This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2005 proceedings.

Performance Enhancement of Heavy Tailed Queueing Systems using a Hybrid Integration Approach

Mostafa H. Dahshan and Pramode K. Verma
Telecommunications Systems Program
Electrical and Computer Engineering
University of Oklahoma—Tulsa
Tulsa, OK 74135, USA
Email: {mdahshan, pverma}@ou.edu

***Abstract*— This paper presents a hybrid integration approach for heterogeneous heavy tailed traffic from multiple classes over a channel of known capacity, where the service times follow lognormal distribution. Using our hybrid integration approach, we formulate a procedure that can minimize delay while improving the channel utilization. We show that this approach yields lower delays than either the integrated or segregated approach described in published literature.**

***Keywords*— Segregation, integration, hybrid integration, heterogeneous traffic, M/G/1 queue, lognormal distribution.**

I. Introduction

Maximizing resource utilization and satisfying user demands are two competing goals for network providers. With better resource utilization, more subscribers can be serviced with fewer facilities. On the other hand, increasing resource utilization beyond some limit degrades the performance to an unacceptable level. In general, traffic with uniform characteristics will experience lower delays when they are integrated into a single channel [1]. However, this is not the case with heterogeneous traffic with disparate characteristics. A number of studies [2]-[6] have shown that a better performance can be achieved by segregating heterogeneous classes into separate channels and allocating exclusive bandwidth to each channel. Since both integration and segregation approaches provide better performance under different conditions, it is inefficient to design the system to operate statically as one or the other. The purpose of this paper is to provide a hybrid approach that combines the advantage of both approaches to provide a level of performance that is better than (or at least equal to) either.

Beginning the early and mid 1990s, with the publication of a few seminal papers [7]-[9] on self similarity in the Internet traffic, research on developing an appropriate model has intensified. Many studies have shown that internet traffic is heavy tailed or long range dependent. This suggests that queueing models such as M/G/1 [10], M/G/∞ [11] and G/G/1 with heavy tailed service times and/or interarrivals are more appropriate than the classical M/M/1 models. Some heavy tailed distributions such as Pareto and lognormal were shown to closely model service times [12]. A recent study has confirmed that the Poisson model characterizing the arrival process is still valid for multiplexed high load traffic [13].

For the M/G/1 queueing model the delay analysis can be implemented by using the Pollaczek-Khintchine's formula [10]. In this paper, we use the M/G/1 model with heavy tailed service times. We use the lognormal distribution for modeling service times because of its analytical tractability (finite variance). The lognormal distribution is widely used as a heavy tailed distribution in published literature [14]. It has also been shown to provide a good fit to modeling file sizes in the World Wide Web [15], and characterizing common network applications such as FTP, TELNET and SMTP [12].

This paper is organized as follows. Section II provides the delay analysis for segregated and integrated systems with an arbitrary number of classes. Section III describes the proposed hybrid integration approach and the procedure to determine the best hybrid grouping. Section IV presents an example to illustrate the hybrid approach. Section V shows simulation results. Section VI presents our conclusions.

II. Delay Analysis for Segregated and Integrated Systems

This section provides the equations required for calculating delay for segregated and integrated systems with n classes. We assume each class generates Poisson traffic with a mean arrival rate of λ_i and mean length of $1/\mu_i$. The distribution of the service times follows lognormal distribution with shape parameter α and scale parameter β where ($\alpha > 0$) [14].

A. Average Delay per Message for Segregated System

Under the segregated scenario, the total capacity C of the link is divided into n channels. The capacity allocated for each channel i is C_i, i=1, 2, …, n. Let $\lambda = \sum_{i=1}^{n} \lambda_i$. The average delay for the segregated system can be calculated as [2]:

$$T_{seg} = \sum_{i=1}^{n} T_i \frac{\lambda_i}{\lambda} \tag{1}$$

In order to calculate T_i, we use Pollaczek-Khintchine's (Kendall's) formula [2], [10]:

$$T = \frac{1}{\mu C} + \frac{\rho^2 + \lambda^2 \sigma^2}{2\lambda (1-\rho)} \tag{2}$$

Let X_i be a random variable that represents the service time for a single class i. Since X_i follows lognormal distribution, we have [16]:

$$E[X_i] = \frac{1}{\mu_i C_i} = e^{0.5\alpha_i^2 + \beta_i} \tag{3}$$

$$\sigma^2 [X_i] = e^{\alpha_i^2 + 2\beta_i} \left(e^{\alpha_i^2} - 1 \right) \tag{4}$$

Substituting in (2) we get:

$$T_i = \frac{1}{\mu_i C_i} + \frac{\lambda_i e^{\alpha_i^2}}{2\mu_i C (\mu_i C_i - \lambda_i)} \tag{5}$$

B. Average Delay per Message for Integrated System

In the integrated system, all classes share the available total capacity C. Let Y be a random variable representing the service time for the integrated system. All what we need to know about Y is the mean and the variance. The mean service time is [2]:

$$E[Y] = \frac{1}{\mu C} = \frac{1}{\lambda C} \sum_{i=1}^{n} \frac{\lambda_i}{\mu_i} \tag{6}$$

The variance of the service time is calculated as [2]:

$$\sigma^2 [Y] = \sum_{i=1}^{n} \frac{\lambda_i}{\lambda} E\left[X_i^2\right] - \sum_{i=1}^{n} \left(\frac{\lambda_i}{\lambda} E[X_i] \right)^2 \tag{7}$$

Thus,

$$\sigma^2 [Y] = \sum_{i=1}^{n} \frac{\lambda_i}{\lambda} e^{\alpha_i^2 + 2\ln(1/\mu C)} - \sum_{i=1}^{n} \left(\frac{\lambda_i}{\lambda} \frac{1}{\mu C} \right)^2 \tag{8}$$

The average delay can now be calculated using (2).

C. Channel Capacity Allocation

In order to provide the lowest possible delay for the segregated system, the capacity of each channel must be allocated optimally. This has been done using the method of Lagrange multipliers [1] for exponentially distributed service times. Because of the simplicity of the delay equation for the M/M/1 queue, it was possible to obtain a closed form solution for the optimum channel capacity. For a general service time distribution, and particularly for the lognormal distribution, using Lagrange multipliers results in a non-polynomial equation that can only be solved numerically. Another study [17] used a perturbation technique for general service time distribution. This technique requires solving high-order equations for practical applications.

To simplify our analysis, but without sacrificing its validity, we distribute the total available channel capacity among each of the segregated channels so that the utilization ρ for each channel is the same. i.e., $\rho=\lambda_i/(\mu_i C_i)$ for all channels. Thus, the capacity allocated for each channel would be

$$C_i = \frac{1}{\rho} \frac{\lambda_i}{\mu_i} \tag{9}$$

We note that if we did assign the channel capacities in an optimal manner, the advantage of the segregated system will be higher; in other words, the segregated advantage in our analysis presents a pessimistic result.

III. THE HYBRID INTEGRATION APPROACH

The basic idea of the hybrid integration approach presented in this paper is grouping classes with comparable characteristics into channels and allocating bandwidth exclusively for each channel. It can be implemented on a "dynamic" basis. For example, a system implementing this approach can determine the best grouping as the traffic characteristics change and configure groupings accordingly. The hybrid system is equivalent to the segregated system if the number of groups equals the number of classes and is equivalent to the integrated system when the number of groups is 1. In the rest of this section, we determine the delay of the hybrid system, calculate the number of groupings and determine the optimal grouping.

A. Delay Calculation

Let K = number of groups and $i \in j$ denotes all classes i in group j. We define $\hat{\lambda}_j$ $\sum_{i \in j} \lambda_i$ and $\hat{C}_j$ $\sum_{i \in j} C_i$ as the combined arrival rate and allocated capacity for group j, respectively (see Fig. 1). Note that C_i is calculated from (9). The average delay for the hybrid system can be calculated in a very similar way to that of the segregated system, as follows:

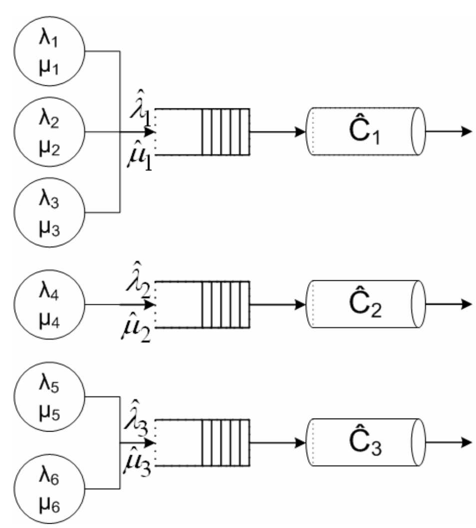

Fig. 1: Conceptual view of the hybrid integration approach

$$T_{hyb} = \sum_{j=1}^{K} \hat{T}_j \frac{\hat{\lambda}_j}{\lambda} \quad (10)$$

$\hat{T}_j$ is calculated for each group using (2). It is required to determine the mean and variance of the service time. Let Z_j be a random variable representing the service time for group j. We have:

$$E\left[Z_j\right] = \frac{1}{\hat{\mu}_j \hat{C}_j} = \frac{1}{\hat{\lambda}_j \hat{C}_j} \sum_{i \in j} \frac{\lambda_i}{\mu_i} \quad (11)$$

$$\sigma^2\left[Z_j\right] = \sum_{i \in j} \frac{\lambda_i}{\hat{\lambda}_j} E\left[X_i^2\right] - \sum_{i \in j} \left(\frac{\lambda_i}{\hat{\lambda}_j} E\left[X_i\right] \right)^2 \quad (12)$$

Thus,

$$\sigma^2\left[Z_j\right] = \sum_{i \in j} \frac{\lambda_i}{\hat{\lambda}_j} e^{\alpha_i^2 + 2\ln\left(1/\hat{\mu}_j \hat{C}_j\right)} - \sum_{i \in j} \left(\frac{\lambda_i}{\hat{\lambda}_j} \frac{1}{\hat{\mu}_j \hat{C}_j} \right)^2 \quad (13)$$

After calculating $\hat{T}_j$ for all groups, the average delay for the system T_{hyb} is calculated using (10).

B. Number of Groupings Calculation

To ensure that the hybrid grouping results in the lowest possible delay, in general, all possible groupings should be examined. A system with n classes can be arranged into a minimum of 1 group (1 possible grouping) and a maximum of n groups (1 possible grouping). Between these two extremes, there are different ways in which groupings can be made. Let A_{nj} = Number of groupings for n classes in j groups, where $1 \le j \le n$. R_n = Total number of groupings for n classes. It follows that

$$R_n = \sum_{j=1}^{n} A_{nj} \quad (14)$$

As stated above, we have

$$A_{n1} = 1 \quad (15)$$

It can be shown that

$$A_{nj} = A_{n-1,j-1} + j \cdot A_{n-1,j} \quad (n > 1, j > 1) \quad (16)$$

Equation (16) can be explained as follows. Let S_{nj} denote the state in which n classes are arranged in j groups. S_{nj} can be reached either from $S_{n-1,j}$ or $S_{n-1,j-1}$. In the first case, there are already j groups and the n^{th} class can only be placed in one of the j groups, which have $A_{n-1,j}$ possible groupings, yielding a total of $j\,A_{n-1,j}$ possibilities. In the second case, the n^{th} class is placed in a new group by itself. The number of groupings in this case is the same as with $j-1$ groups, i.e. $A_{n-1,j-1}$. Table I shows the number of groupings for different number of classes.

TABLE I
NUMBER OF CLASSES n AND THE CORRESPONDING NUMBER OF POSSIBLE GROUPINGS Rn.

n	4	8	10	25
R_n	15	877	115971	4.68×10^{18}

The number of possible groupings increases rapidly with the number of classes. In order to reduce the number of groupings to be examined, the input classes are arranged in ascending order (according to their mean service times) obviating the need to examine a grouping such as {{1,3},{2}} because it will not satisfy the optimality criterion. Fig. 2 shows possible groupings for an ordered set of up to 4 classes. Each level of the tree shows the possible groupings for its corresponding number of classes. It can be seen from Fig. 2 that the number of groupings for 4 classes is 2^3. In general, since groupings follow a binary tree, there will be 2^{n-1} possible groupings for n classes arranged in ascending order.

C. Optimal Hybrid Grouping

With a moderate number of classes, testing all possible 2^{n-1} groupings might be acceptable. For a large number of classes, the number of comparisons can be reduced by observing the following: From Fig. 2, if the grouping {{1,2}} has a lower delay than {{1},{2}} for instance, then it can be shown using (10) that any grouping in lower levels containing {{1},{2}} has a higher delay than its corresponding grouping containing {{1,2}} and hence need not to be examined. This can be demonstrated as follows:

$$T\left(\{1\},\{2\}\right) = \frac{\lambda_1}{\lambda_1+\lambda_2} T_1 + \frac{\lambda_2}{\lambda_1+\lambda_2} T_2 \quad (17)$$

$$\begin{aligned} T\left(\{1\},\{2\},\{3\}\right) &= \frac{\lambda_1}{\lambda_1+\lambda_2+\lambda_3} T_1 + \frac{\lambda_2}{\lambda_1+\lambda_2+\lambda_3} T_2 + \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3} T_3 \\ &= \left(\frac{\lambda_1}{\lambda_1+\lambda_2} T_1 + \frac{\lambda_2}{\lambda_1+\lambda_2} T_2 \right) \frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2+\lambda_3} + \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3} T_3 \\ &= \frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2+\lambda_3} T\left(\{1\},\{2\}\right) + \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3} T_3 \end{aligned} \quad (18)$$

$$T\left(\{1,2\},\{3\}\right) = \frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2+\lambda_3} T\left(\{1,2\}\right) + \frac{\lambda_3}{\lambda_1+\lambda_2+\lambda_3} T_3 \quad (19)$$

By comparing (18) and (19) we can see that

$$T\left(\{1,2\},\{3\}\right) < T\left(\{1\},\{2\},\{3\}\right) \Leftrightarrow T\left(\{1,2\}\right) < T\left(\{1\},\{2\}\right) \quad (20)$$

Generalizing this observation, let i = the tree level, H_i = an array with the groupings generated in level i. We can now state the following algorithm:

1. Start with $i \leftarrow 1$. $H_i \leftarrow \{\{1\}\}$
2. Calculate T_{hyb} for groupings in H_i

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2005 proceedings.

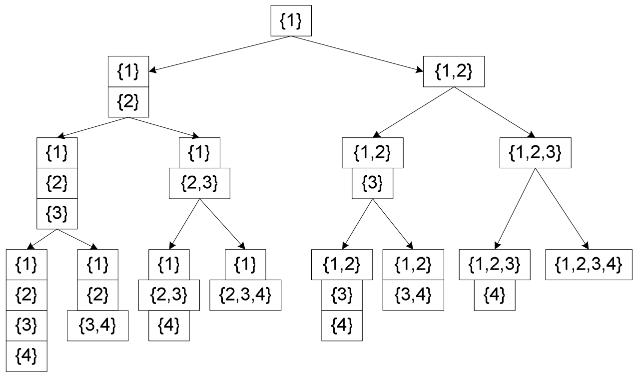

Fig. 2: Hybrid grouping for 1,2,3 and 4 orderd classes

3. $L \leftarrow$ minimum T_{hyb} grouping in H_i
4. If $i = n$ then return L and stop
5. Generate the left child (segregated) and right child (integrated) of L. For all other groupings, generate only the right child. Add generated groupings to H_{i+1}
6. $i \leftarrow i+1$. Go to step 2.

It can be easily seen that the total number of comparisons needed using this algorithm is $\sum_{i=1}^{n} i$ or $n(n+1)/2$, a significant saving over 2^{n-1} for large n.

IV. Numerical Results

In this section, we plot the average delay as a function of the disparity in the mean service time, using the standard deviation σ_L as a measure for disparity.

Let $1/\mu = (1/\lambda)\sum_{i=1}^{n}(\lambda_i/\mu_i)$, then

$$\sigma_L = \sqrt{\sum_{i=1}^{n}\frac{\lambda_i}{\lambda}\left(\frac{1}{\mu_i}-\frac{1}{\mu}\right)^2} \tag{21}$$

Let $C = 2190$. The arrival rates and mean service times are varied such that ρ is held constant at 0.9. In the analysis, 6 classes were used in 9 datasets, each dataset corresponds to a different value of σ_L. The shape parameter value α=0.5 was used for all classes. Input values and output results are summarized in Table II and Fig. 3.

From the results we see that the hybrid approach has a better performance than both integrated and segregated approaches over the entire range of σ_L considered. The *integ penalty* in Table II is the relative additional capacity required to achieve the same delay as the hybrid approach using a single integrated channel. This shows the savings in the bandwidth with the hybrid approach.

V. Simulation Results

To support the analytical results, we used the Network Simulator ns-2 [18] to provide simulation based results. The network topology used was as follows. For the integrated system, all classes were attached to one source node which was connected by a single link to the sink node. For the segregated system, each class was attached to a different source node and all of these nodes were connected to the sink node by separate

TABLE II

INPUT AND OUTPUT DATA OF THE HYBRID INTEGRATION ANALYSIS AND SIMULATION

	1	2	3	4	5	6	7	8	9
λ_1	328.5	328.5	328.5	328.5	328.5	328.5	328.5	328.5	328.5
$1/\mu_1$	1	1	1	1	1	1	1	1	1
λ_2	203.4	135.6	101.7	81.36	67.8	58.114	50.85	45.2	40.68
$1/\mu_2$	2	3	4	5	6	7	8	9	10
λ_3	49.725	24.863	14.207	9.945	7.1036	5.3757	4.1438	3.3712	2.7625
$1/\mu_3$	4	8	14	20	28	37	48	59	72
λ_4	8.2286	3.0316	1.44	0.78904	0.48814	0.32542	0.22857	0.16696	0.12632
$1/\mu_4$	7	19	40	73	118	177	252	345	456
λ_5	47.782	11.945	4.4923	2.0857	1.1159	0.65782	0.41582	0.27765	0.19331
$1/\mu_5$	11	44	117	252	471	799	1264	1893	2719
λ_6	23.874	4.536	1.3622	0.53554	0.2495	0.13083	0.07479	0.045666	0.029371
$1/\mu_6$	19	100	333	847	1818	3467	6065	9933	15444
σ_L	4.0533	11.313	21.612	34.986	51.57	71.389	94.526	120.98	150.83
T_{int}	0.02377	0.099115	0.29594	0.70784	1.4577	2.6973	4.6113	7.4132	11.353
T_{int}sim	0.023952	0.10421	0.28378	0.78075	1.7216	2.5391	5.168	7.218	8.5412
T_{seg}	0.055331	0.071983	0.081031	0.086485	0.090318	0.09311	0.095264	0.096943	0.098315
T_{seg}sim	0.055638	0.074282	0.076316	0.091318	0.10035	0.093241	0.10936	0.10835	0.099833
T_{hyb}	0.021521	0.03445	0.048559	0.057357	0.064355	0.070923	0.077595	0.083429	0.087065
T_{hyb}sim	0.022477	0.036266	0.047197	0.061249	0.074649	0.068415	0.082825	0.10196	0.089299
Hybrid Arrgmt	{{1, 2, 3}, {4, 5, 6}}	{{1, 2, 3}, {4, 5, 6}}	{{1, 2, 3}, {4, 5, 6}}	{{1}, {2, 3}, {4, 5, 6}}	{{1}, {2, 3}, {4, 5, 6}}	{{1}, {2, 3}, {4, 5, 6}}	{{1}, {2, 3}, {4, 5, 6}}	{{1}, {2, 3}, {4, 5}, {6}}	{{1}, {2}, {3}, {4, 5}, {6}}
Integ Penalty	%0.998443	%15.5841	%35.9215	%65.8728	%103.218	%146.319	%194.076	%247.637	%310.083

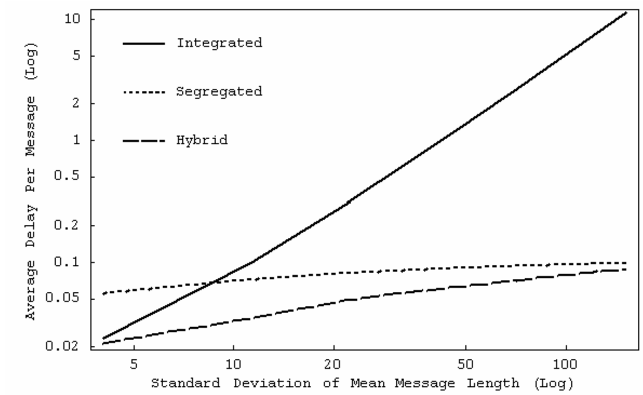

Fig. 3: Delay of the integrated, segregated and hybrid systems with different disparities (analytical)

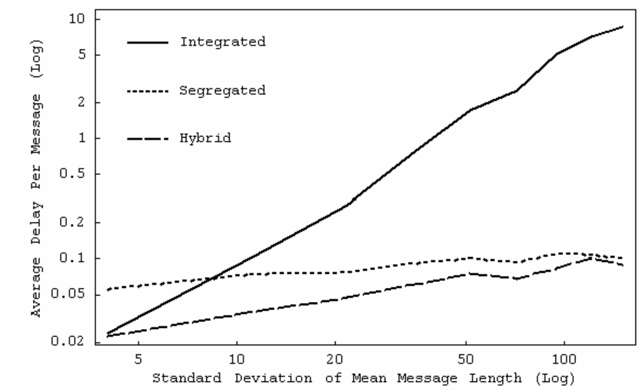

Fig. 4: Delay of the integrated, segregated and hybrid systems with different disparities (simulation)

links. For the hybrid system, all classes in each group were attached to one source node and all these nodes were connected to the sink node by separate links.

Classes were defined as UDP agents. For each agent, an exponential random variable was defined for interarrival time and a lognormal random variable was defined for the packet size. The mean values and parameters used were the same as those used in the analysis of Section IV. For each agent, a *send_packet* subroutine was defined that will generate a packet of size determined from the *packet_size* random variable. The packet is transmitted to the sink node and the subroutine will schedule itself to run again after a time period determined from the *interarrival_time* random variable.

The simulation was run for 120 minutes for each of the 9 datasets used in Table II. At the end of the simulation, a script is run that would calculate the average delay per packet from the output trace file generated by the simulation. The simulation results are shown in Table II as T_{seg}sim, T_{int}sim and T_{hyb}sim and are plotted in Fig. 4. We found that the simulation results closely match the analytical results.

VI. Conclusions

This paper has introduced hybrid integration as an approach for enhancing the performance of heavy tailed queueing systems with lognormal service times, over the integrated and segregated approaches. We have shown that the hybrid approach outperforms or at least matches the best solution given by either the integrated or the segregated approach. We have also validated the analytical results by simulation and demonstrated a close match between the two.

References

[1] Kleinrock, L., "Communication Nets," *McGraw-Hill*, New York, 1964.
[2] Verma, P.K. and Rybczynski, A.M., "The Economics of Segregated and Integrated Systems in Data Communication with Geometrically Distributed Message Length," *IEEE Trans. Communications*, pp. 1844-1848, Nov. 1974.
[3] Bobby N.W. and Mark, J.W., "Performance Analysis of Burst Switching for Integrated Voice/Data Services," *IEEE Trans. Communications*, vol. 36, No. 3, pp. 282–297, Mar. 1988.
[4] Kim, Y.H. and Un, C.K., "Performance Analysis of Statistical Multiplexing for Heterogeneous Bursty Traffic in ATM Network," *IEEE Trans. Communications*, vol. 42, No. 2/3/4, pp.745-753, Feb./ Mar./ Apr. 1994.
[5] Su, C. and de Veciana, G., "Statistical Multiplexing and Mix-Dependent Alternative Routing in Multiservice VP Networks," *IEEE/ACM Trans. Networking*, vol. 8, No. 1, pp. 99-108, Feb. 2000.
[6] Eberle, H. and Gura, N., "Separated High-bandwidth and Low-latency Communication in the Cluster Interconnect Clint," *Proc. 2002 ACM/IEEE Conf. Supercomputing*, Baltimore, Maryland, Nov. 2002.
[7] Leland, W., Taqqu, M., Willinger, W. and Wilson, D., "On the Self-Similar Nature of Ethernet Traffic (Extended Version)," *IEEE/ACM Trans. Networking*, vol. 2, pp. 1-15, Feb. 1994.
[8] Crovella, M. and Bestaros, A., "Self-Similarity in World-Wide Web Traffic: Evidence and Possible Causes," *Proc., ACM Sigmetrics Conf. on Measurement and Modeling of Computer Systems*, May 1996.
[9] Paxson, V. and Floyd, S., "Wide-Area Traffic: The Failure of Poisson Modeling", IEEE/ACM Trans. on Networking, vol. 3 pp. 226-224, 1995.
[10] Boxma, O.J. and Cohen, J.W., "The M/G/1 Queue with Heavy-Tailed Service Time Distribution," *IEEE Journal on Selected Areas in Communications*, vol. 16, No. 5, pp. 749-763, Jun. 1998.
[11] Parulekar, M. and Makowski, A.M. "M|G|∞ input processes: a versatile class of models," *INFOCOM '97, Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proc. IEEE*, vol.2, pp. 419-426, Apr. 1997
[12] Paxson, V., "Empirically Derived Analytic Models of Wide-Area TCP Connections", *IEEE/ACM Trans. on Networking*, vol. 2, No. 4, pp. 316-336, Aug. 1994.
[13] Cao, J., Cleveland, W., Lin, D. and Sun, D., "Internet Traffic Tends Toward Poisson and Independent as the Load Increases," *Nonlinear Estimation and Classification*, eds. C. Holmes, D. Denison, M. Hansen, B. Yu, and B. Mallick, Springer, New York, 2002.
[14] Sigman, K., "Appendix: A Primer on Heavy-Tailed Distributions," *Queueing Systems*, 33, pp. 261-275, 1999.
[15] Downey, A.B., "The Structural Cause of File Size Distributions," *Proc. Ninth International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems*, pp. 361-370, 2001.
[16] Weisstein, E.W., "Log Normal Distribution," From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/LogNormalDistribution.html.
[17] Pollett, P.K., "Residual Life Approximations in General Queueing Networks". *Elektronische Informationsverarbeitung und Kybernetik*, vol. 2, pp. 41-54, 1984.
[18] The Network Simulator: ns-2, http://www.isi.edu/nsnam/ns.

