Studies on the Impact of Long-Term Correlation on Computer Network Performance: Part II - Transport-layer Modeling

Hany D. Alsaialy and John A. Silvester

CENG 98-27

Department of Electrical Engineering - Systems
University of Southern California
Los Angeles, California 90089-2562
(213-740-4579)
December 1998

Studies on the Impact of Long-Term Correlation on Computer Network Performance: Part II - Transport-layer Modeling

Hany D. Alsaialy John A. Silvester
<alsaialy, silvester>@usc.edu

Department of Electrical Engineering-Systems
University of Southern California
Los Angeles, CA 90089-0251

CENG 98-27

Abstract

Several research studies have shown the omnipresence of Long-Range Dependence (LRD) in Local Area Network (LAN) traffic, Wide Area Network (WAN) traffic, and Variable-Bit-Rate (VBR) video traffic [BSTW95], [LTWW94], [LTWW95].

In this paper, we present a mechanism for the modeling of the Internet's World Wide Web (WWW or Web) traffic based on the superposition of independent and identically distributed (i.i.d.) point processes. We describe the Superposition of Fractal Renewal Processes (Sup-FRP) traffic model used for modeling Web request arrivals. We show that neglecting the correlation found in real Web requests will lead to inaccurate performance evaluations. We show the performance impact of correlated Web requests on throughput and average response time, and compare our findings with previously reported results.

1 Introduction

Simulation of computer networks is considered an efficient tool used in analysis and performance evaluation. In today's computer networks (e.g. the Internet), a range of users generate different types of traffic, and expect different types of response from the underlying network (i.e. each with a different set of rules to specify their QOS). In traffic modeling for simulation it is, hence, important to choose the best models that mimic the behavior of real users to ensure a correct analysis from a simulation point of view. This enables us to correctly predict the performance of a network prior to its deployment or redesign.

The concept of Long-Range-Dependence (LRD) initially reported in [LTWW94] opened new challenges in the engineering of computer network traffic. For many years researchers relied on Markovian models for network performance prediction, however, these models are inherently Short-Range Dependent (SRD). Other recent studies also revealed the self-similarity or "fractal" nature of streams collected over the Internet [Ryu98]. But what are the implications of LRD found in these streams from an engineering point of view, and why it is important to find new models that capture these features? We summarize the answer as follows:

1. The cell/packet loss probability in network queues decays faster when Markovian types of traffic models are used compared to comparable "fractal" models (i.e. with similar first and second order statistics).

2. The distribution of asymptotic queue size (i.e. maximum queue occupancy) decays faster in Markovian models.

These are some of the reasons encouraging researchers to gain a better understanding and find appropriate mathematical models that capture the behavior of real computer networks streams.

We just mentioned the effects of using inaccurate performance models (e.g. Markovian), the effects in a real network from a user perspectives are increased response times caused by either the very long queues, or, for reliable protocols (e.g. TCP), the retransmissions of dropped packets. For protocols such as ATM a degradation in the QOS (e.g. increased cell-loss rate) will result.

Having traffic models that correctly mimic real computer network traffic streams, allows us to generate a variety of synthesized streams that can be used in simulation; So that we may obtain correct analysis and performance evaluation of computer networks.

We classify the modeling and simulation of computer networks into two main categories: i) Link-level, and ii) Application or Transport-level. In the former, models are constructed and model-parameters are matched from a previously collected stream or group of streams. These models can be later used to generate synthesized streams in simulation. The latter approach is similar, however, the modeling is performed above the link-layer. In other words, while in the former method we usually analyze long traces of data at the packet or cell level, in the former approach we look at either user behavior, or more generally, higher-layer protocols (e.g. file size distributions, request lengths, etc.). In this paper, we focus on the latter aspect, specifically, we analyze the behavior of the popular World Wide Web (also known as the Web or WWW).

2 Related Work

Several studies have attempted to understand how self-similarity arises in computer networks. In [PKC96] the ON/OFF model described in [WTSW95] was used simulate multiple client-server sessions emulating the behavior of Web traffic. [CB96] suggested that the reasons for traffic self-similarity can be attributed to the heavy-tailed nature of file size distributions available on the Web. [PF94] showed similar results modeling FTP bursts. Other results focused on the modeling via Cumulative-Distribution Functions (CDFs) estimated by empirical results, see [DJ91], [Edd96], [Mah97]. [Mah97] estimated several empirical CDFs to model HTTP traffic. The CDFs capture parameters of Web client/server behavior such as HTTP request/reply lengths, documents sizes and user think time. All of these studies require the use of a TCP algorithm to get packet-level results since the modeling is performed above the transport layer.

Before proceeding, we identify alternatives to model-driven simulations, and describe their major drawbacks:

1. Trace-driven simulation:

   One way to simulate real traffic is by using a trace or a group of traces collected from real computer networks. The three main drawbacks to this approach are:

   (a) A trace represents only a particular instance of "*history*".

   (b) The simulation is limited by the "*length*" of the trace.

   (c) We may need a large number of traces to accurately verify the simulation results.

2. Empirical-driven simulation:

   In this method, an empirical cumulative-distribution-function (CDF) is defined from a collection of real traffic traces, and hence used to generate random events in the simulation. This method too has its drawbacks which we identify briefly in the following two points:

   (a) In many cases the empirical CDF captures only first-order statistics (distribution). However, higher order statistics may be important.

   (b) Values of the random variables are "*bounded*" by the minimum and maximum values used to generate the empirical CDF. Large values, however, are not unlikely and have higher probability than previously thought (e.g. heavy-tailed distribution).

Even though empirical driven simulations can be considered of a somewhat more realistic approach than trace-driven simulations, the drawbacks mentioned suffice to motivate researchers to look for alternative methods in their simulations to insure realistic results.

The use of validated traffic models for simulation not only simplifies the simulation itself but can improve the simulation by giving more realistic results and permitting efficient and accurate network performance evaluation under a wide variability of scenarios. We briefly point out some advantages associated to this approach:

1. Virtually no limit in the *length* of the simulation.

2. A wider variability of scenarios can be simulated since traffic models can be adjusted to given input parameters with specific traffic characteristics.

3 Issues and Traffic Model Proposed

In previous sections, we briefly mentioned that today's computer networks include a wide variety of heterogeneous traffic types. These traffic not only differ in nature, but users differ in the quality of service (QOS) they expect. It is of little practical use to propose a generic traffic model; we, therefore, focus on traffic generated by the popular World Wide Web (WWW or Web), with emphasis on the request arrival process.

3.1 Web Workload traffic (for request arrivals and volumes)

Traffic generated by the Web is considered the leading source of backbone network traffic found in today's Internet [Mah97]. WWW uses the Hypertext Transfer Protocol (HTTP) as its application layer protocol. HTTP uses the Transmission Control Protocol (TCP) as its transport layer protocol which is a *guaranteed* delivery protocol.

While users of this type of traffic tolerate delays (compared to other Internet users such as users of real-time applications), there is little, if any, tolerance to loss of data. This implies that the network can be operated at wider ranges of utilization so long as data is not lost (this is, of course, guaranteed by TCP). We emphasize the possibility of operating the network at relatively high utilization levels (recall there is an inverse relationship between utilization and delay).

From analyzing a NASA and a Berkeley Web trace, [Ryu98] found that the request arrivals are well modeled by a superposition of fractal renewal processes (Sup-FRP). Compared to previous related work (e.g. [PKC96]), we show that ignoring second-order statistics in the Web arrival process will lead to inaccurate performance results in terms of response-time and packet loss.

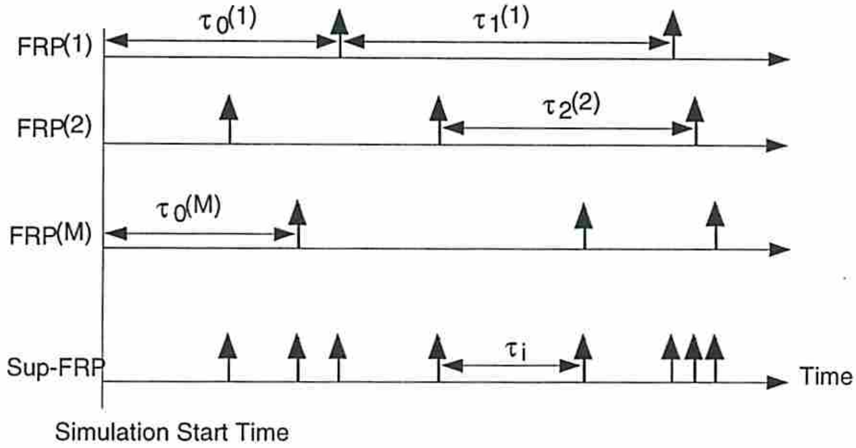

Figure 1: The Sup-FRP process shown graphically. Each FRP is i.i.d. (M=3 in the figure).

3.2 The Superposition of Fractal Renewal Processes Model (Sup-FRP)

The Sup-FRP is a process generated from "M" independent and identically distributed (i.i.d.) fractal renewal processes (FRPs). Each FRP is defined by the following pdf:

$$p(t) = \begin{cases} \gamma A^{-1} e^{\frac{-\gamma t}{A}} & \text{for } t \leqslant A \\ \gamma e^{-\gamma} A^{\gamma} t^{-(\gamma+1)} & \text{for } t > A \end{cases} \tag{1}$$

With $1 < \gamma < 2$. The parameter A serves as a threshold between exponential behavior and power-law behavior of interarrival times. Figure 1 shows a graphical realization of the Sup-FRP.

In the Appendix we describe the algorithm used to generate the Sup-FRP arrival process in detail.

To fully describe the sup-FRP model we need the following three values to be known:

1. γ :the shape of the pdf.
2. A :the cut-off value of the pdf.
3. M :the number of i.i.d. FRPs.

What is usually known, or can be estimated, from a given stream of traffic are the following three parameters:

1. H :the Hurst parameter defining the degree of self-similarity.
2. λ :the average arrival rate.
3. T_0^{α} :the onset-time (i.e. the level of aggregation where the fractal behavior starts, see [RL96], [RL98]).

There is a relationship between traffic and model parameters, the following set of equations defines this relationship. Details can be found in [RL96], [RL98]:

$$\gamma = 2 - \alpha \tag{2}$$

$$H = \frac{\alpha + 1}{2} \tag{3}$$

$$\lambda = M\gamma[1 + (\gamma - 1)^{-1}e^{-\gamma}]^{-1}A^{-1} \tag{4}$$

$$T_0^{\alpha} = 2^{-1}\gamma^{-2}(\gamma - 1)^{-1}(2 - \gamma)(3 - \gamma)e^{-\gamma}[1 + (\gamma - 1)e^{\gamma}]^2 A^{2-\gamma} \tag{5}$$

It is fairly simple to solve this system of equations to derive the Sup-FRP parameters. From Eq. (3) we can find α. Solving Eq. (2) we get γ. Next Eq. (5) yields A, and finally we obtain the value of M from Eq. (4).

3.3 The Sup-FRP Match the Arrival Process

From the analysis of two Web traces from NASA and Berkeley, Ryu [Ryu98] matches the IDC curve (Index of Dispersion for Counts, also known as the Fano factor F(T)) of the Web request arrival process and the Sup-FRP process. Figure 2 verifies this match. The IDC is defined as the variance of the number of arrivals in a given time window of width T divided by the mean number of arrivals in T. Note that for a Poisson point process (i.e. exponentially distributed interarrivals), the IDC curve value is 1 over the entire range of time scales. Recall that both the mean and variance of the Poisson process are identical, and the resulting aggregate process is still Poisson [Kle75].

We, therefore, use the Sup-FRP process to model the Web request arrival process.

3.4 Heavy-Tailed Distributions: A Note

A random variable X follows a heavy-tailed distribution if the its complementary distribution (also known as the *survivor* distribution) has the form [AFT98]:

$$P[X > x] \sim x^{-\alpha}, \qquad \text{as } x \longrightarrow \infty,\ 0 < \alpha < 2 \tag{6}$$

An example is the Pareto distribution with probability density function:

$$f(x \mid \alpha, k) = \alpha k^{\alpha} x^{-(\alpha+1)} \qquad \alpha, k > 0,\ x \geqslant k \tag{7}$$

There are several properties associated with heavy-tailed distributions such as infinite mean and variance. The mean of the pareto pdf is:

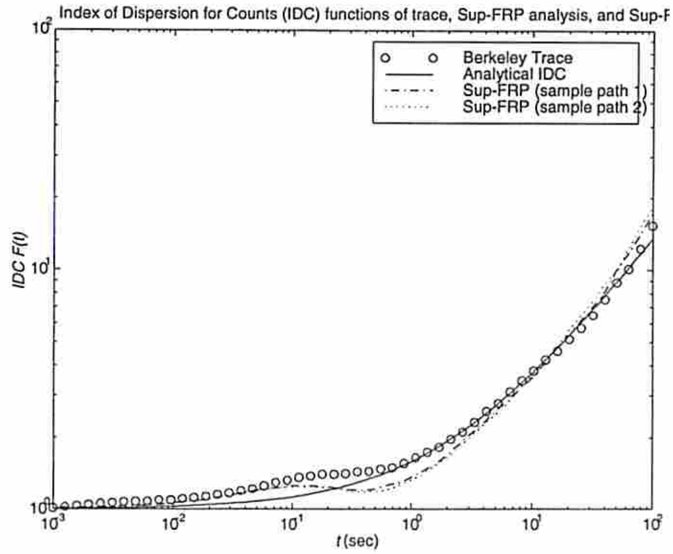

Figure 2: IDC match between the Sup-FRP and the Web request arrival process [Ryu98].

$$E(X) = \int_k^\infty x f\,(x \mid \alpha, k)\,dx = \frac{\alpha k}{\alpha - 1} \tag{8}$$

The variance is:

$$Var(X) = E(X^2) - E(X)^2 = \frac{\alpha k^2}{(\alpha-2)(\alpha-1)^2} \tag{9}$$

Hence, for $\alpha \leqslant 2$, the distribution has infinite variance, and for $\alpha \leqslant 1$, the distribution has also infinite mean.

The value H (the Hurst parameter), and the parameter α of the Pareto pdf are related by the equation[PKC97],[AFT98]: $H = (3 - \alpha)/2$. Therefore, for LRD (i.e. $.5 < H < 1$) we require $1 < \alpha < 2$. For $\alpha > 2$ the process is SRD (i.e. $H < .5$).

4 Network Model

The network model we use in our simulation is a fairly simple one. Figure 3 shows our simple two node topology.

The reason for this simple model is twofold; from a simulation point of view it allow us to investigate the behavior of the transport layer (TCP) and eliminate the effects of the routing protocol. From a modeling point of view it enables us to gain a better understanding of the effects of using the proposed arrival process. As shown in Figure 3, the request arrival process at node $G2$ initiates an FTP sessions. Note that an HTTP request may initiate several FTP sessions. The sup-FRP model proposed will be used to model the initiation (i.e. request arrival) of an FTP, not the arrival of an HTTP request. In other words, we model the FTP requests seen from the Web server at $G2$.

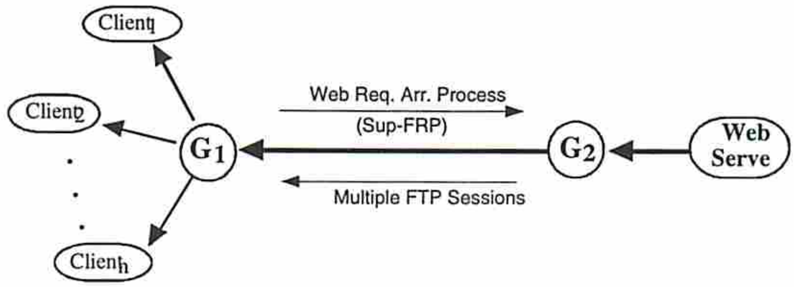

Figure 3: The two node simulation topology.

In our analysis we study the performance of the downstream traffic ($G2 \rightarrow G1$) for a wide range of buffer sizes. We also study network performance at different utilization levels (e.g. at different link speeds). By fixing the arrival process and changing the available capacity we investigate the effects of response time seen by the user.

4.1 Simulation Environment

To study the performance impact of the proposed traffic model on TCP[1], we used the LBNL Network Simulator (ns) [MF98]. ns is an event-driven simulator derived from S. Keshav's REAL network simulator. We modified ns by adding an implementation of the Sup-FRP model.

To measure performance, we recorded throughput for each FTP session:

$$Throuput = \frac{\textit{File Size (bits)}}{\textit{File Transmission Time (simulated seconds)}} \tag{10}$$

As described in [PD96], the throughput can be thought of the *achieved* bandwidth, compared to the *available* bandwidth.

For each simulation run, we average throughput over all FTP sessions. For a given scenario (i.e. combination of arrival process and file size distribution), we find the averaged throughput at different levels of network utilization ρ in the range $[0.1, 0.9]$. Utilization is found by:

$$\rho = \frac{\lambda F}{C} \tag{11}$$

where, λ = Average arrival rate (requests/sec), F = Average file size (bits), and C = Link capacity (bps).

To vary ρ, we adjusted the value of C and fixed λ, and F for a given run.

5 Simulation Results

We performed several experiments for different arrival processes (e.g. Exponential, and Sup-FRP) and file size distributions (e.g. Exponential, and Pareto). For a given experiment, we kept the

[1] Our focus here is to verify the general impact of the arrival process on network performance. We will, therefore, use only one of the available TCP flavours, namely, "Tahoe" TCP.

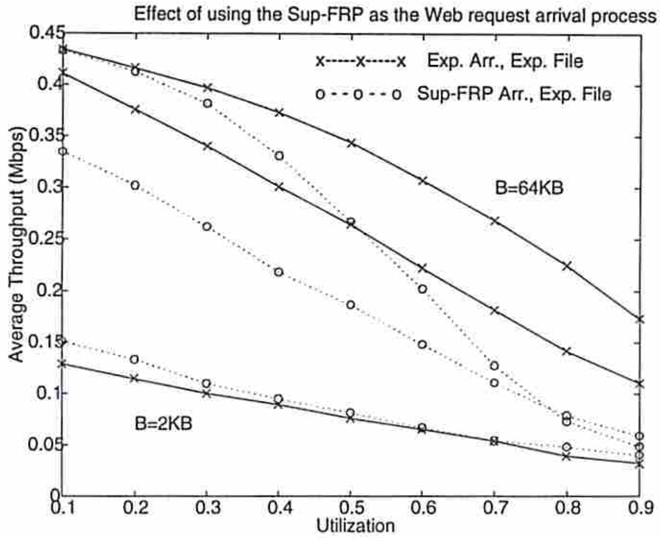

Figure 4: Effects from correlated arrivals get pronounced as the buffer size increrase (we used B=2KB, 4KB, and 64KB).

average arrival rate and mean file size equal to ensure a parsimonious comparison, in other words, whether the interarrival request was exponentially distributed or followed the Sup-FRP process, we used the same average arrival rate (for the Sup-FRP process we adjusted the onset-time (T_0^α) $\approx [.1, .3]$, and the Hurst parameter (H) $\approx .8$). The same for file sizes. Table 1 summarize values of link capacity (C) for different values of utilization (ρ). We used an average arrival rate (λ) of 10 arrivals/sec[2], and average file size (F) of 9.375KB. The $G2 \rightarrow G1$ link-delay was fixed at 20ms. In each simulation run, we generated 3000 arrivals (i.e. approximately 5 simulated minutes).

Utilization "ρ"	.1	.2	.3	.4	.5	.6	.7	.8	.9
Capacity "C" (Mbps)	7.5	3.75	2.5	1.875	1.5	1.2	1.07	.9375	.8333

Table 1: Network Utilization and equivalent Link Capacity.

In Figure 4 we see the effect of using correlated arrivals. We used two different arrival processes, one uncorrelated (exponentially distributed interarrivals) and the other correlated (Sup-FRP arrivals). To evaluate the effects of the arrival process on throughput and eliminate effects due to heavy-tailed file size pdf, we used exponentially distributed file sizes. We performed several simulation runs with buffer size of 2KB, 4KB, and 64KB. As the buffer size increase, we observe a pronounced effect of Sup-FRP on average throughput. It is interesting to see that strongly correlated arrivals not only lowered the achieved throughput, but had a stronger effect at higher utilization levels; we attribute this result to: a) increased buffer occupancy at higher utilization, and b) increased number of packets dropped, hence, increased transmissions times due to the retransmission of lost packets.

[2]The low arrival rate (e.g. $\lambda \cong 10$) was motivated to lower the number of superimposed FTP sessions. We believe that similar results consistent with our findings will be achieved for larger-scale simulations (e.g. $\lambda \cong 1000$).

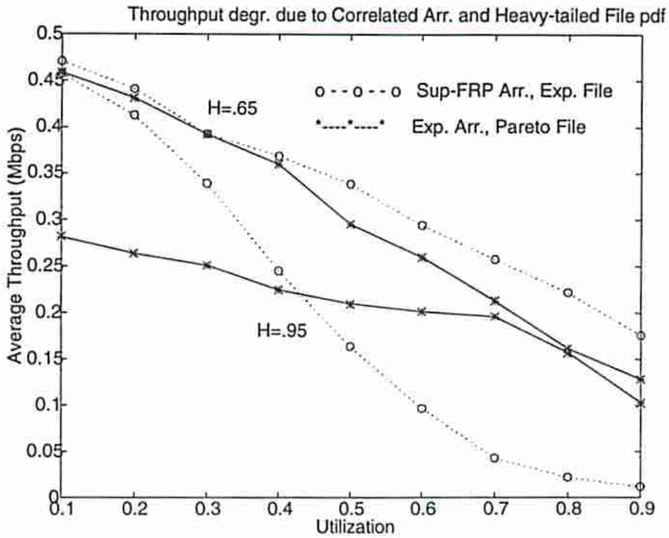

Figure 5: Effects of changing the Hurst parameter of the Arrival process and File size pdf on Throughput.

In the next experiment we investigated the performance impact using both a correlated arrival process, and a heavy-tailed file size distribution. As reported in [CB96], traffic self-similarity may be attributed to the heavy-tailed nature of file sizes found in the Web. Even though several studies (e.g. [PKC96]) relied on the findings of [CB96], we emphasize the importance of characterizing the nature of files *requested* (in contrast to files *found* or observed for multiple requests). In [AFT98] it is argued that the presence of caching in the Web has the effect of making the set of transmitted files relatively insensitive to the set of files requested, and distributionally similar to the set of available files.

In Figure 5, we compare the performance impact due to a heavy-tailed file size distribution and correlated request arrivals for different values of H (the Hurst parameter).

We observe the performance impact when file sizes follow a heavy-tailed distribution (e.g. Pareto), and notice that the effect is mainly at lower utilization levels (ρ in the range $[0.1, 0.5]$). The arrival process, as previously observed has a predominant effect at higher utilization levels (ρ in the range $[0.5, 0.9]$). Figure 6 summarizes the main results (for $H = .8$).

Since the observed Web request arrival process is well matched by the Sup-FRP process [Ryu98], and previous studies reported that Web file sizes are well modeled by the pareto pdf [PKC96], [Mah97], [AFT98]; we analyze the performance impact using the Sup-FRP process to simulate the Web request arrival process, and the Pareto pdf to simulate Web volumes. Figure 7 reveals an enormous degradation in performance due to both phenomena.

In Figure 8, we plot the average FTP session delay recorded. The Sup-FRP resulted in longer delays compared to delays predicted by both an uncorrelated arrival process (e.g. exponentially distributed interarrivals), and heavy-tailed file size distribution (e.g. Pareto). Our intuition attributes this observation to: a) overall increased queue lengths with the Sup-FRP arrival process, and b) increase in the number of dropped packets for most of the FTP sessions with the Sup-FRP arrival process. We also argue that the lower delays observed at higher utilization for Pareto distributed file sizes may be due to the high variability in the file sizes and increased number of

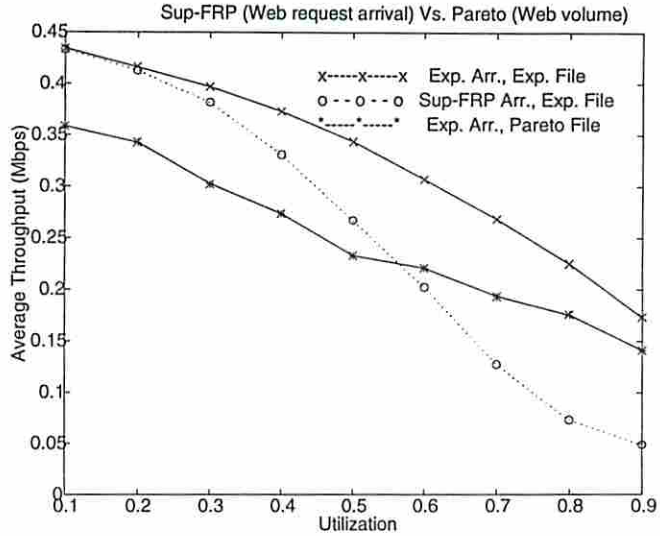

Figure 6: File size distribution impact performance at low utilization levels. The arrival process impact performance at high utilization level.

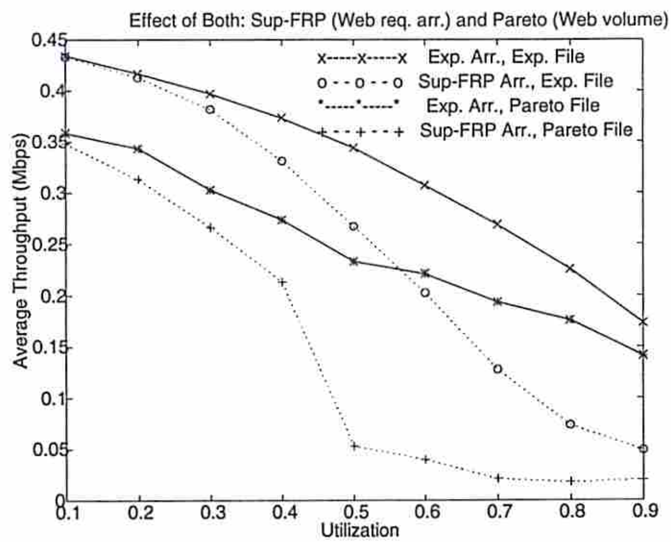

Figure 7: Both a strongly correlated arrival process and heavy-tailed file pdf result in high performance degradation.

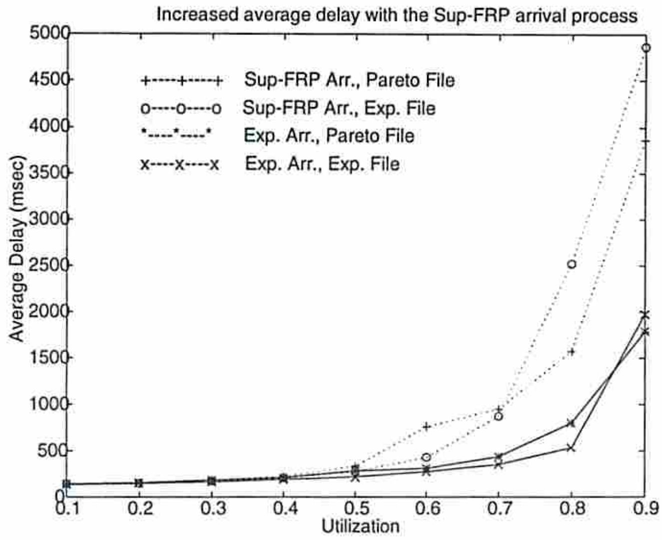

Figure 8: Average file transmission time increase with the Sup-FRP arrival process.

small files.

To investigate the results of Figure 8 we plot the delay variance for all FTP sessions. Figure 9 shows the variance plotted on a log scale. We observe that using a heavy-tailed file size pdf introduces high variability at lower utilization levels, compared to the Sup-FRP process which affects utilization at higher levels.

In Table 2, we summarize our results. We describe the effects of correlated arrivals (e.g. Sup-FRP) and the heavy-tailed distributed file sizes (e.g. Pareto) on throughout and response time. We compare the effects of both phenomena with uncorrelated arrivals and exponentially distributed file sizes.

<<Effects>>	Correlated Arrivals (Sup-FRP)	Heavy-Tailed File Dist. (Pareto)
Average Throughput	degradation at higher util. levels ($\rho \cong [0.5, 0.9]$).	degradation at lower util. levels ($\rho \cong [0.1, 0.5]$).
Average Response Time	increases faster from lower utilization levels ($\rho \cong .65$).	only little incr. at lower util. due to incr. no. of small files.
Response Time Variance	increased at higher utilization levels ($\rho \cong [0.5, 0.9]$).	higher over the entire range of utilization.

Table 2: Effects of correlated arrivals and heavy-tailed file size pdf on performance: Summary of Results.

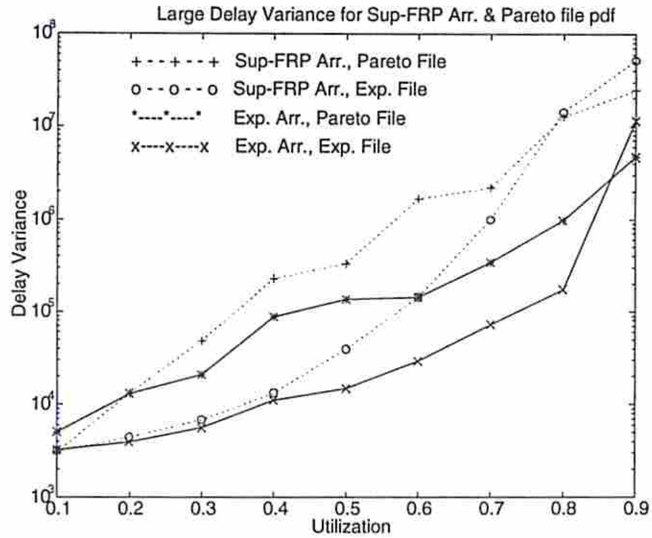

Figure 9: Variance in file transmission time increase with both the Sup-FRP arrival process and Heavy-tailed file pdf.

6 Conclusion

We analyzed the performance impact in Internet Web traffic due to correlated arrivals and heavy-tailed file size distribution. We used the Sup-FRP process proposed in [Ryu98] for modeling the Web request arrival process and compared the performance impact with uncorrelated arrivals (i.e. exponentially distributed interarrivals) and heavy-tailed file size pdf (e.g. Pareto). From several experiments, we show that ignoring the correlation found in the Web request arrival process will lead to inaccurate performance analysis.

As reported in [PKC96], [AFT98], and [Mah97], we verified that the heavy-tailed nature of files found in the Web lead to a performance degradation, furthermore, we show that the network performance impact due to heavy-tailed file size distribution is mostly at low utilization level. We also show that correlated Web requests (e.g. Sup-FRP) affects performance at higher utilization level.

Since Web users are tolerant to delays (compared to users of real-time application), we believe it is not unlikely to operate networks at high utilization levels; Therefore, we believe that the arrival process has an impact on performance and needs to be well characterized.

Another issue we believe deserves investigation is to compare transmission time distributions with the Sup-FRP arrival process, and both exponential and Pareto file size pdfs. [AFT98] shows that there does not seem to be strong sample correlation between file sizes and transmission times. We, therefore, believe that the arrival process may be the reason for the observed heavy-tailed nature in the distribution of transmission times. In addition, in order to fully validate the Sup-FRP model, we propose to compare our results to empirical studies previously reported (e.g. [Mah97], [DJ91]).

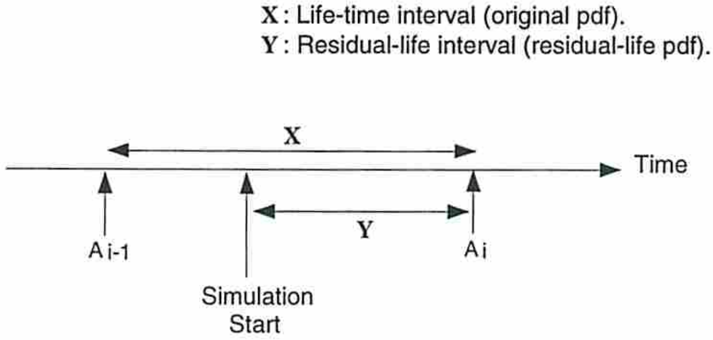

Figure 10: Life-time and Residual-life interval.

7 Appendix: Derivation of the Sup-FRP

We present a complete derivation for the Sup-FRP process described in Section 3.2.

The interarrival time of each Fractal-Renewal-Process (FRP) of Figure 1 is defined by the pdf:

$$p(t) = \begin{cases} \gamma A^{-1} e^{\frac{-\gamma t}{A}} & \text{for } t \leqslant A \\ \gamma e^{-\gamma} A^{\gamma} t^{-(\gamma+1)} & \text{for } t > A \end{cases} \tag{12}$$

We find the cumulative-distribution-function (CDF) for the interarrival time:
for $t \leqslant A$

$$F(t) = \int_0^t \gamma A^{-1} e^{\frac{-\gamma t}{A}} dt = 1 - e^{\frac{-\gamma t}{A}} \tag{13}$$

similarly for $t > A$

$$F(t) = \int_0^A \gamma A^{-1} e^{\frac{-\gamma t}{A}} dt + \int_A^t \gamma e^{-\gamma} A^{\gamma} t^{-(\gamma+1)} dt = 1 - e^{-\gamma} A^{\gamma} t^{-\gamma} \tag{14}$$

Using the inverse method and solving both equations we get:

$$T = \begin{cases} \frac{-A}{\gamma} \ln U & U \geq e^{-\gamma} \\ A e^{-1} U^{\frac{-1}{\gamma}} & U < e^{-\gamma} \end{cases} \tag{15}$$

Since the continuous pdf of (12) is not the exponential pdf (i.e. is not memoryless), we need to derive the distribution of the *residual-life* to model the first interval τ_0 of each FRP, see Figures 1 and 10. From [Kle75], the pdf of the residual-life is defined as follow:

$$p_{residual}(t) = \frac{1 - F(t)}{E(t)} \tag{16}$$

$$E(t) = \int_0^A t\gamma A^{-1} e^{\frac{-\gamma t}{A}} dt + \int_A^\infty t\gamma e^{-\gamma} A^\gamma t^{-(\gamma+1)} dt = \frac{A}{\gamma(\gamma-1)}(e^{-\gamma} + \gamma - 1) \tag{17}$$

the pdf of the residual life is then:

$$p_{residual}(t) = \begin{cases} \frac{\gamma(\gamma-1)e^{\frac{-\gamma t}{A}}}{A(e^{-\gamma}+\gamma-1)} & \text{for } t \leqslant A \\ \frac{\gamma(\gamma-1)e^{-\gamma}A^\gamma t^{-\gamma}}{A(e^{-\gamma}+\gamma-1)} & \text{for } t > A \end{cases} \tag{18}$$

As before, we find the CDF for the residual-life interval:
for $t \leqslant A$

$$F_{residual}(t) = \int_0^t \frac{\gamma(\gamma-1)e^{\frac{-\gamma t}{A}}}{A(e^{-\gamma}+\gamma-1)} = \frac{\gamma-1}{e^{-\gamma}+\gamma-1}(1 - e^{\frac{-\gamma t}{A}}) \tag{19}$$

similarly for $t > A$

$$\begin{aligned} F_{residual}(t) &= \int_0^A \frac{\gamma(\gamma-1)e^{\frac{-\gamma t}{A}}}{A(e^{-\gamma}+\gamma-1)} dt + \int_A^t \frac{\gamma(\gamma-1)e^{-\gamma}A^\gamma t^{-\gamma}}{A(e^{-\gamma}+\gamma-1)} dt \\ &= \frac{1}{e^{-\gamma}+\gamma-1}\left(\gamma - 1 + e^{-\gamma} - \gamma e^{-\gamma} A^{\gamma-1} t^{1-\gamma}\right) \end{aligned} \tag{20}$$

Using the inverse method and solving both equations we get:

$$T_{residual} = \begin{cases} -\gamma^{-1} A \ln[U + (U-1)(\gamma-1)^{-1} e^{-\gamma}] & V \geq 1 \\ AV^{\frac{1}{1-\gamma}} & V < 1 \end{cases} \tag{21}$$

where,

$$V = \frac{1 + (\gamma-1)e^\gamma}{\gamma} U \tag{22}$$

Hence, as shown in Figure 1, to generate Sup-FRP arrivals, we use $T_{residual}$ Eq. (21) to find $\tau_0^{(i)}$, and T Eq. (15) to find $\tau_j^{(i)}$, $\forall i = 1, 2, ..., M$ and $j > 0$.

Acknowledgments

The first author gratefully thank Dr. Bo Ryu from the Hughes Research Labs (HRL) for the initial work that lead to this report, and the partial support of the HRL Labs, Malibu, CA. The first author would also like to acknowledge the support of the Institute of Public Administration (IPA), Riyadh, Saudi Arabia.

References

[AFT98] R. J. Adler, R. E. Feldman, M. S. Taqqu, "A Practical Guide to Heavy Tails," Birkhauser, pp. 3-25, 1998.

[BSTW95] J. Beran, R. Sherman, M. S. Taqqu, and W. Willinger, "Long-Range Dependence in Variable-Bit-Rate Video Traffic," IEEE Trans. Commun., vol. 43, no. 2/3/4, pp. 1566-1579, Feb./March/April 1995.

[CB96] M. Crovella and A. Bestavros, "Self-similarity in World Wide Web traffic: Evidence and possible causes," In Proceedings of the ACM SIGMETRICS Conference on Measurement & Modeling of Computer Systems, pp. 160-169, Philadelphia, PA, May 1996.

[DJ91] Peter B. Danzig and Sugih Jamin, "tcplib: A library of TCP internetwork traffic characteristics," Technical Report USC-CS-91-495, Computer Science Department, University of Southern California, Los Angeles, CA, 1991.

[Edd96] Rusty Eddy, "HTTP analysis of IP level traces, Feb. 1996, Available at http://catarica.usc.edu/eddy/http-traffic/http-traces.html.

[Kle75] Leonard Kleinrock, "Queueing Systems, Volume I: Theory," John Wiley & Sons, 1975.

[LTWW94] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the Self-Similar Nature of Ethernet Traffic (Extended Version)," IEEE/ACM Trans. Networking, vol. 2, no. 1, pp.1-15, Feb. 1994.

[LTWW95] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "Self-Similarity in High-Speed Packet Traffic: Analysis and Modeling of Ethernet Traffic Measurements," Statistical Science, vol. 10, no. 1. pp. 67-85, 1995.

[Mah97] Bruce A. Mah, "An Empirical Model of HTTP Network Traffic," Proceedings of INFOCOM '97, Kobe, Japan, April 1997.

[MF98] S. McCanne and S. Floyd, "ns - Network Simulator," Available at http://www-mash.cs.berkeley.edu/ns/.

[PD96] Larry L. Peterson & Bruce S. Davie, "Computer Networks, A System Approach," Morgan Kaufmann Publishers, 1996.

[PF94] V. Paxson and S. Floyd, "Wide Area Traffic: The Failure of Poisson Modeling," In Proc. ACM SIGCOMM '94, pp. 257-268, 1994.

[PKC96] K. Park, G. T. Kim, and M. E. Crovella, "On the relationship between file sizes, transport protocols, and self-similar network traffic," in Proceedings of the Fourth International Conference on Network Protocols (ICNP '96), pp. 171-180, October, 1996.

[PKC97] K. Park, G. T. Kim, and M. E. Crovella, "On the Effect of Traffic Self-Similarity on Network Performance," CSD-TR 97-024, Dept. of Computer Sciences, Purdue University, July 1997.

[RL96] B. K. Ryu and S. B. Lowen, "Point process approaches to the modeling and analysis of self-similar traffic: Part I - Model constriction. In Proc. IEEE INFOCOM '96, San Francisco, CA, 1996.

[RL98] B. K. Ryu and S. B. Lowen, "Point process models for self-similar network traffic, with applications" Stochastic Models (M. Neuts, Ed.), vol. 14, Nr. 3, pp. 735-761, 1998.

[Ryu98] Bo Ryu, "Modeling and Simulation of Broadband Satellite Networks: Part II - Traffic Modeling," submitted to IEEE Communications Magazine, 1998.

[WTSW95] W. Willinger, M. Taqqu, R. Sherman, and D. Wilson, "Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level," In Proc. ACM SIGCOMM '95, pp. 100-113, 1995.