Long Range Dependence in the Aggregate Flow of TCP Controlled Elastic Sessions: An Investigation via the Processor Sharing Model*

Anurag Kumar, K.V.S. Hari, R. Shobhanjali and *Srikumar Sharma*
Dept. of Electrical Communicaton Engg.
Indian Institute of Science, Bangalore, 560 012, INDIA
email: anurag, hari @ece.iisc.ernet.in

Abstract

We consider elastic sessions sharing a single bottleneck link; the sessions are flow controlled by TCP. Sessions arrive randomly, in a Poisson process, request the transfer of a randomly chosen volume of data (from some file size distribution), and then depart after completion of the transfer. We first consider the Processor Sharing (PS) model for this problem, and study the tail behaviour of the autocovariance of the aggregate flow process in the link. We show that if the file size distribution is Pareto, with parameter α, then the process is long range dependent with Hurst parameter $\frac{3-\alpha}{2}$. This latter result is similar to the one previous researchers have obtained by using the $M/G/\infty$ model, which is inappropriate when the sessions are bottleneck constrained.

We then study the same scenario on a network test-bed in which TCP enforces the bandwidth sharing. For Pareto distributed file sizes, with tail parameter $1 < \alpha < 2$, we find that (i) the measured Hurst parameter of the aggregate traffic in the link matches well with the analytical formula when α is close to 1, and (ii) buffer loss is excessive, leading to lower session throughput than predicted by the PS model. We find that for file size distributions with finite second moment the PS model predicts fairly well the average throughput per session, and there is little sensitivity to the file size distribution.

*To appear in NCC 2000, New Delhi, January 2000.

1 Introduction

Several measurements on traffic in the Internet have shown that packet flow processes have the property of long range dependence (LRD) ([13], [14]). Thus these processes are not well modelled by the Poisson process, or even a Markov modulated process with a small number of states. Since LRD has important consequences for congestion phenomena, and hence network dimensioning, there have been efforts to explain long range dependence via analytical models. For elastic sessions (which are our concern in this paper) the origins of LRD have been traced to heavy tailed file sizes ([6]). If several sources transmit such files, alternating between file transmission at a *fixed rate* and idleness, then it has been shown that the aggregate flow process is LRD. In the Internet, however, session throughputs are constrained by network bottlenecks. The TCP protocol roughly attempts to share the network bandwidth fairly between contending flows. Hence there is a need to understand the LRD behaviour of aggregate flows in the Internet, keeping adaptive rate controlled bandwidth sharing in mind.

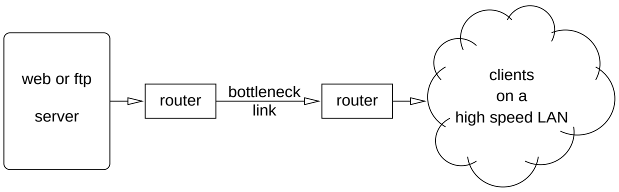

Figure 1: The network scenario considered in this paper.

In this paper we consider the network scenario shown in Figure 1. The clients are downloading files from the server across the bottleneck link. Requests for file transfer arrive in a Poisson process; the request sizes are independently and identically distributed with some distribution. We study this scenario using an analytical model, and an experimental test-bed.

We first use the processor sharing (PS) queue (see [12]) as a model for ideal bandwidth sharing on the bottleneck link. Hence one of the themes of this work is also to investigate the PS queue as a model for TCP controlled bandwidth sharing (see [15] for such use of the PS model). The busy-idle process of the PS queue then models the aggregate packet flow in the network (we justify this view in Section 2). The service time distribution in the queue corresponds to the distribution of file sizes requested by the clients. We consider the stationary version of the busy-idle process, and, using Laplace Transforms and a Tauberian theorem, we show that if the file size distribution is Pareto with tail parameter α, then the process is LRD with Hurst parameter $H = \frac{3-\alpha}{2}$. If the file size distribution has finite second moment, then the process is not LRD. The above formula for the Hurst parameter was also obtained earlier, but in the context of the M/G/∞ model, which does not model bottleneck constrained bandwidth sharing (see [14]).

We then use a test-bed to study how well the above theoretical results, obtained from the PS model, hold when the sessions are flow controlled by TCP. We also study how well the PS model predicts the TCP controlled session throughputs. In our network test-bed three Linux PCs are used, one as the web server, one acting as the clients, and one serving as a serial link emulator. As in the analytical model, the simulated transfer requests arrive in a Poisson process, and request the transfer of files with independent and identically distributed (i.i.d.) sizes.

Our experimental results show that for Pareto file size distribution with α close to 1, the departure process of the bottleneck queue has an H parameter that matches well with the formula $H = \frac{3-\alpha}{2}$. Further, the H parameter decreases with increasing α.

We find that the PS model predicts the session throughputs fairly well when the file size distributions have a finite second moment. Further, the average per session throughput is only slightly sensitive to file size distribution[1]. For Pareto file size distribution, however, and for the tail parameter α close to one, the PS model overestimates the session throughputs, the error becoming larger as α approaches 1. We trace these observations to the fact that TCP performance is sensitive to packet loss. We find that with heavy tailed file sizes, there is greater buffer loss, and hence throughput suffers.

The most important related work is that of Heyman et al [9]. The authors study a model comprising several hosts connected to a backbone link via low speed access links. Each host alternates between downloading files and "thinking". An analytical model is developed and it is shown that the session throughputs are insensitive to the file size distributions. There is, however, no analytical study of LRD in the network traffic. Another related reference is [3], where the authors provide asymptotic approximations for bandwidth dimensioning using the model in [9].

Our work demonstrates the importance of considering heavy tailed file sizes in network buffer sizing for TCP controlled elastic traffic, and relates the parameter of the Pareto file size distribution to the Hurst parameter of the traffic in the network link.

The paper is organised as follows. In Section 2 we justify the use of the PS model, and

[1]Note that the (customer) average session throughput is sensitive to file size distribution even for the PS model; it is the time average fair bandwidth share that is insensitive (see Section 4).

relate the aggregate network traffic to the busy-idle process of the PS queue. In Section 3 we analyse the stationary busy-idle process and develop results on the tail behaviour of its autocovariance function. In Section 4 we provide the experimental results from a network test-bed.

2 The Processor Sharing Model: Ideal Bandwidth Sharing

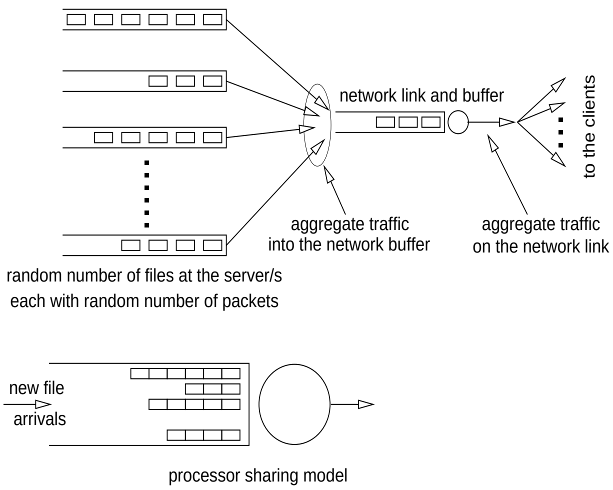

Figure 2: The actual flow of traffic in the network, and the processor sharing model.

Figure 2 shows a schematic of how traffic flows in the network shown in Figure 1. The clients generate file transfer requests. After the requests reach the server, we can view the files to be "queued" at the server, waiting as they are gradually transferred to the clients. In the figure, the files have been shown broken up into fixed length packets. With ideal bandwidth sharing, when there are n files, each file transfers its packets at $\frac{1}{n}$th of the link rate. The network link carries the superposition of these rate controlled flows from each file transfer. Thus, considering a *fluid* model, we can model the file transfer process, with ideal bandwidth sharing, by a processor sharing queue model. The unsent fragments of the files *in the server*, constitute the "customers" queued in the PS model (see the lower part of Figure 2).

Notice, from the above discussion, that the aggregate flow in the network will be related to the *departure process* of the PS model. There are, however, two aggregate flow processes in our single link "network" (see Figure 2), one that flows into the link buffer, and one that flows out of the buffer on to the link. One can reasonably argue that the process entering the network buffer, and the network buffer's behaviour, relate to a rough approximation of the implicit feedback based fair bandwidth sharing, and are thus "internal" to this fair sharing mechanism, whereas the process on the network link is akin to the departure process of the PS model. The link alternates between carrying traffic at the full link rate, and idle periods, and therefore the flow process on it would correspond to the busy-idle process of the PS model. In the results of our test-bed experiments, we will estimate the H parameter of this output process and compare the estimates with what is predicted by the PS model.

The PS model assumes ideal bandwidth sharing, and does not explicitly model the crude manner in which TCP attempts to adaptively share the link bandwidth. In the actual network discrete packets flow and these need to be queued in the network buffers. With TCP controlling the bandwidth sharing, there will be rate mismatches between the packet arrival rate into the link and the link service rate. Hence buffers can fill up and lose packets. Thus, if specific file size distributions result in greater packet loss, then one would expect that the throughputs with TCP will not be the same as predicted by the PS model.

In the next section we show that with Pareto distributed file sizes the PS model predicts that the aggregate traffic is LRD. In Section 4 we examine how well this result holds in an experimental test-bed where the flows are TCP controlled. We also find that the PS model overestimates the throughputs for Pareto file sizes, whereas it is a good approximation for file size distributions with finite second moment.

3 Long Range Dependence of the Aggregate Bit Rate Process

It is well known that the busy-idle process of an M/G/1 PS queue is the same as that of any M/G/1 queue with a nonidling ("work conserving") service discipline. Consider the busy-idle process of such an M/G/1 queue. Let $\{X(t), t \geq 0\}$ denote a stochastic process that is 1 when the queue is busy, and 0 when the queue is idle. We know (see, for example, [11]) that $\{X(t)\}$ is an alternating renewal process [4].

Let λ denote the rate of the Poisson arrival process. This corresponds to the arrival rate of file transfer requests (see Figure 2). Let B denote the random variable for the work brought by a customer, and $\tilde{b}(s)$ its Laplace Stieltjes transform (LST)[2]. In the network

[2]We follow the convention that if Z is a random variable then EZ is its expectation, $Z(\cdot)$ denotes its cumulative distribution function, $z(\cdot)$ its density (if it exists), $\tilde{z}(s)$ the LST of $Z(\cdot)$, and $\tilde{Z}(\cdot)$ the Laplace

context, this work is the file size to be transferred divided by the bit rate of the link. We assume throughout that $EB < \infty$. Let G denote the busy period random variable, and $\tilde{g}(s)$ its LST. Finally, as usual, we let $\rho = \lambda EB$.

Except where explicitly noted, we consider the *stationary* version of the process $\{X(t), t \geq 0\}$. It is well known that $\{X(t)\}$ is characterised as follows: $X(0) = 1$ with probability ρ, and $X(0) = 0$ otherwise. If $X(0) = 1$, then the first 1-period of the process has duration G_e, the excess life distribution of G; i.e., $G_e(t) = \frac{1}{EG}\int_0^t (1 - G(u))du$. On the other hand if $X(0) = 0$, then the first 0-period is exponentially distributed with parameter λ. After the first period is obtained in this way, the process alternates between the states 0 and 1, staying in state 0 for a duration distributed as $Exponential(\lambda)$, and staying in state 1 for a duration distributed as $G(\cdot)$.

Define

$$p_e(t) = Prob(X(t) = 1 \mid X(0) = 1)$$

and

$$p_o(t) = Prob(X(t) = 1 \mid X(0) = 1 \text{ and } t = 0 \text{ is the start of an } \textit{ordinary} \text{ busy period})$$

3.1 The Autocovariance Function of $\{X(t)\}$

As discussed in Section 2, we wish to study the autocovariance function $r(\tau)$ of $\{X(t)\}$; i.e.,

$$r(\tau) = EX(t)X(t+\tau) - \rho^2 \tag{1}$$

Clearly, since $X(t) \in \{0, 1\}$,

$$r(\tau) = \rho(p_e(\tau) - \rho) \tag{2}$$

A renewal argument yields the following equation (the superscript c on a distribution denotes the complementary distribution)

$$p_e(t) \quad = \quad G_e^c(t) + \int_0^t \int_0^{t-u} \lambda e^{-\lambda v} \, p_o(t - (u+v)) \; dv \; dG_e(u) \tag{3}$$

where $p_o(t)$ is characterised by the following equation

$$p_o(t) \quad = \quad G^c(t) + \int_0^t \int_0^{t-u} \lambda e^{-\lambda v} \, p_o(t - (u+v))dv \; dG(u) \tag{4}$$

Transform (LT) of $Z(\cdot)$.

Taking Laplace Transforms in Equations 4 and 3, we obtain

$$\begin{aligned}
\tilde{p}_o(s) &= \frac{1-\tilde{g}(s)}{s} + \left(\frac{\lambda}{s+\lambda}\right) \tilde{p}_o(s) \cdot \tilde{g}(s) \\
\tilde{p}_e(s) &= \frac{1-\tilde{g}_e(s)}{s} + \left(\frac{\lambda}{s+\lambda}\right) \tilde{p}_o(s) \cdot \tilde{g}_e(s)
\end{aligned} \tag{5}$$

Obtaining $\tilde{p}_o(s)$ from Equation 5, and noting that $\tilde{g}_e(s) = \frac{1-\tilde{g}(s)}{sEG}$, we obtain

$$\tilde{p}_e(s) = \frac{1}{s}\left(1 - \frac{1-\tilde{g}(s)}{EG(s+\lambda(1-\tilde{g}(s)))}\right) \tag{6}$$

Define $x(s) = 1-\tilde{g}(s)$. From Equation 1, and the fact that $EG = \frac{EB}{1-\rho}$, we get the Laplace transform for $r(t)$ as

$$\tilde{r}(s) = \frac{\rho(1-\rho)}{s}\ \left(1 - \frac{x(s)}{EB(s+\lambda x(s))}\right) \tag{7}$$

Our aim is to now to study the properties of $r(t)$ for large t via its Laplace transform.

In order to apply certain Tauberian theorems (to be stated later) we first need to show that $r(t)$ is ultimately nonnegative, in the sense that there exits t_0 such that $r(t) \geq 0, t \geq t_0$. Define $q_e(t) = 1 - p_e(t)$. Take $X(0) = 1$, and let $\{T_k, k \geq 1\}$ denote the epochs at which busy periods (or 1-periods) start in the process $\{X(t)\}$. Let $m_{e,1}(t)$ denote the renewal function (see [18]) of the delayed renewal process $\{T_k, k \geq 1\}$ ("delayed" because the first 1-period has the distribution $G_e(\cdot)$. The mean renewal time in this renewal process is $EG + (1/\lambda)$. It is then easily shown that

$$p_e(t) = 1 - \frac{1}{\lambda}\, \dot{m}_{e,1}(t) \tag{8}$$

where, $\dot{m}_{e,1}(t)$ is the renewal density. Taking transforms, we have

$$\tilde{p}_e(s) = \frac{1}{s} - \frac{1}{\lambda}\, \tilde{m}_{e,1}(s) \tag{9}$$

where $\tilde{m}_{e,1}(s)$ is the LST of $m_{e,1}(t)$. It can be shown that

$$\tilde{m}_{e,1}(s) = \frac{\lambda\, x(s)}{sEG(s+\lambda\, x(s))} \tag{10}$$

Notice that Equations 10 and 9 are consistent with Equation 6.

We show that $r(t)$ is ultimately nonnegative by showing that $\dot{m}_{e,1}(t)$ increases to its limit $\frac{1}{(1/\lambda)+EG}$. For this we proceed as follows. Noting that $\tilde{m}_{e,1}(s)$ is the LST of $m_{e,1}(t)$, from Equation 10 we have

$$\begin{aligned} s\tilde{m}_{e,1}(s) &= \frac{\lambda\,(1-\tilde{g}(s))}{EG(s+\lambda(1-\tilde{g}(s)))} \\ &= \frac{\lambda(EGs-o(s))}{EG(s+\lambda(EGs-o(s)))} \\ &= \frac{1-\frac{o(s)}{sEG}}{\frac{1}{\lambda}+EG\left(1-\frac{o(s)}{sEG}\right)} \end{aligned}$$

If the $o(s)$ term can be shown to be positive then it is clear that, ultimately,

$$\dot{m}_{e,1}(t)\uparrow\frac{1}{(1/\lambda)+EG}$$

In both the cases we study below (finite EB^2 and Pareto distribution) this turns out to be true. Hence, ultimately, $p_e(t)\downarrow\rho$, which implies (see Equation 2) that $r(t)$ is ultimately nonnegative.

3.2 Two Tauberian Theorems

To study the asymptotics of $r(t)$ from its Laplace transform, we need slight generalisations of two Tauberian theorem from Widder [16, Theorem 2.2, page 197; Theorem 5.3, page 209]. The generalisations are in replacing the nonnegativity of the function by its ultimate nonegativity.

Theorem 1 *If*

1. $\tilde{f}(s)=\int\limits_0^\infty e^{-st}f(t)\;dt, 0<x<\infty$
2. $f(t)\geq 0$ *for* $t\geq\; t_0$
3. $\lim\limits_{s\to 0^+}\tilde{f}(s)=A$

then

$$\int\limits_0^\infty f(u)du=A$$

□

Theorem 2 *If* [3]

$$
\begin{aligned}
&1.\quad \tilde{f}(s) = \int\limits_0^\infty e^{-st} f(t)\ dt, 0 < x < \infty \\
&2.\quad f(t) \geq 0 \quad for \quad t \geq\ t_0 \\
&3.\quad \tilde{f}(s) \sim \frac{A}{s^\gamma}, \qquad s \to 0^+, \qquad for\ some\ \gamma > 0
\end{aligned}
$$

then

$$\int\limits_0^t f(u)du \sim \frac{At^\gamma}{\Gamma(\gamma+1)} \qquad t \to \infty$$

□

The proofs of these results are simple derivations from the results of Widder[16]. We provide the proof of Theorem 2 in Appendix A.

3.3 File Sizes with Finite Second Moment

Theorem 3 *If* $0 < EB^2 < \infty$ *then the autocovariance function* $r(t)$ *is summable, i.e.,*

$$\int_0^\infty r(u)\ du < \infty$$

.

Proof: We have Equation 7

$$\tilde{r}(s) = \frac{\rho(1-\rho)}{s} \left(1 - \frac{x(s)}{EB(s+\lambda x(s))}\right)$$

In an attempt to apply Theorem 1, we consider

$$\lim_{s\to 0+} \frac{1}{s}\left(1 - \frac{x(s)}{EB(s+\lambda x(s))}\right) = \lim_{s\to 0+} \frac{EBs - (1-\rho)\ x(s)}{sEB(s+\lambda x(s))}$$

Applying L'Hospital's rule twice, we get

$$
\begin{aligned}
\lim_{s\to 0+} \frac{EBs - (1-\rho)\ x(s)}{sEB(s+\lambda x(s))} &= \frac{EB - (1-\rho)x'(s)\ |_{s\to 0+}}{EB(s+\lambda x(s)) + sEB(1+\lambda x'(s))\ |_{s\to 0+}} \\
&= \frac{-(1-\rho)x''(s)\ |_{s\to 0+}}{EB(1+\lambda x'(s)) + EB(1+\lambda x'(s)) + sEB(\lambda x''(s))\ |_{s\to 0+}}
\end{aligned}
$$

[3]We use the usual notation $a(t) \sim b(t), t \to t_0$ to mean $\lim_{t\to t_0} \frac{a(t)}{b(t)} = 1$.

Since we are given that $EB^2 < \infty$, the equation continues as

$$\begin{aligned}
&= \frac{(1-\rho)\ EG^2}{2EB\ \ (1/(1-\rho))} \\
&= \frac{(1-\rho)^2}{2EB}\frac{EB^2}{(1-\rho)^3} \\
&= \frac{EB^2}{2EB(1-\rho)}
\end{aligned}$$

where the formula for EG^2 is taken from Kleinrock, Vol I [11]. Hence

$$\lim_{s \to 0^+} \tilde{r}(s) = \frac{\rho EB^2}{2EB}$$

Further, with $0 < EB^2 < \infty$, the argument at the end of Section 3.1 shows that $r(t)$ is ultimately nonnegative. It follows, from Theorem 1 that

$$\int_0^\infty r(u)\ du = \frac{\rho EB^2}{2EB}$$

Notice that the expression has the familiar form of the mean residual service time in the M/G/1 queue. □

Since $r(t)$ is summable, it follows that when the file size distribution has a finite second moment, then the aggregate process of (ideally) flow controlled elastic flows cannot be long-range dependent (see [2]).

3.4 Pareto Distributed File Sizes

We now consider file size distributions given by

$$1 - B(x) = \begin{cases} 1 & \text{for } \ 0 \leq x \leq 1 \\ x^{-\alpha} & \text{for } \ x \geq 1 \end{cases}$$

where $1 < \alpha < 2$. For this distribution we have $EB = \frac{\alpha}{\alpha - 1}$, and $EB^2 = \infty$. It is easily seen that $\tilde{b}(s)$ is given by

$$\tilde{b}(s) = \alpha \cdot s^\alpha \ \ \Gamma(-\alpha, s)$$

where, $\Gamma(\cdot, \cdot)$ is the incomplete Gamma function[4]. Further, by using the series expansion of the incomplete Gamma function (see [8, page 941]), we get

$$\begin{aligned}
\tilde{b}(s) &= \alpha s^\alpha \left\{ \Gamma(-\alpha) - \sum_{n=0}^{\infty} \frac{(-1)^n \ \ s^{-\alpha+n}}{n!\ (-\alpha+n)} \right\} \\
&= \alpha s^\alpha\ \Gamma(-\alpha) - \alpha \sum_{n=0}^{\infty} \frac{(-1)^n \ \ s^n}{n!\ (-\alpha+n)} \qquad (11)
\end{aligned}$$

[4] $\Gamma(a, y) = \int_y^a \ e^{-u} u^{a-1}\ du$

For the M/G/1 busy period we have the relationship [11]

$$\tilde{g}(s) = \tilde{b}(s + \lambda(1 - \tilde{g}(s))$$

Using Equation 11, and recalling that $x(s) = 1 - \tilde{g}(s)$, we obtain

$$\begin{aligned}
\tilde{g}(s) &= 1 - \alpha \sum_{n=1}^{\infty} \frac{(-1)^n}{n!\,(-\alpha+n)}\ (s + \lambda(1 - \tilde{g}(s))^n + \alpha\Gamma(-\alpha)(s + \lambda(1 - \tilde{g}(s)))^{\alpha} \\
0 &= x(s) - \alpha \sum_{n=1}^{\infty} \frac{(-1)^n}{n!\,(-\alpha+n)}\ (s + \lambda x(s))^n + \alpha\Gamma(-\alpha)(s + \lambda x(s))^{\alpha} \\
\frac{x(s)}{s + \lambda x(s)} &= \alpha \sum_{n=1}^{\infty} \frac{(-1)^n}{n!\,(-\alpha+n)}\ (s + \lambda x(s))^{n-1} - \alpha\Gamma(-\alpha)(s + \lambda x(s))^{\alpha-1} \\
1 - \frac{x(s)}{EB(s + \lambda x(s))} &= (\alpha - 1)\left(\Gamma(-\alpha)(s + \lambda x(s))^{\alpha-1} - \sum_{n=2}^{\infty} \frac{(-1)^n (s + \lambda x(s))}{n!\,(-\alpha+n)}\right)
\end{aligned}$$

It follows, from Equation 7, that

$$\tilde{r}(s) = \frac{\rho(1-\rho)(\alpha-1)}{s}\ \left(\Gamma(-\alpha)(s + \lambda x(s))^{\alpha-1} - \sum_{n=2}^{\infty} \frac{(-1)^n\ (s + \lambda x(s))^{n-1}}{n!\,(-\alpha+n)}\right) \tag{12}$$

In order to apply Theorem 2, we establish the following Lemma.

Lemma 1 *For* $1 < \alpha < 2$,

$$\tilde{r}(s) \sim \rho(1-\rho)^{(2-\alpha)}(\alpha-1)\Gamma(-\alpha)\ s^{-(2-\alpha)} \qquad s \to 0^+$$

Proof: With Equation 12 in mind, we consider

$$\frac{s^{(2-\alpha)}}{s}\ \left(\Gamma(-\alpha)(s + \lambda x(s))^{\alpha-1} - \sum_{n=2}^{\infty} \frac{(-1)^n\ (s + \lambda x(s))^{n-1}}{n!\,(-\alpha+n)}\right) =$$
$$\Gamma(-\alpha)\ \left(\frac{s + \lambda x(s)}{s}\right)^{(\alpha-1)} \left(1 - \sum_{n=2}^{\infty} \frac{(-1)^n\ (s + \lambda x(s))^{n-\alpha}}{n!\,(-\alpha+n)\ \Gamma(-\alpha)}\right)$$

Since $1 < \alpha < 2$, writing $x(s) = EGs - o(s)$, as $s \to 0^+$ we get the right hand side of the equation to converge to

$$\Gamma(-\alpha)\ (1 + \lambda EG)^{(\alpha-1)}$$

which, since $EG = EB/(1-\rho)$, becomes

$$\Gamma(-\alpha)\ (1-\rho)^{(1-\alpha)}$$

Hence the result follows. $\square$

We need to also establish that $r(t)$ is ultimately nonnegative. To apply the approach at the end of Section 3.1 we need to show that the $o(s)$ term there is postive. To show this we proceed from the series expansion of $x(s)$:

$$x(s) \;=\; -\alpha\,\Gamma(-\alpha)\,(s+\lambda x(s))^{\alpha} + \alpha\sum_{n=1}^{\infty}\frac{(-1)^n}{n!\,(-\alpha+n)}\,(s+\lambda x(s))^n$$

Taking only the lower order terms (noting that $\rho = \lambda\frac{\alpha}{\alpha-1}$, and $x(s) = EGs + o(s)$),

$$x(s) - \rho x(s) - \left(\frac{\alpha}{\alpha-1}\right)s \;=\; -\alpha\Gamma(-\alpha)\,\left(\frac{s}{1-\rho}\right)^{\alpha} + o(s^2)$$

It follows that

$$\begin{aligned} x(s) &= \left(\frac{\alpha/(\alpha-1)}{1-\rho}\right)s - \frac{\alpha\Gamma(-\alpha)}{(1-\rho)^{(\alpha+1)}}\,s^{\alpha} + o(s^2) \\ &= sEG - \frac{\alpha\Gamma(-\alpha)}{(1-\rho)^{(\alpha+1)}}\,s^{\alpha} + o(s^2) \end{aligned}$$

Since $1 < \alpha < 2$ and $\frac{\alpha\Gamma(-\alpha)}{(1-\rho)^{(\alpha+1)}} > 0$ we are done, and $r(t)$ is ultimately nonnegative.

We can now apply Theorem 2 to conclude that

$$\int_0^t r(u)\,du \sim \frac{\rho(1-\rho)^{(2-\alpha)}(\alpha-1)\Gamma(-\alpha)}{\Gamma(3-\alpha)}\,t^{(2-\alpha)} \qquad t\to\infty \tag{13}$$

We can thus conclude that, as $t \to \infty$, the autocovariance function $r(t)$ behaves as $t^{-(\alpha-1)}$, and hence (with $1 < \alpha < 2$) the process $\{X(t)\}$ is long range dependent with Hurst parameter $H = \frac{3-\alpha}{2}$. Note that this formula is the same as that reported in [5], and in [14] (see also [17]), where it was obtained by considering the M/G/∞ model. The M/G/∞ model assumes that the network flows are constrained by access links, or the clients, or by some other means, and the flows do not interact due to bandwidth sharing in the network link. The processor sharing model, however, captures the interaction between flows that are constrained by a bottleneck link. It is interesting that the same formula for the Hurst parameter carries over to this more complex bandwidth sharing situation.

4 Experimental Results from a Network Test-Bed

4.1 The Network Test-Bed

Figure 3 shows the experimental setup that we have used to study the network scenario of Figure 1. One machine generates file downloads from another machine that runs a web

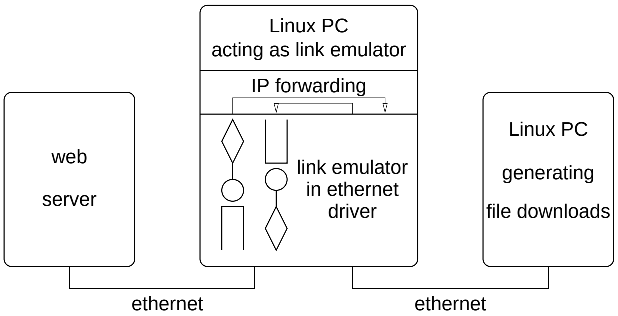

Figure 3: Experimental setup for emulation of the scenario in Figure 1. The Linux PC in the middle emulates a wide area link. The buffer sizes, link speeds, and propagation delays are configurable.

server. All the traffic between these machines passes (via Ethernet) through a third Linux PC that acts as a link emulator. In the generic Ethernet driver of this machine we have introduced some software (see [1]) that buffers packets in either direction, and then releases them at a configurable speed. In general, this emulator can also introduce a "propagation" delay, and introduce random losses; but we do not exercise these features in our experiments reported in this paper.

4.2 Experimental Procedure

We wish to experiment with file sizes of various distributions. For any given distribution, we generate a set of 1000 random numbers from that distribution, and taking these as a 1000 file sizes, we create one file of each size and store it in the server machine. On the client machine a program repeatedly requests for the transfer of a file randomly chosen from among the 1000 files on the server. Between successive requests, the program "sleeps" for an exponentially distributed time, thus modelling the Poisson "arrival" epochs of file transfer requests.

The link speed in the emulator is set to the desired bit rate (128Kbps in all the results reported here). We take the propagation delay to be zero. The buffer size on the emulator is also set to the desired value (32KB for the results reported here).

The client machine runs a TCP packet capture program that records the TCP/IP header, and the receipt epoch, of every packet delivered to the clients. Note that these receipt epochs correspond to departure epochs from the network queue in Figure 2. Using these packet traces we can study the second order characteristics of the network departure process.

4.3 H Parameter of the Aggregate Departure Processes

From the packet trace provided by the packet capture program we obtain the time series of the aggregate number of bytes in the network link departure process in 0.1 second intervals. Let $\{B_k\}$ denote this time series. Since the aggregation intervals are of fixed length, the $\{B_k\}$ process is proportional to the rate process. Hence, $\{B_k\}$ corresponds to the process $\{X_k\}$ that was studied in Section 2. We estimate the H parameter of $\{B_k\}$, and compare it with that predicted by the theory based on the PS model. The variance-time plot technique is used to estimate H (see [2]).

We show the results from five such experiments in Tables 1, 2, 3, 4, and 5. The file size distributions for these tables are, respectively, exponential, gamma (order 2), and Pareto with $\alpha = 1.1, 1.2, 1.6$, and with mean 10KB. In order to get the average of the 1000 sample files for Pareto with $\alpha = 1.1$ to be close to 10KB, we had to use a distribution with mean 12KB; this is because there is significant mass in the tail of this distribution.

Each table shows the configured link load (arrival rate $\times$ mean file size), the measured H parameter at the link output, the number of packets lost, the number of bytes lost, the measured average file transfer throughput, the measured link occupancy, and the number of files transferred. The measured link occupancy is different from the configured load for two reasons: the link carries header bytes (this causes the link occupancy to be higher than that configured); when the request arrival rate is high (the interarrival times are small) the UNIX *sleep* utility is inaccurate (this causes the actual offered load to be smaller than that configured).

We infer from the analysis in Section 2 that for exponential and gamma distributions we expect the H parameter to be 0.5, irrespective of the load, and for Pareto, with parameter α we expect $H = \frac{3-\alpha}{2}$. We observe from Tables 1 and 2, that for exponential and gamma distributions H is about 0.5 for load 0.1, and increases to about 0.7 for load 0.9. From Table 3, we observe that for Pareto with $\alpha = 1.1$ the H parameter varies between 0.96 and 0.93, as against the theoretical value of 0.95. From Tables 4 and 5, we see that for Pareto, and $\alpha = 1.2$, the measured H parameter is in the range 0.94 and 0.89 (the theory yielding 0.9), and for $\alpha = 1.6$ the measured H parameter is in the range 0.77 to 0.85 (the theory yielding 0.7).

At the very least, it is clear from these measurements that the H parameter of the departure process depends on the file size distribution. The closeness of match between the theoretical and the measured Hurst parameters for $\alpha = 1.1$ is striking, and supports the application of the PS model results in Section 2 to this situation. For Pareto file size distributions, the H parameter does decrease with increasing α, but the H values are higher than predicted by theory. For file size distributions with finite second moments, the H

Offered load (as configured)	H	packets lost	bytes lost	Throughput (Kbps)	Measured load	Total No. of files transferred
0.1	0.49	0	0	13.11	0.10	555
0.2	0.57	11	16654	12.50	0.21	1136
0.3	0.58	17	24341	11.80	0.32	1689
0.4	0.61	30	43388	11.16	0.41	2216
0.5	0.60	61	86738	10.34	0.51	2742
0.6	0.60	124	184370	9.30	0.60	3248
0.7	0.59	299	441749	8.35	0.69	3701
0.8	0.67	870	1279408	6.71	0.79	4250
0.9	0.70	1951	2848090	4.93	0.88	4666

Table 1: Exponential file size distribution

Offered load (as configured)	H	packets lost	bytes lost	Throughput (Kbps)	Measured load	Total No. of files transferred
0.1	0.54	0	0	13.30	0.10	547
0.2	0.49	0	0	13.25	0.21	1112
0.3	0.53	0	0	12.76	0.31	1659
0.4	0.53	0	0	12.19	0.41	2191
0.5	0.57	5	6742	11.46	0.51	2701
0.6	0.58	16	24224	10.53	0.60	3193
0.7	0.55	39	56014	9.38	0.69	3684
0.8	0.64	137	199986	8.18	0.78	4125
0.9	0.70	897	1310032	6.29	0.86	4556

Table 2: Gamma file size distribution

parameter is close to 0.5 only for low loads, but increases for higher loads.

Offered load (as configured)	H	packets lost	bytes lost	Throughput (Kbps)	Measured load	Total No. of files transferred
0.1	0.96	63	90305	10.81	0.13	579
0.2	0.94	104	150280	10.67	0.23	1180
0.3	0.94	161	228636	10.05	0.31	1754
0.4	0.93	263	379832	8.91	0.42	2318
0.5	0.95	648	907225	6.99	0.59	2851
0.6	0.95	863	1221730	5.41	0.71	3360
0.7	0.95	1420	2008299	4.24	0.80	3873
0.8	0.94	2006	2831271	3.38	0.85	4408
0.9	0.93	2305	3251933	2.76	0.89	4764

Table 3: Pareto file size distribution, $\alpha = 1.1$

While the Hurst parameter is a useful, parsimonious quantification of long range dependence in processes, it is also known that the estimation of H is fraught with difficulties. We have chosen a simple estimation technique (the variance-time plot), for which there is the

Offered load (as configured)	H	packets lost	bytes lost	Throughput (Kbps)	Measured load	Total No. of files transferred
0.1	0.94	42	60858	11.68	0.13	565
0.2	0.89	94	138727	11.62	0.21	1157
0.3	0.91	163	235380	10.78	0.31	1718
0.4	0.91	256	365646	9.96	0.40	2265
0.5	0.94	549	788322	7.71	0.59	2781
0.6	0.94	928	1318624	6.33	0.69	3285
0.7	0.95	1497	2131505	4.81	0.79	3796
0.8	0.94	2065	2942136	3.80	0.85	4312
0.9	0.92	2896	4123049	2.97	0.90	4748

Table 4: Pareto file size distribution, $\alpha = 1.2$

Offered load (as configured)	H	packets lost	bytes lost	Throughput (Kbps)	Measured load	Total No. of files transferred
0.1	0.78	11	16654	13.65	0.11	555
0.2	0.79	64	96043	13.04	0.21	1133
0.3	0.81	128	189610	12.22	0.31	1688
0.4	0.77	170	255783	11.50	0.40	2216
0.5	0.83	398	586872	9.93	0.54	2738
0.6	0.85	622	911215	8.88	0.63	3243
0.7	0.85	914	1344548	7.59	0.72	3737
0.8	0.85	1463	2158562	6.10	0.81	4253
0.9	0.85	2580	3794063	4.59	0.88	4695

Table 5: Pareto file size distribution, $\alpha = 1.6$

problem of determining where on the plot the asymptotic behaviour "starts". Even more sophisticated estimators, such as the maximum likelihood based Whittle estimator, are known to give widely different results for the same data depending on the modelling assumptions made (see [2, page 118-119]). Further investigations are needed to better reconcile the above test-bed results with the theory.

4.4 Average File Transfer Throughput

We first need to define some more notation. Consider several files being transferred over the network, indexed by $k = 1, 2, \ldots$. For the kth file, let V_k denote the volume of data (say in Bytes), and let W_k denote the time taken to transfer the file. At time t let $N(t)$ denote the number of transfers in progress. In the PS model, $\{N(t)\}$ is the "queue" length process. Then the average file transfer throughput is defined as

$$\tau = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \frac{V_k}{W_k} \tag{14}$$

For the PS model, letting (V, W) denote the stationary random vector of the file volume and the corresponding transfer time, it is clear that

$$\tau = E\left(\frac{X}{W}\right)$$

Even for the simple M/G/1 PS model, calculation of τ is difficult. However, in an experiment, estimation of τ via its definition in Equation 14 is straightforward.

Analytically, a more tractable measure is the average active session bandwidth share (see also [9]) defined as follows

$$\sigma = \lim_{t\to\infty} \frac{\frac{1}{t}\int_0^t \frac{1}{N(u)} I_{\{N(u)\geq 1\}}\, du}{\frac{1}{t}\int_0^t I_{\{N(u)\geq 1\}} du} \tag{15}$$

Letting $\pi(n)$ denote the stationary probability of n file transfers being in progress, it is easily seen that

$$\sigma = \frac{1}{\rho} \sum_{n=1}^{\infty} \frac{\pi(n)}{n}$$

For an M/G/1 PS queue, we have (independent of the particular file size distribution)

$$\pi(n) = (1-\rho)\rho^n$$

It is then easily seen that,

$$\sigma = \frac{1}{\rho} \sum_{n=1}^{\infty} \frac{\pi(n)}{n} = \frac{1-\rho}{\rho} \ln \frac{1}{1-\rho} \tag{16}$$

In general, for a given M/G/1 PS model, τ and σ are different, and whereas σ is insensitive to file size distribution, τ is sensitive (see [10]). We have found, however, that these measures are quite close to each other. Thus we will use τ when displaying experimental results, and σ for the theoretical results from the PS model.

Figure 4 shows τ and σ normalised to the link rate, and plotted against the link occupancy (ρ). The solid curve shows σ obtained for the PS model, using Equation 16. The other curves show τ obtained from the test-bed, for various file size distributions; with reference to Tables 1, 2, 3, 4, and 5, we are basically plotting column 5 versus column 6. The average file size is 10KB, and the distributions studied are exponential, gamma (order 2), uniform, and Pareto (with $\alpha = 1.1, 1.2, 1.6$). We notice that whereas for exponential, gamma, and uniform, the throughputs obtained from the PS model are close to those obtained from the experiment, for Pareto file sizes the throughputs become progressively smaller as α approaches 1. We note that the droop in throughput for low loads is probably because of a measurement problem. We are using the throughput measurements provided by the Linux operating system; when the throughputs are large, the file transfer times are small, and the granularity of the time measurement affects the results.

The above throughput results should be viewed in light of the packet loss data shown in Tables 1, 2, 3, 4, and 5. Notice that packet loss rates are much higher for Pareto file size distributions as compared to the light tailed distributions exponential and gamma. TCP reacts to packet losses by dropping its window, and hence the sending rate; thus packet losses are detrimental to TCP throughput.

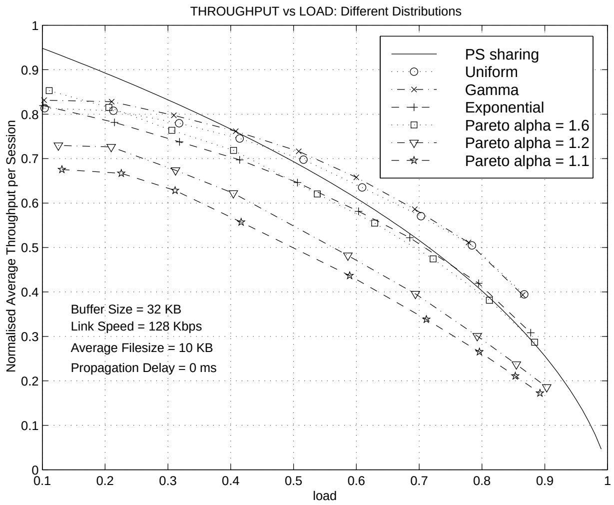

Figure 4: Average fair bandwidth share per session from the PS model, and average session throughput from the test-bed, plotted versus the link load.

5 Conclusions

We have studied the Processor Sharing queue as a model for TCP controlled bandwidth sharing at a single link. We related the departure process of the PS queue to the departure process of the TCP controlled flows from the link. We have shown that for service time distributions with finite second moment, the departure process of the PS queue does not display long range dependence. For Pareto distributed service times, with tail parameter α, however, the departure process of the PS model is LRD, with $H = \frac{3-\alpha}{2}$.

We then conducted experiments on a test-bed with actual TCP code and a link emulator. We made the following observations:

- For Pareto file size distribution with $\alpha = 1.1$ the measured H parameter is close to 0.95, as predicted by the theory.
- For other file size distributions the measured H parameter does not match very well with theory, but as the Pareto tail parameter, α, increases the H parameter decreases.
- The measured throughputs for Pareto distributed file sizes are sensitive to the parameter α, the throughput decreasing as the distribution becomes more heavy tailed.
- There are significantly more packet losses with heavy tailed file size distributions, than with light tailed ones.

In conclusion, we can say that whereas the Processor Sharing model is a useful analytical tool for analysing bandwidth sharing for elastic sessions in packet networks, and it has some predictive capability even for TCP controlled flows, the PS model needs to be used judiciously. An important research effort will be to provide corrections to the results obtained from the PS model in order to better predict the performance of TCP controlled bandwidth sharing (see, for example, [9]).

A Appendix: Proof of Theorem 2

Proof: Define $g(t) = f(t + t_0)$. We have $g(t) > 0$ for $t > 0$. Also

$$\begin{aligned}
\tilde{g}(s) &= \int_0^\infty e^{-st}\, g(t)dt \\
&= e^{+st_0} \int_0^\infty e^{-s(t+t_0)}\, f(t+t_0)\, dt \\
&= e^{st_0} \int_{t_0}^\infty e^{-su}\, f(u)du \\
&= e^{st_0} \Big(\tilde{f}(s) - \int_0^{t_0} e^{-su}\, f(u)du\Big)
\end{aligned}$$

Now consider

$$\begin{aligned}
\frac{\tilde{g}(s)}{(A/s^\gamma)} &= \frac{e^{st_0}\tilde{f}(s)}{(A/s^\gamma)} - \frac{e^{st_0} \int_0^{t_0} e^{-su} f(u)du}{(A/s^\gamma)} \\
&\longrightarrow_{s\to 0^+} \quad 1
\end{aligned}$$

Hence applying Theorem 5.3 from Widder [16], we have

$$\int_0^t g(u)du \sim \frac{At^\gamma}{\Gamma(\gamma+1)} \quad t \to \infty$$

Now consider, for t large enough,

$$\begin{aligned}
\frac{\int_0^t f(u)du}{\frac{At^\gamma}{\Gamma(\gamma+1)}} &= \frac{\int_0^{t_0} f(u)du + \int_{t_0}^t f(u)du}{\frac{At^\gamma}{\Gamma(\gamma+1)}} \\
&= \frac{\int_0^{t_0} f(u)du}{\frac{At^\gamma}{\Gamma(\gamma+1)}} + \frac{\int_0^{t-t_0} g(v)dv}{\frac{A(t-t_0)^\gamma}{\Gamma(\gamma+1)} \frac{t^\gamma}{(t-t_0)^\gamma}} \\
&\longrightarrow_{t\to\infty} \quad 1
\end{aligned}$$

Hence

$$\int_0^t f(u)du \sim \frac{At^\gamma}{\Gamma(\gamma+1)}$$

$\square$

References

[1] A. Anvekar, "WALE: A Wide Area Link Emulator on a Linux PC," *ERNET Project Report*, No. ERNET-IISc-1999.1, Indian Institute of Science, Bangalore, June 1999. See also http://ece.iisc.ernet.in/netlab/

[2] J. Beran, *Statistics of Long-Memory Processes*, Chapman and Hall, New York, 1994.

[3] A.W. Berger and Yaakov Kogan, "Dimensioning Bandwidth for Elastic Traffic in High-Speed Data Networks," manuscript submitted for publication, 1998.

[4] D.R. Cox, *Renewal Theory*, Methuen Science Paperback, London, 1967.

[5] D.R. Cox, "Long-Range Dependence: A Review", in: *Statistics An Appraisal*, ed. H.A. David and H.T. David, Iowa State University Press, pp 55-74, 1984.

[6] M.E. Crovella and A. Bestavros, "Self-Similarity in World-Wide Web Traffic: Evidence and Possible Causes," *ACM SIGMETRICS '96*, pp 160-169, 1996.

[7] A. Erramilli, O. Narayan, and W. Willinger, "Experimental Queueing Analysis with Long-Range Dependent Packet Traffic," *IEEE Transactions on Networking*, Vol. 4, No. 2, pp 209-223, April 1996.

[8] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press, Orlando, Florida, 1990.

[9] D.P. Heyman, T.V. Lakshman, A.L. Neidhardt, "A New Method for Analyzing Feedback-Based Protocols with Applications to Engineering Web Traffic over the Internet," *Performance Evaluation Review*, Vol. 25, No. 1, *Proc. ACM SIGMETRICS '97*, pp 24-38, 1997.

[10] A.A. Kherani, "Rate Based Congestion Control for Ephemeral Best-Effort Sessions in Packet Networks," Master of Engg. Project Report, ECE Department, Indian Institute of Science, Bangalore, January 1999.

[11] L. Kleinrock, *Queueing Systems: Volume 1*, John Wiley, New York, 1976.

[12] L. Kleinrock, *Queueing Systems: Volume 2*, John Wiley, New York, 1976.

[13] W.E. Leyland, M.S. Taqqu, W. Willinger, and D.V. Wilson, "On the Self-Similar Nature of Ethernet Traffic (Extended Version)," *IEEE Transactions on Networking*, Vo.l 2, No. 1, pp 1-15, February 1994.

[14] V. Paxson and S. Floyd, "Wide Area Traffic: The Failure of Poisson Modelling", *IEEE/ACM Transactions on Networking*, Vol. 3, No. 3, pp 226-244, 1995.

[15] J.W. Roberts and L. Massoulie, "Bandwidth Sharing and Admission Control for Elastic Traffic," *ITC Specialists Seminar*, Japan, 1998.

[16] D.V Widder, *An Introduction to Transform Theory*, Academic Press, New York, 1971.

[17] W. Willinger, V. Paxson, and M.S. Taqqu, Self-Similarity and Heavy Tails: Structural Modelling of Network Traffic, in: *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, Robert Adler, Raisa Feldman, Murad S. Taqqu, editors, Birkhauser, Boston, 1998.

[18] R.W. Wolff, *Stochastic Processes and the Theory of Queues*, Prentice Hall, New Jersey, 1989.