

# A flow-based model for Internet backbone traffic

Chadi Barakat, Patrick Thiran, Gianluca Iannaccone, Christophe Diot, Philippe Owezarski

**Abstract**—Our goal is to design a traffic model for uncongested IP backbone links that is simple enough to be used in network operation, and that is protocol and application agnostic in order to be as general as possible. The proposed solution is to model the traffic at the flow level by a Poisson shot-noise process. In our model, a flow is a generic notion that must be able to capture the characteristics of any kind of data stream. We analyze the accuracy of the model with real traffic traces collected on the Sprint IP backbone network. Despite its simplicity, our model provides a good approximation of the real traffic observed in the backbone and of its variation. Finally, we discuss three applications of our model to network design and management.

## I. INTRODUCTION

Modeling the Internet traffic is an important issue. It is unlikely that we will be able to understand the traffic characteristics, predict network performance (e.g., for QoS guarantees or Service Level Agreement definition), or design dimensioning tools without analytical models. The successful evolution of the Internet is tightly coupled to the ability to design simple and accurate models.

The objective of this work is to design a traffic model that can be used in routers or by network administrators to assist in network design and management. Such a model needs to be simple, i.e., it has to be fast to compute and to rely on simple parameters that can easily be acquired by a router. Currently, network operators have very basic information about the traffic. They mostly use SNMP [8] that provides average throughput information over 5 minutes intervals. An analytical model could provide more accurate information on the traffic. It could be used in various applications such as detection of anomalies (e.g., denial of service attacks or link failures), prediction of traffic growth, or assessment of the impact on the network traffic of a new customer or of a new application. Consequently, a second important characteristic of the model we want to design is to be protocol and application agnostic: it needs to be general enough to evaluate link throughput independently of the application nature and of the transport mechanism.

Modeling the traffic at the packet level has proven to be very difficult since traffic on a link is the result of a high level of

multiplexing of numerous flows whose behavior is strongly influenced by the transport protocol and by the application. In addition, monitoring the traffic at the packet level becomes critical at OC-192 and above link speed.

Recently, a new trend has emerged, which consists in modeling the Internet traffic at the flow level (see [3] and the references therein). A flow here is a very generic notion. It can be a TCP connection or a UDP stream (described by source and destination IP addresses, source and destination port numbers, and the protocol number), or it can be a destination address prefix (e.g., destination IP address in the form a.b.0.0/16). Flows arrive at random times and share the available bandwidth in the network according to certain rules. Using Processor Sharing queues [18], it is possible to get an idea about the response time of a flow and about the distribution of the flows that are active at a certain time in the network. From a simplicity standpoint, it is much easier to monitor flows than to monitor packets in a router. Tools such as NetFlow already provide flow information in Cisco routers<sup>1</sup>.

In this paper, we propose a model that relies on flow-level information to compute the total (aggregate) rate of data observed on an IP backbone link. We are interested in capturing the dynamics of the traffic at short timescales (i.e., in the order of hundreds of milliseconds). For the purpose of modeling, the traffic is viewed as the superposition (i.e., multiplexing) of a large number of flows that arrive at random times and that stay active for random periods. As explained earlier, a flow is a generic notion that must be able to capture the characteristics of any kind of data stream.

In contrast to other works in the literature (e.g., [3], [5], [16]), we choose to model a link that is *not* congested (congestion possibly appears elsewhere on the flow path). This assumption is valid, and in fact is the rule, for backbone links that are generally over-provisioned (i.e., the network is designed so that a backbone link utilization stays below 50% in the absence of link failure [12]). It is driven by our main objective, which is to provide a link dimensioning tool usable in backbone network management.

The contribution of this work is the design of a flow-based Internet traffic model using simple mathematical tools (Poisson shot-noise). Thanks to the notion of *shots* we introduce in the purpose of modeling flow transmission rates, our model is able to compute the total rate of data in the backbone using flows' characteristics (i.e., arrivals, sizes, durations). Once the model is introduced, the paper focuses on its confrontation to real data collected on the Sprint IP backbone network. This confrontation illustrates the efficiency of the model in computing the traffic in the backbone and its variation. We then discuss three possible applications of our model, namely (i) the dimen-

C. Barakat (chadi.barakat@sophia.inria.fr) is with INRIA, Sophia Antipolis, France. This work was done while he was working with EPFL. P. Thiran (patrick.thiran@epfl.ch) is with EPFL, Lausanne, Switzerland. G. Iannaccone (gianluca@sprintlabs.com) and C. Diot (cdiot@sprintlabs.com) are with Sprint ATL, Burlingame, CA. P. Owezarski (owe@laas.fr) is with LAAS-CNRS, Toulouse, France.

<sup>1</sup><http://www.cisco.com/warp/public/732/Tech/netflow>

sioning and provisioning of the backbone, (ii) the short-term prediction of the total rate of data on a backbone link, and (iii) the generation of backbone traffic for network simulation tools. In addition, our model opens the door to many future works on a simpler modeling and better understanding of Internet backbone traffic.

In the next section, we survey the related literature and position our contribution. Section III describes the traces we use throughout the paper for the validation of our model. In Section IV, we present our model and we analyze its performance in Section V. In Section VI, the model is confronted to the real traces. We discuss three different applications of the model in Section VII. Conclusions and perspectives on our future work are presented at the end of the paper.

## II. RELATED WORK

Different works (e.g., [9], [13], [19], [22]) have analyzed the Internet traffic and have shown that it is self-similar over large time scales. Self-similarity, or equivalently long-range dependence, has been considered as a revolution against the short-range dependence supposed by classical Markovian models. The latter models under-estimate the burstiness of Internet traffic and result in a wrong evaluation of network performance.

Self-similarity has been explained by the presence of some heavy-tailed distributions in the network. [19] explains the self-similarity in LAN traffic by the multiplexing of ON/OFF sources with heavy-tailed distributions for the durations of the ON and/or OFF periods. [22] explains the self-similarity in WAN traffic by the heavy-tailed distribution of transfer durations. The M/G/ $\infty$  queue with a heavy-tailed distribution for service times is used to generate self-similar traffic. [9] explains the self-similarity in web traffic by the heavy-tailed distribution of file sizes. [13] shows that the process of arrivals of TCP connections is self-similar, and explains this phenomenon by the heavy-tailed distribution of times between the beginning of TCP connections in a WEB session.

The other body of the literature (e.g., [3], [5], [16]) studies fairness issues by modeling Internet traffic at the flow level. The main objective is to show how the capacity of the network is shared among the different flows, or equivalently, to compute the response times of flows. It is difficult to achieve such an objective with models at the packet level, as the arrivals of packets to a router interface are strongly dependent on the transport protocol involved in the transmission.

Flow-based models consider congested links that constrain the bandwidth allocated to flows. Processor sharing queues [18] are used to model such links. [3] uses a processor sharing queue to study how the bandwidth of congested links is distributed among large flows (elephants) and small flows (mice). In [3], an M/G/ $\infty$  model is also proposed for the number of active flows on a non-congested backbone link. It coincides with a very particular case of our model where all flows would have exactly the same rate. In [5], a multi-class processor sharing queue is used to compute the queue length and the packet loss probability in an Active Queue Management buffer crossed by TCP flows of different sizes. The average response time of a TCP flow is obtained. A processor sharing queue is also used in [16] to study the response time of a TCP flow on a congested link.

Date	Length	Avg. Link Utilization
Nov 8th, 2001	7h	243 Mbps
Nov 8th, 2001	10h	180 Mbps
Nov 8th, 2001	6h	262 Mbps
Nov 8th, 2001	39h 30m	26 Mbps
Sep 5th, 2001	10h	136 Mbps
Sep 5th, 2001	7h	187 Mbps
Sep 5th, 2001	16h	72 Mbps

TABLE I

SUMMARY OF OC-12 LINK TRACES

Once the utilization of the link reaches 100%, all the TCP flows are supposed by [16] to reduce their rates and to increase them linearly until the link becomes again fully utilized, and so on. Note that all the above flow-based models make the assumption that flows arrive according to a homogeneous Poisson process.

Our model is different from the above works in that (i) it is designed for non congested links and as a consequence, is specific to backbone links, (ii) it uses any flavor of flow definition to model the variation and the correlation of the traffic, and (iii) it focuses on the variation of the traffic, a performance measure of particular interest for network engineering (i.e., provisioning, SLA definition, anomaly detection, etc.).

## III. MEASUREMENT TESTBED

We consider data collected from OC-12 (622 Mbps) links on the Sprint IP backbone. The monitored links are over-provisioned so that the link utilization does not exceed 50% in the absence of link failures. In short, the infrastructure we use to collect packet traces consists of passive monitoring systems that tap optical links between access routers and backbone routers (see [12] for details on the monitoring infrastructure). Every packet on those links is timestamped and its first 44 bytes are recorded to disk.

In this paper, we present data from 7 different internal POP (Point-Of-Presence) links collected on September 5th and November 8th 2001 in three different POPs of the backbone. Table I provides a summary of the traces. The traces have different link utilizations (average rates ranging from 26 Mbps to 262 Mbps), resulting in different trace lengths.

We divide each trace into 30 minutes intervals. We tried various intervals and we found that 30 minutes is a good compromise in term of (i) keeping the arrival process stationary, and (ii) giving enough points for the analysis of our model. We discuss later in more details the consequence of this analysis interval on our observations.

We apply the model to each interval and we validate its efficiency in computing the traffic. We focus on the first two moments of the total data rate, namely the mean and the variance. Considering the variance in addition to the mean allows a better characterization of backbone traffic. As we will see, the variability of the traffic on some links of the backbone can be as high as 30% compared to the mean. The importance of the first two moments of the traffic in dimensioning backbone links will be illustrated in Section V-E.

For each interval, we measure the coefficient of variation of the total rate  $\rho_R$  (standard deviation divided by the mean), and we compare it to the value given by the model. Our model only requires information on flows, which we derive from the traces (e.g., average arrival rate of flows).

In the measurements, we use two definitions of “flow”:

- 1) Flow defined by *5-tuple*, which is a stream of packets having the same source and destination IP addresses, same source and destination port numbers, and same protocol number.
- 2) Flow defined by *prefix*, which is a stream of packets having the same /24 destination address prefix (i.e., only the 24 most significant bits of the destination IP address are taken into account).

In both cases, the size of a flow is measured in bytes, while the duration is equal to the time difference between the first and the last packet of the flow. In order to identify the end of a flow, we use a fixed timeout of 60 seconds: if the timeout expires before recording any additional packet, the flow is considered completed. A flow made of only one packet is discarded (the duration would be zero), and that packet is not counted for the purpose of the variance of the measured total rate. Flows that belong to 30 minutes intervals are split over the intervals they overlap. We found that this artificial splitting affects only a small number of flows, as shown in Figure 1. The graph on the left-hand side shows the cumulative number of flows that arrive during one 30 minutes interval. We use the second definition of flow (i.e., /24 prefix) for this graph, since the splitting of flows has more impact with this definition than with the first one (durations of flows are longer in average). The second graph is a zoom around 0 of the first one. The divided flows result in a fast increase in the number of flows at the beginning of the interval. The arrival rate remains pretty constant throughout the 30 minutes interval, except for the first 0.4 seconds, where we count only around 15,000 extra flows that are the continuation of flows started in the previous interval, out of a total of 680,000 flows. We consider therefore that the splitting of flows on these intervals has a nonzero, yet marginal effect on the arrival process, and in order to keep the model tractable, we do not correct for these effects.

As we mentioned in the Introduction, our model can operate with any definition of flow. The definitions we consider in this paper are no more than two interesting examples, corresponding to two different aggregation levels.

#### IV. THE MODEL

In this section, we describe the model (Poisson shot-noise) used for data flows arriving on a backbone link (Figure 2). It is based on the following two assumptions.

*Assumption 1:* Flow arrivals follow a homogeneous Poisson process of finite rate  $\lambda$ .

This assumption can be relaxed to more general processes such as MAPs (Markov Arrival Processes) [1], or non homogeneous Poisson processes, but we will keep working with it for simplicity of the analysis. Poisson might be the right model if we consider recent findings by [2], [6] about the process of flow arrivals and packet arrivals in the backbone of the Internet.

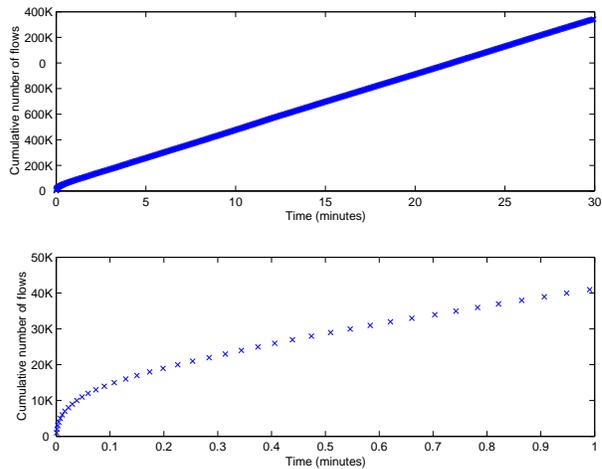


Fig. 1. Cumulative number of flows during one 30 minutes interval

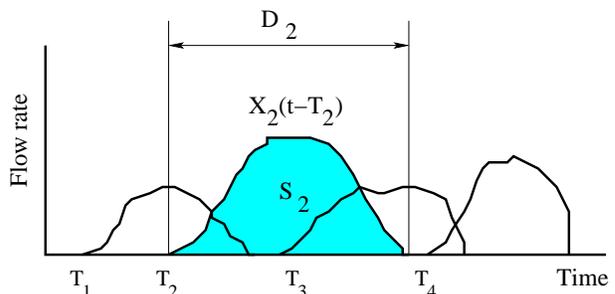


Fig. 2. Traffic modeled as a multiplexing of flows represented by “shots”

When multiplexing a large number of flows from many different sources, flows’ inter-arrival times are closer to those of a Poisson process. The latter property is known to apply to aggregates at the session level [13], [20], [22]. Note that since our model does not depend on a particular definition of flow, one can group packets into sessions that have Poisson arrivals, and apply the model at the session level.

We computed the distribution and auto-correlation of the flow inter-arrival times on the collected traces. We found that they are close to those of a homogeneous Poisson process having the same rate. We show the results for one 30 minutes interval in Figures 3 and 4. The other 30 minutes intervals provide similar results. These two figures correspond to the two definitions of flow, respectively. In each figure, the graph on the left-hand side shows the quantile-quantile plot (qq-plot) of flow inter-arrival times, and the graph on the right-hand side shows their coefficient of auto-correlation for different lags. For the qq-plot, the x-axis corresponds to the quantiles of the measured inter-arrival times, and the y-axis indicates the corresponding quantiles of the exponential distribution. We chose qq-plots rather than histograms, because a qq-plot provides a stricter test on the tail of the distributions, where a possible disagreement between the tested distribution and the exponential one is the most likely. The fit in distribution and the low-level of correlation are clear from the graphs, which confirms that the flow arrival process on our traces is close to Poisson.

Denote by  $T_n$ ,  $n \in \mathbb{Z}$ , the arrival time of the  $n$ -th flow, by  $S_n$  its size (e.g., in bits), and by  $D_n$  its duration (e.g., in seconds).

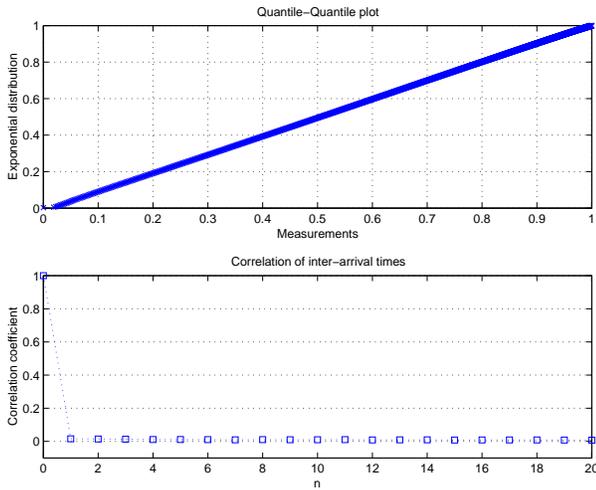


Fig. 3. Distribution and auto-correlation of inter-arrival times  $\{T_{n+1} - T_n\}$  for 5-tuple flows

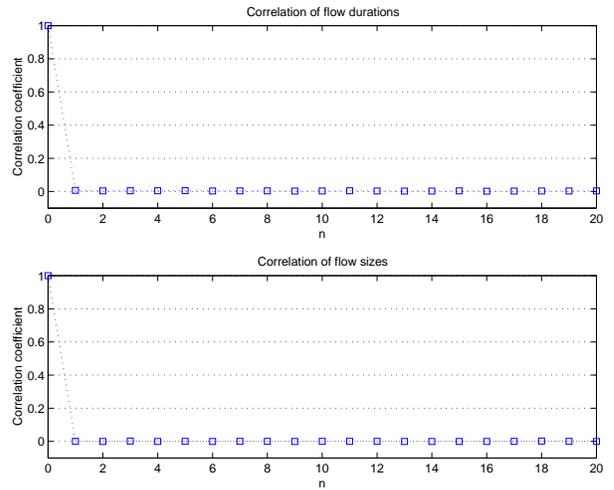


Fig. 5. Correlation of sequences  $\{S_n\}$  and  $\{D_n\}$  for 5-tuple flows

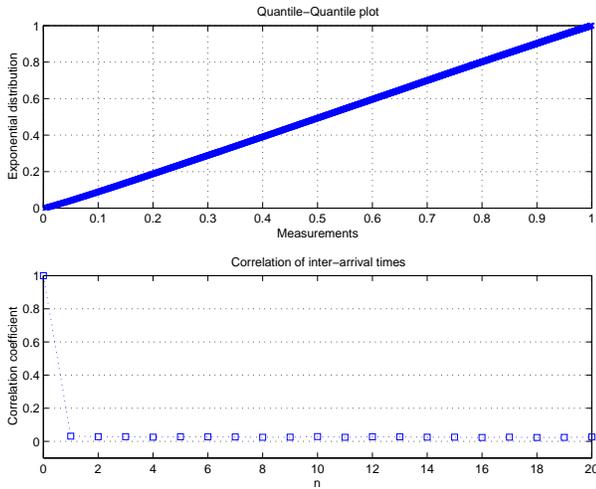


Fig. 4. Distribution and auto-correlation of inter-arrival times  $\{T_{n+1} - T_n\}$  for /24 prefix flows

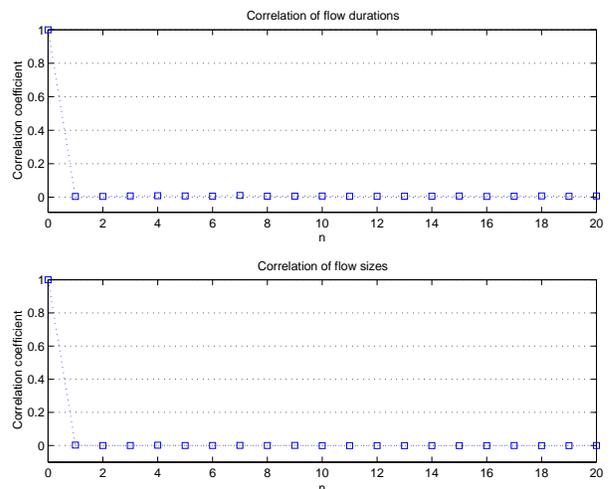


Fig. 6. Correlation of sequences  $\{S_n\}$  and  $\{D_n\}$  for /24 prefix flows

A flow is called *active* at time  $t$  when  $T_n \leq t \leq T_n + D_n$ . Define  $X_n(t - T_n)$  as the rate of the  $n$ -th flow at time  $t$  (e.g., in bits/s), with  $X_n(t - T_n)$  equal to zero for  $t < T_n$  and for  $t > (T_n + D_n)$ . In other words,  $X_n(t - T_n)$  is zero if flow  $n$  is not active at time  $t$ . We call  $X_n(\cdot)$  the *flow rate function* or *shot*.  $X_n(\cdot)$  depends on  $S_n$ ,  $D_n$  and on the dynamics governing the flow rate. For example, for TCP flows, the dynamics of the flow rate is a function of the dynamics of the window size, which in turn is a function of the round-trip time of the TCP connection, and of the features of the packet loss process [1], [7], [10], [21]. Our second assumption on  $X_n(\cdot)$  is as follows.

*Assumption 2:* Flow rate functions are independent of each other and identically distributed.

The assumption on the independence of flow rate functions is based on the following facts: (i) The link we consider is a backbone link kept under-utilized by engineering rules. It does not therefore experience congestion, and so it does not introduce dependence among the flow rate functions. (ii) The flows sharing this link have a large number of different sources and destinations, and use many different routes before being multi-

plexed on the backbone link. The assumption of identical distribution can be relaxed by introducing multiple classes (based on transport protocol, flow size, or any other metric). We keep however a single class in this paper, hence  $\{X_n(\cdot)\}$  are iid (independent and identically distributed). A direct consequence of Assumption 2 is that sequences  $\{S_n\}$  and  $\{D_n\}$  also form iid sequences, although for the same  $n$ ,  $S_n$  and  $D_n$  are obviously correlated: the larger  $S_n$ , the larger  $D_n$  (in general). Finally, we assume that  $\mathbb{E}[D_n]$  is finite.

We computed the auto-correlation of sequences  $\{S_n\}$  and  $\{D_n\}$  on our traces. We found indeed that these sequences exhibit little correlation. The result is illustrated in Figures 5 and 6, where we show the auto-correlation coefficients of the two sequences for one 30 minutes interval, using our two definitions of flow. In each figure, the left-hand side corresponds to the sequence  $\{S_n\}$ , and the right-hand side to the sequence  $\{D_n\}$ . The auto-correlation drops quickly to zero after lag-0.

Define  $R(t)$  as the total rate of data (e.g., in bits/s) on the modeled link at time  $t$ . It is the result of the addition of the

rates of the different flows. We can then write

$$R(t) = \sum_{n \in \mathbb{Z}} X_n(t - T_n). \quad (1)$$

This model is a *Poisson shot-noise process* [4], [11], where the term “shot” is synonymous here of “flow rate function”. In the particular case where  $X_n(t - T_n) = 1_{\{t \in [T_n, T_n + D_n]\}}$ , that is, where shots are rectangles of height 1 and length  $D_n$ , the process (1) is the number of clients found at time  $t$  in an M/G/∞ queue [17], if clients are identified with flows. We will allow however for “shots” with a more general shape than a rectangle of height 1, and we will see in this paper that this is indeed essential to characterize the total data rate on backbone links.

We will use two alternative approaches to compute the moments of  $R(t)$ . The first one uses queuing theory to compute the Laplace Stieltjes Transform (LST) of  $R(t)$  in stationary regime (we always assume that we have reached the stationary regime, which exists for finite  $\lambda$  and  $\mathbb{E}[D_n]$ ). The expression of the LST of  $R(t)$  allows us to compute all moments of  $R(t)$ , as well as its first order distribution. We focus in this paper on the first two moments (average and variance).

In the second approach, we use elements from the theory of Poisson shot-noise (Campbell’s theorem [4], [11]) to compute, in addition to the first two moments of  $R(t)$ , its spectral density and its auto-covariance function.

In summary, our model allows us to completely characterize the data rate on a backbone link based on the following inputs:

- 1) The arrival rate of flows  $\lambda$ .
- 2) The distributions of flow sizes  $\{S_n\}$  and flow durations  $\{D_n\}$ .
- 3) The shot function  $X_n(\cdot)$ .

For the particular shapes of the shot presented in Figure 7, we will see that with only three parameters, our model is able to compute the average and the variation of the backbone traffic. In summary:

- The average total rate of the traffic is given by the two parameters  $\lambda$  and  $\mathbb{E}[S_n]$  (Corollary 1).
- The variance of the total rate (i.e., burstiness of the traffic) is given by the two parameters  $\lambda$  and  $\mathbb{E}[S_n^2/D_n]$  (Section V-C and V-D), with a multiplicative factor function of the shot shape.

## V. PERFORMANCE ANALYSIS

### A. LST and moments of the total rate

Let  $N(t)$  denote the number of active flows at time  $t$ . Assumptions 1 and 2 imply that the total rate  $R(t)$  at time  $t$  is the sum of a random number  $N(t)$  of iid random variables which are the rates of active flows. Assumption 1 ensures that there is no correlation among the instants of arrivals of active flows. In this section, we compute the LST of  $R(t)$  which we denote as  $\tilde{R}(w) = \mathbb{E}[e^{-wR(t)}]$ ,  $\text{Re}(w) \geq 0$ . Our main result is stated in the following theorem.

*Theorem 1:* For  $w \in \mathbb{C}$  and  $\text{Re}(w) \geq 0$ , the LST of the total rate is

$$\tilde{R}(w) = \exp \left( \lambda \mathbb{E} \left[ \int_0^{D_n} e^{-wX_n(u)} du \right] - \lambda \mathbb{E}[D_n] \right).$$

**Proof:** The rate of a flow found active at time  $t$  is equal to  $X_n(t - T_n)$ . Let  $\tilde{X}^0(w) = \mathbb{E}^0 [e^{-wX_n(t - T_n)}]$ ,  $\text{Re}(w) \geq 0$ , denote its LST. We use superscript 0 to indicate that the expectation is computed under the condition that flow  $n$  is active at time  $t$ :  $\mathbb{E}^0 [e^{-wX_n(t - T_n)}] = \mathbb{E} [e^{-wX_n(t - T_n)} | t \in [T_n, T_n + D_n]]$ . Define  $\tilde{N}(z) = \mathbb{E} [z^{N(t)}]$ , for  $z \in \mathbb{C}$  and  $\|z\| \leq 1$ , as the Probability Generating Function (PGF) of the random variable  $N(t)$ . Using [17, Equation II.34], the LST of  $R(t)$  at a certain point  $w \in \mathbb{C}$ , with  $\text{Re}(w) \geq 0$ , can be obtained by evaluating the function  $\tilde{N}(z)$  at  $z = \tilde{X}^0(w)$ :

$$\tilde{R}(w) = \tilde{N}(\tilde{X}^0(w)). \quad (2)$$

In the following, we simplify this expression of the LST of  $R(t)$  and write it as a function of the random variables  $S_n$ ,  $D_n$ , and the rate function  $X_n(\cdot)$ .

We start by computing the PGF of  $N(t)$ , which is the number of clients found at time  $t$  in an M/G/∞ queuing model (Section IV). The distribution of  $N(t)$  in the stationary regime is given by (see e.g., [15], [17])

$$\mathbb{P}\{N(t) = k\} = \frac{(\lambda \mathbb{E}[D_n])^k}{k!} e^{-\lambda \mathbb{E}[D_n]}, \quad k \in \mathbb{N},$$

where  $\lambda \mathbb{E}[D_n]$  is the load of the queue. As  $\lambda \mathbb{E}[D_n]$  is finite, the system is stable and the number of active flows does not grow to infinity with probability 1. It follows that

$$\begin{aligned} \tilde{N}(z) &= \sum_{k=0}^{\infty} \mathbb{P}\{N(t) = k\} z^k \\ &= \exp(\lambda \mathbb{E}[D_n](z - 1)). \end{aligned} \quad (3)$$

Next, we simplify  $\mathbb{E}^0 [e^{-wX_n(t - T_n)}]$  by getting rid of the instant of arrival  $T_n$  and of the superscript 0. We start by the instant of arrival. The arrival process being Poisson, the starting time of a flow known to be active at time  $t$  is uniformly distributed between  $t - D_n$  and  $t$ , and is independent of the starting times of the other active flows [17, Section 2.5]. We can thus write that

$$\begin{aligned} \tilde{X}^0(w) &= \mathbb{E}^0 [e^{-wX_n(t - T_n)}] \\ &= \mathbb{E}^0 \left[ \frac{1}{D_n} \int_0^{D_n} e^{-wX_n(u)} du \right]. \end{aligned} \quad (4)$$

To get rid of the superscript 0, we use a technique similar to that used in queuing theory to compute the distribution of the residual service time of the client in the server when a new client arrives to the queue (e.g., [17, Section 5.2]). We recall that the superscript 0 represents the fact that the flow is active at time  $t$ . Given that the time  $t$  is arbitrary, a flow active at time  $t$  has in average a longer duration than an arbitrary flow. Let  $f^0(d)$  and  $f(d)$  denote the probability density functions of the duration of a flow active at time  $t$  and of the duration of an arbitrary flow, respectively.  $f^0(d)$  is obtained by scaling  $f(d)$  with the duration of the flow in a way to give more weight to large values of  $D_n$ . We have the following relation between the two densities [17, Section 5.2]:  $f^0(d) = f(d) \frac{d}{\mathbb{E}[D_n]}$ . Using this equation in (4), and substituting the result in (2) then in (3), we get the

expression of the LST of  $R(t)$  stated in the theorem. The expectations in the theorem are computed for an arbitrary flow, and not for a flow we found active at time  $t$ .  $\square$

By differentiating with respect to  $w$  and then setting  $w$  to 0, the LST in Theorem 1 can give us all the moments of the total rate in the stationary regime. We have the following main results.

*Corollary 1:* The average of the total rate is  $\mathbb{E}[R(t)] = \lambda \mathbb{E}[S_n]$ .

This expected expression for  $\mathbb{E}[R(t)]$  is very simple, and it does not need any knowledge of the distribution of flow durations. It states that the average total rate is equal to the average arrival rate of flows times the average amount of data brought by each flow.

*Corollary 2:* The variance of the total rate is  $V_R = \lambda \mathbb{E} \left[ \int_0^{D_n} X_n^2(u) du \right]$ .

The variance of the total rate is the second important performance measure an ISP needs to know in order to properly dimension the links of its network. A backbone link has to be provisioned so as to absorb the average of the total rate as well as its variations. In contrast to the average, our model tells us that the variance of the total rate is a function of the durations of flows and their rate functions. This requires some assumptions (or more information) on the dynamics of flow rate. Later, we will provide approximations of the variance of  $R(t)$  for some particular rate functions.

*Corollary 3:* Let  $Y_i = \int_0^{D_n} X_n^i(u) du$ . The computation of the  $k$ -th moment of the total rate requires the expectations  $\mathbb{E}[Y_i]$ ,  $i = 1, 2, \dots, k$ .

## B. Spectral density and auto-correlation of the total rate

The second approach for computing the moments of  $R(t)$  is to write (1) as the convolution of a train of Dirac pulses located at times  $T_n$ ,  $U(t) = \sum_{n \in \mathbb{Z}} \delta(t - T_n)$ , with function  $X_n(\cdot)$ :  $R(t) = (X_n \star U)(t) = \int_{-\infty}^{+\infty} X_n(t - s)U(s)ds$ , where  $\star$  denotes the convolution operator. As  $X_n(\cdot)$  and  $U(\cdot)$  are independent because of our assumptions in Section IV, we can compute the first and second order moments of  $R(t)$  using Campbell's theorem [4], [11], which is an alternative approach to obtain Corollary 1 and 2. It also provides the spectral density  $S_R(f)$  of the centered process  $R(t) - \mathbb{E}[R(t)]$ , which is  $S_R(f) = \lambda \mathbb{E} \left[ \left| \tilde{X}(f) \right|^2 \right]$ ,  $\tilde{X}(f)$  being the Fourier transform of  $X_n(\cdot)$ :  $\tilde{X}(f) = \int_{-\infty}^{+\infty} X_n(u) e^{-2\pi j f u} du$ . Inverting this expression for the spectral density into the time domain, we obtain the following result for the auto-covariance function of the total rate.

*Theorem 2:* The auto-covariance function of the total rate is given by

$$\begin{aligned} C_R(\tau) &= \mathbb{E}[R(t - \tau)R(t)] - \mathbb{E}[R^2(t)] \\ &= \lambda \mathbb{E} \left[ \mathbf{1}_{\{D_n > |\tau|\}} \int_0^{D_n - |\tau|} X_n(u) X_n(u + |\tau|) du \right]. \end{aligned}$$

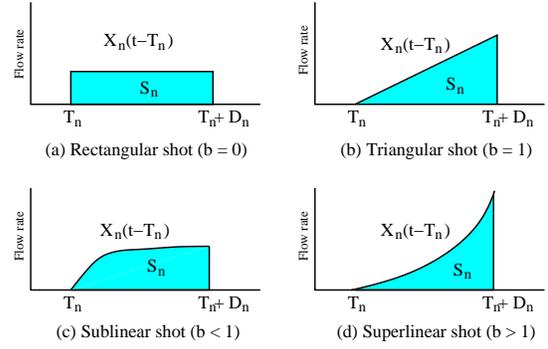


Fig. 7. Simple models for shots

**Proof:** The spectral density  $S_R(f) = \lambda \mathbb{E} \left[ \left| \tilde{X}(f) \right|^2 \right]$ , is by definition the Fourier Transform of the auto-covariance function  $C_R(\tau)$ . Using Inverse Fourier Transform, we write

$$\begin{aligned} C_R(\tau) &= \int_{-\infty}^{+\infty} S_R(f) e^{2\pi j f \tau} df \\ &= \lambda \mathbb{E} \left[ \int_{-\infty}^{+\infty} |\tilde{X}(f)|^2 e^{2\pi j f \tau} df \right] \\ &= \lambda \mathbb{E} \left[ \int_0^{D_n} \int_0^{D_n} X_n(u) X_n(v) \delta(\tau + u - v) du dv \right]. \end{aligned}$$

For  $\tau > 0$ , the double integral in the last expectation is equal to 0 when  $D_n \leq 0$  and to  $\int_0^{D_n - \tau} X_n(u) X_n(u + \tau) du$  otherwise. Combining this result with the fact that  $C_R(\tau)$  is an even function, leads to the expression stated in the theorem.  $\square$

For the particular case  $\tau = 0$ , we clearly retrieve, as expected, the expression of the variance of  $R(t)$  given by Corollary 2. Note that this approach can be extended to non Poisson arrivals using [4].

## C. Two particular shot shapes

To compute the moments of the total rate of order higher than 2, we need to model the shot (Corollary 3). Before moving to more general models, let us examine the two particular cases shown in Figure 7a and 7b.

*1) Rectangular shots:* First, we consider the case where the rate of a flow is constant and equal to  $S_n/D_n$  (which gives the rectangular shot of length  $D_n$  and height  $S_n/D_n$  of Figure 7a). Here, the random variable  $Y_i$  (Corollary 3) is equal to  $S_n^i/D_n^{i-1}$ , and we need therefore to compute all expectations  $\mathbb{E}[S_n^i/D_n^{i-1}]$  in order to fully characterize the total rate. The average of the total rate requires the expression of  $\mathbb{E}[S_n]$ . The variance of the total rate requires the expression of  $\mathbb{E}[S_n^2/D_n]$ , and so on. In particular, Corollary 2 yields that the variance of  $R(t)$  is equal to  $V_R = \lambda \mathbb{E} \left[ \frac{S_n^2}{D_n} \right]$ .

The rectangular assumption is the simplest one; the only generalization from an M/G/ $\infty$  model is the height of the ‘‘shot’’ which is now variable. With this assumption, we only capture the variation of the total rate caused by the variation of

$N(t)$  and by the variation of the ratio  $S_n/D_n$ . In fact, a Poisson shot-noise with rectangular shots always under-estimates the variance of the actual process.

*Theorem 3:* Among all possible flow rate functions, the rectangular flow rate function achieves the lowest variance  $V_R$  of the total rate.

**Proof:** Consider the following quantity which is clearly positive over all the set of flow rate functions:

$$\lambda \mathbb{E} \left[ \int_0^{D_n} \left( X_n(u) - \frac{S_n}{D_n} \right)^2 du \right] \geq 0.$$

It can be expanded as

$$\begin{aligned} & \lambda \mathbb{E} \left[ \int_0^{D_n} \left( X_n(u) - \frac{S_n}{D_n} \right)^2 du \right] \\ &= \lambda \mathbb{E} \left[ \int_0^{D_n} \left( X_n^2(u) - 2 \frac{S_n}{D_n} X_n(u) + \left( \frac{S_n}{D_n} \right)^2 \right) du \right] \\ &= \lambda \mathbb{E} \left[ \int_0^{D_n} X_n^2(u) du \right] - \lambda \mathbb{E} \left[ \frac{S_n^2}{D_n} \right], \end{aligned}$$

so that  $\lambda \mathbb{E} \left[ \int_0^{D_n} X_n^2(u) du \right] \geq \lambda \mathbb{E} \left[ \frac{S_n^2}{D_n} \right]$ .

This result holds for any flow rate function  $X_n(\cdot)$ . The first term is the variance  $V_R$  of the total rate for a general flow rate function  $X_n(\cdot)$  (Corollary 2). The second term is the variance of the total rate when flow rates are modeled by rectangles (Section V-C.1). This concludes the proof.  $\square$

We know that flow rates are dynamic, especially with a protocol such as TCP which adapts the rate of a flow to network conditions. Better results could be obtained with other flow rate functions that capture a part of, if not all, this dynamics. We will give an example in the next sections. One must not interpret these particular shots as representative of a particular protocol. They can be seen as attempts to capture some dynamics of flow rates, while keeping our model simple and general.

2) *Triangular shots:* Another assumption is to consider that the rate of a flow linearly increases with time (Figure 7b). This assumption is inspired from the dynamics of TCP transfers that form a large majority of the flows in IP backbones [12]. For a flow of size  $S_n$  and of duration  $D_n$ , the rate is assumed to increase linearly from zero to  $2S_n/D_n$ , with a mean equal to  $S_n/D_n$ . At a time  $t$  between  $T_n$  and  $T_n + D_n$ , we can write  $X_n(t - T_n) = (2S_n/D_n^2)(t - T_n)$ . Hence, for  $i \geq 1$ ,  $Y_i = \frac{2^i}{i+1} S_n^i / D_n^{i-1}$ . Again, we need to compute all the expectations  $\mathbb{E} \left[ S_n^i / D_n^{i-1} \right]$  in order to fully characterize the total rate. For the variance of  $R(t)$  we have  $V_R = \frac{4\lambda}{3} \mathbb{E} \left[ \frac{S_n^2}{D_n} \right]$ . As expected from Theorem 3, the variance is larger than in the rectangular case (by a multiplicative factor 4/3).

#### D. Measurement-based derivation of flow rate functions

The rectangular and the triangular assumptions are just two examples among many. One can always consider other approximations of flow rate functions using  $S_n$  and  $D_n$  (log, square root, exponential, etc.).

There are two methods to determine the function  $X_n(\cdot)$ . The first method is based on *modeling* the individual rate function  $X_n(\cdot)$ . For some transport protocols like TCP, it is indeed possible to capture the dynamics of the protocol and to derive an optimal ‘‘shot’’  $X_n(\cdot)$  from it. We are following this approach for TCP flows in a parallel work. The advantage of our flow-based model is to reduce the modeling effort of an aggregate rate of heterogeneous flows to that of a single flow rate, provided this modeling can be achieved with sufficient accuracy.

In this paper however, we want to remain as general as possible, and hence we must resort to the second method, based on *measurements*. It consists in fitting a parametric model of the shot  $X_n(\cdot) = x_\theta(\cdot)$ , where  $x_\theta(\cdot)$  is an *a priori* chosen function parameterized by a parameter vector  $\theta$ , which must satisfy the constraint

$$\int_0^{D_n} x_\theta(u) du = S_n. \quad (5)$$

Vector  $\theta$  is then computed to minimize some error functional between the experimental value of the distribution (or some moments of  $R(t)$ ), and the value computed by Theorem 1 (or Corollary 2 or 3). Since in this paper we focus on the variance, we compute  $x_\theta(\cdot)$  so that

$$\hat{V}_R = \lambda \mathbb{E} \left[ \int_0^{D_n} x_\theta^2(u) du \right], \quad (6)$$

where  $\hat{V}_R$  is the actual empirical variance of the measured aggregate rate.

Since we have two equations (5) and (6), we need therefore two parameters:  $\theta = (a, b)$ . A simple function is a power function  $x_\theta(u) = au^b$ , illustrated in Figure 7c and 7d. It includes, as particular cases, the rectangular ( $b = 0$ ) and the triangular ( $b = 1$ ) shots. Solving (5) yields that  $a = (b+1)S_n/D_n^{b+1}$ , and plugging this value in (6) we get  $\hat{V}_R = \lambda \frac{(b+1)^2}{2b+1} \mathbb{E} \left[ \frac{S_n^2}{D_n} \right]$ . We deduce an estimate of  $b$ , based on the measurement of  $\hat{V}_R$  (and clearly, as before, of  $\lambda$  and  $\mathbb{E} \left[ \frac{S_n^2}{D_n} \right]$ ), which is  $b = \alpha - 1 + \sqrt{\alpha^2 - \alpha}$ , with  $\alpha = \hat{V}_R / (\lambda \mathbb{E} \left[ \frac{S_n^2}{D_n} \right])$  (note that because of Theorem 3,  $\alpha \geq 1$ ). Of course, the introduction of a larger number of parameters allows to fit  $x_\theta(\cdot)$  to more moments than simply  $\hat{V}_R$ . We will use this expression of  $b$  in Section VI.

#### E. Distribution of the total rate

By inverting the LST of  $R(t)$  given in Theorem 1, we can obtain the probability density function (pdf) of the total rate. The difficulty with such a method is that it requires the LST of  $R(t)$ , which in turn requires the knowledge of the distribution of the sizes of flows and of their durations.

If we only dispose of the first and second moments of  $R(t)$ , we can resort to the Central Limit Theorem to approximate the pdf of the total rate by a Gaussian pdf:  $f_R(r) = \frac{1}{\sqrt{2\pi V_R}} e^{-\frac{(r - E[R])^2}{2V_R}}$ . This approximation is justified by the large number of active flows at a given time on a backbone link [12].

The Gaussian approximation is simple, and at the same time very useful to dimension the links of a backbone network. For

example, it tells us that during 70% of time, the total rate is between  $\mathbb{E}[R] - \sqrt{V_R}$  and  $\mathbb{E}[R] + \sqrt{V_R}$  [15]. It also gives us an approximation of the tail of the total rate (i.e., the probability that  $R(t)$  exceeds a certain level). Using this approximation, the bandwidth of a link of the backbone can be chosen so that congestion is avoided. Let  $A(\epsilon)$  be the normal quantile function defined by  $\mathbb{P}\{V > A(\epsilon)\} = \epsilon$ , with  $0 < \epsilon < 1$  and  $V$  a normal random variable of mean 0 and variance 1. Suppose that we want to choose the bandwidth of the link ( $L$ ) in such a way that congestion occurs in less than  $100 * \epsilon\%$  of time, i.e.,  $\mathbb{P}\{R(t) > L\} < \epsilon$ . The Gaussian approximation tells that  $L$  has to be set larger than  $\mathbb{E}[R(t)] + A(\epsilon) * \sqrt{V_R}$ . One can use large deviations techniques [23] to find a better approximation of the tail of the total rate.

### F. Moments of $R(t)$ and averaging interval

In reality, the total measured rate  $R_m(t)$  at a certain time  $t$  is computed by averaging and sampling the volume of data (e.g., number of bytes) that cross the backbone link during a short time interval  $\delta$  around  $t$ :

$$R_m(t) = \frac{1}{\delta} \int_{k\delta}^{(k+1)\delta} R(s) ds,$$

with  $t \in [k\delta, (k+1)\delta)$ ,  $k \in \mathbb{Z}$ .  $\delta$  denotes the length of the averaging and sampling period. The measured rate appears thus as a piecewise constant function, with segments of length  $\delta$ . It amounts to convolve the instantaneous rate  $R(t)$  by a linear filter of impulse response  $1_{\{0 \leq t < \delta\}}$  before taking the sample. Except for the first one, the moments of  $R_m(t)$  depend on  $\delta$ : the longer the averaging interval, the smoother the total rate<sup>2</sup>. The Wiener-Khinchine relations [14] imply indeed that the spectral densities of  $R_m(t)$  and  $R(t)$  are linked by  $S_{R_m}(f) = S_R(f) (\sin(\pi f \delta) / \pi f \delta)^2$ , from which we deduce that the variance of  $R_m(t)$ , i.e., the measured variance, is

$$\hat{V}_R = \frac{2}{\delta} \int_0^\delta (1 - \tau/\delta) C_R(\tau) d\tau. \quad (7)$$

Since  $C_R(\tau) \leq V_R$ , this expression is always smaller than  $V_R$ . The scaling factor between  $V_R$  and  $\hat{V}_R$  requires the knowledge of  $C_R(\tau)$ , which is given in Theorem 2. Clearly, if  $C_R(\tau)$  does not decrease too rapidly in  $[0, \delta]$ , both variances will remain close to each other. Consequently, we do not take into account the averaging of the data rate in the model, but we rather keep  $\delta$  small so that  $C_R(\tau)$  remains close to  $C_R(0) = V_R$  in  $[0, \delta]$ . By taking  $\delta$  small, the instant at which a flow arrives or leaves the backbone within a single averaging interval does not have a big impact on the total rate. This is because the volume of data transmitted by a flow during an averaging interval is small compared to the size of the flow, or equivalently, the averaging interval is small compared to the duration of the flow.  $V_R$  can then be safely used as an approximation of  $\hat{V}_R$ , which models the variance of the measured samples of the total rate. Note that one can always compute the exact value of  $\hat{V}_R$  by plugging the expression of  $C_R(\tau)$  (Theorem 2) in (7).

<sup>2</sup>The only exception is when the traffic is strongly self-similar (Hurst parameter equal to 1) [19]: increasing the averaging interval will not reduce the burstiness of the total rate.

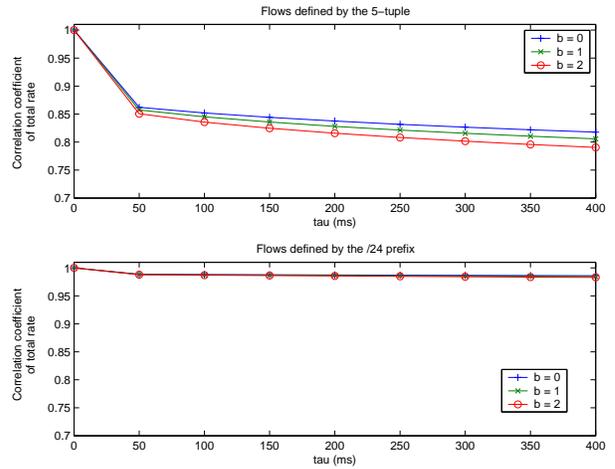


Fig. 8. Auto-correlation coefficient of the total rate computed by Theorem 2

Before using our model, an ISP has to choose a value  $\delta$  of the averaging interval. It can be the longest busy period (i.e., period where the utilization of the link is 100%) allowed by the ISP. It is also the interval below which the ISP does not care about the congestion of the network, possibly because this short-term congestion is absorbed by the buffers at the inputs of links. If the chosen value  $\delta$  is small enough so that the auto-covariance function  $C_R(\tau)$  slowly decreases in  $[0, \delta]$ ,  $V_R$  can be used as an approximation of the measured variance, otherwise  $\hat{V}_R$  has to be computed and used. In what follows, we will choose as averaging interval the (average) round-trip time of flows (200 ms), since we know that most of the flows take more than one round-trip time to end. Our choice is also motivated by the fact that TCP flows update their transmission rates approximately once per round-trip time. Recall that the averaging interval is a parameter that can be set by the ISP to any other value than the round-trip time, depending on the maximum burstiness it tolerates at the inputs of the links of its backbone.

We checked the decrease of the autocorrelation coefficient  $C_R(\tau)/V_R$  in  $[0, \delta]$  by applying Theorem 2 to the traces. We consider the three shots that correspond to  $b = 0$ ,  $b = 1$ , and  $b = 2$ . Figure 8 shows the results for one 30 minutes trace. The graph on the left-hand side corresponds to 5-tuple flows, and the one on the right-hand side corresponds to /24 prefix flows. We plot the results for  $\tau \in [0, 400]$ ms. Indeed, the auto-correlation coefficient slowly decreases, especially for /24 prefix flows due to their longer durations.

### G. Complexity of the model

Our model requires few parameters to characterize the backbone traffic. To compute the average traffic, we need the arrival rate of flows  $\lambda$  and their average size  $\mathbb{E}[S_n]$  (Corollary 1). The variance of the total rate requires one more parameter: the expectation of the ratio between the square of the size of a flow and its duration  $\mathbb{E}[S_n^2/D_n]$  (for the shots of Figure 7). The average and the variation of the traffic can therefore be computed with only three parameters:  $\lambda$ ,  $\mathbb{E}[S_n]$ , and  $\mathbb{E}[S_n^2/D_n]$ .

In this paper, we compute the parameters of the model offline. We infer their values from statistics on the processes  $\{S_n\}$  and  $\{D_n\}$ . The computation is simple and it only requires an

averaging over the different samples of the processes. An implementation of the model would require an online computation of these parameters with, for example, an Exponentially Weighted Moving Algorithm, such as the one used by TCP to estimate the average round-trip time.

Suppose that we have a tool that gives us the size and duration of a flow as soon as it ends. Let  $\hat{S}$  denote the estimate of  $\mathbb{E}[S_n]$ . When the tool indicates the departure of a flow of size  $s$ , the estimate  $\hat{S}$  can be updated in the following way  $\hat{S} \leftarrow (1 - \alpha)\hat{S} + \alpha s$ , with  $0 < \alpha < 1$ . The smaller  $\alpha$ , the slower the reaction of the estimator to a change in the average flow size. A similar algorithm can be applied to estimators of  $\lambda$  and  $\mathbb{E}[S_n^2/D_n]$ .

We keep the problem of the online estimation of the parameters of our model for future research. Our main objective in this paper is to validate the model and to show its usefulness for provisioning and managing IP networks. Given that our model requires few parameters, we believe that it simpler (in term of computation cost and implementability in an operational environment) than a packet level model that would provide the same information about the traffic.

## VI. EXPERIMENTAL VALIDATION

In this section we validate our model using the traces collected on the Sprint IP backbone, and presented in Section III.

We compare the real coefficient of variation of the total rate  $\rho_R = \frac{\sqrt{\hat{V}_R}}{\mathbb{E}[R]}$ , with the results obtained from our model  $\rho_R = \frac{\sqrt{\lambda \mathbb{E}[\int_0^{D_n} X_n^2(u) du]}}{\lambda \mathbb{E}[S_n]}$ , when the input data (i.e., flow arrival rate  $\lambda$  and the expectation of  $S_n^2/D_n$ ) are directly derived from the traces. Samples of the total rate are computed using averaging intervals of 200 ms. This is comparable with the average round-trip time we measure on these links (Section V-F).

Even if Assumptions 1 and 2 hold, the measurement process introduces two differences with the model of Section V. We already addressed these two differences.

- The first difference is the averaging and sampling of the measured rate at a periodicity of 200 ms, which will lead to an experimental value of variance  $\hat{V}_R$  smaller than the variance of the instantaneous rate  $V_R$ , as explained in Section V-F. We have indeed observed on experimental data that the longer the averaging interval, the smaller  $\hat{V}_R$ . Therefore, we expect to find a few occurrences of an empirical value  $\hat{V}_R$  smaller than the lower bound obtained with a rectangular shot (Theorem 3). A better matching would be obtained by comparing the experimental variance to the value  $\hat{V}_R$  given by (7), instead of comparing it to the variance predicted by Corollary 2. This will complicate a little the analysis since it will require the computation of the integral in (7), using the value of  $C_R(\tau)$  given in Theorem 2. A better matching would also be obtained when the averaging interval tends towards 0.
- The second one is the splitting of flows located on the boundaries of the 30 minutes intervals. As we explained in Section III, the number of these flows is very small compared to the total number of flows that arrive in the intervals, and the splitting has therefore a negligible impact.

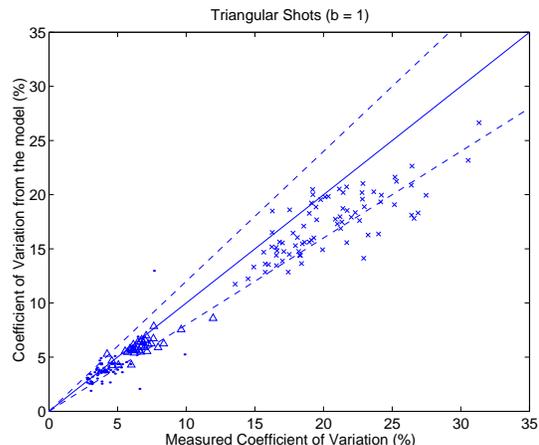


Fig. 9. Coefficient of variation of the total rate with triangular shots and flows defined by the 5-tuple

These two sources of errors are unavoidable: the first one because traffic is packet-based and not fluid, so that the measurements must be averaged over intervals of some minimal length, and the second one because we need to divide the trace in intervals short enough to keep the arrival process stationary and to reduce the volume of data to manipulate.

### A. Results

In this section we do not present results on the first moment of the total rate, since they are computed by our model and by measurements in exactly the same way.

In Figures 9 and 10 we compare the coefficient of variation computed via measurements with that given by our model with triangular ( $b = 1$ ) and parabolic ( $b = 2$ ) shots. These results refer to the first definition of flow using the 5-tuple. Each point in the figures correspond to a 30 minutes interval. A cross indicates that the average rate during that interval is below 50 Mbps; a triangle is used for those intervals with an average rate between 50 and 125 Mbps; the dots are used for rates above 125 Mbps. The x-axis shows the measured coefficient of variation of the total rate, while the y-axis shows the coefficient of variation given by the model. A point on the diagonal crossing the figures represents a perfect match between the model and the measurements. The two dashed lines identify the bounds for an error in the estimate of 20%.

There is a good match between the model and the measurements with better results for the parabolic case; the triangular case often under-estimates the real coefficient of variation since it does not capture all the dynamics of flow rate.

The above two figures show three clusters of points, that can be easily distinguished. The interpretation is simple and is related to the fact that we are collecting traces on many diverse links, with three main different utilization levels (Section III). As we will explain in Section VII-A, backbone traffic becomes smoother when the arrival rate of flows  $\lambda$  increases. An increase in the arrival rate of flows is the main responsible for the increase in the utilization among the links, since it is safe to assume that the average file size is the same on all links of the backbone (Corollary 1). Links with higher utilization (above 125 Mbps) exhibit very low variation, and, thus contribute to

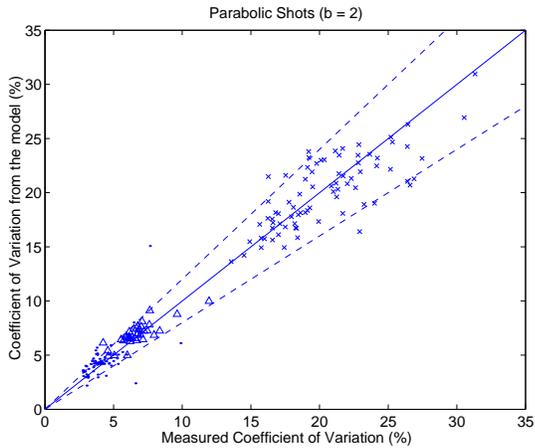


Fig. 10. Coefficient of variation of the total rate with parabolic shots and flows defined by the 5-tuple

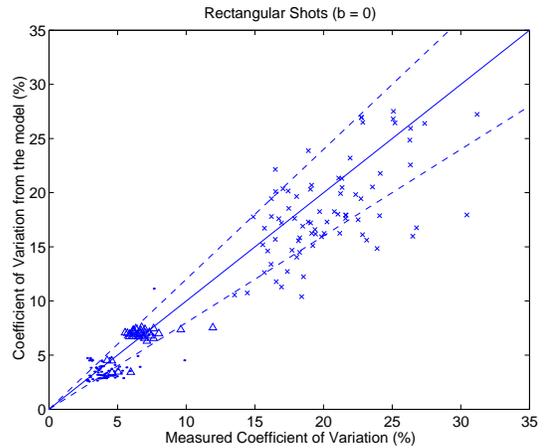


Fig. 12. Coefficient of variation of the total rate with rectangular shots and flows defined by destination address prefix

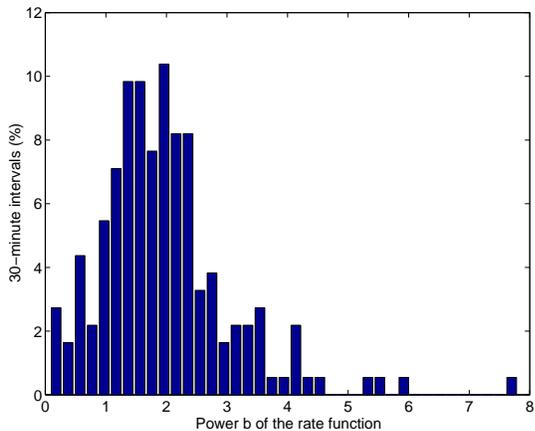


Fig. 11. Power  $b$  of flow rate functions with flows defined by the 5-tuple

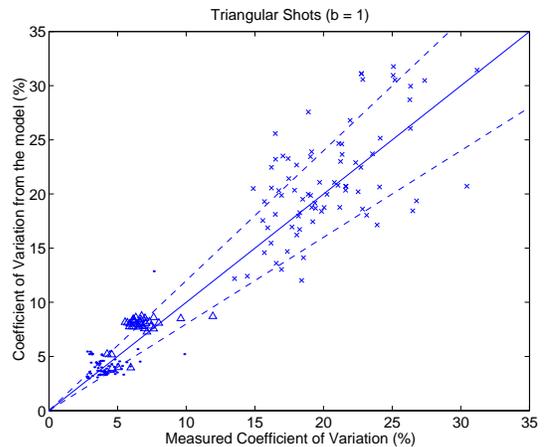


Fig. 13. Coefficient of variation of the total rate with triangular shots and flows defined by destination address prefix

the first cluster of points at the bottom-left corner of the figures. Those links with a medium utilization (between 50 Mbps and 125 Mbps) are represented in the cluster in the middle. Finally, the links with the lowest utilization (below 50 Mbps) exhibit the highest traffic variability (around 30%), and yield the cluster of points on the right-hand side of the figures.

In Section V-D, we explained how the optimal power  $b$  can be computed from a trace so that the variance of the total rate given by our model matches that given by measurements. For the different 30 minutes traces, we compute this optimal power and we plot its histogram in Figure 11. The average value of  $b$  over all the traces is equal to 1.98, which means that parabolic shots are in average the most suited to model traffic when flows are defined by the 5-tuple (from variation point of view). We are currently working on the interpretation of the difference in the value of  $b$  among the traces. A possible reason could be the difference in file sizes: small files require a large value of  $b$  due to the slow start phase of TCP, and large files require a small value of  $b$  due to the slow window increase in TCP congestion avoidance mode.

Figures 12 and 13 provide the coefficient of variation for the second definition of flow based on destination address prefixes. We plot the cases with rectangular shots ( $b = 0$ ) and triangular shots ( $b = 1$ ). The use of rectangular shots seems to be able

to capture the variability of the traffic aggregate at the level of destination address prefixes. This is probably due to the fact that such a level of aggregation “dilutes” the impact of specific transport protocol mechanisms on the total rate.

This result shows that our model can estimate the total rate and its variance independently of the protocol or application characteristics. The ability of defining a flow through the destination prefix greatly reduces the complexity of a possible implementation. Indeed, on our traces, the number of flows of which a router would need to keep track is reduced on average by one order of magnitude when using a  $/24$  destination prefix. A straightforward extension to this flow definition would be the use of “routable” prefixes (i.e., prefixes present in the forwarding table of the router) to define flows. Such an extension would result in an additional decrease of the burden for the router given the level of flow aggregation (with  $/8$  and  $/16$  prefixes, for example) that could be achieved. Moreover, it would give the possibility to use flow information in conjunction with routing information in order to infer the variation of the rate over all the links of the same network from the measurement of only a subset of the links. We discuss later the issue of combining our model with routing information.

## VII. APPLICATIONS OF THE MODEL

We present three possible applications of our model to highlight the role that such a model may have in the engineering and management of an IP backbone network.

### A. Network dimensioning and provisioning

This is the first and the most intuitive application of our flow-based model. Suppose that an ISP collects statistics on flow sizes, flow durations, and flow arrivals (for example with tools such as Cisco NetFlow). With this sole information, the ISP is able to compute the moments of the total rate, in particular its average and its variance. This way, the ISP would have more detailed information than that provided by SNMP (one of the problems of SNMP is that it does not capture traffic variation at short time scales).

The information on flows can be collected on the link we want to monitor. It can also be collected at the edges of the backbone. Combined with the routing information in the edge routers, this will give us information on flows on each link of the backbone. Our model can then be used to compute the traffic on the links of the backbone, by only monitoring the edges.

The detailed information provided by our model on the traffic helps to dimension backbone links, as explained in Section V-E. Rate variation at short time scales are very useful in the definition of the buffer size and in the evaluation of the maximum queuing delay. In the case we collect information on flows at the edges, our model can help in routing flows in the backbone, with the objective to optimize the utilization of the available resources.

Computing the traffic in the backbone using information on flows is not the only application of our model to network dimensioning and management. The model enables to study the impact of a change on backbone link utilization by other variables, such as the distribution of flow sizes, the distribution of flow durations, the flow arrival rate, etc. This study is very useful for upgrading backbone links. For example, we can evaluate the effect caused by a new application resulting in large transfer sizes, or by an increase in the number of users in the congested access networks, resulting in longer flow durations. We can also evaluate the impact on the traffic of a change in the shot shape  $X_n(\cdot)$ , which may follow a change in the application or in the transport protocol.

We illustrate this latter application by the following example, which shows the impact of a change in the flow arrival rate  $\lambda$  on the traffic, and hence on the dimensioning of the backbone. Consider the case when the joint distribution of flow sizes and flow durations is stationary over long time intervals, and does not depend on the flow arrival rate<sup>3</sup>. Suppose that the ISP sets the bandwidth of its links to  $\mathbb{E}[R(t)] + A(\epsilon)\sqrt{V_R}$ , where  $A(\epsilon)$  is the  $\epsilon$ -quantile of the centered and normalized total rate  $R(t)$ , i.e., the value such  $\mathbb{P}\{R(t) > (\mathbb{E}[R(t)] + A(\epsilon)\sqrt{V_R})\} = \epsilon$ ,  $0 < \epsilon < 1$ .  $\epsilon$  is the congestion probability. The moments of  $R(t)$  in this expression of the bandwidth are given by our

<sup>3</sup>In the other case, a model has to be developed for the rest of the Internet, to evaluate the impact of a change in the arrival rate of flows on the joint distribution of flow sizes and flow durations. We will address this problem in a future research.

model (Corollary 1 and 2). The function  $A(\epsilon)$  can be computed using the Gaussian approximation of Section V-E, which gives for example  $A(0.05) = 1.96$ . When the arrival rate of flows increases, the bandwidth of the backbone links has to be increased as well, since the first and second moments of  $R(t)$  increase with  $\lambda$ . However, while the first moment of  $R(t)$  increases as  $\lambda$ , the standard deviation of  $R(t)$  increases as  $\sqrt{\lambda}$ . This indicates that the coefficient of variation of  $R(t)$  decreases as  $1/\sqrt{\lambda}$ . Concretely, this means that the traffic in the backbone becomes smoother and smoother when more and more flows are multiplexed. The consequence of this smoothing is that the ISP does not need to scale the bandwidth of its links linearly with  $\lambda$ . (S)He can gain in bandwidth by accounting for the smoothing of the traffic.

### B. Prediction of the total rate

The fact that our model gives the auto-covariance function of the total rate  $C_R(\tau)$ , allows us to propose predictors for its future values. To get a small error, the prediction has to be done on intervals of the same order of magnitude as the average flow duration. Prediction over longer time intervals is difficult as traffic becomes uncorrelated.

Predicting the future values of the total rate would help an ISP in the real-time management of the resources of its backbone. For example, an ISP can take the decision to route any new flow on a different path when the predicted total rate exceeds the available bandwidth, and to re-use the old path when the predicted total rate drops below the available bandwidth. This will prevent the congestion of the network and will ensure an efficient utilization of resources on the main paths.

The auto-covariance function of the total rate is computed from statistics on flows using (2). The prediction of the total rate after a certain time  $\tau$  in the future can be done using the present and the past values of the total rate itself. It can also be done using the present and the past values of the number of active flows  $N(t)$ . We will only consider the first kind of predictors in this section, however the second kind of predictors is a straightforward extension of the first one.

In this section we will consider one possible predictor for the total rate: the Moving Average predictor [14]. Let  $\tau$  denote the time between two predictions, and let  $R_k$  denote the total rate sampled at  $k\tau$ :  $R_k = R(k\tau)$ , with  $k \in \mathbb{Z}$ . We denote the estimate of  $R_k$  made by the predictor by  $\hat{R}_k$ . We take the prediction of the total rate as a linear function of the last  $M$  values of the process  $\{R_k\}$ :  $\hat{R}_{k+1} = \sum_{l=0}^{M-1} a_l R_{k-l}$ . Then, we look for the coefficients  $a_l$  that minimize the mean square error between  $\hat{R}_k$  and  $R_k$ , i.e.,  $\mathbb{E}[(\hat{R}_k - R_k)^2]$ . The desired coefficients are the solution of the following set of  $M$  linear equations, called normal equations in linear prediction theory [14],

$$\sum_{m=0}^{M-1} a_m r_R(m-l) = r_R(l+1), \quad l = 0, 1, \dots, M-1. \quad (8)$$

$r_R(l)$  is the lag- $l$  auto-correlation function of the process  $\{R_k\}$ , which is here  $r_R(l) = C_R(l\tau) + (\mathbb{E}[R(t)])^2$ , with  $C_R(l\tau)$  given by Theorem 2.

The value of  $M$  depends on how much the total rate is correlated; a significant correlation requires a large value of  $M$  to obtain a good prediction. But the value of  $M$  also depends on the number of samples available for the estimation of the auto-correlation functions of  $\{R_k\}$ . If the number of samples is small, the estimated auto-correlation functions will contain an important error. A large value of  $M$  may give a worse prediction error than a small value, since the number of lags of the auto-correlation function involved is larger in the former case. To choose an appropriate value for  $M$ , we increase it starting from one, then we take the lowest value that precedes an increase in the mean square error.

A natural approach for rate prediction is to compute the auto-correlation function  $r_R(l)$  and the prediction coefficients  $a_m, 0 \leq m \leq M-1$ , directly from the past samples of the measured rate  $\{R_k\}$ , instead of using Theorem 2, which requires additional information on flows. The advantage brought by our model is its ability to make use of a larger number of samples (since it uses samples of flow sizes and flow durations to compute the auto-correlation functions), and hence of a larger value of  $M$ , than a direct prediction from  $\{R_k\}$ . This is helpful when the prediction interval  $\tau$  gets large, as in this case the number of samples of the measured process  $\{R_k\}$  rapidly decreases, making a prediction using the samples of the process  $\{R_k\}$  not accurate. A solution could be to take samples of the process  $\{R_k\}$  over longer periods, but this is not feasible because of the stationarity problem.

We apply our predictor on a 30 minutes trace taken on one of the OC-12 links described in Table I. Similar results were found on the other traces.

First, we find the optimal value of  $M$ , and the optimal prediction coefficients  $a_m, 0 \leq m \leq M-1$ , using our model in the case of triangular shots. Then, we find the optimal Moving Average predictor using the measured process  $\{R_k\}$ . The optimal value of  $M$  is computed as explained above, and the optimal coefficients of the predictor are computed using (8). For each predictor, we compute the prediction error that we define as  $\sqrt{\frac{\mathbb{E}[(\hat{R}_k - R_k)^2]}{\mathbb{E}[R_k]}}$ . We show the results in Table II for different values of the prediction interval  $\tau$ . We deduce that our model allows a good prediction of the total rate; the error is close to what we obtain when we do the prediction using the samples of the total rate itself. For large  $\tau$ , our model behaves even better than using the measured process  $\{R_k\}$ , since the number of samples of the process  $\{R_k\}$  becomes small. In opposite to using the measured process  $\{R_k\}$ , our model relies in the computation of the auto-correlation function  $r_R(l)$  on the number of flows, which does not change when the prediction interval  $\tau$  changes. The graphs in Figure 14 better illustrate the results of the prediction for  $\tau = 10$ . The dashed lines plot the total rate measured on the link every 10 seconds. The thick lines show the prediction of the rate using the two techniques.

### C. Generation of backbone traffic

The third application of our model is the generation of backbone traffic. We want to use statistics on flows to generate a traffic (e.g., in a simulation tool) that looks like the real Internet traffic. We want both traffic to possess the same statistical

		Prediction interval $\tau$ (s)				
		2	5	10	30	60
Using $\{R_k\}$	M	6	2	2	2	3
	error (%)	3.9	4.2	4.68	4.74	5.61
Using the model	M	6	4	5	3	3
	error (%)	3.93	3.88	4.01	4.55	5.21

TABLE II  
PERFORMANCE OF OUR MODEL IN PREDICTING THE TOTAL RATE

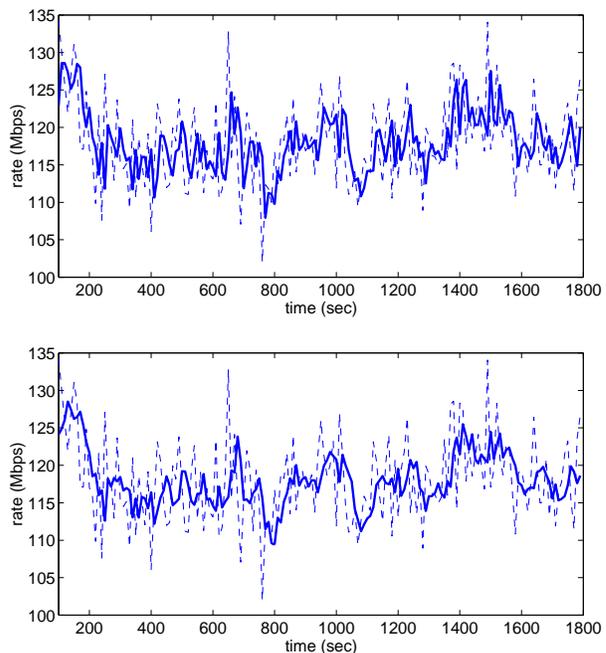


Fig. 14. Prediction of backbone traffic using our model (top) or rate measurements (bottom)

properties, for example, the same second order moments. The simplest method is to generate flows as a homogeneous Poisson process, and then, for each flow, to pick up a size  $S$  and a duration  $D$  using some distribution, and finally to transmit data at a constant rate equal to  $S/D$  during a time  $D$ . The intensity of the Poisson process and the distribution used to generate  $S$  and  $D$  are obtained from the statistics we have on the real traffic.

The traffic generated in this way models correctly the backbone traffic only when shots (i.e., flow rate functions) are really rectangles. We get a different traffic in the other cases. To improve the quality of the generated traffic, one has to transmit data during the lifetime of flows using a better model for shots, for example, using the power  $b$  model,  $b \neq 0$ . Here resides the main contribution of our model. It introduces a new component into traditional traffic models, that is, the shot  $X_n(\cdot)$ . In addition to  $\lambda$ ,  $S_n$  and  $D_n$ , one has to find the optimal value of  $X_n(\cdot)$  to be used for the transmission of data during the lifetime of flows. We discussed this issue in Section V-D. The optimal  $X_n(\cdot)$  depends on which property of the total rate we want to capture. The shot that captures the variance of the total rate would be different from the one that captures the auto-correlation of the total rate, or from the one that captures the

distribution of the total rate. Our model opens the door to many future research works on the characterization of the appropriate shot to be used to simulate Internet traffic.

### VIII. CONCLUSIONS

We proposed a traffic model for uncongested backbone links that is simple enough to be used in network operation and engineering. The model relies on Poisson shot-noise. With only 3 parameters ( $\lambda$ , arrival rate of flows,  $\mathbb{E}[S_n]$ , average size of a flow, and  $\mathbb{E}[S_n^2/D_n]$ , average value of the ratio of the square of a flow size and its duration), the model is able to find good approximations for the average of the total rate (the throughput) on a backbone link and for its variations at short timescales. The model is designed to be general so that it can be easily used without any constraint on the definition of flows, nor on the application or the transport protocol.

We are working on various extensions of the present work. We stated in the paper a result for the auto-covariance function of the total rate. Using this result, we are investigating the correlation of Internet traffic and its relation with the flow arrival process and the distributions of flow sizes and flow durations. We are further investigating the three applications of our model sketched in the paper. We are also studying the gain of introducing classes of flows with a different shot for each class. This will solve the problem when the flow rate functions do not have the same distribution. Finally, we are evaluating the worthiness of considering more complex flow arrival processes than Poisson. The challenge is to improve our evaluation of the total data rate without much increasing the complexity of the model. We want the model to be usable with current management tools used by ISPs.

### REFERENCES

- [1] E. Altman, K. Avrachenkov, and C. Barakat, "A stochastic model for TCP/IP with stationary random losses", *ACM SIGCOMM*, September 2000.
- [2] S. Ata, M. Murata, and H. Miyahara, "Analysis of network traffic and its application to design of high-speed routers", *IEICE Transactions on Information and Systems*, vol. E83-D, pp. 988-995, May 2000.
- [3] S. Ben Fredj, T. Bonald, A. Proutiere, G. Regnie and J. Roberts, "Statistical Bandwidth Sharing: A Study of Congestion at Flow Level", *ACM SIGCOMM*, August 2001.
- [4] P. Brémaud and L. Massoulié, "Power spectra of general shot noises and Hawkes point processes with a random excitation", *Journal of Applied Probability*, to appear.
- [5] T. Bu and D. Towsley, "Fixed Point Approximation for TCP behavior in an AQM Network", *ACM SIGMETRICS*, Jun. 2001.
- [6] J. Cao, W. Cleveland, D. Lin and D. Sun, "Internet traffic tends to Poisson and independent as the load increases", *Bell Labs Tech Report*, 2001.
- [7] N. Cardwell, S. Savage and T. Anderson, "Modeling TCP Latency", *IEEE INFOCOM*, March 2000.
- [8] J. Case, M. Fedor, M. Schoffstall and J. Davin, "A Simple Network Management Protocol (SNMP)", RFC 1157, May 1990.
- [9] M. Crovella and A. Bestavros, "Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes", *IEEE/ACM Transactions on Networking*, vol. 5, no. 6, pp. 835-846, Dec. 1997.
- [10] V. Dumas, F. Guillemin and P. Robert, "Limit results for Markovian models of TCP", *Globecom*, April 2001.
- [11] D. Daley and D. Vere-Jones, "An introduction to the theory of point processes", *Springer-Verlag*, 1988.
- [12] C. Fraleigh, S. Moon, C. Diot, B. Lyles and F. Tobagi, "Packet-Level Traffic Measurements from a Tier-1 IP Backbone", Sprint ATL Technical Report TR01-ATL-110101, November 2001.
- [13] A. Feldmann, "Characteristics of TCP connection arrivals," in Self-Similar Network Traffic and Performance Evaluation (K. Park and W. Willinger, eds.), John Wiley, 2000.
- [14] S. Haykin, "Modern filters", *Macmillan publishing company*, 1989.
- [15] R. Jain, "The art of computer systems performance analysis", *Wiley*, 1991.
- [16] A.A. Kherani and A. Kumar, "Performance Analysis of TCP with Non-persistent Sessions", *Workshop on Modeling of Flow and Congestion Control*, September 2000.
- [17] L. Kleinrock, "Queueing Systems, Vol. I: Theory", *Wiley*, 1975.
- [18] L. Kleinrock, "Queueing Systems, Vol. II: Computer Applications", *Wiley*, 1976.
- [19] W. Leland, M. Taqq, W. Willinger and D. Wilson, "On the self-similar nature of Ethernet traffic", *ACM SIGCOMM*, September 1993.
- [20] C. Nuzman, I. Saniee, W. Sweldens and A. Weiss, "A Compound Model for TCP Connection Arrivals", *ITC workshop*, Sep. 2000.
- [21] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP Throughput: a Simple Model and its Empirical Validation", *ACM SIGCOMM*, September 1998.
- [22] V. Paxson and S. Floyd, "Wide-Area Traffic: The Failure of Poisson Modeling", *IEEE/ACM Transactions on Networking*, vol. 3, no. 3, pp. 226-244, June 1995.
- [23] A. Weiss, "An introduction to large deviations for communication networks", *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 6, pp. 938-952, October 1995.