Uday Pratap Singh

Roll No - 06212322

A Performance Model and Analysis of Heterogeneous Traffic with Heavy Tails

November 20, 2007

by

Imad Antonios, Lester Lipsky University of

Connecticut Computer Science and Engineering

Abstract

Several research efforts have recently attempted to characterize the effects of heavy-tailed traffic on router performance, while the effects of light-tailed traffic have for long been understood. Since general web traffic originates from heterogeneous sources, the study of traffic mixing is of importance as it can reveal the degree to which heavy-tailed traffic can be handled by networking infrastructure before router delay becomes unacceptable. We present a model for heterogeneous traffic sources, where each is an ON/OFF process with an exponential OFF time and an arbitrary ON-time distribution. We consider the case of a 2-source process where one has exponential ON time and the other power tailed and present some analysis results showing the significance of traffic mixing on router performance.

Introduction

One of the goals of the network performance modeling research has been to characterize the conditions under which router performance becomes unacceptable, linking performance to the presence of heavy-tailed transmission times, for which the term heavy-tailed traffic is used. This research has assumed that the traffic generated by all sources is heavy-tailed, which does not accurately reflect on the nature of the traffic generated by hosts. Web traffic generally consists of a confluence of independent streams that may not all necessarily be heavy-tailed. In this paper, we present a performance model for heterogeneous sources along with some analysis. The arrival process is represented as a Markov Modulated Poisson Process (MMPP) and is formulated in a Linear Algebraic Queueing framework that allows for an analytic solution of the model. The following section presents an overview of the model and includes foundational elements for the modeling approach used here, and provides insights for interpreting our results.

Model Overview and Background

The model we present in this paper is of N independent, heterogeneous ON/OFF sources . The aggregated traffic is then processed as input to a router that is modeled as a queue with exponential service rate,V. Sources have an OFF time that is exponentially distributed with mean time z, and an ON time that is represented as an arbitrary matrix-exponential, or ME, distribution with the vector-matrix pair $\langle p_i, B_i \rangle$. During an ON period, a source transmits packets in a poisson stream at a rate of λ_{p_i} , which is the peak source transmission rate. This can be viewed as a hardware or system constraint. Previous work with a model of identical sources revealed that performance blowups at the router are attributed to the heavy-tailed ON-time of the sources. First we will discuss the Truncated Power Tail distribution and its significance for traffic modeling. A discussion of the trucated power tail - distribution and N-Burst model and its analysis are the subject discussed below.

Truncated Power-Tail Distribution(TPTD)

In our traffic modeling is a Matrix Exponential representation of a power-tail distribution. This is motivated by the discovery that the ON-time (or Burst Size). Distribution must be heavy-tailed to account for the long-range dependence of the traffic measured. There is a good reason to believe that such functions are involved in network traffic. First of all, long-range dependent arrival patterns in ON-OFF systems can be caused by processes with PT ON-times. Secondly, it has been shown in numerous places that the file size distribution in many facilities, including those transmitted over the Internet, tend to have PT behavior. In this section, we briefly describe a family of heavy-tailed distributions, called Truncated Power-Tail distributions (For a full treatment of TPT), which are phase distributions, and thus can be used in analytic models of Markovian-like systems. Power- Tail (PT) distributions with exponent α , have the following properties:

$$E(X^l) = \int_{x=0}^{\infty} x^l f(x) dx = \begin{cases} = \infty, & for \;\; l \geq \alpha. \\ < \infty, & for \;\; l < \alpha. \end{cases}$$

where X is the r.v. with pdf $f(x)$. Reliability functions with tail behavior (behavior for large x) of the form:

$$R(x) = P(X \geq x) \rightarrow \frac{c}{x^{\alpha}}, \;\; (c > 0)$$

are the simplest functions having this property, and are called Power-Tail distributions. They have also been referred to as heavy-tailed, long-tailed, Pareto, or Levy functions. We now briefly describe a hyper-exponential function (which is matrix exponential) that in the limit becomes a PT distribution, and that we have used very successfully in standard queueing models. Consider the following family of functions with random variables, X_T.

$$R_T(x) = \frac{1-\theta}{1-\theta^T} \sum_{j=0}^{k-1} \theta^j exp(-\mu x/\gamma^j),$$
$$where \;\; 0 < \theta < 1 \;\; and \;\; 1 < \gamma$$

It can be shown that

$$\lim_{T \to \infty} E(X_T^l) = \infty \;\; for \;\; l \geq \alpha$$

The N-Burst Model

The N-Burst model, which was introduced to, serves as a powerful descriptor of the behavior of heavytailed traffic, and provides the means for carrying out detailed performance analysis using a linear algebraic approach to queueing theory. The model mimics the behavior of N identical and independent hosts that intermittently put data onto a telecommunications line leading to a router. Each of the hosts is an ON/OFF process represented as an MMPP with a power-tail ON time and an exponential OFF time. A source transmits packets at a rate of $\mathcal{K}$, contributing to the total packet generation rate of λ, which is simply $N \times \mathcal{K}$, During an ON period, the source transmits packets at a rate of λ_p. The burst parameter,b, is defined to be the fraction of the time that a source is OFF, namely OFF (ON + OFF) . This parameter is instrumental. This parameter is instrumental in characterizing the degree of burstiness of a traffic source, in that as b approaches 0 traffic resembles a poisson stream, whereas when b approaches the other limit of 1 traffic is presented for transmission in a very short period of time and the source is idle the majority of time. Keeping $\mathcal{K}$ constant, as b goes from 0 to 1, we can observe that the packet arrival rate during an ON period grows unbounded. One of the most significant results of this work is the discovery of the conditions that cause blowups in packet delay at the router. Performance blowup turned out to be a function of the router utilization and the burst parameter, while its magnitude is dependent on the range of the Power-Tail function and the α parameter. In the model, the router is represented as an exponential server with an average service rate of V . With the arrival process being an MMPP, the system becomes an SM/M/1 queue, which can be solved analytically. The utilization ratio of the router, ρ, is simply $N \times \mathcal{K}/V.$

For a process with one source, or the 1-burst model, a blowup point is located where $b = 1 - \rho$ or equivalently where $\lambda_p = V$. This indicates that it is sufficient for the peak rate of the source to exceed the router rate in order for a blowup to occur, irrespective of the average rate at which a source is submitting packets for transmission. If there are 2 or more sources, the situation becomes more complicated. An N-Burst process gives rise to N blowup points, each with increasing delay severity as b increases. The value of λ_p, for each blowup is given by:

$$\lambda_p(i \mid N) = \frac{V - (N-i)\mathcal{K}}{i}.$$

Since $\mathcal{K}$ = $(1-b)\lambda_p$ and $\rho = N\mathcal{K}/V$, the V value of blowup point can be given as

$$b(i \mid N) = \frac{V - N\mathcal{K}}{V - (N-i)\mathcal{K}} = \frac{1-\rho}{1-(N-i)\rho/N}.$$

Conclusion

From above presented a performance model for N Truncated power tail distribution and N-burst model. The model was expressed using a linear algebraic formulation, which captures the phenomenology of events in a system of 2- heterogeneous ON/OFF sources. We then derived an exact solution for the model endpoints, which define the performance bounds for a TPTD N-burst model. Our results show that the inclusion of non-PT sources improves router performance in two respects: (1) blowup occurs for higher values of the burst parameter compared to the N-Burst model,(2) the magnitude of the blowup is dampened in accordance with the relative contribution of non-PT sources to the total traffic.