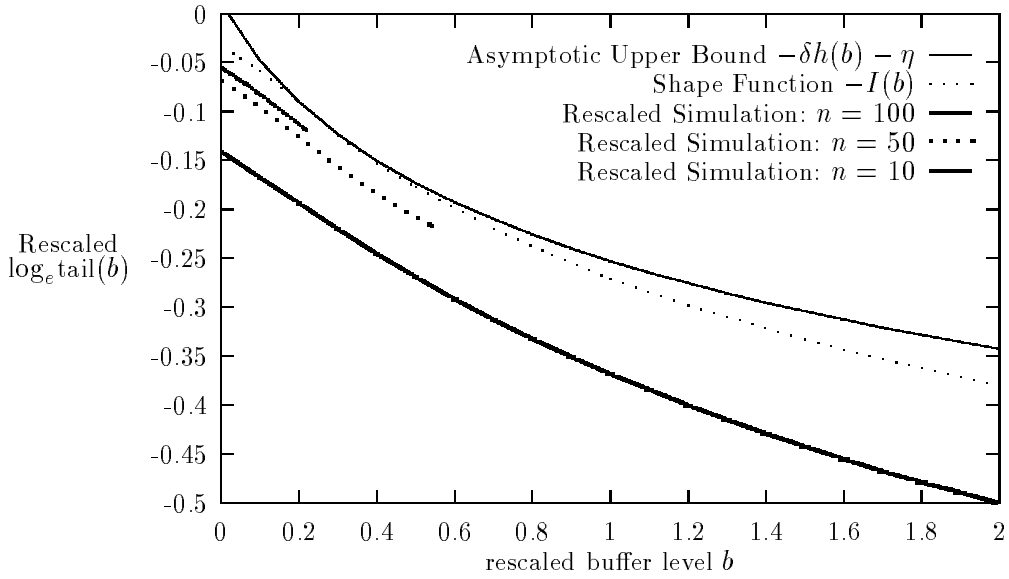

Figure 1: Rescaled tail probabilities: comparison of bounds, asymptotics, and simulations.

Application to traffic with a delay constraint. $\mathcal{I}(0)$ depends on the session duration only through its mean. We now use a calculation of $\mathcal{I}(B)$ to investigate the amount of additional statistical multiplexing gain available by shared buffering of many sources, and its dependence on properties of the session duration beyond the mean. We will also compare the gain obtained with that obtained by increasing the service rate rather than buffer size. One particular case of interest is for applications with a delay constraint in which arrivals delayed by an amount greater than some Δ are no longer useful and are hence dropped. This constrains the buffer size to be (no larger than) $B = \Delta/S$, with the probability of overflow p to be suitably small. For this reason we will want to calculate $\mathcal{I}(B)$ for small B, rather than use the large-buffer approximations derived from (1.5). The delay $\Delta = B/S = nb/(ns) = b/s$ is independent of the aggregation level n at a given load. Consequently, a design algorithm could be as follows: for a given traffic model with arrival rate r and session distribution G, and targets of utilization ρ, maximum delay Δ and buffer overflow rate p, then

1. Select s for required utilization $\rho = r\mathbf{E}[\sigma]/s$
2. Select b for required maximum delay $\Delta = b/s$
3. Select an aggregation level n for the required overflow probability $p = e^{-nI(b)}$.

End-to-end considerations. We have not attempted to model end-to-end performance. However, our single node result can be used as a component of end-to-end delay if, for example,

- The system modeled is an aggregator at the edge of the network and there is no traffic shaping at interior nodes over timescales longer than the unit time: then add 1 per hop to Δ in order to get total delay; or

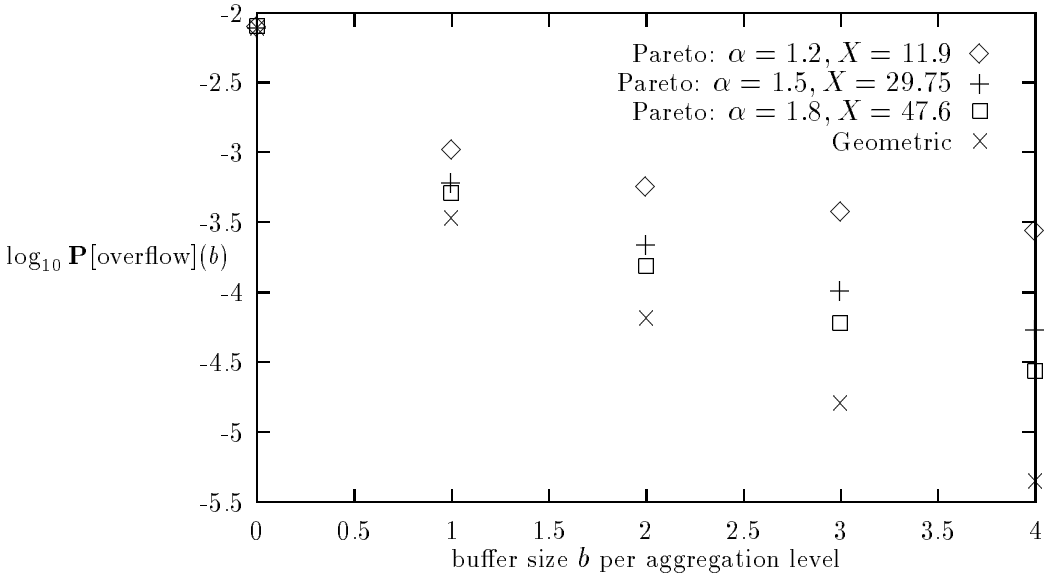

Figure 2: Estimated overflow probabilities from small buffers for three Pareto session duration distributions with common mean, aggregation level n=200. Geometrically distributed durations with the same mean included for comparison.

- any delays $\Delta_i > \tau$ occurring at an interior point i of the network occur independently of those at the aggregator with small probability p_i. Then delays greater than $\max\{\Delta, \max_i \Delta_i\}$ will occur with probability no more than roughly $p + \sum_i p_i$.

Numerical comparisons. We investigate the effects on the shape function I of varying the distribution G of the session duration while keeping the mean duration constant. As an example, we take mean session duration $\mathbf{E}[\sigma] = 60$. Session detail is constant: $z_t = \bar{z} = 1$. Treating n as an aggregation level, we write the aggregate session arrival rate as $R = nr$, where $r = 6 \times 10^{-3}$ as a nominal session arrival rate per source. This gives a mean activity rate per source of 0.36. We take $S = ns$ with $s = 0.5$ giving an offered load of $r\mathbf{E}[\sigma]/s = 0.72$. This doubles the mean utilization per source as compared with allocating unit service rate per source.

In Figure 2 we display the shape function estimate of the (common) logarithm of the overflow probability. This is done for three parameterizations of Pareto distributions of session length, with ccdf $G^c(x) = (1 + x/X)^{-\alpha}$ and mean length 60 ticks, together with the same for geometrically distributed session length of the same mean. (For geometric distributions, the scaling function v is linear, but the same general variational form (2.29) is used the shape function in such cases; see [6]). We set the aggregation level to $n = 200$. We show values of the estimate of log-overflow probability $-\mathcal{I}(nb) \log_{10} e$ and b in unit steps from 0 to 4 cells. This corresponds to delays $\Delta = B/S = b/s$ from 0 to 8 time units.

An alternative way to achieve a target overflow probabilities is to increase S and thereby decrease the offered load. By applying (5.1) we can show that the smallest overflow probability displayed in Figure 2 (roughly $10^{-5.3}$ for $b = 4$ and exponential duration) can be achieved by reducing the arrival

Distribution	b=0	b=1	b=2	b=3	b=4
Pareto, α=1.2	1	109	192	270	347
Pareto, α=1.5	1	73	116	155	191
Pareto, α=1.8	1	64	99	128	156
Geometric	1	51	72	88	102

Table 1: For each point in Figure 2, time τ over which arrivals cause exceedence of level nb, asymptotic as $n \to \infty$.

rate r, and hence the utilization, by about 20%. The methods given here enable the determination of relative performance impact of increasing buffering as against increasing rate allocation. But other quantities, for example relative cost of resources, would be factors in determining the extent to which each method would be used in practice.

Time-scale of applicability of the model. The use of the $M/G/\infty$ model has the potential drawback that it does not model correlations between the occurrences of sessions. For example, Poissonian session arrival processes have been reported for some telnet and ftp session arrivals in IP networks [32], although not for other applications. A different study [17] has reported better fits by other models even in the former cases. However, as is well-known from large deviation theory, the time τ at which the infimum occurs in the variational expression (2.22) for $I(b)$ is (asymptotically as $n \to \infty$) the precise time scale over which the buffer occupancy builds to a level nb. So if τ is found to be less than the typical time T between sessions, we can be confident of the application, since correlations between session arrivals should not, even if present, impinge upon this estimate. Conversely, $\tau > T$ indicates that we have crossed the threshold into burst-level queueing of fluid from sessions: the model may no longer be appropriate.

For the parameters chosen above, the mean time between sessions is $T = r^{-1} \approx 167$ units. In Table 1 we display τ for each point plotted in Figure 2. Observe that T increases for longer tails (i.e. lower α) under constant mean length.

References

[1] J. Abate and W. Whitt, Calculating transient characteristics of the Erlang loss model by numerical transform inversion, *Stochastic Models*, submitted.

[2] V. Anantharam (1988). How large delays build up in a GI/G/1 queue. *Queueing Systems*,**5:**345–368.

[3] V. Anantharam (1995). On the Sojourn Time of Sessions at an ATM Buffer with Long-Range Dependent Input Traffic, *Proceedings of the 34th IEEE Conference on Decision and Control, December 1995.*

[4] N.H. Bingham, C.M. Goldie and J.L. Teugels (1987). *Regular Variation*, Encyclopedia of Mathematics and its Applications. Vol 27. Cambridge University Press, Cambridge.

[5] A.A. Borovkov (1976). *Stochastic processes in queueing theory*, Springer 1976.

[6] D.D. Botvich and N.G. Duffield (1995). Large deviations, economies of scale, and the shape of the loss curve in large multiplexers. *Queueing Systems.* **20:** 293-320.

[7] O. Boxma (1996). Fluid queues and regular variation. Preprint.

[8] C.-S. Chang (1994). Stability, queue length and delay of deterministic and stochastic queueing networks. *IEEE Trans. on Automatic Control*, **39**:913–931.

[9] G.L. Choudhury and W. Whitt (1996). Long-tail buffer-content distributions in broadband networks. Preprint, AT&T Laboratories.

[10] J.W. Cohen (1972). On the tail of the stationary waiting-time distribution and limit theorems for the M/G/1 queue. *Ann. Inst. H. Poincaré,* **8:**255–263

[11] C. Courcoubetis and R. Weber (1996). Buffer overflow asymptotics for a switch handling many traffic sources. *J. Appl. Prob.* **33:** 886–903

[12] D.R. Cox (1984). Long range dependence: a review," *Statistics: An Appraisal*, H.A. David & H.T. David, Eds., The Iowa State University Press, Ames (IA), 1984, pp55–74.

[13] A. Dembo and O. Zeitouni (1993). *Large Deviation Techniques and Applications.* Jones and Bartlett, Boston-London.

[14] N.G. Duffield (1996). On the relevance of long-tailed durations for the statistical multiplexing of large aggregations. *Proceedings of the 34*$^{\text{th}}$ *Annual Allerton Conference on Communication, Control and Computation*, October 2-4, 1996.

[15] N.G. Duffield (1996). Economies of scale in queues with sources having power-law large deviation scalings. *J. Appl. Prob.*, **33:** 840–857.

[16] N.G. Duffield and N. O'Connell (1995). Large deviations and overflow probabilities for the general single-server queue, with applications, *Math. Proc. Cam. Phil. Soc.*, **118**:363–374.

[17] A. Feldmann (1995). On-line call admission for high-speed networks. Ph.D. Thesis, School of Computer Science, Carnegie Mellon University.

[18] D. Heath, S. Resnick, G. Samorodnitsky (1996). Patterns of buffer overflow in a class of queues with long memory on the input stream. Preprint, Cornell University.

[19] J.Y. Hui (1988). Resource allocation for broadband networks. *IEEE J. Selected Areas in Commun.* **6**:1598–1608

[20] I. Iscoe, P. Ney and E. Nummelin (1985). Large deviations of uniformly recurrent Markov additive processes. *Adv. in Appl. Math.* **6:**373–412

[21] P.W. Glynn and W. Whitt (1994). Logarithmic asymptotics for steady-state tail probabilities in a single-server queue. In: *Studies in Applied Probability* Eds. J. Galambos and J. Gani, *Journal of Applied Probability, Special Volume* **31A** 131–159

[22] P. Jelenkovic and A. Lazar (1996). Multiplexing On-Off Sources with Subexponential On Periods, CTR Technical Report #457-96-23, Center for Telecommunications Research, Columbia University.

[23] G. Kesidis, J. Walrand and C.S. Chang (1993). Effective bandwidths for multiclass Markov fluids and other ATM Sources. *IEEE/ACM Trans. Networking*, **1**:424-428.

[24] W.E. Leland, M.S. Taqqu, W. Willinger and D.V. Wilson (1993). On the self-similar nature of Ethernet traffic. *ACM SIGCOMM Computer Communications Review* **23** 183-193

[25] N. Likhanov, B. Tsybakov and N.D. Georganas (1995). Analysis of an ATM buffer with self-similar ("fractal") input traffic, *Proceedings of Infocom95,* Boston, April 1995, pp. 985–992.

[26] B.B. Mandelbrot and J.W. Van Ness (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review*, **10**:422–437.

[27] I. Norros (1994). A storage model with self-similar input. *Queueing Systems*, **16**:387–396

[28] A.G. Pakes (1975). On the tails of waiting-time distributions. *J. Appl. Prob.* **12**:556–564

[29] M. Parulekar and A. Makowski (1996). Tail probabilities for a multiplexer with self-similar traffic. *Proc. IEEE INFOCOM'96, San Francisco, March 26-28, 1996,* pp1452–1459.

[30] M. Parulekar and A. Makowski (1996). Tail probabilities for $M/G/\infty$ input processes(I): preliminary asymptotics. Preprint, University of Maryland.

[31] M. Parulekar and A. Makowski (1996). $M/G/\infty$ input process: a versatile class of models for network traffic. Preprint, University of Maryland.

[32] V. Paxson and S. Floyd (1995). Wide-area traffic: the failure of Poisson modeling. *IEEE/ACM Transactions on Networking*, **3**:226-244.

[33] R.T. Rockafellar (1970) *Convex Analysis.* Princeton University Press, Princeton.

[34] A. Simonian and J. Guibert (1994). Large deviations approximation for fluid queues fed by a large number of on-off sources. *Proceedings of ITC 14, Antibes, 1994* 1013–1022.