Long Range Dependence in vBNS ATM Cell Level Traffic*

Ronn Ritke†and Mario Gerla
UCLA – Computer Science Department,
405 Hilgard Ave., Los Angeles, CA 90024
ritke@cs.ucla.edu, gerla@cs.ucla.edu

Abstract

Long Range Dependent (LRD) network traffic does not behave like the traffic generated by the Poisson model or other Markovian models. The main difference is that LRD traffic increases queueing delays by several oders of magnitude with respect to Poisson traffic due to the burstiness over many time scales. LRD traffic has been measured in different types and sizes of networks, for different applications (eg. WWW) and different traffic aggregations. Since LRD behaviour is not rare nor isolated, accurate characterization of LRD traffic is very important in order to predict performance and to allocate network resources. The Hurst parameter is used to describe the degree of LRD and the burstiness of the traffic. In this paper we introduce a methodology for measuring LRD parameters in an ATM environment and apply it to analyze segments from a vBNS ATM cell level network traffic trace. A variance-time plot is created and the Hurst parameter is computed for each segment.

Key words: Hurst parameter, Long Range Dependence, vBNS, ATM.

*This research is supported by the National Science Foundation through grant NCR-9613876 and grant ANI-9805436.

†Ronn Ritke is the corresponding author, Fax: 1 310 8257578

1 Introduction

Accurate characterization of Internet traffic is very important for precise modeling and network design decisions. Modeling of Internet traffic is based on traffic measurements and the resulting models often serve as input for simulations. The results of simulations are used for a number of network design decisions. For many years the Poisson model was widely used to model Internet traffic, but in the last few years new characteristics have been discovered in Internet traffic. Long Range Dependence (LRD) has been discovered in LANs [1] [2] , WANs [3] and MANs [4]. It has also been discovered in different services and applications: Aggregate traffic [1] [2] , World Wide Web [5] [6] , Variable-Bit-Rate (VBR) video traffic [7] and different types of computer networks: Ethernet [1] [2] , ISDN [8] and ATM [4]. The traffic with the LRD property is more bursty than traffic generated with the Poisson model. The Poisson model is Short Range Dependent and does not accurately model LRD traffic [3]. In comparison to LRD traffic, the use of the Poisson model (or other Markovian models) results in overly optimistic queueing performance. The queue length distribution decays much more slowly for LRD traffic. The queueing delay rises dramatically with increasing LRD [10] and the Hurst parameter quantifies this long range dependence.

Leland, Taqqu, Willinger and Wilson [1] compared a current model (a compound Poisson process) with actual network traffic. They found that after aggregation over the seconds time scale the compound Poisson process is very smooth. In contrast, real traffic did not smooth out and is bursty over many time scales (self-similar). They argue that the Hurst parameter quantifies the degree of self-similarity and can be used as a measure of the burstiness of the traffic.

A number of segments from a real ATM traffic trace was used to check for LRD at the ATM cell level in vBNS traffic, leading to results similar to [4].

The rest of the paper is organized as follows: Section 2 contains the ATM traffic trace information. The definition for Long Range Dependence is given in Section 3. Section 4 contains the definition for and generation of the time series used in this paper. The computation of the Hurst parameter and a discussion on the variance-time plot and examples of its use in estimating the Hurst parameter with ATM traffic is covered in Section 5. An example of the impact of LRD traffic is presented in Section 6. Lastly, Section 7 summarizes the findings.

2 Traffic Traces

A. vBNS ATM Cell Level Traffic Trace

The ATM cell level traffic trace was taken on Friday March 19 1999 using the OC12MON monitor on the vBNS at Downer's Grove (DNG) in Chicago. The trace represents the network traffic on the vBNS IP over ATM network (http://www.vbns.net/). While [4] found LRD in ATM cell level traffic, the ATM traces came from an experimental network (BAGNet) which was not regulary used and generally ran at very low utilization. In contrast, the vBNS ATM trace we used is from a network that is used regularly and run at higher utilizations. The trace we used represents over 10 percent utilization of an OC12 link. The resulting trace was processed to obtain the format needed to test for LRD (arrival time and cell length for each cell). The original ATM trace is an aggregate traffic trace. Due to the large number of arrivals in a short time period (over 256 million cells in 28 minutes) the original ATM trace was divided into segments of size 200M. ATM traffic trace segment information is summarized in Tables 1 and 2.

Trace Name	Start Date	Start Time	Time Span (Sec)	Arrivals
atm200MSeg1	3/19/99	12:26:15.920	50.5832	8000000
atm200MSeg2	3/19/99	12:26:43.007	50.5634	8000000
atm200MSeg3	3/19/99	12:31:50.841	55.7397	8000000
atm200MSeg4	3/19/99	12:37:17.771	56.3392	8000000

Table 1: vBNS ATM Cell Level Traffic Trace Segment Information.

Trace Name	Total Bytes	Bytes/Sec	Bytes/Arrival	Arrival/Sec
atm200MSeg1	4.42577E+08	8.74949E+06	53	158155
atm200MSeg2	4.42577E+08	8.75290E+06	53	158217
atm200MSeg3	4.42577E+08	7.94006E+06	53	143524
atm200MSeg4	4.42577E+08	7.85557E+06	53	141997

Table 2: Detailed vBNS ATM Cell Level Traffic Trace Segment Information.

3 Definition of Self-similar and LRD processes

Our approach is to define LRD following the definitions given in [1] [9].

Let $X = (X_t : t = 0, 1, 2, \ldots)$ be a covariance stationary stochastic process with mean μ, variance σ^2 and autocorrelation function $r(k), k \geq 0$. Assume $r(k)$ is of the form

$$r(k) \sim k^{-\beta}, \quad as \quad k \to \infty \tag{1}$$

where $0 < \beta < 1$.

For each m=1,2,3,..., let $X^{(m)} = (X_t^{(m)} : t = 1, 2, 3, \ldots)$ denote the new covariance stationary time series obtained by averaging the original series X over non-overlapping blocks of size m, i.e.,

$$X_t^{(m)} = (X_{tm-m+1} + \cdots + X_{tm})/m, \quad t \geq 1 \tag{2}$$

The process X is called (exactly) second-order self-similar if for all $m = 1, 2, 3, \ldots,\ var(X^{(m)}) = \sigma^2 m^{-\beta}$ and

$$r^{(m)}(k) \sim r(k), \quad k \geq 0 \tag{3}$$

The process X is called (asymptotically) second-order self-similar if for all k large enough,

$$r^{(m)}(k) \to r(k), \quad as \quad m \to \infty \tag{4}$$

The key property of this class of self similar processes is the fact that the covariance does not change under block aggregation and change in time scale. The relationship between the Hurst parameter and β is $H = 1 - \beta/2$. Note that here $1/2 < H < 1$, since $0 < \beta < 1$. A self similar process with $1/2 < H < 1$ (i.e., $\beta < 1$) is long range dependent (LRD).

Since $\beta < 1$ the function $\sum_k r(k) = \sum r^{-\beta} = \infty$. By contrast, a short-range dependent process (eg. Poisson Process) has fast decaying autocorrelation function (i.e., $\beta > 1$), hence, $\sum_k r(k) < \infty$. The Hurst parameter is thus a key indicator of LRD behavior. One immediate consequence of LRD behavior is that the traffic exhibits burstiness across many time scales (see Figure 1).

Having introduced the definitions for the Hurst parameter and LRD, in the following section we define an important time series and its relationship to the original traffic trace.

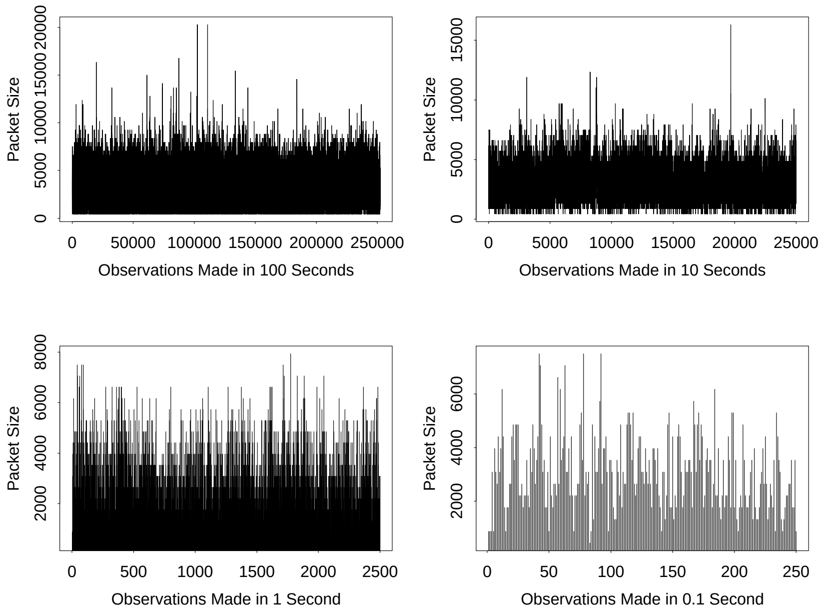

Figure 1: Traffic Trace Showing Burstiness Over Many Time Scales.

4 Time Series Definition and Generation

This paper examines vBNS ATM cell level traffic to determine if it is LRD. Generally, a traffic trace is characterized by two variables: time (of arrival of a cell) and length (of the cell). ATM cell lengths are fixed at 53 bytes. Thus, from this trace a time series with only one variable is generated in order to estimate its Hurst parameter. There are several methods for generating such single variable time series from data traces. Researchers from

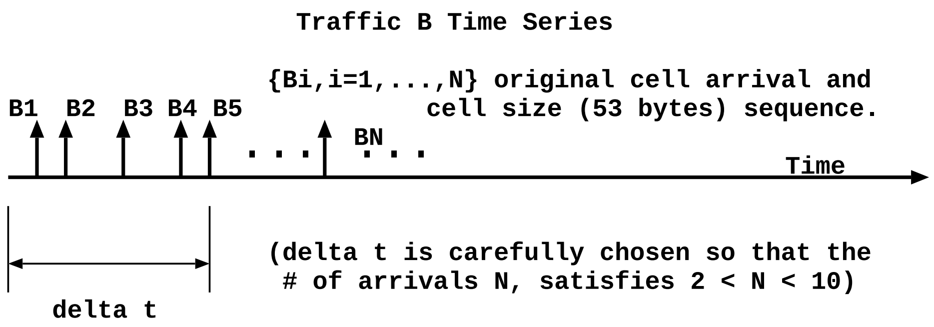

Figure 2: ATM Time Series Diagram.

Bellcore [1] [2] have proposed the {Traffic B} time series. To compute {Traffic B}, we choose a time interval Δt which typically contains between 2 and 10 arrivals (see Figure 2). Within non-overlapping time intervals of size Δt we sum the number of bytes $\overline{B}_i$ arriving in each interval Δt_i and get the time series $\{\text{Traffic B}\} = \{\overline{B}_i, i = 1, 2, 3, \ldots\}$. We derive the same synthetic traffic trace ({Traffic B}) from the real trace to see if the traffic is LRD. Next we examine the degree of self-similarity of this trace by calculating the Hurst value. The relationship of the time series to the cell arrival times and cell lengths is displayed in Figure 2. The time series captures different aspects of the traffic trace.

5 Hurst Parameter Estimation and Computation

A. Estimation of the Hurst Parameter

We present two different views of the Hurst parameter: one is a visual view (variance-time plot), the other is a computed view (see next subsection). To visually estimate the Hurst parameter, we plot $var(X^{(m)})$ as a function of m. The variance-time plot draws the variance vs. m in a log-log scale, which shows the slowly decaying variance of a self-similar series. If the input traffic has the LRD property, the curve should be linear (for large m) with slope larger than -1. The 'Reference' line on the variance-time plot (Figure 3) represents the slope of the line of $\beta = 1$, that is $var(X^{(m)}) = m^{-1}$, then H = 1/2. This reference line corresponds to a poisson process, with no long range dependence (LRD). Any line with a slope less than 0 and greater than this reference line exhibits LRD and has an H parameter value $1/2 < H < 1$. The variance-time plots for the {Traffic B} time series derived from each of the ATM trace segments are shown in Figure 3. The captions represent the curves top down on the graph. By inspection of the variance-time plot, it is apparent that the curves on all 4 of the plots have an H value greater than 1/2 and less than 1, demonstrating that the curves show the property of Long Range Dependence.

B. Computation of the Hurst Parameter

The Hurst parameter was also computed using Least-Squares Curve Fitting [11], leading to an analytic equation for the curve. The resulting equation is in the form of $y = -\beta x + b$ where $-\beta$ is the slope of the curve. The Hurst value is then computed using the relation $H = 1 - \beta/2$. Table 3 contains the computed Hurst values of each {Traffic B} for each ATM traffic trace sement used in this paper. Notice that all 4 computed Hurst values show that the segments are LRD.

Traffic Trace	{Traffic B}
atm200MSeg1	0.7060
atm200MSeg2	0.7001
atm200MSeg3	0.8076
atm200MSeg4	0.7323

Table 3: Hurst values for the vBNS ATM trace segments.

6 LRD Impact

Vern Paxson and Sally Floyd [3] wrote a paper on the failure of the Poisson Model in modeling a number of types of network traffic. In spite of this the Poisson Model is still widely used for traffic modeling. I include Figure 4 in order to show the difference bewteen the impact of LRD ATM traffic at higher network utilizations as compared to a synthetic M/M/1 trace which uses the values from 20M of the ATM cell level trace atm200MSeg1. Notice that as network utilization increases there is also an increase in the difference in average queue delay

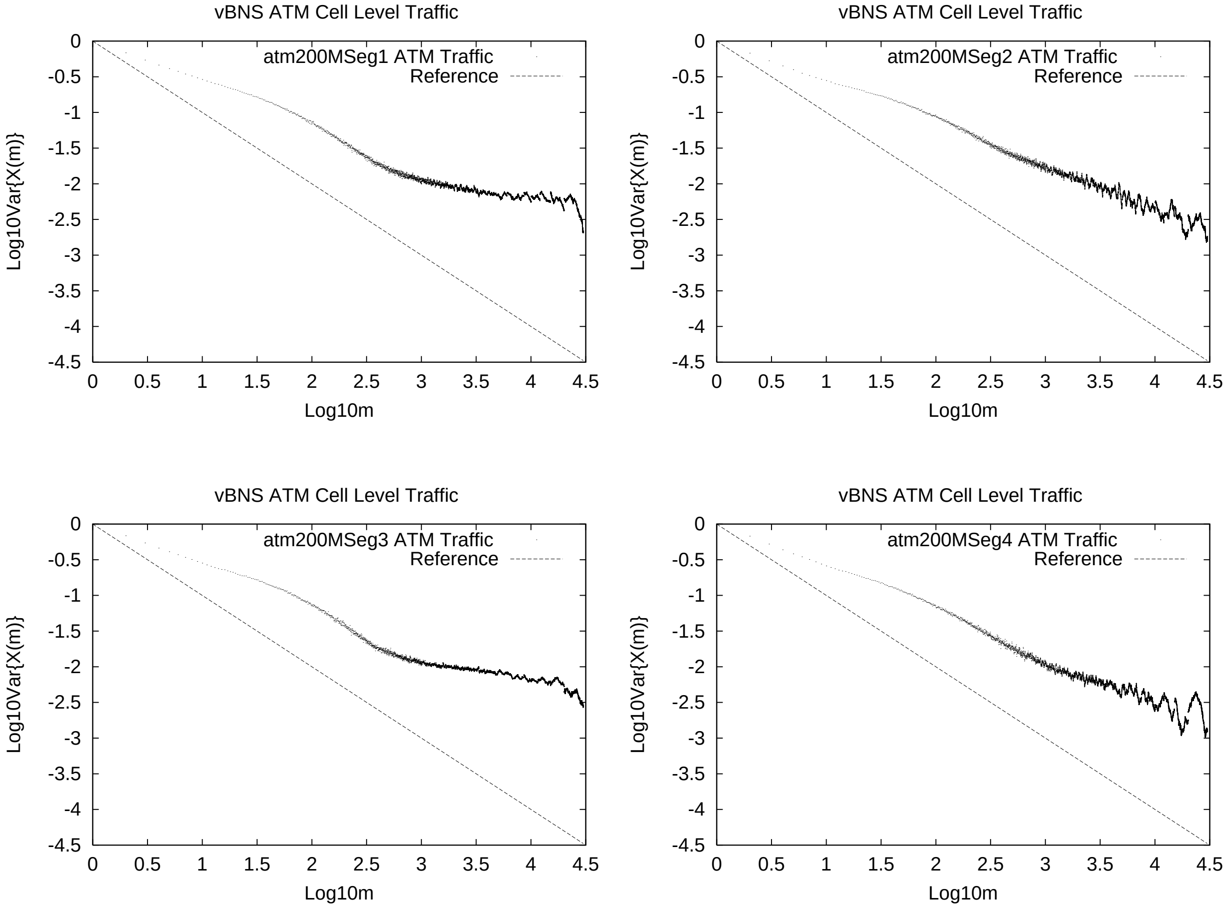

Figure 3: Variance-time plots for the vBNS ATM trace segments.

for the LRD ATM trace as compared to the M/M/1 trace.

7 Conclusion

Segments from a real ATM traffic trace were used. The results show that vBNS ATM cell level traffic is Long Range Dependent (LRD). In particular, the variance-time plots and the computed Hurst values for all 4 segments confirm the fact that they are LRD. For example, the {Traffic B} time series derived from the ATM trace has a Hurst value H = 0.7. Recall that for an LRD process, $1/2 < H < 1$.

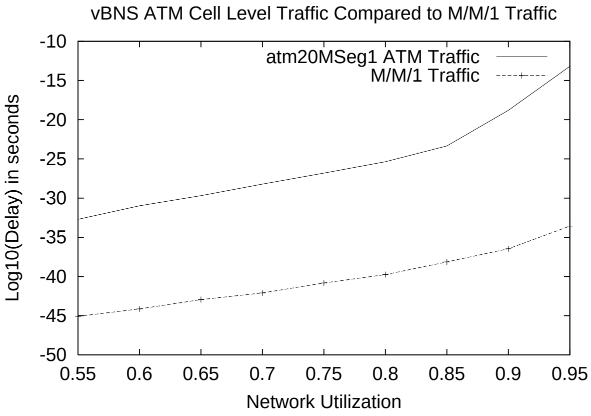

Figure 4: Average Queue Delay for LRD ATM and M/M/1 traffic.

8 *Acknowledgments*

Authors would like to give special thanks to Greg Miller and Joel Apisdorf from MCI Worldcom for their help in obtaining the vBNS ATM cell level trace, and thanks to Joe Albowicz and Xiao Chen for their help with two of the graphs.

References

[1] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, 1994: On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Trans. on Networking*, **2**, no. 1, pp. 1-15, Feb. 1994.

[2] W. Willinger, M. S. Taqqu, M. S. Sherman, and D. V. Wilson, 1996: Self-similarity through high-variability: Statistical analysis of ethernet LAN traffic at the source level (extended version). *IEEE/ACM Trans. on Networking*, **5**, pp. 71-86, December 1996.

[3] Vern Paxson and Sally Floyd, 1995: Wide-Area Traffic: The Failure of Poisson Modeling. *IEEE/ACM Trans. on Networking*, **3**, no. 3, pp. 226-244.

[4] Walter Willinger, Siddhartha Y. Devadhar, et. al., 1997: Measuring ATM Traffic Cell-by-Cell: Experiences and Preliminary Findings from BAGNet. *PMCCN '97*, International Conference on the Performance and Management of Complex Communication, Tsukuba, Japan, November 17-21, 1997.

[5] Mark E. Crovella and Azer Bestavros, 1995: Explaining World Wide Web Traffic Self-Similarity. *Technical Report TR-95-015, Boston University, August 1995.*

[6] Mark E. Crovella and Azer Bestavros, 1996: Self-Similarity in World Wide Web Traffic Evidence and Possible Causes. *1996 ACM SIGMETRICS Int. Conf. Measurement, Modeling of Computer System*, pp. 160-169, May 1996.

[7] Jan Beran, Robert Sherman, M. S. Taqqu and W. Willinger, 1995: Long-Range Dependence in Variable-Bit-Rate Video Traffic. *IEEE/ACM Trans. on Communications*, **2**, no. 4, pp. 1566-1579.

[8] H. J. Fowler and W. E. Leland, 1991: Local area network traffic characteristics, with inplications for broadband network congestion management. *IEEE J. Select. Areas Commun.*, **9**, pp. 1139-1149.

[9] W. Willinger, M. S. Taqqu, et. al., 1995: Self-Similarity in High-Speed Packet Traffic: Analysis of Modeling of Ethernet Traffic Measurements. *Statistical Science*, **10**, no. 1, pp. 71-86, December 1995.

[10] A. Erramilli, O. Narayan, and W. Willinger, 1996: Experimental queueing analysis with long-range dependent packet traffic. *IEEE/ACM Trans. on Networking*, **4**, no. 2, pp. 209-223, April 1996.

[11] Trishor S. Trivedi, 1982: Probability & Statistics with Reliability, Queueing, and Computer Science Applications. *Prentice-Hall Inc.*, Englewood Cliffs New Jersey, pp. 543-545.