The University of Kansas

Information and Telecommunication Technology Center

Technical Report

Empirical Study of a Queueing System with fBM Traffic

Yasong Jin, Soshant Bali, Tyrone Duncan,
Dahlia Al-ammar and Victor S. Frost

ITTC-FY2005-TR-30540-02

October 2004

Project Sponsor:
National Science Foundation

Copyright © 2004:
The University of Kansas
2335 Irving Hill Road, Lawrence, KS 66045-7612
All rights reserved.

Empirical Study of a Queueing System with fBM Traffic

Abstract

In packet networks, the rate of congestion events has been proposed as a metric of quality of service (QoS) [2]. We studied a queueing system with fractional Brownian input. A method to compute the expected rate of congestion events is justified and evaluated.

I. Introduction

The queueing model discussed here is the fractional Brownian (fBM) model, proposed by Norros [5] (Figure 1). The input of the queue is modeled by fBM, which has long range dependence. The output is deterministic.

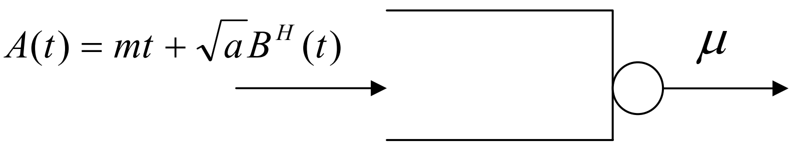

Figure 1

Suppose that the queue is empty at t=0, i.e., $Q(0) = 0$, then the queue length, $Q(t)$, is given by $Q(t) = A(t) - \mu t - \min_{0\le s\le t}(A(t) - \mu s)$ (1)

where $A(t)$ is the arrival process and $A(t) = mt + \sqrt{a}B^H(t)$,

m : mean input rate (bps), μ : service rate (bps)

a : variance (bit^2), $B^H(t)$: standard fBM with parameter H.

Let c be the difference between m and μ, then $c = \mu - m$, and $Q(t)$ can be written as

$$Q(t) = \sqrt{a}B^H(t) - ct - \min_{0\le s\le t}(\sqrt{a}B^H(s) - cs) \qquad (2)$$

Obviously, $Q(t)$ is a random process. The definition of a congestion event follows [2]. Shown in Figure 2, a congestion event is defined to occur at time t_b if t_b is the first time the process $Q(t)$ reaches a given level b following the end of the previous busy period containing a congestion event. Note that in a congestion episode, $Q(t)$ can cross b several times. τ_b denotes the time between two congestion events. Its expectation, $E\tau_b$, (or the expected rate of congestion events, $1/E\tau_b$) is a good QoS metric.

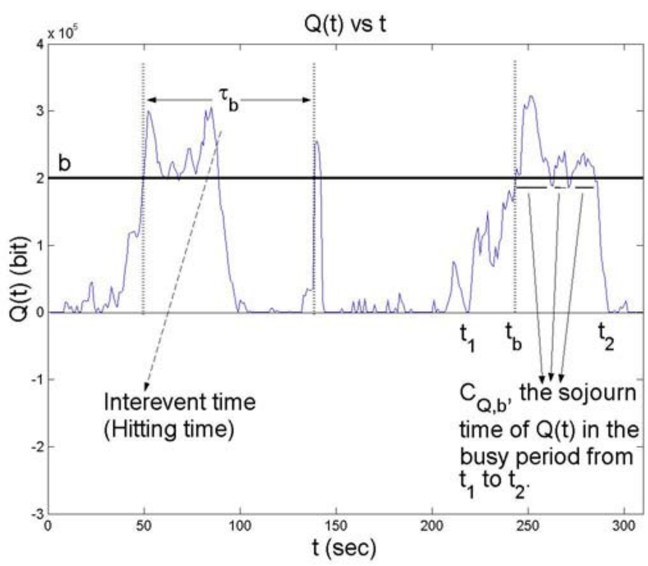

Figure 2

II. Poisson clumping method

Following [1], the Poisson clumping method is used to approximate $E\tau_b$, that is,

$$E\tau_b \approx EC_{Q,b} / P(Q \geq b) \tag{3}$$

where $EC_{Q,b}$ is the mean time that $Q(t)$ stays above level b in a busy period, and $P(Q \geq b)$ is the probability that the queue length is greater than b. In [1] the author warns that the Poisson clumping method may not be used to the traffic which has long range dependency. But from our studies, the method works well for the fBM traffic. Table 1 shows a comparison of the Poisson clumping approximations with the simulation results.

Table 1: Applicability of the Poisson clumping method

$m = 2Mbps$, $a = 4\times10^{11} bit^2$, $H = 0.79$, $Load = m/\mu = 0.7$

b (bit)	$E\tau_b$ (sec) (measured in simulations)	$EC_{Q,b}$ (sec) (measured in simulations)	$P(Q \geq b)$ (measured in simulations)	Poisson clumping approximations $EC_{Q,b} / P(Q \geq b)$	Relative Error between $E\tau_b$ and the Poisson clumping results
1000000	98.96	3.34	0.033663	99.22	0.26%
1200000	146.37	3.91	0.026691	146.49	0.08%
1400000	195.16	4.25	0.021744	195.46	0.15%
1600000	269.19	4.79	0.017766	269.61	0.16%
1800000	343.10	5.04	0.014656	343.89	0.23%
2000000	442.95	5.44	0.012249	444.12	0.27%
2200000	550.32	5.68	0.010295	551.71	0.25%
2400000	727.40	6.34	0.008686	729.87	0.34%
2600000	853.58	6.37	0.007447	855.40	0.21%
2800000	995.84	6.44	0.006446	999.07	0.32%
3000000	1230.16	7.01	0.005681	1233.94	0.31%
3200000	1416.56	7.35	0.005084	1445.79	2.06%
3400000	1910.72	8.79	0.004507	1950.43	2.08%
3600000	2106.69	8.57	0.003983	2151.76	2.14%
3800000	2282.25	8.11	0.003480	2330.33	2.11%
4000000	2650.36	8.37	0.003093	2705.69	2.09%
4200000	2934.34	8.28	0.002764	2995.80	2.09%
4400000	3160.06	8.01	0.002482	3227.09	2.12%
4600000	3572.26	7.85	0.002153	3646.41	2.08%
4800000	3863.44	7.51	0.001879	3996.98	3.46%
5000000	4056.62	6.81	0.001623	4195.38	3.42%

The mean interevent time ($E\tau_b$), the tail probability ($P(Q \geq b)$) and the mean sojourn time for different b ($EC_{Q,b}$) were measured and the corresponding approximation results were calculated. It is obvious from Table 1 that the approximations are very close to the simulation results, i.e., $E(\tau_b) \approx EC_{Q,b} / P(Q \geq b)$. For different parameters, we observed the same phenomena.

III. Properties of a busy period

To apply the Poisson clumping approximation, we need to find $EC_{Q,b}$ and $P(Q \ge b)$. Following [3], we can use the asymptotic tail probability to approximate $P(Q \ge b)$. Now the problem reduced to computing the mean sojourn time, $EC_{Q,b}$. We will analyze the busy periods and present a way to approximate $EC_{Q,b}$.

A busy period, which is from t_1 to t_2 (Figure 3), can be divided into two parts, t_1 to t_b and t_b to t_2. Note that $Q(t) > 0$, for $t \in (t_1, t_2)$. t_b is the first time that $Q(t)$ reaches b. Let $t_{0b} = t_b - t_1$ and $t_{b0} = t_2 - t_b$.

Simulations were conducted to study the empirical properties of t_{0b} and t_{b0}.

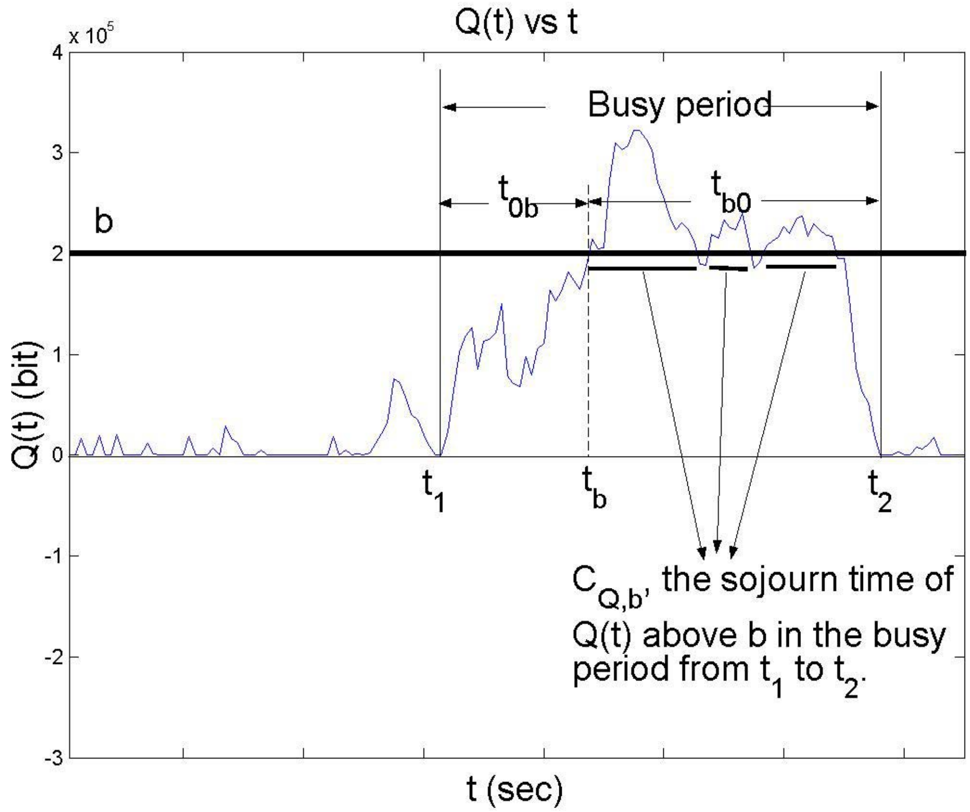

Figure 3

(i). Empirical property of t_{0b}

The values of t_{0b} cluster around its mean. For different loads and H values, we always have $Et_{0b}/std(t_{0b})>1$. As shown in Figure 4-6, the dotted curve, $Et_{0b}/std(t_{0b})$, is always above the line y=1.

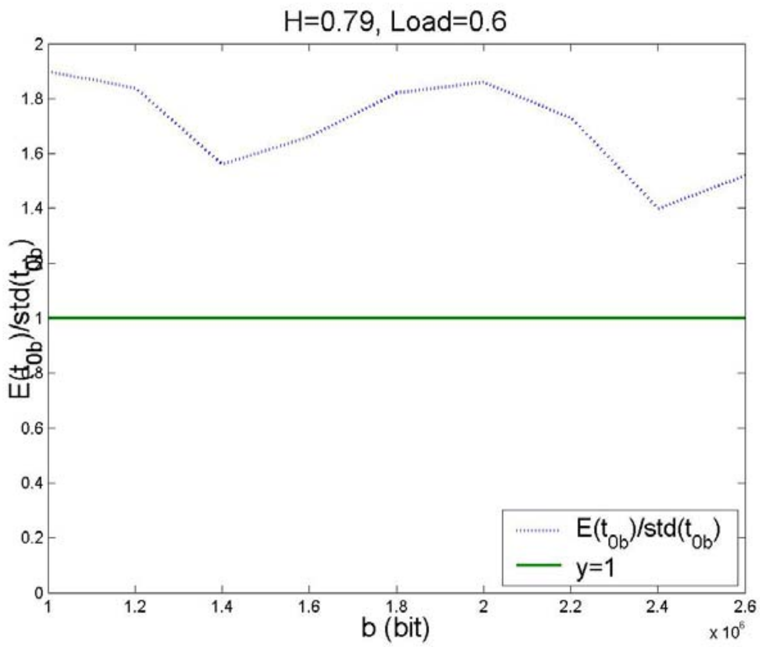

Figure 4

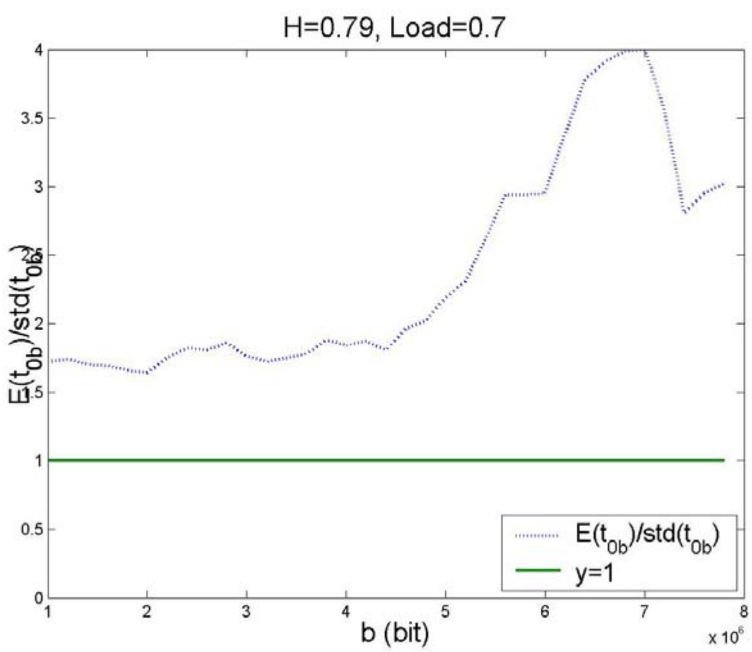

Figure 5

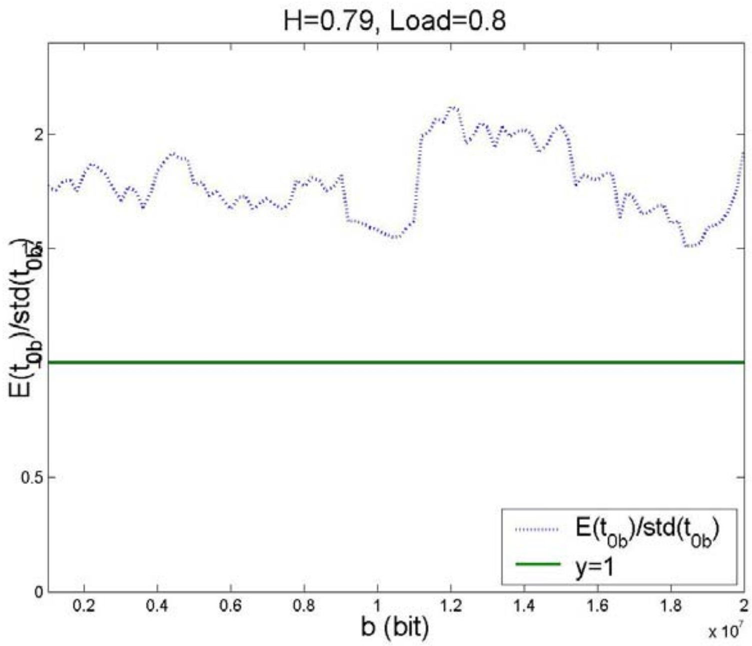

Figure 6

(ii). Empirical property of t_{b0}

For large b, we have $Et_{b0} / std(t_{b0}) > 1$ and most values of t_{b0} are less than $2Et_{b0}$. From Figure 7-9, we can see that as b increases, the curve, $Et_{b0} / std(t_{b0})$, crosses y=1 and stays above it. In other words, as the congestion event becomes rare, we have $Et_{b0} / std(t_{b0}) > 1$.

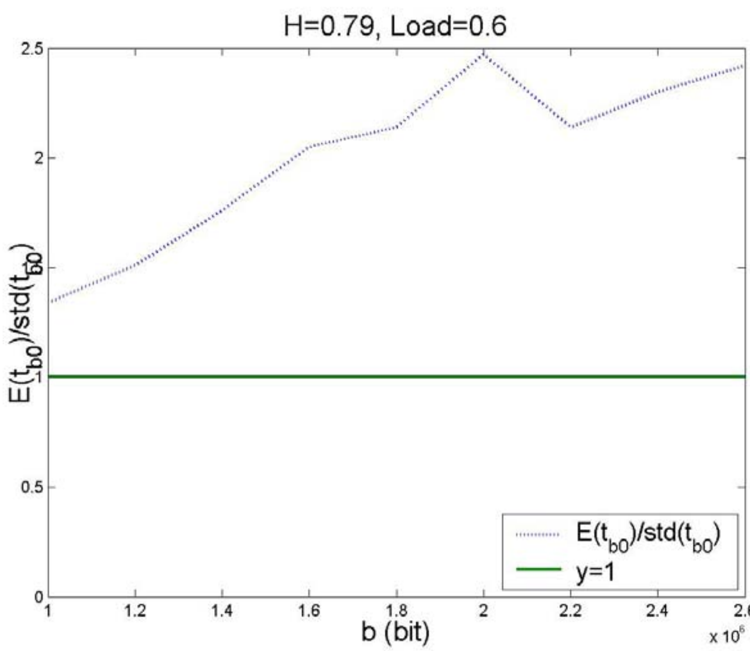

Figure 7

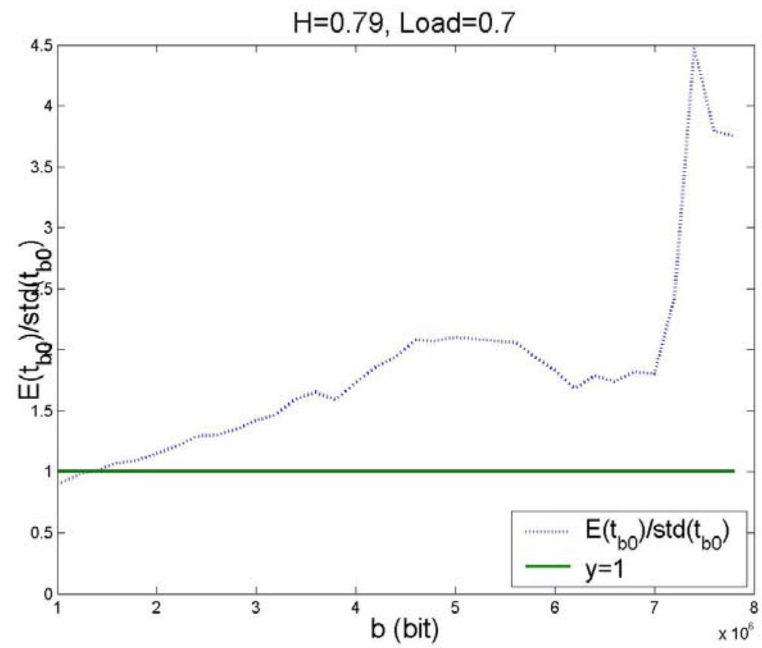

Figure 8

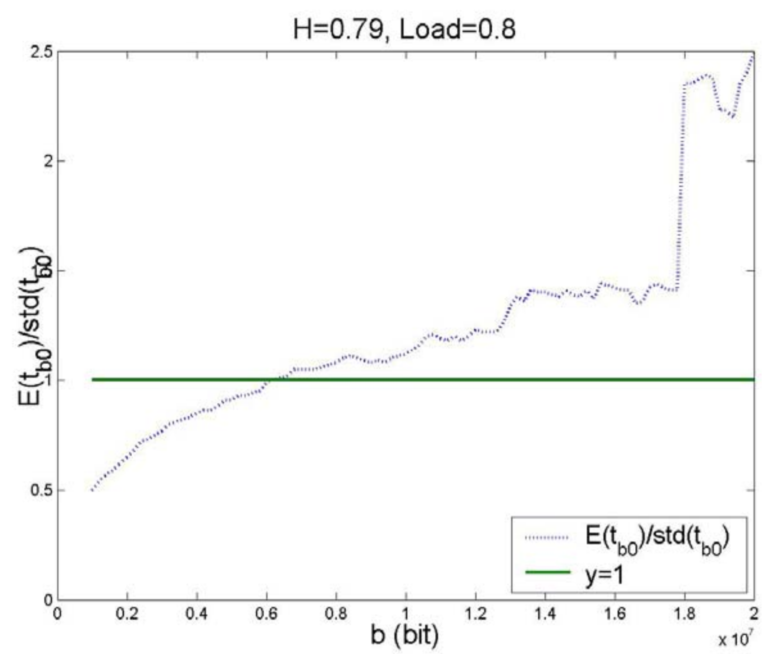

Figure 9

IV. An approximation of the busy period

A busy period is from t_1 to t_2 (Figure 3). t_b is the first time that the queue reaches level b, and t_2 is the first time that the queue returns to 0 after t_b. Note that $Q(t_1)=0$, $Q(t_b)=b$ and $Q(s)>0, s\in(t_1,t_2)$,

From The queue length formula (2), it is easy to derive that

(i) $\sqrt{a}B^H(t_1)-ct_1 = \inf_{0\le s\le t_2}(\sqrt{a}B^H(s)-cs)$

(ii) $\sqrt{a}B^H(t_b)-ct_b-(\sqrt{a}B^H(t_1)-ct_1)=b$

For $t\in[t_b,t_2)$,

$$\begin{aligned}
Q(t) &= \sqrt{a}B^H(t)-ct-\inf_{0\le s\le t}(\sqrt{a}B^H(s)-cs)\\
&= \sqrt{a}B^H(t)-ct-(\sqrt{a}B^H(t_1)-ct_1)\text{, from (i)}\\
&= \sqrt{a}B^H(t)-ct-(\sqrt{a}B^H(t_b)-ct_b)+(\sqrt{a}B^H(t_b)-ct_b)-(\sqrt{a}B^H(t_1)-ct_1)\\
&= \sqrt{a}B^H(t)-ct-(\sqrt{a}B^H(t_b)-ct_b)+b\text{, from (ii)}\\
&= \sqrt{a}[B^H(t)-B^H(t_b)]-c(t-t_b)+b
\end{aligned}$$

If t_b is a fixed constant, $\{Q(t),\ t\in[t_b,t_2)\}$ is equivalent to $\{\sqrt{a}B^H(t)-ct+b$, $t\in[0,t_2-t_b)\}$ in distribution. Roughly speaking, locally $Q(t)$ performs like a fractional Brownian motion with negative drift in $[t_b,t_2)$ given the condition that $Q(t)$ increases from 0 to b in $[t_1,t_b]$. This motivates us to approximate the busy period of $Q(t)$ with a new process $\{X(t)=b+\sqrt{a}B^H(t)-ct, t\in[0,s_2]\}$, given that $X(-s_1)=0$, $X(0)=b$. $[-s_1,0]$ and $[0,s_2]$ of $X(t)$ correspond to $[t_1,t_b]$, $[t_b,t_2]$ of $Q(t)$ respectively. Also s_1 and s_2 correspond to t_{0b} and t_{b0} respectively. From section III, we know that the values of t_{0b} $(=t_b-t_1)$ concentrates on its mean value and $E(t_{0b})/std(t_{0b})\ge 1$. So we assume that s_1 is a constant. s_2 is the first time that $X(t)$ reaches 0, that is, $s_2=\min_{t\ge 0}\{t: X(t)\le 0\}$ (Figure 10).

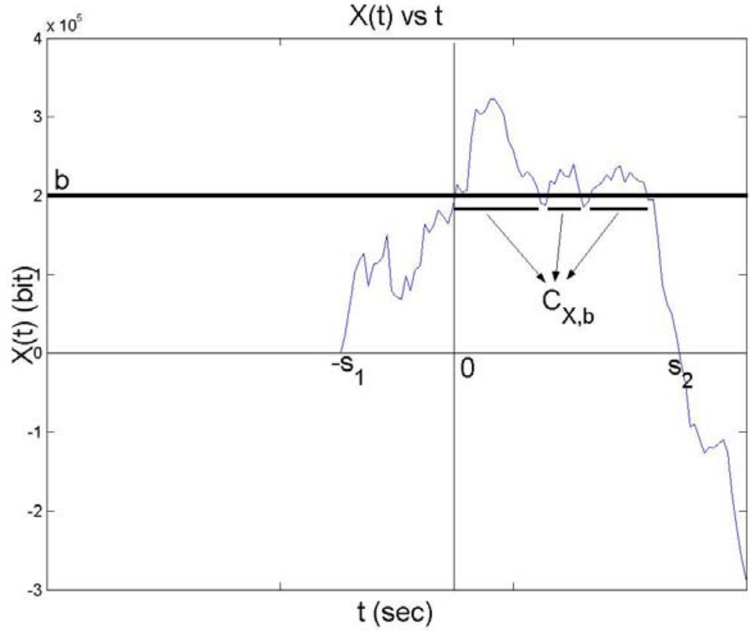

Figure 10

For the random variable s_2, we have $P(s_2 \le t) = P(\inf_{0\le s\le t} X(s) \le 0)$. It is difficult to derive its distribution. But it is easy to find a lower bound,

$$P(s_2 \le t) = P(\inf_{0\le s\le t} X(s) \le 0) \ge P(X(t) \le 0)$$

Considering the condition that $X(-s_1) = 0$, $X(0) = b$, we obtain that $B^H(0) = 0$, $\Delta B^H(s_1) = B^H(0) - B^H(-s_1) = \frac{b + cs_1}{\sqrt{a}}$.

So $P(s_2 \le t) \ge P(X(t) \le 0) = P\{B^H(t) \le \frac{ct - b}{\sqrt{a}} \mid \Delta B^H(s_1)\}$

For $t \ge 0$, $(B^H(t) \mid \Delta B^H(s_1))$ is conditional Gaussian with parameters $\mu(t, s_1)$ and $\sigma^2(t, s_1)$, where

$$\mu(t,s_1) = \frac{(t+s_1)^{2H} - t^{2H} - s_1^{2H}}{2s_1^{2H}} \Delta B^H(s_1),\ \sigma^2(t,s_1) = t^{2H} - \frac{[(t+s_1)^{2H} - t^{2H} - s_1^{2H}]^2}{4s_1^{2H}}$$

Thus,

$$P(s_2 \le t) \ge P(X(t) \le 0) = \Phi(\frac{(ct-b)/\sqrt{a} - \mu(t,s_1)}{\sqrt{\sigma^2(t,s_1)}}).$$

Define a random variable T, such that, $P(T \le t) = P(X(t) \le 0)$. So we have

$$P(s_2 \le t) \ge P(T \le t) = \Phi(\frac{(ct-b)/\sqrt{a} - \mu(t,s_1)}{\sqrt{\sigma^2(t,s_1)}}) \qquad (4)$$

As people found in [4], [5], this lower bound is a good approximation of $P(s_2 \le t)$ when the load is light and b is large. Besides the distribution, we also compared the expectations of both random variables. Empirically, $Es_2 \le ET$. When the load is light and b is large, ET is close to Es_2, as shown in the following figures.

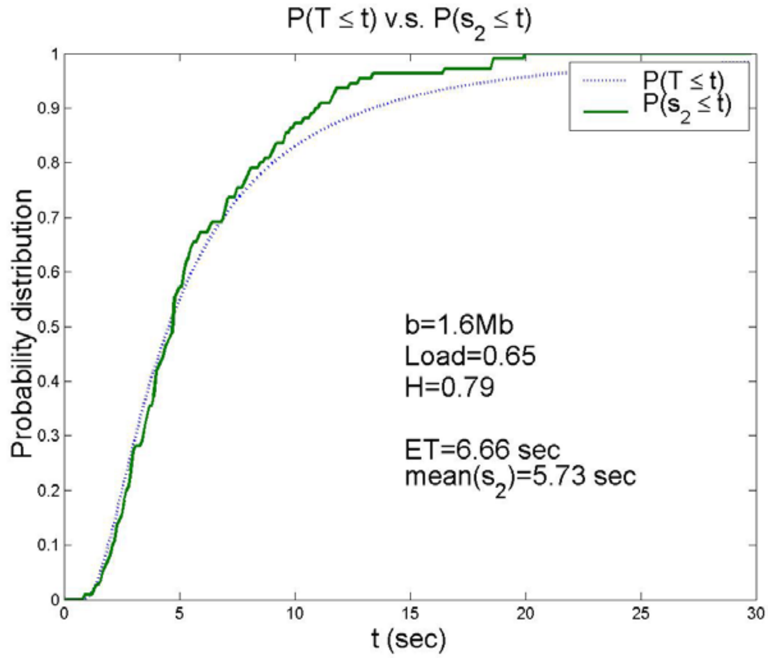

Figure 11

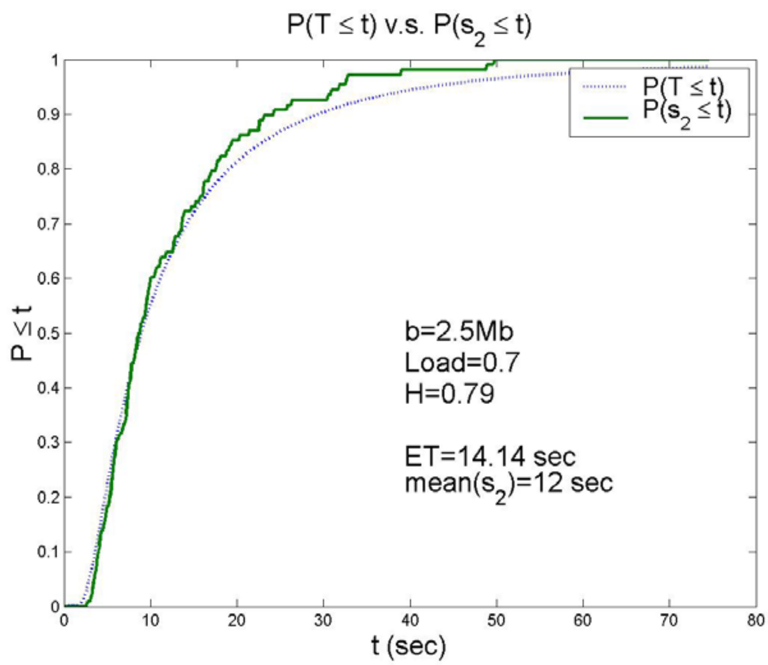

Figure 12

V. Computation of the mean sojourn time $EC_{Q,b}$

Let $C_{Q,b}$ denote the sojourn time of $Q(t)$ above a threshold b in a busy period. Here we present a way to approximate the mean value of $C_{Q,b}$, i.e. $EC_{Q,b}$.

Since we approximate the busy period with $X(t)$, we will approximate $C_{Q,b}$ with $C_{X,b}$, the sojourn time of $X(t)$ above a threshold b. Thus we have

$$EC_{Q,b} \approx EC_{X,b} \tag{5}$$

Given s_1, $C_{X,b}$ can be expressed as $C_{X,b}(s_1) = \int_0^{s_2} 1_{[b,\infty)}(X(t))dt$.

Thus, $$EC_{X,b}(s_1) = E\int_0^{s_2} 1_{[b,\infty)}(X(t))dt \tag{6}$$

From section III, we know that $Es_2 / std(s_2) > 1$ and most values of s_2 are less than $2Es_2$. But the values which are larger than $2Es_2$ affect $EC_{X,b}$ greatly. So for simplicity, we choose $2Es_2$ as a balance point, to replace s_2 in (6). Then,

$$\begin{aligned} EC_{X,b}(s_1) &\approx E\int_0^{2Es_2} 1_{[b,\infty)}(X(t))dt = \int_0^{2Es_2} P(X(t) \ge b)dt \\ &= \int_0^{2Es_2} P(b + \sqrt{a}B^H(t) - ct \ge b \mid \Delta B^H(s_1))dt \\ &= \int_0^{2Es_2} P(B^H(t) \ge \frac{ct}{\sqrt{a}} \mid \Delta B^H(s_1))dt \\ &= \int_0^{2Es_2} [1 - \Phi(\frac{ct/\sqrt{a} - \mu(t,s_1)}{\sqrt{\sigma^2(t,s_1)}})]dt \end{aligned} \tag{7}$$

However Es_2 is still difficult to compute. But in section IV, we know that for rare events (load ≤ 0.8 and b is large) ET is close to Es_2. Thus we will replace Es_2 with ET. Even though Es_2 is overestimated by ET (empirically $Es_2 \le ET$), as we will see that the trend of the computed hitting time is not very sensitive to this overestimation.

Replacing Es_2 by ET in (7), we obtain,

$$EC_{X,b}(s_1) \approx \int_0^{2ET} [1 - \Phi(\frac{ct/\sqrt{a} - \mu(t,s_1)}{\sqrt{\sigma^2(t,s_1)}})]dt \tag{8}$$

$EC_{X,b}$ is a function of s_1. As shown in Figure 13, $EC_{X,b}$ has a unique minimum as s_1 varies. We use the minimum as the value of $EC_{X,b}$ and combine with (5), then

$$EC_{X,b} = \min_{s_1 > 0} EC_{X,b}(s_1) \approx EC_{Q,b} \tag{9}$$

If more knowledge of s_1 is available, we can get better results. For example, if the expectation of s_1 is known, we can compute $EC_{X,b}$ as $EC_{X,b} = EC_{X,b}(Es_1) \approx EC_{Q,b}$ (Figure 16).

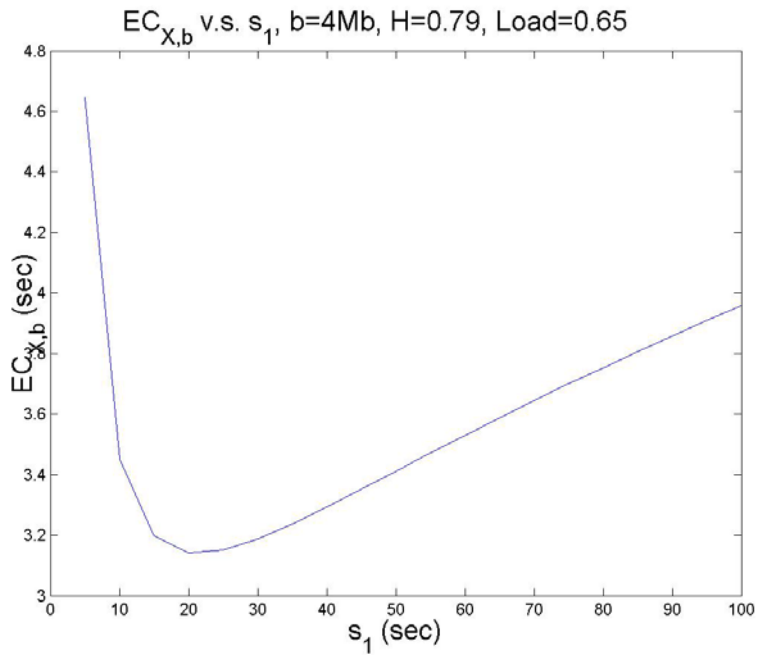

Figure 13

Combining (9) and (3), we obtain $E\tau_b$, the average time between congestion events. Although several approximations are applied in the procedure, the results are satisfying. In the next section, we will evaluate this method and show that the approximation results can follow the trend of the simulation results.

VI. Evaluation

Based on the above analysis there are two ways of predicting the average time between congestion events: 1) the reciprocal of the tail of the queue fill probability distribution, $1/P(Q>b)$, and 2) the approach derived above.

Simulations were conducted to compare the two methods and the simulation results. $B^H(t)$ is generated with the method in [6] and is used to generate the arrival process $A(t)$. The parameters m and a are fixed to be $2\times10^6 bit/s$ and $4\times10^{11} bit^2$ respectively, which are the representative of the LAN traffic [4]. H and $\rho = m/\mu$ are varied to modify the long range intensity and the traffic load.

Figures 14-16 provide a representative comparison of the predictions based on the method described above.

The difference between the simulation results and the approximation developed here primarily arises from the choice of the value of s_1 and the error from the approximation for the probability $P(Q>b)$ [3]. In Figure 16, simulation was used to optimize the selection of s_1.

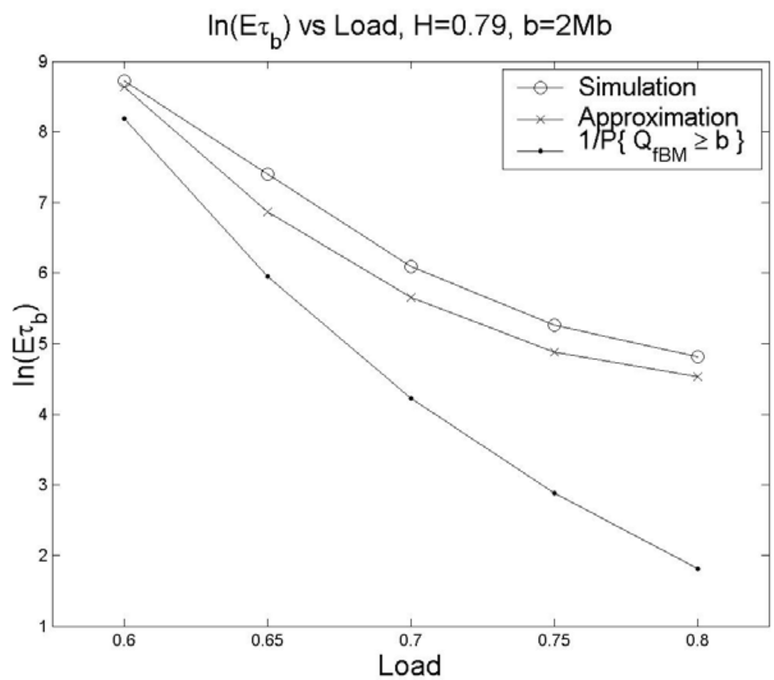

Figure 14

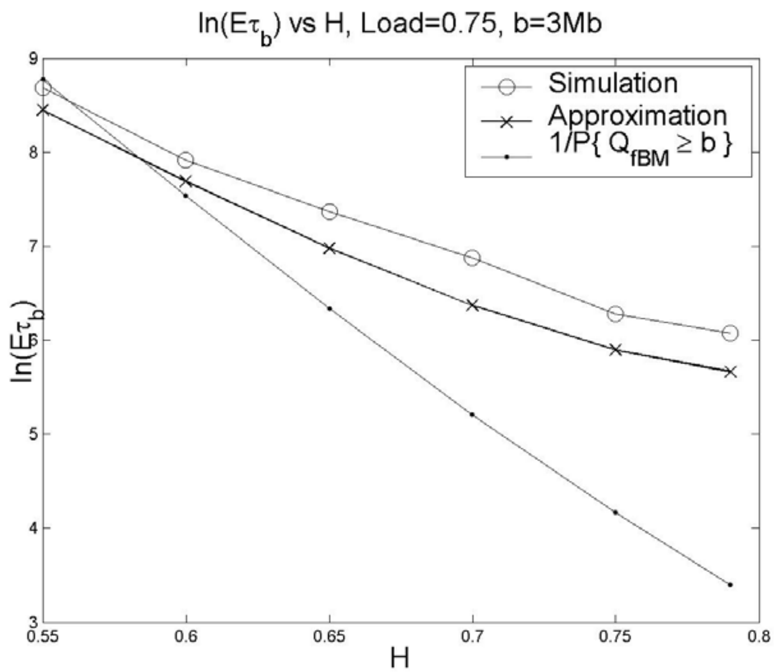

Figure 15

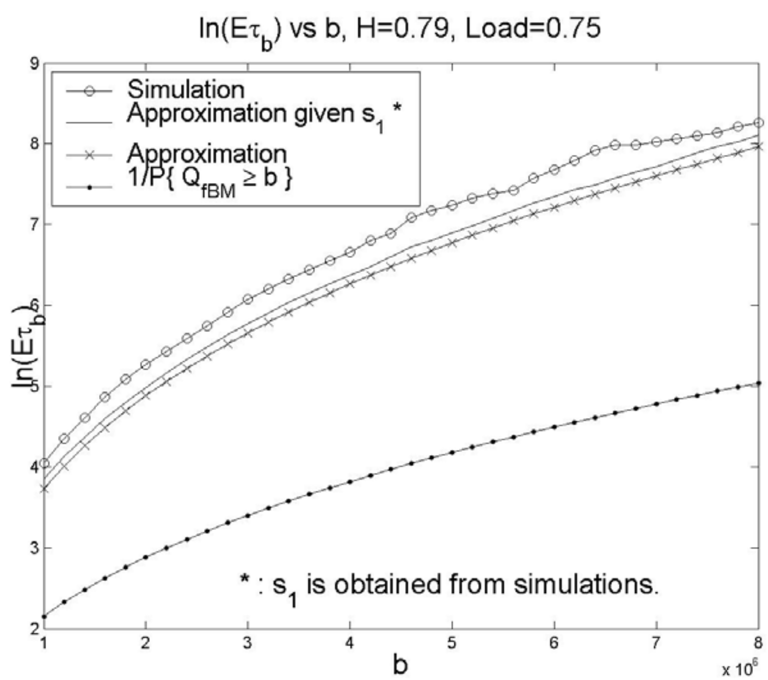

Figure 16

Figure 16, "Approximation given s_1" is calculated in the following way $E\tau_b \approx EC_{Q,b} / P(Q \geq b) \approx EC_{X,b}(Es_1) / P(Q \geq b)$, where Es_1 is measured in simulations. It shows that if we have improved knowledge of s_1, we can gain better results.

VII. Conclusions

A predictor for the time between congestion events is derived based on the Poisson clumping method. Although several approximations are applied in the derivation, the results can follow the trend of the average interevent times obtained from simulations. Our work shows that the reciprocal of the tail probability, $1/P(Q>b)$, is a poor indicator for the interevent times in most cases, especially for large H.

References

[1] D. Aldous, "Probability Approximations via the Poisson Clumping Heuristic", Springler-Verlang, New York, 1989.

[2] V. S. Frost, "Quantifying the Temporal Characteristics of Network Congestion Events for Multimedia Services", IEEE Transactions on Multimedia, Vol 5, No. 3, pp 458-465, Sep. 2003.

[3] O. Narayan, "Exact asymptotic queue length distribution for fractional Brownian traffic", Advances in Performance Analysis, vol. 1, no. 1, pp39-63, 1998.

[4] J. Roberts, U. Mocci, J. Virtamo (editors), "Broadband Network Teletraffic Performance Evaluation and Design of Broadband Multiservice Networks : Final Report of Action COST 242", Springer, Berlin, 1996.

[5] I. Norros, "A storage model with self-similar input", Queueing Systems 16, pp387-396, 1994.

[6] J. Beran, "Statistics for Long-Memory Processes", Monographs on Statistics and Applied Probability, Chapman and Hall, New York, NY, 1994.