THE IMPACT OF LONG RANGE DEPENDENCE ON CELL LOSS IN AN ATM WIDE AREA NETWORK

Attila Vidács, Sándor Molnár, Géza Gordos
High Speed Networks Laboratory, Dept. of Telecommunications and Telematics
Technical University of Budapest, Sztoczek u. 2., H-1111, Budapest, Hungary
Tel: (361) 463 3889, Fax: (361) 463 3107, E-mail: {vidacs, molnar}@ttt-atm.ttt.bme.hu

István Cselényi
Broadband Team, Telia Research AB, Rudsjöterrassen 2., S-13680, Haninge, Sweden
Tel: (468) 707 5038, Fax: (468) 707 5596, E-mail: Istvan.I.Cselenyi@telia.se

ABSTRACT

This paper reports our results on the effects of long range dependent (LRD) traffic on cell loss in an Asynchronous Transfer Mode (ATM) Wide Area Network (WAN). An LRD and queueing analysis of a large number of ATM traffic traces have been performed. The traces were measured by a custom-built measurement tool which is able to record more than 8 million consecutive cell arrivals. The external shuffling method has been applied to reveal the effects of short and long range correlations on cell loss. The analysis considers different buffer sizes and load situations with a wide range of shuffling block sizes. The results indicate that there is an upper time scale which limits the range of correlations of interest and a discussion of how to find this time scale is presented.

1 INTRODUCTION

The analysis of several traffic measurements taken from different high speed networks (see [9, 11, 12, 16] and references therein) revealed interesting traffic properties which are usually termed as *fractal properties*. One of the most interesting phenomenon among these is the *long range dependence* (LRD).

There are many different approaches to capture LRD including models based on self-similarity (e.g. Fractional Brownian Motion)[14], nonstationarity (e.g. level shifting models)[5] or complex traditional models (e.g. hierarchical Markovian models) [1]. The inherent nature of real traffic is a topic of many present researches [5, 9, 13, 15] but from a very engineering point of view, we do not care about how well a traffic model matches the properties of real data. Instead, we are focusing on *finding appropriate practical methods for relevant traffic engineering problems* like buffer dimensioning or call admission control. In this paper we contribute some new results to buffer dimensioning and cell loss estimation by finding important practical properties based on our queueing studies.

There are different concerns about whether LRD is an important traffic characteristics for cell loss estimation or not [4, 6]. We investigate this question and try to identify the relevant correlation time scales of actual measured traffic. We have used a large amount of long traffic traces taken during a trial on the Swedish ATM wide area network and performed LRD, queueing and shuffling analysis in order to investigate

- the effect of LRD on cell loss,
- the relevance of different time scales on cell loss,
- the effect of different buffer sizes on relevant time scales, and
- the effect of different loads on relevant time scales.

2 ATM MEASUREMENTS

LAN interconnection is one of the most popular services provided by Telia, the Swedish network operator, on its ATM wide area network. Apart from business customers, different parts of the Swedish University Network (SUNET) are also attached to the Swedish ATM WAN. The aggregated traffic on the SUNET were analyzed during summer 1996 in the framework of a common trial between the SUNET community and Telia Research. The LAN traffic of universities in the northern region, around Uppsala are connected to an FDDI backbone which is connected via R1, R2 routers and a 34 Mbps PDH link to the ATM backbone in Stockholm (Figure 1). This network joins the northern LANs of SUNET to the

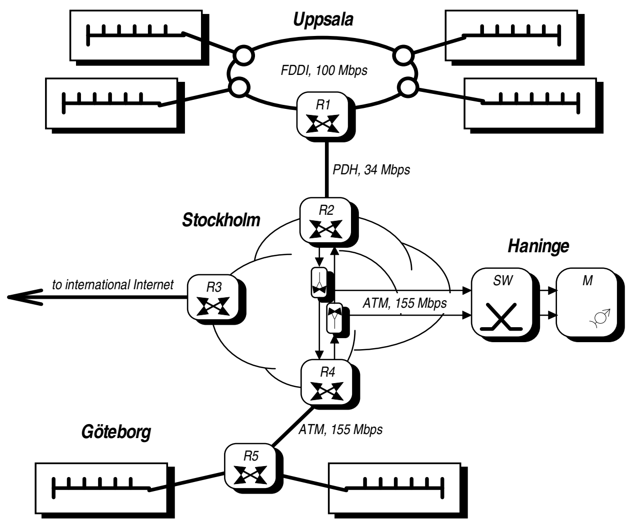

Figure 1: The configuration of measurements on the SUNET.

international Internet backbone and to the southern university networks around Göteborg. A CBR connection with 90 000 cps (38.16 Mbps) cell rate was established on the SDH link between the routers R4 and R5 for the trial. The measurements reported here were performed on the connections between Uppsala and Göteborg. ATM traffic streams were duplicated by means of optical splitters avoiding impacts on original traffic flows. The duplicated traffic streams were routed on dedicated links to Telia Research in Haninge, where almost one hundred traffic traces were collected with more than 8 million cell arrivals in each trace using a non-commercial custom built measurement instrument developed in the RACE Parasol project [10].

These connections used Telia's Guaranteed Traffic Class thus the influence from other traffic in the network was negligible. A good assumption is that the traffic was an ordinary mix of common Internet traffic types such as HTTP, FTP, telnet, chat, IPphone etc.

3 LRD ANALYSIS

Let $\{X_i, i = 1, 2, \ldots\}$ be a covariance-stationary process (e.g., the sequence of cell interarrival times) with mean μ, variance σ^2 and autocorrelation function

$$\rho(i-j) = E[(X_i - \mu)(X_j - \mu)]/\sigma^2.$$

The long range dependence property can be defined as [3]:

Definition 1 *$\{X_i\}$ is called long range dependent, if there exists a real number $\alpha \in (0,1)$ and a constant $c_1 > 0$ such that*

$$\lim_{k\to\infty} \rho(k)/[c_1 k^{-\alpha}] = 1. \qquad (1)$$

If we use the parameter $H = 1 - \alpha/2$ instead of α, long range dependence occurs for $1/2 < H < 1$. The higher the value of H, the stronger the long range dependence is.

Note, that Eq.(1) is an asymptotic definition, it determines only the rate of convergence as the lag tends to infinity, but it does not specify the correlations for *any* fixed finite lags. However, this definition cannot be directly applied in practice. From an engineering point of view, we call a (necessarily finite) sample path long range dependent, if there are significant correlations present for (arbitrarily) large time scales.

To estimate the Hurst-parameter H, the R/S and variance-time analysis [3] were performed for 45 data sets. Briefly, the variance-time method uses the fact that for a long range dependent process the variance of the aggregated processes $\{X_k^{(m)}\}$ defined by

$$X_k^{(m)} = m^{-1}(X_{(k-1)m+1} + \cdots + X_{km}), \quad k \geq 1, \qquad (2)$$

scales as

$$\mathrm{var}\{X^{(m)}\} \sim c_2 m^{2H-2}, \qquad (3)$$

as $m \to \infty$ with some constant $c_2 > 0$.

The R/S method is based on the followings. Given an empirical time series $\{X_i : i = 1, \ldots, N\}$ of length N, the rescaled adjusted range statistics $R(d)/S(d)$ for a number of values d is given by $R(d) = \max\{0, W_1, W_2, \ldots, W_d\} - \min\{0, W_1, W_2, \ldots, W_d\}$, with $W_k = (X_1 + X_2 + \cdots + X_k) - k\overline{X}(d)$, $(k = 1, 2, \ldots, d)$, $S^2(d)$ is the sample variance, and $\overline{X}(d)$ is the sample mean. For a long range dependent data the R/S statistics is characterized by

$$E\{R(d)/S(d)\} \sim c_3 d^H, \qquad (4)$$

with constant $c_3 > 0$ as $d \to \infty$ and $0.5 < H < 1$.

The obtained values of H are plotted on Figure 2. To get more information about the Hurst-parameter of the measured traffic, H is plotted against the average load of the traces. As can be seen on the figure, H varies within the range $(0.8, 1)$ and does not depend significantly on the load. Figure 3 reveals an interesting phenomenon, namely, that there is a knee-point in the R/S diagram that separates two linear regions of sample points. Since LRD is an asymptotic property, the linear region to the right (sample points marked with an 'x') was used to determine the estimate of H. The same knee-point was found in all the data sets. The origin of this behaviour could be revealed by examining the burstiness structure of the data more deeply.

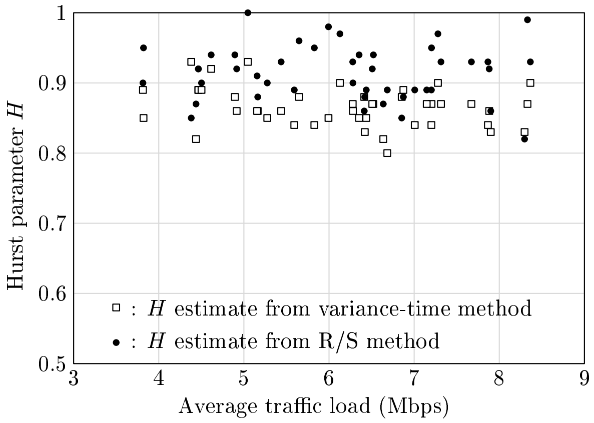

Figure 2: Estimated values of H as a function of traffic load.

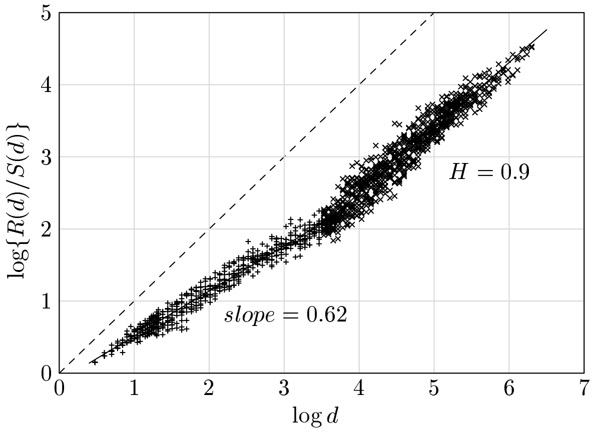

Figure 3: R/S values plotted against the logarithm of block size.

4 RELEVANCE OF TIME SCALES IN QUEUEING

An important question is what is the impact of LRD on queueing. Several engineering issues, such as buffer dimensioning and traffic control, are related to this question which makes it extremely important. There are two opposing viewpoints on this problem. One claim is that the queueing performance is determined by the time scale of busy periods of the queues and there is no practical impact of correlations above this time scale [6, 7, 8, 15]. The contradicting claim is based on several studies [4, 9] and states that LRD is one of the main characteristics of the traffic with significant effect on queueing behaviour. In this section we present some experimental results in order to clarify this question.

4.1 QUEUEING SET-UP

In the performance analysis we used the queueing set-up shown in Figure 4. The pre-recorded traces were taken from our SUNET database as described in section 2. The correlation structure of the original data was artificially modified by a 'shuffler' (see next section). The functionality of this shuffler was investigated by computer simulation in our experiment. The QoS measures under investigation were the complementary queue length distribution and the cell loss ratio (CLR).

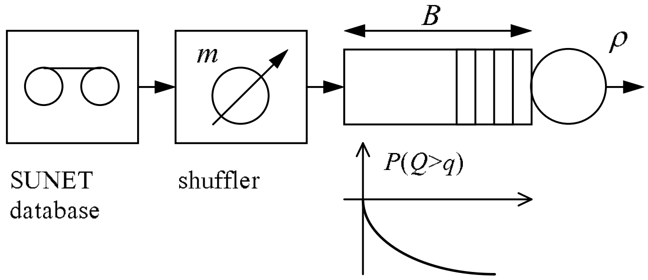

Figure 4: Queueing set-up.

4.2 EXTERNAL SHUFFLING

In our analysis, external shuffling was used as a tool to modify the correlation structure of the measured traffic traces. Using the terminology of [2, 4], we call 'external shuffling' the following method. First, divide the sequence of interarrival times[1] into blocks of size m. For a measured traffic trace containing N cell arrivals, there are N/m such blocks. Then the order of the blocks is shuffled, while preserving the cell sequence inside each block. Thus, for different values of m we preserve the short-range correlations (up to lag m) while eliminating the long-range correlations (beyond lag m).

Therefore by plotting the complementary queue length distributions for the original and the shuffled traces with different block sizes we have a simple tool to investigate the effect of short-term and long-term correlations in queueing (see Figure 5).

4.3 QUEUEING PROPERTIES OF LRD INPUT

In case of LRD traffic input, the tail of the complementary queue length distribution decays slower than exponentially. The uppermost solid line related to the original traffic trace in Figure 5 shows this phenomenon. (The buffer size was set to 4000 cells and the service rate was chosen so that the utilization ρ was 0.7. The shuffling block size m is shown

[1] Beside the permutation of a sequence of interarrival times, one could perform the same shuffling on the sequence of the number of arrivals in consecutive time slots [2].

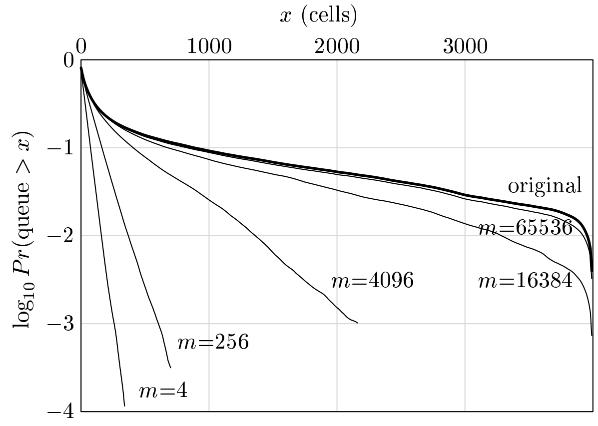

Figure 5: Complementary queue length distributions.

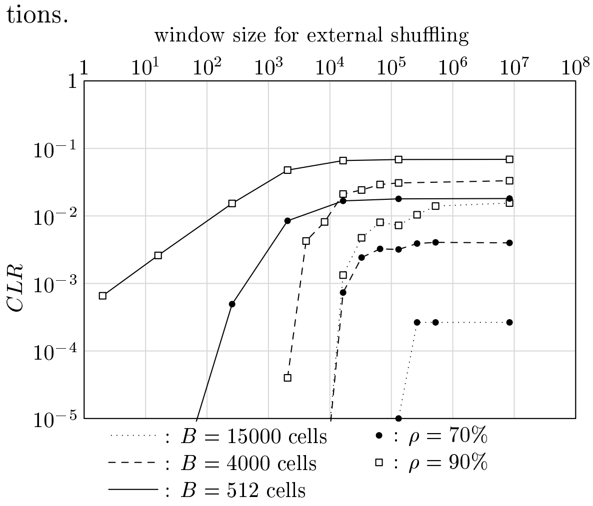

Figure 6: Cell loss ratio with external shuffling.

on the figure for each simulated curve.) By shuffling the input traffic, we can observe that if we increase the block size the curve approaches the distribution curve of the original trace. It means that we can find the time scale of the relevant correlations, i.e., the block size of shuffling where the distribution is approximately the same as for the original case, which are important for queueing. It can also be concluded that beyond that time scale there is no significant effect of correlations. In our case this time scale is in the range of 4–5 seconds (block size of 10^5) for load 0.7. This finding is in accordance with the results reported in [6].

4.4 IMPACTS OF LRD ON CELL LOSS

The queue length distribution in the infinite buffer case gives (almost) always an upper bound on the cell loss ratio (CLR) for the finite buffer case. In practice, the latter is of interest. To investigate the effects of LRD on CLR, we performed simulation studies with the measured traces as input for three different buffer sizes (512, 4000 and 15000 cells). The rate of service was set to obtain utilizations 0.7 and 0.9. The results are plotted on Figure 6.

We can observe in Figure 6 that *there is an upper time scale determined by the buffer size and the load where there is no effect of correlations on cell loss if we go beyond that time scale.* For example, in our experiments above cell lag 10^5 (approximately 5 seconds) the cell loss curves are practically constant even for the large (15000 cells) buffer case. It supports our finding derived from Figure 5, namely, that there is a time scale which determines the biggest lag of correlations which has effect on the cell loss. However, this upper time scale seems to be dependent on many parameters. First, it depends on the buffer size, i.e. the bigger the buffer the bigger the upper time scale. For moderate and large buffers there is a sharp cut-off in cell loss as a function of the shuffling block size but for small buffers the appearance of this upper time scale is not very pronounced. Secondly, this upper time scale is also dependent on the load, the curves are shifting left as the load increases although the shifting of cut-off lag is not significant. In our experiments we can observe that the cut-off lag is always above the buffer size but typically not more than one decade above the buffer size. These findings give us simple practical engineering rule of thumb for estimating the range of relevant correlation time scale. A comprehensive study of these investigations with many traffic traces and queueing conditions is performed in our research but the detailed presentation of these results is beyond the scope of this paper.

5 CONCLUSION

We reported a traffic measurement experiment on SUNET ATM WAN. The recorded long cell traces that cover many time scales were used for LRD analysis and to investigate the question how LRD behaviour influences cell loss in different queueing environments. The impacts of correlations of different time scales were analysed by using external shuffling with different block sizes.

Our basic finding is that there is an upper time scale which determines the range of correlation of interest from a cell loss point of view. This time scale sensitively depends on the buffer size and slightly depends on the load but roughly speaking it is not bigger than ten times the buffer size. By using this simple engineering rule we have a fast approximation about the time scales of queueing interest. The detailed analysis and the development of an accurate approximation of the upper time scale is the topic of our present and future research.

ACKNOWLEDGEMENT
The reported ATM WAN measurements were carried out by the CARAT laboratory of Telia Research in Sweden. We would like to thank Nils Björkman, Urban Hansson, Alexander Latour-Henner and Aziz Miah for the measurement data and their contribution to this paper.

References

[1] A.T. Andersen, A. Jensen, and B.F. Nielsen. Modeling and performance study of packet-traffic with self-similar characteristics over several time-scales with Markovian Arrival Processes (MAP). *Proc. 12th Nordic Teletraffic Seminar*, pages 269–283, Espoo, Finland, August, 1995.

[2] A.T. Andersen and B.F. Nielsen. On the implications of certain random permutations of interarrival times or counts in a point process. *COST 257TD(020) technical document*, Rome, Italy, January, 1998.

[3] J. Beran. *Statistics for long-memory processes*, Chapman & Hall, 1994.

[4] A. Erramilli, O. Narayan, and W. Willinger. Experimental queueing analysis with long-range dependent packet traffic. *IEEE/ACM Trans. on Networking*, 4(2):209–223, 1996.

[5] M. Grasse, M.R. Frater, and J,F. Arnold. Origins of long-range dependence in variable bit rate video traffic. *Proc. 15th Int. Teletraffic Congress—ITC15*, pages 1397–1388, Washington, D.C., USA, 1997.

[6] M. Grossglauser and J-C. Bolot. On the relevance of long-range dependence in network traffic. *Proc. on SIGCOMM'96*, pages 15–24, CA, USA, 1996.

[7] D.P. Heyman, T.V. Lakshman. What are the implications of long-range dependence for VBR-video traffic engineering? *IEEE/ACM Transactions on Networking*, 4(3):301–317, 1996.

[8] K. Kobayashi and Y. Takahashi. The tail probability of a Gaussian fluid queue under finite measurement of input processes. *Proc. Int. Conference on the Performance and Management of Complex Communication Networks—PMCCN'97*, pages 57–72, Tsukuba, Japan, November, 1997.

[9] W.E. Leland, M.S. Taqqu, W. Willinger, and D.V. Wilson. On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, 2(1):1–15, 1993.

[10] O. Melteig. Introduction to the PARASOL Project. *The 9th Nordic Teletraffic Seminar*, Norwegian Telecom, Research Department, August, 1990.

[11] S. Molnár, I. Cselényi, N. Björkman, C.G. Prentz, and M. Boda. ATM traffic measurements and analysis on a real testbed. *Proc. 10th Int. Teletraffic Congress Specialist Seminar*, pages 237–250, Lund, Sweden, September, 1996.

[12] S. Molnár and A. Vidács. On modeling and shaping self-similar ATM traffic. *Proc. 15th Int. Teletraffic Congress—ITC15*, Washington, D.C., USA, 1997.

[13] S. Molnár, A. Vidács, and A.A. Nilsson. Bottlenecks on the way towards fractal characterization of network traffic: Estimation and interpretation of the Hurst parameter. *Proc. Int. Conference on the Performance and Management of Complex Communication Networks—PMCCN'97*, pages 125–144, Tsukuba, Japan, November, 1997.

[14] I. Norros. A storage model with self-similar input. *Queueing Systems*, 16:387–396, 1994.

[15] B.K. Ryu and A. Elwalid. The importance of long-range dlel94gro96lel94ependence of VBR video traffic in ATM traffic engineering: Myths and realities. Proc. on SIGCOMM'96, pages 3–14, CA, USA, 1996.

[16] W. Willinger, S. Devadhar, A.T. Heybey, R. Sherman, M. Sullivan, and J.R. Vollaro. Measuring ATM traffic cell-by-cell: Experiences and preliminary findings from BAGNet. *Proc. Int. Conference on the Performance and Management of Complex Communication Networks—PMCCN'97*, pages 91–110, Tsukuba, Japan, November, 1997.