1 Introduction

A useful way to characterize the arrival of traffic at an output queue of a router is to measure packet arrivals per unit time. In this approach, the traffic is represented as an integer valued time-series $\{x_1, x_2, ...\}$ where x_j represents the number of packets that arrived at the queue during interval j. In modeling queuing behavior, it is generally assumed that the values, x_j, are identically distributed according some distribution $F_x(\,)$. Early models of traffic arrival arrival processes also commonly assumed that x_j and x_k were statistically independent for $j \neq k$. That is, the values, x_j, were independent and identically distributed, $(i.i.d)$. The foundational work of Leland Leland, Taqqu, Willinger, and Wilson [?], demonstrated that the assumption of independence was unjustified. It has since been established that the statistical dependence observed in most in LAN and WAN traffic arrival processes approximates long-range dependence. That is, $\sum_k r(k) = \infty$, where $r(k)$ represents the autocorrelation (normalized covariance) of elements that are k units apart in the time-series. The practical consequence of long-range dependence is the occurrence of both long bursts in which the arrival counts, x_j, are are considerably larger that the mean of $F_x(\,)$ and compensating bursts in which the number of arrivals is smaller than the mean.

A number of studies have shown that under heavy loads and with unbounded queue capacity, long-range dependence in the arrival process produce significantly larger mean queue lengths than occur when arrival counts are independent. However, the queue capacity of real routers is clearly finite. The objective of this study is to investigate and report upon the interaction self-similarity and bounded queue capacity. We will show that when queue capacity is bounded, both short-range dependent and long-range dependent arrival processes produce similar mean queue lengths. Perhaps more surprisingly, similar drop rates are also observed.

2 Methodology

The results presented here are obtained via discrete event simulation of a single-server queuing system. Details of the implementation of the simulator can be found in [?]. The distribution, $F_x(\,)$, of the traffic arrival process that is simulated is the empirical distribution of the Bellcore workload described in [?]. Elements of the synthetic time series $\{x_1, x_2, ...\}$ represent the number of packet arrivals in time units of 0.125 seconds. The mean number of arrivals per unit time is 39.25. The length of each simulated run is 32768 time units which corresponds to an elapsed real time of 4096 seconds or a total of about 1.3 million arrivals. The stochastic dependence in the arrival process is based upon an exactly self-similar model in which

$$r(k) = (1/2)[(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}].$$

The value of the Hurst parameter, H, is also derived from the Bellcore workload and is 0.845. To compare the effects of long and short-range dependence the actual autocorrelations present in the synthetic traffic are:

$$r_N(k) = r(k), k <= N$$

$$r_N(k) = 0, k > N$$

Synthesis of time-series having both a prescribed empirical distribution function and a prescribed autocorrelation function is a non-trivial procedure and is described in [?].

3 Results and Conclusions

Results of the simulations are summarized in tables 1, 2, and 3. Each reported queue length and drop rate is the mean value obtained from 32 replications of the simulation. Confidence intervals were computed at the 90% level for all reported results. For the results reported in table 1 the widest of these was 8.33085 ∓ 0.01141 which was obtained for the r_{256} workload at a utilization of 0.75. Note that the results obtained for the r_8 and r_{256} are virtually identical. However, at high utilization the results obtained with the r_2 workload differ by a small but statistically significant degree.

Util	$r_2(k)$	$r_8(k)$	$r_{256}(k)$
0.25	0.33988	0.33974	0.33972
0.30	0.47574	0.47538	0.47558
0.35	0.68873	0.68857	0.68833
0.40	1.02978	1.03453	1.03330
0.45	1.55548	1.56147	1.56282
0.50	2.27131	2.28420	2.28626
0.55	3.17375	3.20075	3.20379
0.60	4.24523	4.29706	4.29841
0.65	5.47872	5.53921	5.54176
0.70	6.82282	6.90137	6.89273
0.75	8.25543	8.32573	8.33085

Table 1: Mean Queue Length: (Max Length = 32)

When the maximum queue length is increased to 2048, significant increases in mean queue lengths of all three synthetic workloads occur. However, as shown in table 2, it is clearly the case here that for high utilizations the extent of the stochastic dependence profoundly affects the mean queue length. It should also be noted that this effect is largely absent until utilization reaches 0.45. It

Util	$r_2(k)$	$r_8(k)$	$r_{256}(k)$
0.25	0.34035	0.34050	0.34035
0.30	0.47816	0.47906	0.47832
0.35	0.70647	0.71187	0.71006
0.40	1.11048	1.16328	1.16937
0.45	1.83439	2.06783	2.17251
0.50	3.08278	3.95428	4.90073
0.55	5.08394	7.26711	11.58443
0.60	8.16683	12.79501	23.38090
0.65	12.77428	22.05752	54.87549
0.70	19.69716	36.28603	99.36962
0.75	30.07873	60.13959	194.38778

Table 2: Mean Queue Length: (Max Length = 2048)

could be argued that a router operating in this range was *already* over capacity.

The number of packets dropped per 1000 packets routed is reported in table 3. Here the widest confidence interval at the 90% level is 65.33344 $\mp$0.34750, again for the r_{256} workload at a utilization of 0.75.

Util	$r_2(k)$	$r_8(k)$	$r_{256}(k)$
0.25	0.01000	0.00938	0.01125
0.30	0.04000	0.03875	0.04531
0.35	0.19969	0.20469	0.19625
0.40	0.80844	0.85250	0.83906
0.45	2.50063	2.66937	2.63125
0.50	6.07906	6.53313	6.41094
0.55	12.00719	12.79937	12.82031
0.60	20.45031	21.93531	21.84719
0.65	31.53937	33.85375	33.82375
0.70	45.11656	48.41406	48.40188
0.75	60.81500	65.45656	65.33344

Table 3: Drops per 1000: (Max Length = 32)

These results are qualitatively consistent with those of table 1 and serve to confirm our our belief that bounded queue capacity negates, to a significant degree, the effects of long-range dependence.