

Long Range Dependence in Communication Networks

Jean Bolot

INRIA-Sophia Antipolis

April 27, 1998

[summary by Philippe Robert]

Abstract

We examine under which conditions the salient long range dependence feature of network traffic must be taken into account in network performance evaluation. We show that “it is all a matter of time scales”. Specifically, when studying the performance of a networking system or an application, many time scales must be taken into account — the time scales in the input traffic, but also the time scales of the system (they show up for example because of finite buffer queues) and the time scales of the performance metric of interest.

1. Introduction

The talk is concerned with the behavior of a buffer of a node in a telecommunication network. Messages are assumed to arrive randomly at this node where they wait in the buffer for transmission. We denote by $(X(t))$ the stochastic process describing the number of messages arrived during the t -th unit of time. The autocorrelation function is defined as

$$r(t) = \frac{E(X(0)X(t)) - E(X(0))E(X(t))}{E(X(0)^2) - E(X(0))^2},$$

where $E(Y)$ denotes the expected value of the random variable Y . The arrival process is said to have mixing properties if $X(0)$ and $X(t)$ are nearly independent when t is large, $r(t)$ is converging to 0 as t goes to infinity. Usually this assumption is satisfied; notice however that the periodic traffics are not mixing.

A simple characterization of the input traffic is provided by the rate at which $r(t)$ tends to 0. Up to now, most of the models analyzed assumed an exponential convergence to 0. This is clearly the case if the $(X(t))$ are i.i.d. random variables, $r(t) = 0$ for all $t \geq 0$. In this case, if the buffer size is infinite, it is known that at equilibrium the number $L(t)$ of messages at time t waiting for transmission has an exponential tail; that is, there exists $\gamma > 0$ such that

$$\lim_{x \rightarrow +\infty} e^{\gamma x} P(L(t) \geq x) = c \in \mathbb{R}_+.$$

This result has practical implications. Because the buffer sizes cannot be infinite, it implies that it is not necessary to design a big buffer because large queues of messages are (exponentially) very unlikely.

However a careful statistical analysis of data collected over a wide variety of networks [4] has provided ample evidence that network traffic processes are not exponentially mixing. In this case we shall say that the traffic exhibits a long range dependence (LRD) behavior. This is the case if $r(t) \sim C/t^\beta$ for some $\beta > 0$. A popular explanation of these LRD traffics is the following: an ON/OFF traffic generated by a single source consists in random burst intervals during which the

source sends many messages alternating with idle intervals. An LRD traffic can be obtained by the superposition of an infinite number of ON/OFF traffics having larger and larger bursts (and idle) intervals.

In some of the (few) models analyzed with this kind of input traffic, it has been shown that if the buffer size is infinite then the number of messages at equilibrium has the following behavior: there exists $\delta > 0$ such that

$$P(L(t) \geq x) \sim \frac{c}{t^\delta}.$$

Notice that for the design of buffer sizes, the situation has changed. It is not clear that a small buffer will be sufficient because large queues of messages are not so unlikely.

Remark. It is possible to give a description of the occurrence of the “rare” events, when the number of messages in the buffer is greater than K , K large. In the case of the exponential decay of the autocorrelation function, it happens gradually, i.e., during some time interval the number of messages increases steadily at rate γ until it reaches K and then it decreases rapidly to 0. In the LRD case, the number of messages is greater than K in one big jump [1, 2].

2. Results

The point of view of the talk is slightly different from the usual presentation of these problems described above. It is stressed that these phenomena must be analyzed with *finite* buffer sizes, instead of guessing the behavior of the finite case from the infinite case. Simulations and measurements of various traffics over the Internet are used to carry out this analysis. The impact of the long range dependence is analyzed through the loss rate of the node, i.e., the average number of messages rejected because of the congestion of the node. The main conclusions of this approach are the following:

- The dependence on the past is limited by the size of the buffer. In other words, it is not necessary to consider a (very) long range dependent traffic to have a realistic traffic input;
- For a long range dependent traffic, increasing the buffer size has less impact than for a short range dependent traffic.

Bibliography

- [1] Asmussen (Søren). – Rare events in the presence of heavy tails. In *Stochastic networks (New York, 1995)*, pp. 197–214. – Springer, New York, 1996.
- [2] Durrett (Richard). – Conditioned limit theorems for random walks with negative drift. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 52, n° 3, 1980, pp. 277–287.
- [3] Grossglauser (Matthias) and Bolot (Jean). – On the relevance of long-range dependence in network traffic. In *Proceedings ACM Sigcomm '96*. – Stanford, CA, September 1996.
- [4] Leland (Will), Taqqu (Murad), Willinger (Walter), and Wilson (Daniel). – On the self-similar nature of ethernet traffic. *IEEE/ACM Transactions on Networking*, vol. 2, n° 1, February 1994, pp. 1–15.