Long Range Dependence in Internet Backbone Traffic

L. Yao, M. Agapie, J. Ganbar, and M. Doroslovacki[♣]
Department of Electrical and Computer Engineering
The George Washington University
801 22nd St., NW
Washington, DC 20052

WorldCom, Inc.
22001 Loudoun County Parkway
Ashburn, VA 20147

Abstract—**We report the first statistical analysis of Internet backbone traffic, based on traces with levels of aggregation 10 times larger and timestamp accuracy 1000 times better than in previous studies. We analyze the first three moments, marginal distributions and correlation structures of packet size, packet inter-arrival time, byte count and packet count, and find that the highly aggregated Internet backbone traffic is still long-range dependent and self-similar. In fact, all time series examined (packet size, inter-arrival time, byte count, packet count) exhibit long-range dependency and self-similarity. In addition to the now-classical analysis at large time-scales (> 100ms), we report the first statistically relevant results on the short-term correlation ([50µs, 10ms]) of byte and packet count processes. We also study the fitness of various analytical models to the traffic traces. The empirical queuing analysis confirms the long-range dependence detected through direct analysis by showing that the queue behavior at high level of aggregation still diverges greatly from that predicted by Poisson model. As expected, statistical multiplexing gains improve the queuing performance, leading to economy of scale.**

I. Introduction

For Internet traffic engineering and core router design it is very important to understand the characteristics of Internet backbone traffic (e.g., the peak-to-mean ratio determines the maximum link utilization and buffer length required to meet certain loss objectives). Since the seminal study of Leland et al. [4], increasing evidence has emerged for the failure of traditional (Poisson-based) models to account for the long-range dependence (LRD) and self-similarity (SS) present in network traffic. In [3, 9], it was recognized that the high variability associated with LRD and SS can greatly deteriorate network performance, creating for instance queuing delays orders of magnitude higher than those predicted by traditional models.

LRD and SS have been reported in various types of data traffic: LAN [3, 4, 9], WAN [1, 6, 15], VBR video [14], SS7 control [12], WWW [2] etc. An overview of the above studies shows link speeds ranging from 10 Mbps (Ethernet) to 622 Mbps (OC-12), link type of WAN “access” (typically connecting a university campus or research lab to an ISP) or LAN, average bandwidths between 1.4 and 42 Mbps, minimum time-scale of 1ms in only one instance [15] – usually above 10ms or 100ms, and at most 6 orders of magnitude for time-scales. Lack of access to high-speed, high-aggregation links, and lack of devices capable of measuring such links have up to now prevented similar studies from being performed on Internet backbone links. In principle, traffic on the backbone could be qualitatively different from the types enumerated above, due to factors such as much higher level of aggregation, traffic conditioning (policing and shaping) performed at the edge, and much larger round-trip-time (RTT) for TCP sessions. Actually, some researches have claimed that aggregating Internet traffic causes convergence to a Poisson limit [1].

In this paper, we study eight traces collected on three different backbone links (OC-12 ATM and OC-48 POS) with mean arrival rates up to 714 Mbps, on time-scales from 50µs to 500s (thus covering 8 orders of magnitude). We analyze mean, variance, skewness, marginal distributions and correlation structures of packet size, packet inter-arrival time, byte count and packet count. As expected, we find that the traffic variability (measured by standard deviation (STD)/mean) decreases as the level of aggregation increases. However, highly aggregated traffic is still LRD and SS, this being shown directly through statistical analysis of the arrival process, and indirectly through its queuing effects. The high timestamp accuracy enables us to include in the analysis time-scales starting at 50 µs (20 times lower than the lowest reported so far). The auto-covariance functions for short lags ([50µs, 100ms]) are found to follow the same known behavior as for the long lags ([100ms, 500s]), i.e., $r(k)$ is hyperbolic, but the shape parameter β turns out to assume different values for the two ranges.

The rest of the paper is organized as follows. Section II gives an overview of eight traces and the device used to collect them. Section III presents statistical tools. Section IV presents the analysis of packet size, packet inter-arrival time, byte count and packet count. In Section V, we study the queuing performance by using the traces as inputs to a first-come-first-served (FCFS) queuing system. Section VI draws the conclusions.

II. Overview Of Measuring Device And Packet Traces

Data was gathered utilizing a device called optical carried monitor (OCxMON), based on a general purpose, x86, high-end PC with network interface cards and software that enables it to collect data from the wire. Frames are mirrored off the fiber optics wire using a passive optical splitter and passed to a specialized driver on the host. The operating system used in OCxMON is Linux (redhat 6.2) with a 2.2.13 kernel. For this study, two different monitoring speeds were

♣ This research was sponsored by WorldCom, Inc.

used: OC12 (622Mbps) and OC48 (2.4Gbps). All systems use identical software but a different driver for each interface.

In this study we were interested in packet arrival time and packet size. For OC12 network interface we utilized a 25 nanosecond clock and for the OC48 interface a 12.5 nanosecond clock. The clock value differs between every interface card due to different hardware. A timestamp was attached to each packet upon arrival at the interface, and then extracted at the application layer before packet information is written to disk. The monitoring system architecture follows libpcap [17] closely with the timestamp addition to the data structures. For maximum efficiency, packet timestamp and size were the only values written to disk. Due to a Unix shell limitation, the maximum file (trace) size was 2G bytes – this explains why the traces corresponding to higher loads have shorter duration (see Table I below).

Packets were collected from three distinct monitoring points at different time of day during several days. Point 1 is an OC12 ATM link between a customer facing aggregation ATM switch and a core router. Point 2 is an exchange point between two regional Autonomous Systems (ASs). Point 3 is an OC48 POS trunk in a regional aggregation hub.

TABLE I SUMMARY OF TRACES

Trace name	Link Speed [Mbps]	Arrival Rate [Mbps]	# of Pkts [M]	Mon. Point	Timestamp accuracy [ns]
OC12_a0	622	28.5	1.9	1	25
OC12_a1	622	33.7	1.18	1	25
OC12_v0	622	84.2	101.7	2	25
OC12_v1	622	108.4	124.0	2	25
OC48_0	2488	321.5	114.7	3	12.5
OC48_1	2488	476.4	114.2	3	12.5
OC48_2	2488	665.2	102.6	3	12.5
OC48_3	2488	714	100.8	3	12.5

III. STATISTICAL MODELS AND METHODS EMPLOYED

A. Skewness

The skewness of a distribution is defined as $\chi = E[(X-\mu)^3]/\sigma^3$, where μ is the mean and σ is the standard deviation. It is used as a measure of the lack of symmetry: a distribution with $\chi>0$ is compressed below mean but has a longer tail above the mean, while for $\chi<0$ the situation is reversed.

B. Goodness-of-fit Tests

Many studies report distributions of traffic random variables summarized by eye. Rarely do they report goodness-of-fit methodologies and results. However, we often find mathematical distributions that appeared to the eye to closely match the distributions of the random variables; but the distributions may not be proved "statistically valid" at a significance level of 5% or even 1% when running goodness-of-fit tests. We use the Kolmogorov-Smirnov (K-S) goodness-of-fit test as a goodness-of-fit metric, which is commonly used with continuous distributions.

C. Correlogram and Variance-Time Plot

There are vast literatures on long-range dependent (LRD) and self-similar (SS) stochastic processes. For a short introduction, the reader is directed to [14], and for in-depth presentations to [13]. We quickly recap some essential concepts in this and the next sections.

Let $\{X(t),\ t\geq 1\}$ be a discrete-time, 2^{nd}-order stationary process, with auto-covariance function $r(k) = cov(X_n, X_{n+k})$. $r(k)$ is estimated as

$$r(k) = \frac{1}{\sigma^2 (N-k)} \sum_{i=0}^{N-k} (X_i - \mu)(X_{i+k} - \mu)\,. \qquad (1)$$

If the covariance sequence is asymptotically hyperbolic ($r(k) \sim ck^{-\beta}$, $0<\beta<1$, c is a constant), we say the process is LRD. The correlogram is a plot of log[$r(k)$] as a function of log(k). For a LRD process, the correlogram is quasi-linear and β can be estimated as the slope of the curve.

A continuous-time process $\{Y(t)\}$ is called SS with Hurst parameter H if, for any factor a, the time-scaled process $Y(at)$ is statistically indistinguishable from the "space"-scaled process $a^H Y(t)$, i.e., all finite-dimensional distributions are identical. In a discrete-time setting, the time-scaling is replaced by aggregating the increment process $X(t) = Y(t) - Y(t-1)$ over non-overlapping blocks of size m: $X^{(m)}(i) = \frac{1}{m} \sum_{t=m(i-1)+1}^{mi} X(t)$. m directly determines the time-scale (by multiplying the time interval of the base series).

In practice, the SS property is used in a twice-weakened form: *2^{nd} order SS* only requires $r(k)=r^{(m)}(k)$ (rather than all finite-dimensional distributions to be equal); *asymptotic SS* only requires statistical equality as $m \to\infty$ (rather than for any m). By combining the above concepts, we obtain the property actually exhibited by all traces reported so far: *2^{nd} order asymptotic SS*. For $H\in(1/2,1)$, a *2^{nd} order asymptotic SS* process is LRD and $\beta = 2-2H$ [13].

Self-similarity (SS) can be detected, and the corresponding value of H can be estimated in many ways [7]. Besides the correlogram, estimators widely used are: rescaled range (R/S), periodogram, Whittle, wavelet and variance-time plot. Making use of the following property of an asymptotically SS process: $\log(Var[X^{(m)}]) \to \log(Var[X]) - \beta\log(m)$ as $m\to\infty$, the variance-time plot shows $\log(Var[X^{(m)}])$ as a function of $\log(m)$. For an SS process, the plot tends to be a straight line with slope $-\beta$.

D. Fractional Brownian Motion Traffic Model

Norros [5] defines the fractional Brownian Motion (fBm) model for an arrival process: $Y(t) = \lambda t + \sigma Z_H(t)$, where $Y(t)$ is the amount of traffic arrived in the time interval $[0,t)$, λ is the mean arrival rate, σ is the variance coefficient (typically defined as $\sigma^2=a\lambda$ with a a constant), and $Z_H(t)$ is the normalized fBm with Hurst parameter H. More specifically, $Z_H(t)$ is a Gaussian random process with zero mean and variance t^{2H}.

The fBm is (strictly) SS, and was proposed as an approximation to the 2nd order asymptotic SS processes encountered in practice. The advantages of using it over other models (e.g., fARIMA) are its parsimony (the model is identified by only three parameters), its SS for finite samples and the existence of an approximation for the queue length distribution [5] – one of a few analytical results available for LRD queuing theory.

IV. Statistical Analysis Of Traces

A. Packet Size

As reported in [8], the marginal distribution of the packet size is a multi-modal distribution with spikes at some commonly used packet sizes, such as 40-byte (TCP ACK) and 1500-byte (Ethernet). Fig. 1 shows the logarithm of the packet size distribution of trace OC48_3. Although the shape decreases slowly at large packet size, it is not a heavy-tailed distribution since the packet size is finite (limited by the Length field in IP header). Table II gives the mean packet size of each trace.

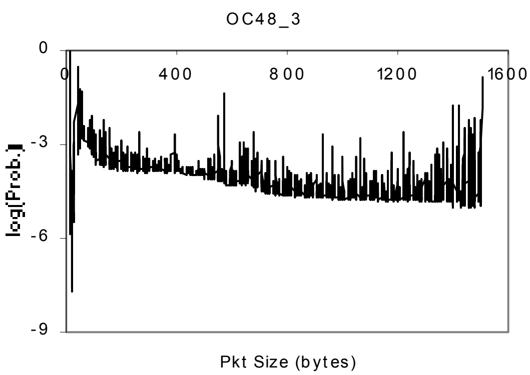

Fig. 1. Logarithm of packet size distribution

TABLE II Statistics Of Packet Size And Packet Interarrival Time

Trace	Mean Pkt Size [Bytes]	Mean Int-arrival Time [µs]	Std of Int-arrival Time [µs]
OC12_a0	421	118.2	1016.1
OC12_a1	583	138.2	1297.8
OC12_v0	555	52.8	54.1
OC12_v1	596	52.7	55.3
OC48_0	474	11.8	51.8
OC48_1	539	9.0	48.7
OC48_2	417	5.0	25.7
OC48_3	452	5.1	14.9

We found that the correlogram of the packet size series satisfies $0<\beta<1$ for all traces, hence the series are LRD. Fig. 2 shows the correlograms for traces OC12_a0 ($\beta=0.08$) and OC48_3 ($\beta=0.55$). At small lags, the correlogram of a high aggregate trace tends to decrease faster than that of a low aggregate trace. However, at large lags, the correlgrams of all traces decrease slowly ($\beta<0.1$).

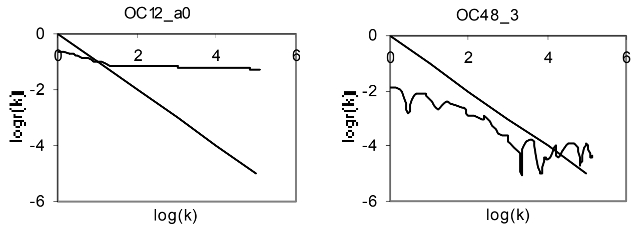

Fig. 2. Correlograms of the packet size series (Straight lines correspond to $\beta=1$)

B. Packet Inter-arrival Time

The packet inter-arrival time distribution typically looks like an exponential distribution or a Weibull distribution as claimed in [1]. For example, Fig. 3 shows the logarithm of the packet inter-arrival time distribution of trace OC48_3. However, even running the K-S test at 1% significant level, we failed to prove it to be statistically valid (for any of the traces). The spikes in the packet inter-arrival distribution are correlated to those in the packet size distribution due to back-to-back packets (as defined in [1]). That is, for back-to-back packets, the packet inter-arrival time is effectively equal to the size of the previous packet divided by the link speed.

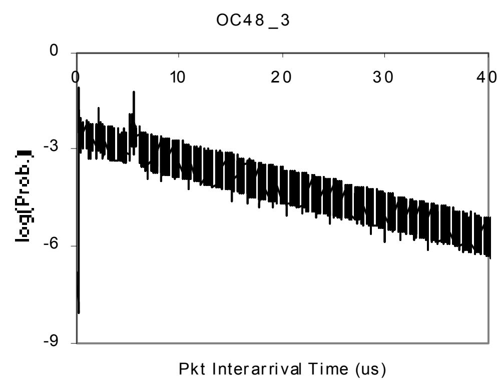

Fig.3. Logarithm of the packet inter-arrival time distribution

The correlogram of the packet inter-arrival time series of every trace satisfies $0<\beta<1$ for all traces, hence the series are LRD. Fig. 4 shows the correlograms for traces OC12_a0 ($\beta=0.20$) and OC48_3 ($\beta=0.45$). At small lags, the correlogram of a high aggregate trace also tends to decrease faster than that of a low aggregate trace. But at large lags, correlograms always decrease slowly ($\beta<0.1$).

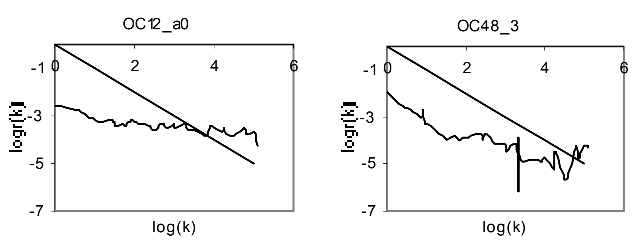

Fig. 4. Correlograms of the packet inter-arrival time serires (Straight lines correspond to $\beta=1$)

Table II lists the mean and standard deviation of the packet inter-arrival time of all traces. As expected, both decrease with increased aggregate level.

C. Byte Count and Packet Count

Since byte counts and packet counts share the same characteristics, we only give here the results for byte count. Table III presents the mean, standard deviation (std), coefficient of variation (std/mean) and skewness (χ) of the byte counts in 10ms time intervals. The std/mean for byte count and packet count decreases with increased aggregate level. A more detailed analysis (not shown) indicates that the variance changes approximately linearly with the mean, but this by no means implies that aggregating traffic causes it to smooth out as quickly as Poisson [1] - for example, the fBm model also predicts the same linear relationship.

The skewness for byte count and packet count shifts from positive to negative values, which implies that the upper tails of the marginal distributions get lighter as the level of aggregation increases. This can be also seen directly on the histograms in Fig. 5.

TABLE III MEAN, STDDEV AND SKEWNESS OF BYTE COUNT IN 10 ms INTERVALS

Trace	Mean (Bytes)	Std (Bytes)	Std/Mean	χ
OC12_a0	35,584	16,213	0.46	1.29
OC12-a1	42,141	18,158	0.43	1.20
OC12_v0	105,196	23,265	0.22	0.51
OC12_v1	135,526	25,750	0.19	0.56
OC48_0	401,859	50,412	0.13	-0.79
OC48_1	595,469	51,940	0.09	-3.47
OC48_2	831,441	91,222	0.11	-2.19
OC48_3	892,533	64,629	0.07	-3.37

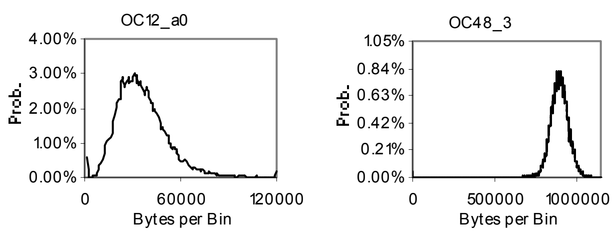

Fig.5. Marginal distributions of byte counts in 10ms time intervals

The marginal distributions of byte counts and packet counts converge to Gaussian with increased aggregation. Indeed, the K-S tests (at 5% significant level) for byte counts in 100ms time intervals have been successful for all OC48 traces while only OC48_2 and OC48_3 have passed the same test for 10ms intervals.

Turning to the correlation structure, we first note that all eight traces exhibit 2nd order asymptotic SS. An example of correlograms converging with increasing timescale is given in Fig. 6. Note that the three correlograms are roughly parallel with a lot of detail conserved and each curve has a slope between -1 and 0, which means $0<\beta<1$.

Since all variance-time plots of OC12 traces are very close to each other, and the same is true for OC48 traces, in Fig. 7 we show only a representative of each. The asymptotic slopes yield regression estimates of $H\in(0.91,\ 0.96)$ (i.e., $\beta\in(0.1, 0.18)$) for the OC12 traces, in accordance with the observation widely reported that highly aggregated traffic has a high degree of SS. For OC48 we get unexpectedly low estimates of $H\in(0.84, 0.89)$ (i.e., $\beta\in(0.22, 0.32)$), but this is because these traces cover a shorter duration, due to an increased number of packets; the highest timescale used is only 200s (compared to 500s for OC12).

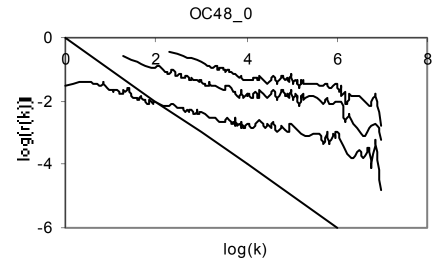

Fig. 6. Correlograms of byte counts series for timescales of 0.1 ms (bottom), 1ms (middle) and 10ms (top). The straight line corresponds to $\beta=1$.

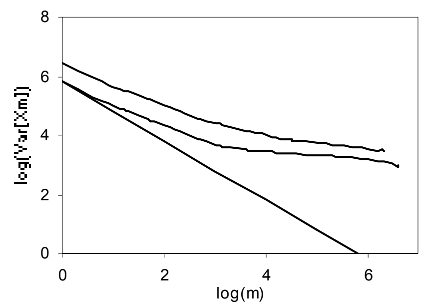

Fig. 7. Variance-time plots of the byte counts for OC12_v1 (bottom) and OC48_1 (top). *1* on x-axis is 50us.The straight line corresponds to $\beta=1$.

It is has been known for a while that the short-term correlation structure of traffic is quite different from the long-term one [16], but limited timestamp accuracy prevented researchers to explore it in detail. We report that all variance-time plots (and correlograms) also exhibit *linear* behavior at ***small*** time-scales, but β for small time-scales (i.e., (0.65, 0.71) for OC12 and (0.55, 0.61) for OC48) is considerably larger than that for large time-scales. Moreover, it seems that the point of transition between the two slopes moves towards the right with increased aggregate level (e.g., from around 100ms for OC12 to above 1s for OC48).

V. EMPIRICAL QUEUING ANALYSIS

We use OC12 and OC48 traces as inputs to a FCFS queuing system with infinite buffer. The simulation results are then compared with theoretical results predicted by M/M/1 model [11] and fBm model [5].

Fig.8 shows that the Poisson-based model predicts queue lengths typically orders of magnitude lower than those resulted from either simulation. At utilizations under 70%, OC12_v0 yields better queuing performance than OC48_3, whereas above 70% the situation is reversed – an illustration of the 'crossover' phenomenon [18], characteristic for queues with SS arrivals. Traffic with higher aggregation will always yield better queuing performance at high utilization, because of the reduced variation (see Section IV.C).

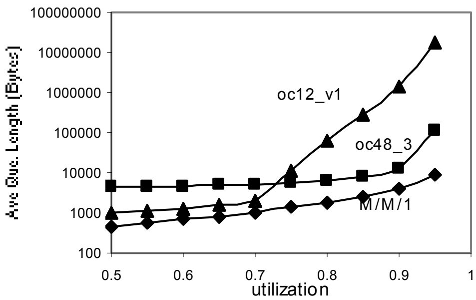

Fig. 8. Average queue lengths predicted by M/M/1 model (bottom), and resulted from simulating a queue fed by an OC12 trace and an OC48 trace

Fig.9 shows the tail probabilities of the queue length distribution from the simulation results using the trace OC48_3 as input to a FCFS queuing system with a constant service rate of 893 Mbps (80% utilization), and from the theoretical results predicted by M/M/1 model and fBm model. For a given queue length, the overflow probability obtained from the simulation is typically several magnitude bigger than that predicted by M/M/1 model, but is close to that predicted by fBm model (i.e., within one magnitude). Thus, we still cannot use the Poisson model as the analytical model for queuing analysis at high aggregate level. However, a good news from the simulation results is that we only need a moderate amount of memory (e.g., < 1ms) to achieve a low loss rate of 10^{-5} with high aggregation.

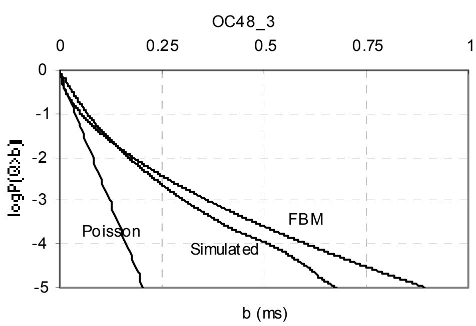

Fig. 9. Tail Probability of Queue Length

VI. Conclusions

Based on packets contained in eight Internet backbone traces with high timestamp accuracy and data rates up to 100s of Mbps, we find that Internet backbone traffic is long-range dependent and self-similar. We prove the Gaussian hypothesis for the marginal distributions of byte count and packet count is statistically valid according to the K-S goodness-of-fit test. However, although the marginal distribution of the packet inter-arrival time looks close to Exponential, the K-S test does not validate this hypothesis.

The auto-correlation for small lags ([50µs, 100ms]) is found to follow the same behavior as known for the large lags ([100ms, 500s]), i.e. *r(k)* is hyperbolic, but the parameter β has different values for the two ranges. The disparity between the two decreases at higher aggregate level, and the 100ms demarcation point moves to the right, in the 1s range.

The traffic variability decreases as the level of aggregation increases, as shown by the decreased std/mean of byte and packet count processes, and by the lightening of the tails of their marginal distributions. With the less variable traffic, we can achieve a low loss rate using a reasonable moderate buffer size (economy of scale), however the Poisson model is still vastly inappropriate for predicting queuing performance.

Acknowledgments

This work would not have been possible without the support of Kevin Thompson, Erik Sherk, David E. McDysan, Howie Xu and Robert Beverly at WorldCom.

References

[1] J. Cao, W. S. Cleveland, D. Lin, D. X. Sun, “Internet traffic tends toward Poisson and independent as the load increases,” in Nonlinear Estimation and Classification, C. Holmes, D. Denison, M. Hansen, B. Yu, B. Mallick - editors, Springer, 2002.

[2] M. E. Crovella and A. Bestavros, “Self-similarity in World Wide Web traffic: evidence and possible causes,” IEEE/ACM Trans. On Networking, Vol. 5, No. 6, pp.835-846, December 1997.

[3] A. Erramilli, O. Narayan and W. Willinger, “Experimental queueing analysis with long-range-dependent packet traffic,” IEEE/ACM Transactions on Networking, Vol. 4, No. 2, April 1994.

[4] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, “On the self-similar nature of Ethernet traffic,” In Proceeding of ACM SIGCOMM’93, pp.183-193, 1993.

[5] I. Norros, “On the use of fractional Brownian motion in the theory of connectionless networks,” IEEE Journal on selected areas in communications, vol. 13, No. 6, August 1995.

[6] V. Paxson and S. Floyd, “Wide-area traffic: the failure of Poisson modeling,” Proceedings of SIGCOMM’94, September 1994.

[7] M.S. Taqqu, V. Teverovsky, and W. Willinger, “Estimators for long-range dependence: an empirical study,” Fractals, Vol. 3, No. 4, pp785-788, 1995.

[8] K. Thompson, G. Miller, and R. Wilder, “Wide-area Internet traffic patterns and characteristics,” IEEE Network Magazine, Nov.-Dec. 1997, pp. 10-23.

[9] W. Willinger, M.S. Taqqu, R. Sherman and D.V. Wilson, “Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level,” IEEE/ACM Trans. on Networking, Vol. 5, No. 1, February 1997.

[10] W. Willinger, V. Paxson and M.S. Taqqu, “Self-similarity and heavy tails: structural modeling of network traffic,” A Practical Guide to Heavy Tails: Statistical Techniques and Applications, Birkhauser, Boston, 1998.

[11] D. Bertsekas and R. Gallager, Data Networks, Prentice Hall, NJ, 1992.

[12] D. Duffy, A. McIntosh, M. Rosenstein, and W. Willinger, “Statistical analysis of CCSN/SS7 traffic data from working CCS subnetworks,” IEEE JSAC 12, 544-551, April 1994.

[13] J. Beran, Statistics for Long-Memory Processes, Chapman & Hall, 1994.

[14] J. Beran, R. Sherman, and W. Willinger, “Long-range dependence in VBR video traffic,” IEEE Trans. on Comm, 43:2/3/4, 1566-79, 1995.

[15] M.T. Lucas, et. al., “Statistical characterization of Widea-area IP traffic,” Proc. of 6th International conference on Computer Communications and Networks, 1997.

[16] V. Paxson, and S. Floyd, “Why we don’t know how to simulate the Internet,” Winter Simulation Conference, 1997.

[17] LBNL’s network research group, Network tools: libpcap, http://www-nrg.ee.lbl.gov/

[18] A. L. Neidhardt and J. L. Wang, “The concept of relevant time scales and its application to queuing analysis of self-similar traffic (or Is Hurst naughty or nice?),” Measure and Model Comp. Systems, 1998.