A flow-based model for TCP traffic in an IP backbone network

Chadi Barakat, Patrick Thiran, Christophe Diot

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

A flow-based model for TCP traffic in an IP backbone network

Chadi Barakat — Patrick Thiran — Christophe Diot

N° 4433

April 2002

THÈME 1

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--4433--FR+ENG

A flow-based model for TCP traffic in an IP backbone network

Chadi Barakat* , Patrick Thiran† , Christophe Diot‡

Thème 1 — Réseaux et systèmes
Projet PLANETE

Rapport de recherche n° 4433 — April 2002 — 30 pages

Abstract: We present a model of TCP flows suited to Internet backbone traffic, where links are usually *not* congested. We characterize the traffic using information on flows, i.e., arrival time, size, and duration. The major contribution of this paper is a model capturing the variation of the transmission rate during a TCP flow solely based on the above flow parameters. This model accounts for the dynamics of TCP congestion window and for the Timeout mechanism. It is independent of the packet loss rate and the round-trip time of the connection. We then model the traffic on a backbone link by aggregating TCP flows. Our model is easy to compute, and we show via simulations that it gives a good approximation of Internet backbone traffic.

Key-words: Poisson shot-noise, IP backbone, TCP flows, traffic modeling, simulations

* Chadi Barakat is with INRIA Sophia-Antipolis, PLANETE research group, email: Chadi.Barakat@sophia.inria.fr. The work has been done while the author was working with EPFL, with a grant from Sprint Labs

† Patrick Thiran is with LCA - EPFL, Switzerland, email: Patrick.Thiran@epfl.ch

‡ Christophe Diot is with Sprint Labs, CA, email: cdiot@sprintlabs.com

Modélisation du trafic TCP dans le backbone

Résumé : Nous présentons un modèle pour le trafic TCP dans un backbone de l'Internet. Les backbones de l'Internet ont la particularité d'être souvent sous-utilisés. Nous caractérisons le trafic dans le backbone à l'aide des informations suivantes sur les flots TCP: les instants d'arrivée des flots, leurs volumes, et leurs durées. La contribution majeure de notre travail est l'élaboration d'un modèle pour le débit d'un flot TCP, utilisant seulement les informations citées avant. Contrairement aux autres modèles dans la littérature, notre modèle ne demande pas la connaissance de la probabilité de perdre un paquet TCP, ni la connaissance du délai aller-retour. Une fois le modèle pour le flot TCP présenté, on modélise le trafic dans le backbone comme étant le résultat du multiplexage des flots TCP. Notre modèle est facile à utiliser, et nous montrons par des simulations qu'il permet une bonne approximation du trafic dans les backbones de l'Internet.

Mots-clés : Poisson shot-noise, backbone, flots TCP, modélisation du trafic, simulation

1 Introduction

Modeling Internet traffic is important to evaluate the performance of a backbone network. Models can assist in network links dimensioning. Models can also be used to plan link upgrades using statistics on the evolution of the users' demand. In case of a fixed demand, the model can predict any change in the total traffic on a link of the backbone network given a change in other parts of the network. For example, an increase in bandwidth on the access network will shorten the duration of flows and hence increase the variability of the traffic in the backbone, which requires an increase in the bandwidth of backbone links to absorb such a variability. Finally, we claim that for operational use, it is more important to model the network traffic for non congested links in order to anticipate and avoid congestion by an appropriate resource provisioning.

A promising approach is to model a link traffic as an aggregation of flows. Given statistics on the arrivals of flows, the sizes of flows, and the durations of flows, one can build a model that predicts the different measures of the traffic (e.g. average, variance) that will result from the multiplexing of all those flows. Different flow-based models exist in the literature (see [8] and the references therein), but most of them have the particularity of only dealing with congested links. These paper mostly focus on fairness issues.

In [7], we introduced a flow-based model for the IP backbone traffic. The main assumption was that the backbone is not congested. Flows are assumed to arrive as a Poisson process. They are characterized by their volume S and their duration D. S and D are two random variables that are supposed to have the same distribution for all flows. S and D are also very easy to compute on current routers. Using queuing theory and Poisson-shot noise processes, we characterized the moments of the total rate $R(t)$ resulting from the superposition of all the flows on one of the backbone links. The computation of the first moment of $R(t)$ is straightforward and does not need any particular assumption. If we denote by λ the average rate at which flows arrive on the link, the first moment of $R(t)$ (or the average total rate) will be equal to λ times the average size of flows[1]. For the higher moments of $R(t)$, we need the function that describes the evolution in time of the transmission rate of a flow. Let $X(t-t_0)$ denote this function for a flow that arrived at time t_0. In the context of Poisson shot-noise processes, $X(t-t_0)$ is the *shot* starting at time t_0. The variance

[1]The unit by which the total rate is expressed is equal to the unit by which the sizes of flows are expressed (packets, bytes, bits) divided by the unit of time.

of the total rate in the steady state has been shown to be equal to

$$V_R = \lambda \mathbf{E} \left[\int_0^D X^2(u) \mathrm{d}u \right]. \tag{1}$$

We also found that the computation of the k-th moment of the total rate requires the computation of all the expectations $\mathbf{E}\left[\int_0^D X^i(u)\mathrm{d}u\right]$, up to $i = k$.

In [7], the model is deliberately kept simple and general to make a very sparse use of the statistics that can be collected from passive measurements on the backbone (only flow arrival times, sizes, and durations) to model all IP flows including UDP and TCP. This crude model prevents us from having an explicit expression of $X(t)$ given S and D, which is needed to compute the higher moments of $R(t)$. We therefore made some a-priori choices for the shape of the shot; either a constant equal to S/D during the duration of a transfer (rectangular shot), or a linearly increasing rate from 0 to $2S/D$ in $[0, D]$ (triangular shot)[2]. These assumptions enabled us to compute the moments of $R(t)$ as a function of the join moments of S and D. For example, the variance of $R(t)$ can be expressed as

$$V_R = K\lambda \mathbf{E}\left[\frac{S^2}{D}\right],$$

where $K = 1$ in case of a rectangular shot, and $K = 4/3$ for a triangular shot.

In this paper, we focus our attention on TCP flows and we propose a new model of the moments of $R(t)$ that is specific to TCP. We motivate this choice by the dominance of TCP as transport protocol in the Internet [14]. On the other hand, we might loose in generality as in the future, another protocol could catch up on TCP and our TCP-based model could become less accurate than the rectangular and triangular models proposed in [7]. However, we believe that the domination of TCP is not compromised yet and that it is worth accounting for TCP control algorithms in our model. Contrary to UDP flows, the variation of the transmission rate of a TCP flow can be predicted given the well known TCP congestion control and congestion avoidance mechanisms. We make the same assumptions as in [7], i.e. that backbone links are not congested and flow arrivals are Poisson. However, losses dues to congestion are considered to occur on the access network only. The analysis can be extended to more general arrival processes, but we keep working

[2]The slope and starting point of $X(t)$ have been chosen in a way that the flow transmits a volume of data equal to S in a time equal to D. The function $X(t)$ is taken equal to zero outside the interval $[0, D]$.

with the Poisson assumption given the good results we get in previous works. We propose models for the computation of the integral of the shot created by a TCP flow throughout its lifetime. From these models, we find expressions for the expectation $\mathbf{E}\left[\int_0^D X^i(u)\mathrm{d}u\right]$, and hence, for the moments of the total rate $R(t)$. For this purpose, we use fluid models inspired from the models proposed in [2, 6, 9, 10, 13].

Our objective is to express the integral of the shot created by a TCP flow as a function of its size and its duration only. When computing these expressions, we account for the initial slow start phase of a TCP connection, its steady state, as well as Timeouts which are frequent at high loss rates [2, 13]. Models in the literature usually characterize a TCP flow using the connection round-trip time and the process of loss events (such as the intensity of loss events, the packet loss rate, etc.) as an implicit indicator of a congestion[2, 9, 13]. The problem with round-trip time and loss events is that they are difficult to measure in reality because (i) the finite duration of flows prohibits any measurement of quantities as the average packet loss rate or the average time between loss events[3], and (ii) the accurate measurement of these parameters requires accurate time source as well as the observation of the two paths of the TCP connection. With our model, the total rate can be approximated using only statistics on flow sizes and flow durations, which are easier to measure (using router embedded tools such as Netflow for example). Another objective for this paper is to propose a simple model for the shot of a TCP flow of size S and duration D that can be used for simulating and generating traffic.

We first focus in Section 2 on long-lived TCP flows. We give close-form expressions for the expectations $\mathbf{E}\left[\int_0^D X^i(u)\mathrm{d}u\right]$. Those expectations are required in the computation of the moments of $R(t)$, the total rate process [7]. We present the analysis for a general order i although we are more interested in the second order statistics (or the variance of $R(t)$ using (1)). The reason is that the analysis is approximately of the same complexity, and so providing expressions for all orders is better. Our expressions are a function of the joint moments of S and D, namely of the expectations $\mathbf{E}\left[S^i/D^{i-1}\right]$, $i \geq 2$. They are not a function of the round-trip times of TCP connections, not of the rate of loss events. This consists our main contribution: the round-trip time and the loss rate, which are the basis of most TCP models (e.g., [2, 9, 13]), are replaced in our expressions by the joint moments of S and D. These expressions hold for any marginal distribution of times between loss events.

[3]These quantities have to be measured on end-to-end and not only in the backbone, where there are generally no losses.

Next, in Section 3, we extend our model to account for the initial slow-start phase in a TCP transfer [9]. We present the analysis for the variance of the total rate, which is given in Equation (1). The same approach can be used to compute the higher moments of the total rate. We focus on the variance of $R(t)$ simply because it is more significant than the higher moments. The initial phase of a TCP transfer has an important impact when the size of the file transferred is short, which is the case with web traffic [9]. TCP has a different behavior during this initial phase than in steady state (exponential instead of linear window increase). We introduce two new parameters to account for this initial phase. These two parameters are S_s, the volume of data transmitted during the initial phase and D_s, the duration of the initial phase. We also use a fluid model for this phase inspired from [4, 9]. Again, the round-trip time of the connection and the packet loss rate are not needed. In Section 4, we combine this model to the steady state model introduced earlier to get a general model of the TCP shot.

At the end of the paper, we present two other models. The first one explains how the shot created by a TCP transfer can be approximated using the round-trip time of the connection and the packet loss rate. With this model, we get rid of D, S_s and D_s. We call this model the "packet model" to distinguish it from our previous model which we call the "fluid model". The advantage of the packet model is that it allows to account for Timeouts, a phenomenon pronounced at high packet loss rates [13]. The second model explains how our fluid model can be corrected to account for Timeouts.

The different models in the paper are validated by simulations using the ns-2 simulator [11]. The paper ends with concluding remarks and perspectives on our future work.

The outline of the paper is as follows. In the next section we present the analysis for the steady state phase of TCP. In Section 3, we present the analysis for the initial slow start phase of a TCP transfer. The general model for the shot is presented in Section 4 together with the simulation results. Sections 5 and 6 contain our last two models, respectively. The paper is concluded in Section 7.

2 Modeling TCP flows in the steady state

Given the distributions of S and D, our objective is to compute the following expectations,

$$\mathbf{E}\left[\int_0^D X^i(u)\mathrm{d}u\right], i \geq 2.$$

The particular expectation for $i = 2$ will be deduced and plugged in (1) to find the variance of the total rate. We consider a particular flow which has size s and duration d. We find first the expectations for this flow, that we sum over all flows to get the expressions of the expectations for an arbitrary TCP transfer. All expectations and probabilities conditioned on the fact that $S = s$ and $D = d$ will be distinguished by the subscript sd. Using this notation and supposing that d is large enough so that the time average of $X^i(t)$ coincides with its expectation (it has been shown in [2] that $X^i(t)$ is ergodic when the process of loss events is), we write

$$\mathbf{E}_{sd}\left[\int_0^D X^i(u)\mathrm{d}u\right] = d\mathbf{E}_{sd}\left[X^i(t)\right]. \qquad (2)$$

The problem is then transformed into the computation of the i-th moment of the transmission rate of a long-lived TCP transfer at an arbitrary time t in the steady state.

We consider a fluid additive-increase multiplicative-decrease model for the steady state of TCP inspired from [2]. We look at the TCP connection in its congestion avoidance mode. We define a loss event as the time at which the transmission rate of the TCP connection (or equivalently the window size) is divided by two. Denote by λ_l the rate of loss events. To simplify the analysis, we suppose that times between loss events are i.i.d., in other words we only consider the marginal distribution of times between loss events and we ignore any correlation. It has been shown in [2] that the correlation of times between loss events has a small impact on the performance of TCP over a wide range of Internet paths. We also suppose that there is no limitation on TCP rate caused by the receiver window or any other factor. It is indeed very difficult to find close-form expressions for TCP performance measures in presence of such a limitation and this is for general processes of loss events [2]. Even for the homogeneous Poisson case, the expressions are complicated and involve many infinite series [3]. Moreover, it is recommended in [1] to conduct all analysis of TCP without this limitation since it will disappear in future TCP versions.

Between loss periods, the rate of TCP increases linearly with a slope α, which is inversely proportional to the square of the round-trip time of the connection [2]. When a loss event appears, the transmission rate of TCP is divided by two. Let X_n denote the rate of TCP just before the n-th loss event. Let T_n be the time at which the n-th loss event appears and let τ_n denote the time between the n-th and the $(n+1)$-th loss events: $\tau_n = T_{n+1} - T_n$. The process $\{\tau_n\}$ is assumed to form an i.i.d. sequence, whose average is denoted by $1/\lambda_l = \mathbf{E}[\tau_n]$, and whose k-th moment is denoted by $\tau^{(k)} = \mathbf{E}\left[\tau_n^k\right]$. The variance of τ_n is thus equal to $\tau^{(2)} - (1/\lambda_l)^2$. For a time t in $(T_n, T_{n+1}]$, we have

$$X(t) = X_n/2 + \alpha(t - T_n).$$

For $t = T_{n+1}$, the previous equation becomes

$$X_{n+1} = X_n/2 + \alpha\tau_n. \tag{3}$$

Using this fluid model, we find the expressions of the moment of TCP transmission rate in the steady state. These expressions are stated in the following theorem.

Theorem 1 *The k-th moment of TCP transmission rate at an arbitrary time in the steady state is equal to*

$$\boldsymbol{E}_{sd}\left[X^k(t)\right] = \sum_{i=0}^{k} \frac{C_k^i}{2^i(k-i+1)}(\alpha/\lambda_l)^{k-i}\hat{\tau}^{(k-i+1)}\boldsymbol{E}_{sd}^0\left[X_n^i\right],$$

where $C_k^i = \frac{k!}{i!(k-i)!}$ is the (k,i) binomial coefficient, $\hat{\tau}^{(k-i+1)} = \lambda_l^{k-i+1}\tau^{(k-i+1)}$ is the $(k-i+1)$-th moment of the times elapsed between loss events divided by the average time between these events, $\boldsymbol{E}_{sd}^0\left[X_n^i\right]$ are the moments of $X(t)$ at loss events, i.e., the moments of $X(t)$ associated with the Palm distribution, which is here the distribution of TCP transmission rates at loss events. The moments $\boldsymbol{E}_{sd}^0\left[X_n^i\right]$ are the solutions of the following set of recurrent equations:

$$\boldsymbol{E}_{sd}^0\left[X_n^i\right] = \sum_{j=0}^{i-1} C_i^j \frac{2^{i-j}}{2^i-1}(\alpha/\lambda_l)^{i-j}\hat{\tau}^{(i-j)}\boldsymbol{E}_{sd}^0\left[X_n^j\right].$$

Proof: Let us put ourselves in the stationary regime of the TCP connection. According to [2], the stationary regime exists and is unique for any initial state. Using the following inversion formula from Palm theory [2, 5], the moment of order k of $X(t)$ at a random time t is given by

$$\mathbf{E}_{sd}\left[X^k(t)\right] = \frac{\mathbf{E}_{sd}^0\left[\int_{T_n}^{T_{n+1}} X^k(u)\mathrm{d}u\right]}{1/\lambda_l} = \frac{\mathbf{E}_{sd}^0\left[\int_0^{\tau_n}(X_n/2+\alpha u)^k\mathrm{d}u\right]}{1/\lambda_l}. \tag{4}$$

The binomial formula tells us that

$$(X_n/2 + \alpha u)^k = \sum_{i=0}^{k} C_k^i \frac{X_n^i}{2^i}\alpha^{k-i}u^{k-i}.$$

Therefore, (4) becomes

$$\begin{aligned}\mathbf{E}_{sd}\left[X^k(t)\right] &= \sum_{i=0}^{k} C_k^i \lambda_l \frac{1}{2^i}\alpha^{k-i}\mathbf{E}_{sd}^0\left[X_n^i \int_0^{\tau_n} u^{k-i}\mathrm{d}u\right] \\ &= \sum_{i=0}^{k} \frac{C_k^i}{2^i(k-i+1)}\lambda_l \alpha^{k-i}\mathbf{E}_{sd}^0\left[X_n^i \tau_n^{k-i+1}\right] \\ &= \sum_{i=0}^{k} \frac{C_k^i}{2^i(k-i+1)}\lambda_l \alpha^{k-i}\tau^{(k-i+1)}\mathbf{E}_{sd}^0\left[X_n^i\right]. \qquad (5)\end{aligned}$$

The later equality uses the fact that X_n and τ_n are independent. This follows from our assumption that the times between loss events form an i.i.d. sequence.

The moments of X_n can be computed using Equation (3). Making a similar development as the one made for obtaining (5), we find

$$\mathbf{E}_{sd}^0\left[X_{n+1}^i\right] = \sum_{j=0}^{i} C_i^j \frac{1}{2^j}\alpha^{i-j}\tau^{(i-j)}\mathbf{E}_{sd}^0\left[X_n^j\right].$$

Since we are working in the stationary regime, we have $\mathbf{E}_{sd}^0\left[X_{n+1}^i\right] = \mathbf{E}_{sd}^0\left[X_n^i\right]$. We bring the term $\mathbf{E}_{sd}^0\left[X_n^i\right]$ from the right hand side of the equation to the left hand side, and we get the following expression of the i-th moment of X_n as a function of lower order moments

$$\mathbf{E}_{sd}^0\left[X_n^i\right] = \sum_{j=0}^{i-1} C_i^j \frac{2^{i-j}}{2^i-1}\alpha^{i-j}\tau^{(i-j)}\mathbf{E}_{sd}^0\left[X_n^j\right]. \qquad (6)$$

By simple multiplications and divisions of Equations (5) and (6) by equal powers of λ_l, we get the system of equations in the Theorem which concludes the proof.

◇

Consequently, we have two systems of equations that give us all the moments of $X(t)$ as a function of two unknowns: the ratio α/λ_l and the normalized moments of times between loss events $\hat{\tau}^{(i)} = \lambda_l^i \tau^{(i)}$. To compute the moment of order k of $X(t)$, we have to know the ratio α/λ_l and the terms $\hat{\tau}^{(i)}$ up to $i = k+1$.

The normalized moments of times between loss events describe the way with which loss events appear. They are a function of the loss event process, and they are independent of the average rate of this process. For example, for equal times between loss events, they are always equal to 1. For exponential times between loss events, they are equal to $i!$. We shall assume in the sequel that these normalized moments are given. We still have to find the expression of the ratio α/λ_l in order to eliminate completely the average rate of loss events and the parameter α (hence the round-trip time) from the expressions of TCP transmission rate moments. This will be done in the next section using the fact that $\mathbf{E}_{sd}[X(t)] = s/d$ (Equation (2)). In the following section, we will focus on the computation of the second moment of $X(t)$ and infer from it the variance of the total rate process $R(t)$.

2.1 First moment of $X(t)$ and the ratio α/λ_l

We solve the system of equations in Theorem 1 for $k = 1$. This gives us

$$\mathbf{E}_{sd}[X(t)] = (\alpha/\lambda_l)\left(1 + 0.5\hat{\tau}^{(2)}\right).$$

This result is similar to the one found in [2]. Using equation (2) which holds since the transfers are assumed to be long, we write $\mathbf{E}_{sd}[X(t)] = s/d$. It follows that

$$\frac{\alpha}{\lambda_l} = \frac{s}{d}\left(1 + 0.5\hat{\tau}^{(2)}\right)^{-1}. \tag{7}$$

We get rid of the ratio (α/λ_l) by using statistics on flow sizes and flow durations as well as information on the normalized second moment of times between loss events that TCP connections encounter in the network. We will use this value of the ratio α/λ_l in the sequel.

2.2 Variance of the total rate

Let $f_{SD}(s,d)$ be the joint probability density of S and D. Using (1) and (2), we write

$$V_R = \lambda \int\int f_{SD}(s,d)\mathbf{E}_{sd}\left[\int_0^D X^2(u)\mathrm{d}u\right]\mathrm{d}s\mathrm{d}d = \lambda\int\int f_{SD}(s,d)d\mathbf{E}_{sd}\left[X^2(t)\right]\mathrm{d}s\mathrm{d}d.$$

We solve the system of equations in Theorem (1) for $k = 2$ and we get

$$\mathbf{E}_{sd}\left[X^2(t)\right] = \frac{1}{3}\left(\frac{\alpha}{\lambda_l}\right)^2\left(2 + 4\hat{\tau}^{(2)} + \hat{\tau}^{(3)}\right).$$

We substitute the ratio α/λ_l by its value given by (7) to obtain

$$\mathbf{E}_{sd}\left[X^2(t)\right] = \frac{1}{3}\left(\frac{s}{d}\right)^2\left(1 + 0.5\hat{\tau}^{(2)}\right)^{-2}\left(2 + 4\hat{\tau}^{(2)} + \hat{\tau}^{(3)}\right).$$

We plug this expression of the second moment of TCP transmission rate in the expression of the variance of $R(t)$ to obtain our following main result

$$V_R = \frac{1}{3}\lambda \left(1 + 0.5\hat{\tau}^{(2)}\right)^{-2} \left(2 + 4\hat{\tau}^{(2)} + \hat{\tau}^{(3)}\right) \mathbf{E}\left[\frac{S^2}{D}\right]. \tag{8}$$

This expression of V_R has the particularity to only depend on the joint moment of flow sizes and flow durations. More precisely, it is a multiple of the expectation $\mathbf{E}\left[S^2/D\right]$. Note that the two expressions of the variance we obtained in [7] and we cited in the introduction are also a multiple of the expectation $\mathbf{E}\left[S^2/D\right]$, but a multiplicative coefficient K which is a priori chosen, and not computed from the distributions of loss events as in (8). Expression (8) is independent of packet loss rates and round-trip times. The higher moments of $R(t)$ can be computed in a similar way and can also be shown to be only dependent on the joint moments of flow sizes and flow durations. The variance of $R(t)$ is also a function of the type of the process of loss events. This is represented by the second and third normalized moments of times between loss events, $\hat{\tau}^{(2)}$ and $\hat{\tau}^{(3)}$. This expression holds for any distribution of times between loss events. We are assuming that the process of loss events has the same type for all TCP connections.

Expression (8) tells us that if the time between loss events is equal (so that $\hat{\tau}^{(k)} = 1$ for all k), the variance of $R(t)$ is equal to $\frac{28}{27}\lambda\mathbf{E}\left[S^2/D\right]$, that is, it is equal to $\frac{28}{27}$ times the variance of $R(t)$ when all shots are assumed to be rectangles. Interestingly enough, if times between losses are exponentially distributed, we get the same variance as in the triangular case, that is, a variance equal to $\frac{4}{3}\lambda\mathbf{E}\left[S^2/D\right]$. In fact, the variance of $R(t)$ is an increasing function of the normalized moments of τ_n. The burstier the process of loss events, the higher the second moment of $X(t)$, and as a result, the higher the variance of $R(t)$.

2.3 Simplified model for TCP flows

Using the result of our modeling, one can find a simple model for the shot created by a long TCP transfer. When used in our previous fluid-based model, this simple model for the shot should give the same characteristics for the process $R(t)$ as the use of an exact model. A simple model for the shot is also useful for the generation of TCP traffic. Instead of mimicking exactly the behavior of TCP when sending packets, one can use a simple model so that the resulting traffic has the same moments as the real TCP traffic.

The complexity of a model for the shot of a TCP flow depends on how the number of high order moments of the process $R(t)$ are required to be computed. If we only need to approximate the first two moments of $R(t)$, we need a model for the shot that gives us the same average and the same variance for $R(t)$ than an exact model. Therefore, we need to find a function $X(t)$ that verifies the following for a TCP flow of size S and of duration D,

$$\int_0^D X(u)\mathrm{d}u = S$$

$$\int_0^D X^2(u)\mathrm{d}u = \frac{1}{3}\left(1+0.5\hat{\tau}^{(2)}\right)^{-2}\left(2+4\hat{\tau}^{(2)}+\hat{\tau}^{(3)}\right)\frac{S^2}{D}.$$

The function $X(t)$ has to be only defined on the interval $[0, D]$. It will be a function of S, D, $\tau^{(2)}$ and $\tau^{(3)}$. Clearly, an infinite number of functions satisfy such requirements. For example, one can take $X(t) = at^b$, and compute the two coefficients a and b using the above two integral equations. When the times between loss events are exponentially distributed (so that $\tau^{(k)} = k!$), we get $b = 1$ and $a = 2S/D^2$. This gives a triangular form for the shot as the one we considered in [7]: the rate of TCP increases linearly from 0 to $2S/D$ in the time interval $[0, D]$. Using this triangular model for the shot, we get the same average and the same variance for the process $R(t)$ as when using a real model for TCP steady state with exponential times between loss events. However, no guarantees can be made about the higher moments of $R(t)$, its marginal distribution, and its correlation function.

As explained in [2, 3, 12], assuming that times between loss events are exponentially distributed leads to good results, especially in wide-area networks where a large number of flows are multiplexed. We adopt this assumption in the rest of the paper. Consequently, we model the shot created by a TCP flow in its steady state as a triangle. Our objective being only to approximate the second order moment of the total rate.

3 Modeling the initial phase of TCP

The above model only accounts for the steady state phase (i.e., the congestion avoidance phase) of a TCP connection. It is valid for long-lived connections where the initial slow-start phase does not have an impact. However, most of transfers in the Internet are of small size [9, 14]. A small transfer starts (and generally terminates) by a slow-start phase that has a different behavior than the congestion avoidance phase. In particular, the increase in the window size is exponential at the beginning of the connection until the first packet loss, which marks the end of this initial phase. We design a model for the slow-start phase in this section. Again, we try not to use packet loss events and round-trip time into our model. This is possible with the two random variables we introduced earlier, S_s and D_s.

Denote by S_s the volume of data transmitted during the initial slow-start phase. Denote by D_s its duration. The initial slow start phase lasts from the beginning of the transfer until the first packet loss (or equivalently the first reduction of the congestion window). The receiver window is assumed to be very large so as not to limit the transmission rate of TCP. The analysis of the initial slow start phase can be easily done in the case of a small receiver window, but since this is not the case for the steady state, we opted to make this assumption during all the duration of the transfer. As in the previous sections, S and D denote the total size and total duration of the flow. Clearly, if the flow does not suffer from losses, we will have $S_s = S$ and $D_s = D$.

The increase in TCP rate in the slow-start phase is known to be exponential [4, 10]. Denote by RTT the round-trip time of the TCP connection and by b a parameter that accounts for delaying ACKs (b=2 if ACKs are delayed, b=1 if not). Using the fluid models in [4, 10], we have during the slow-start phase

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = \frac{X(t)}{bRTT}.$$

This differential equation holds for any unit used to express transmission rates. Let M denote the size of a TCP packet (Maximum Segment Size in TCP implementations). The TCP connection starts usually with a window of one packet. Then, $X(0) = M/RTT$, and the expression of $X(t)$ in the slow-start phase is therefore

$$X(t) = \frac{M}{RTT} e^{t/(bRTT)}. \tag{9}$$

Since we know that

$$\int_0^{D_s} X(u)\mathrm{d}u = S_s,$$

we can solve this integral equation for the round-trip time of the connection as a function of S_s and D_s, to obtain

$$RTT = \frac{D_s}{b\ln((S_s + Mb)/(Mb))}. \tag{10}$$

We substitute this expression for the round-trip time in (9), which then becomes

$$X(t) = \frac{Mb\ln((S_s + Mb)/(Mb))}{D_s} e^{\ln((S_s+Mb)/(Mb))t/D_s}.$$

This gives the expression of the rate of TCP in the slow-start phase as only a function of S_s and D_s; the round-trip time disappeared from this expression. The rate of TCP at the end of the slow-start phase is equal to

$$X(D_s) = \frac{S_s + Mb}{D_s} \ln\left(\frac{S_s + Mb}{Mb}\right).$$

It is now very simple to compute any power of $X(t)$ and find its integral in the slow-start phase as a function of S_s and D_s. Then, we sum over all the values of S_s and D_s to find the expectations $\mathbf{E}\left[\int_0^{D_s} X^k(u)\mathrm{d}u\right]$ that we need for the computation of the moments of the total rate $R(t)$. Combined with our previous model for the steady state, this will give us the moments of the total rate as a function of λ, and the joint moments of S_s, D_s, S and D. We will show how this can be done for the variance of the total rate in the next section.

4 General fluid model for TCP flows

4.1 The model

We have TCP flows that arrive at an average rate λ. The shot of a TCP flow is characterized by the four random variables: S_s, D_s, S and D. We make the same assumption as in [9] that the connection gets into the steady state once the initial slow start phase is over. As explained above, we consider a triangular model for the shot of a TCP flow in its steady state. This yields the following general model for the shot of TCP including the initial and steady states,

$$X(t) = \begin{cases} \frac{Mb\ln((S_s+Mb)/(Mb))}{D_s} e^{\ln((S_s+Mb)/(Mb))t/D_s} & 0 \leq t < D_s \\ \frac{2(S-S_s)}{(D-D_s)^2}(t - D_s) & D_s \leq t < D \end{cases}$$

This model for the shot is better illustrated in Figure 1. We recall that this model, and particularly the steady state part, only approximates the first two moments of $R(t)$.

According to (1), we integrate the square of $X(t)$ between 0 and D and we sum over all the values of the four random variables of our model to get the variance of $R(t)$. Thus,

$$V_R = \lambda\mathbf{E}\left[\int_0^{D_s} X^2(u)\mathrm{d}u\right] + \lambda\mathbf{E}\left[\int_{D_s}^{D} X^2(u)\mathrm{d}u\right].$$

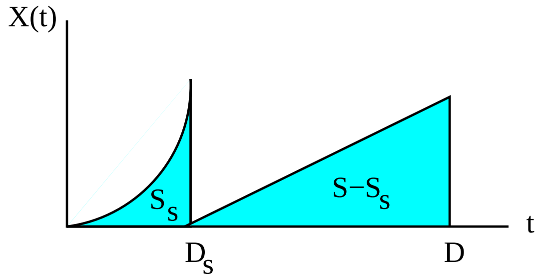

Figure 1: A model for the shot created by a TCP flow

The second term on the right of the equation corresponds to the steady state phase and is equal to $(4/3)\lambda\mathbf{E}\left[(S-S_s)^2/(D-D_s)\right]$. The first term can be computed using the expression of $X(t)$ in the slow-start phase. This gives us the following expression for the variance of the total rate,

$$V_R = \lambda\mathbf{E}\left[\ln\left(\sqrt{\frac{S_s+Mb}{Mb}}\right)\frac{S_s(S_s+2Mb)}{D_s}+\frac{4}{3}\frac{(S-S_s)^2}{D-D_s}\right]. \qquad (11)$$

For TCP flows that do not suffer from losses, the second term vanishes and V_R becomes

$$V_R = \lambda\mathbf{E}\left[\ln\left(\sqrt{\frac{S+Mb}{Mb}}\right)\frac{S(S+2Mb)}{D}\right].$$

4.2 Validation

We validate our general model for TCP shot via simulation. We modify `ns-2` to be able to measure the values of S_s, D_s, S and D for a TCP transfer. Using the statistics on these four random variables, we study how well our expressions for V_R (using different models for the shot) approximate the real variance of the total rate. The real variance of the total rate is computed from the samples of $R(t)$ obtained by averaging the number of packets that cross the bottleneck link over intervals of the same order of the round-trip time. Such an averaging (or sampling) interval is the most appropriate to capture the dynamics of TCP rate, and hence the dynamics of the total traffic. Instead of the variance, we plot the coefficient of variation of $R(t)$ expressed in percents (i.e., we plot $(100 * \sqrt{V_R}/\mathbf{E}\left[R(t)\right])$. The average value of $R(t)$ has been shown to be equal to $\lambda\mathbf{E}\left[S\right]$ [7].

We consider different scenarios. Each scenario studies the sensitivity of the results to a different parameter of the model. Our scenarios consist in a set of TCP transfers arriving according to a Poisson process of rate λ. The transfers cross first a link of 10 Mbps, then they cross another link of 10 Mbps where their packets are dropped according to a Bernouilli process; with a dropping probability p. Note that since the link is under-utilized, these losses actually occur on the access or exit links. The transfers use the New Reno version of TCP, implement the delay ACK mechanism (hence $b = 2$), and have a very large receiver window. We set the Maximum Segment Size to 500 bytes. The parameters we vary are the distribution of the sizes of the transfers (S), the arrival rate of the transfers (λ), the packet drop probability (p), and the round-trip times (RTT) of the TCP connections. All TCP connections in a particular simulation have the same round-trip time, and the averaging interval of the total rate $R(t)$ is set to this round-trip time. We will choose the values of the different parameters of the model (particularly λ and $\mathbf{E}[S]$) so that the 10 Mbps link remains under-utilized. This is the main assumption we made in this paper. Therefore, the simulated networks do not enqueue packets, and the round-trip time for a TCP connection is equal to the two-way propagation delay.

For each scenario, we run 10 simulations of 1000 seconds each and we average the coefficient of variation of $R(t)$ over them. The confidence intervals are very small, so we do not plot them.

4.2.1 Impact of file size on variance

We set the round-trip time to 80 ms, the packet drop probability to 3% and the arrival rate of TCP transfers to 1 transfer per second. Files to be transmitted in a simulation have all the same size. We vary this size from 10 Kbytes to 1 Mbytes and we run 10 simulations for each file size. We compute the coefficient of variation of the total rate given by simulation (by sampling the total rate over 80 ms intervals). We also compute the coefficient of variation of the total rate given by three models for the shot: the rectangular model, the triangular model, and our TCP fluid model proposed in this paper. Figure 2 shows the results. The TCP fluid model gives a good approximation of the variance of the total rate for all sizes of TCP transfers. It gives better results than the triangular model in case of TCP transfers of small size. The reason for that is that the former model accounts for the fast rate increase during the slow start phase. In case of TCP transfers of large size, the TCP fluid model gives similar results to the triangular model since the initial slow start phase becomes negligible. The rectangular model results in the lowest coefficient of variation since it ignores any variation of rate during the lifetime of a TCP flow.

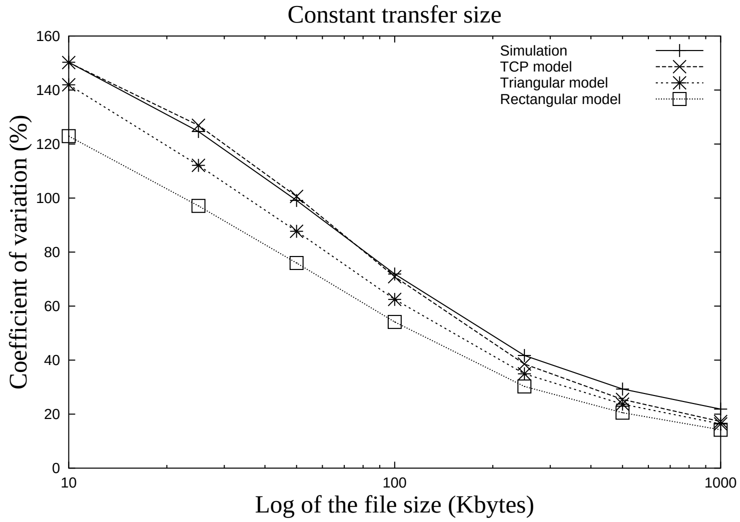

Figure 2: Variance of $R(t)$ versus file size

4.2.2 Impact of arrival rate on variance

We consider the impact of the arrival rate of TCP transfers (or equivalently the load of the network) on the accuracy of the model. We set the round-trip time to 80 ms and the packet loss rate to 3%. We vary the sizes of flows for each simulation run. When a file is to be generated, we pick randomly a real number between 1 and 3 and we generate a file of size in bytes equal to 10 power the picked number. This leads to average file size of 215 Kbytes.

We vary the arrival rate of transfers from 1 and 5 transfers per second and we plot the coefficient of variation of the total rate given by simulation and by the above three models. The results are plotted in Figure 3. First, we notice that the coefficient of variation of the total rate decreases with λ since it is inversely proportional to $\sqrt{\lambda}$. Second, the TCP fluid model gives a better approximation of the real coefficient of variation than the triangular model. The results at high arrival rates are not very representative since the 10 Mbps link starts then to be congested. In the following two scenarios, we will choose an arrival rate equal to 3 transfers per second, since with this value for λ, the probability that the total rate reaches 10 Mbps is very small.

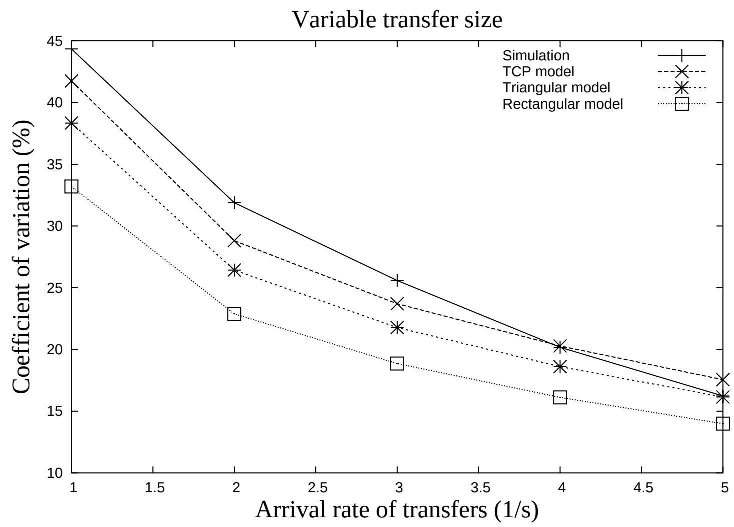

Figure 3: Variance of $R(t)$ versus arrival rate of transfers

4.2.3 Impact of packet loss rate on variance

In this section, we vary the packet loss rate. A change in the packet loss rate impacts the durations of transfers, and hence the variance of the total rate. We consider five values for the packet loss rate: 1%, 2.5%, 5%, 7.5%, and 10%. As in the previous section, the size of a TCP transfer is generated randomly between 10 Kbytes and 1 Mbytes for each simulation run. The round-trip time is set to 80 ms and the transfers' arrival rate to 3 transfers per second. Figure 4 shows how the coefficient of variation of the total rate varies with the packet loss rate and how well our models approximate this coefficient. The increase in packet loss rate stretches the TCP transfers, and hence, decreases the coefficient of variation of the total rate which becomes smoother and smoother. This smoothness of the total rate continues until a certain point where the increase in loss rate does not have an impact on the coefficient of variation. The results given by the three models plotted figure 4 do not account for such stabilization of the coefficient of variation, and continue their decrease proportionally to the packet loss rate.

The stabilization of the simulated coefficient of variation at high loss rates is caused by Timeouts, which become frequent in this region [13]. Timeouts result in the TCP flow switching between silent phases and transmission phases. Even

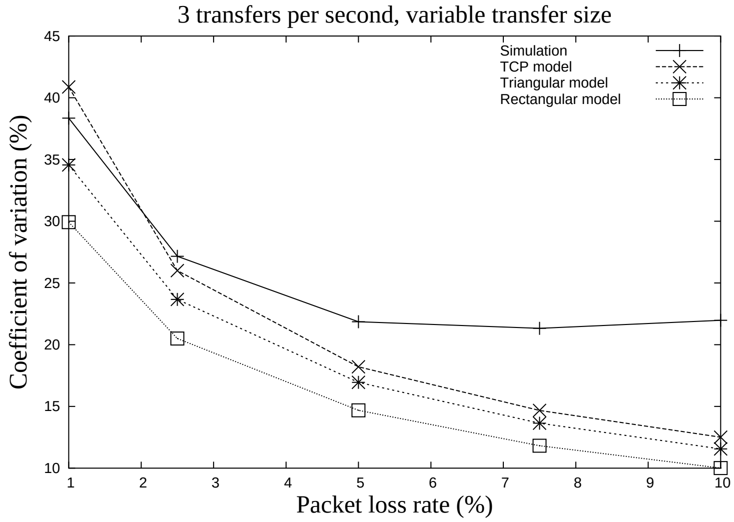

Figure 4: Variance of $R(t)$ versus packet loss rate

though the duration of the transfer becomes longer, the burstiness of the flow caused by Timeouts yields to the maintaining of the coefficient of variation at a constant value. Our fluid model does not account for Timeouts (it assumes that the flow continuously transmits packets during its lifetime). Hence, it is not able to capture this plateau of the coefficient of variation. In Section 6, we will present a method to account for Timeouts in our fluid model for TCP flows.

4.2.4 Impact of round-trip time on variance

Varying the round-trip time also has an impact on the TCP rate. The longer the round-trip time, the longer the duration of the transfer. We study here the impact of the round-trip time on the coefficient of variation of $R(t)$ and on the efficiency of our models. We set the packet loss rate to 3% and the arrival rate of TCP transfers to 3 transfers per second. The size of a transfer is generated randomly between 10 Kbytes and 1 Mbytes as explained above.

We increase the round-trip time (together with the sampling interval of the total rate) from 40 ms to 200 ms. All connections in the same simulation run have the same round-trip time. We plot the coefficient of variation of the total rate as a function of the round-trip time. We find the results in Figure 5. The TCP fluid

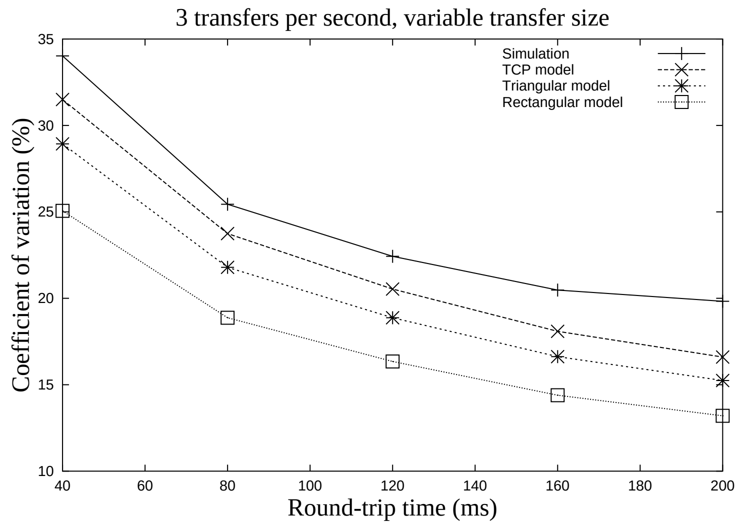

Figure 5: Variance of $R(t)$ versus round-trip time

model gives a good approximation of the real coefficient of variation over all the values of round-trip time we are considering. We also notice that the coefficient of variation of the total rate decreases with the round-trip time since the durations of flows become longer.

5 An alternative method for modeling TCP flows

We present an alternative method for modeling TCP flows. The previous method only uses statistics on flow sizes and flow durations to model the shot created by TCP. To model the slow start phase, it was necessary to introduce two additional parameters that are the volume of data transmitted until the first loss, and the time taken until the first loss. With the present method, we only use three parameters: the size of the TCP transfer S, the packet loss probability p, and the average round-trip time RTT. S corresponds to the volume of data given by the application to the TCP connection to transmit (not including retransmissions). These are the three parameters widely used in the literature to model TCP transfers (e.g., [9]). The advantage of this second method is that it allows to account for Timeouts in TCP transfers. There is no way to model Timeouts without the knowledge of packet loss probability. Another advantage is that we only need to measure two parameters (p and RTT) instead of three (S_s, D_s and D). However, as explained in the introduction, this second method has the problem of measuring p and RTT, which are more difficult to measure than the sizes of flows and their durations. We present this second method for completeness of the study. It will serve as an introduction to the next section, where we introduce the effect of Timeouts into our fluid model.

5.1 The model

Consider a TCP transfer of size $S = s$. Without loss of generality, we suppose that s is measured in packets. Suppose that the packets of the transfer are dropped in the network with a probability p and that the round-trip time of the connection is RTT. M denotes the Maximum Segment Size. We will develop a model inspired from those in [9] and [2] to characterize the shot created by the TCP transfer.

We will only focus on the computation of the variance of $R(t)$. This requires the computation of the following expectation for the TCP transfer

$$\mathbf{E}_{spr}\left[\int_0^d X^2(u)\mathrm{d}u\right], \tag{12}$$

with D the duration of the transfer. The subscript spr means that we are conditioning on the fact that the transfer has a size s, a loss probability p and a round-trip time RTT. Clearly, this expectation will be a function of p, RTT and s. We sum then over all the transfers and we multiply by λ to find the variance of $R(t)$ according to (1).

The idea is simple. First, we integrate $X^2(u)$ in the slow start phase, and we compute the average value of S_s. This gives us the number of packets to transmit in the steady state of the connection which equals $s - \mathbf{E}_{spr}[S_s]$. Then, we use the model for the steady state phase of TCP in [2] to compute the integral of $X^2(u)$ in the steady state. The computation is done in the following two sections.

5.1.1 Computation for the slow start phase

We have to distinguish between two cases. The first case is when no packets of the TCP transfer are lost. The second case is when at least one packet of the TCP transfer is lost. Let p_l denote the probability that at least one packet is lost. We have,

$$p_l = 1 - (1-p)^s. \tag{13}$$

When no packets are lost, all the s packets of the transfer are transmitted in the slow start phase (we are working under the assumption that the receiver window is set to a very large value). Using the expression of RTT in (10), we can get the following value for the duration of the transfer,

$$D = bRTT \ln\left(\frac{s+b}{b}\right).$$

The expectation in (12) will be equal to

$$\mathbf{E}_{spr}\left[\int_0^D X^2(u)\mathrm{d}u \,\middle|\, \text{no loss}\right] = M^2 \frac{s(s+2b)}{2bRTT}.$$

Consider now the case when at least one packet is lost. In this case, S_s packets are transmitted in the slow start phase with $S_s \leq s$. We can find the average value of S_s using the packet loss probability. Denote this average by $\overline{S_s} = \mathbf{E}_{spr}\left[S_s \,|\, \text{loss}\right]$. It is equal to

$$\begin{aligned} \overline{S_s} &= \frac{p}{p_l} + 2\frac{(1-p)p}{p_l} + \ldots + s\frac{(1-p)^{s-1}p}{p_l} \\ &= \frac{1}{p} - \frac{s(1-p)^s}{1-(1-p)^s} \end{aligned} \tag{14}$$

With a more complex computation, we can also find the second moment of the number of packets transmitted in the slow start phase, given that there is at least one loss in the transfer. Denote this second moment by $\overline{S_s^2} = \mathbf{E}_{spr}\left[S_s^2 \,|\, \text{loss}\right]$. One can use $(\overline{S_s})^2$ as an approximation, but we will use the exact value,

$$\overline{S_s^2} = \frac{2}{p^2} - \frac{1}{p} - \frac{(1-p)^s(ps^2+2s)}{p(1-(1-p)^s)}. \tag{15}$$

Again, we use Equation (10) to find D_s, the duration of the initial slow start phase in case of losses, as a function of RTT and the moments of S_s. The expectation of the integral of $X^2(u)$ in the slow start phase is equal to,

$$\mathbf{E}_{spr}\left[\int_0^{D_s} X^2(u)\mathrm{d}u \,\middle|\, \text{loss}\right] = M^2 \frac{\overline{S_s^2} + 2b\overline{S_s}}{2bRTT}.$$

We still have to compute the expectation of the integral of $X^2(u)$ in the steady state in case of losses. Denote by $\overline{X^2}$ the second moment of the transmission rate of TCP in the steady state. Then, we write

$$\mathbf{E}_{spr}\left[\int_{D_s}^{D} X^2(u)\mathrm{d}u \,\middle|\, \text{loss}\right] = M^2(\overline{D} - \overline{D_s})\overline{X^2}.$$

$\overline{D} - \overline{D_s}$ is the average time required to transmit $s - \overline{S_s}$ packets in the steady state. Let $\overline{X}$ denote the average of the transmission rate of TCP in the steady state, then $\overline{D} - \overline{D_s} = (s - \overline{S_s})/\overline{X}$. We combine the different expectations above to get the following expression for the expectation of the integral of $X^2(t)$ through all the duration of the transfer,

$$\mathbf{E}_{spr}\left[\int_0^D X^2(u)\mathrm{d}u\right] = M^2\left((1-p_l)\frac{s(s+2b)}{2bRTT} + p_l\left(\frac{\overline{S_s^2}+2b\overline{S_s}}{2bRTT} + \frac{(s-\overline{S_s})\overline{X^2}}{\overline{X}}\right)\right).$$

$\overline{S_s}$ and $\overline{S_s^2}$ are given by (14) and (15), respectively. $\overline{X}$ and $\overline{X^2}$ will be computed in the next section as a function of p. We still have to sum over all the values of s, p, RTT, and multiply by the flow arrival rate λ to find the variance of the total rate $R(t)$.

5.1.2 Moments of TCP transmission rate in the steady state

We use here the model presented in [2]. We make again the assumption that times between loss events are exponentially distributed. Consider first that there is no Timeouts. According to the model in [2], we have

$$\overline{X} = \frac{1}{RTT}\sqrt{\frac{2}{bp}}, \tag{16}$$

$$\overline{X^2} = \frac{4}{3}(\overline{X})^2 = \frac{8}{3RTT^2}\frac{1}{bp}.$$

To account for Timeouts, we have to divide these two moments by a term $A(p, RTT) = \big(1 + p\overline{X}Q(p)Z(p)\big)$ (see [6] for details). This term, which is larger than one, corresponds to the ratio of the total transfer time (including Timeouts) and the total transmission time (excluding Timeouts). $Q(p)$ and $Z(p)$ have been computed in [13] and are equal to

$$\begin{aligned} Q(p) &= \min\left(1, \frac{(1-(1-p)^3)(1+(1-p)^3(1-(1-p)^{w-3}))}{1-(1-p)^w}\right), \\ Z(p) &= 4RTT\frac{1+p+2p^2+4p^3+8p^4+16p^5+32p^6}{1-p}, \end{aligned} \tag{17}$$

with

$$w = \frac{2+b}{3b} + \sqrt{\frac{8(1-p)}{3bp} + \left(\frac{2+b}{3b}\right)^2}.$$

In summary, $Q(p)$ denotes the probability that a congestion event results in Timeout, and $Z(p)$ denotes the average duration of a Timeout period. As we notice, the numerator and the denominator in the expression of the variance have to be divided by A, and so this factor disappears from the formula. We will later detail on the factor A, when we introduce a correction for our fluid model for the shot of TCP (the model using the four parameters S_s, D_s, S and D).

5.2 Validation

We consider the four scenarios in Section 4.2. For each scenario, we plot the coefficient of variation of the total throughput in a separate figure as a function of the corresponding parameter (i.e., S, λ, p and RTT). This gives us the Figures 6, 7, 8 and 9. Each figure contains the results from simulation, the results from our fluid model for the shot, and the results from our present packet model for the shot. We see that the packet model for the shot gives mainly better results when the packet loss probability is high. In the other regions, its performance is comparable to that of the fluid model, if not worse. The gain we get with the packet model in case of high loss rates is simply due to the fact that it accounts for Timeout intervals, during which the TCP flows are silent and not transmitting packets. The fluid model assumes that TCP sources are transmitting during all the durations of the transfers. Hence, it under-estimates the second moment of TCP transmission rate and consequently the coefficient of variation of the total rate.

6 Correction of the fluid model for Timeouts

As we saw above, the fluid model presents a problem in case of frequent Timeouts, or equivalently, in case of high packet loss rates. The problem comes from the fact that the fluid model assumes that a TCP source is always transmitting during the duration of the transfer D. In reality, it is only transmitting out of Timeout intervals. Instead of D, one has to use the time during which the TCP source is actually transmitting. However, this time is difficult to measure. We will provide in this section an approximation of this time using the function $A(p, RTT)$ introduced in the previous section.

Consider the expression of the variance in Equation (11). The first term in the expectation remains the same since it corresponds to the transitory phase during which no Timeouts occurs. The second term however has to be corrected. In this term, $D - D_s$ denotes the duration of the steady state phase. This duration has

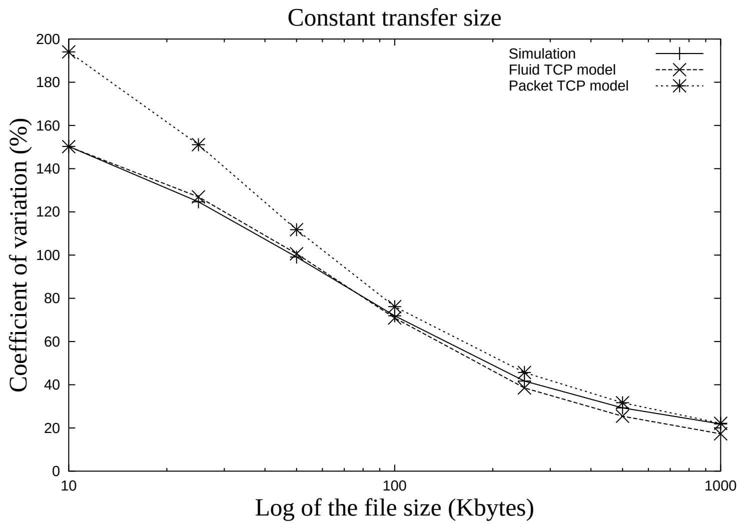

Figure 6: Variance of $R(t)$ versus file size

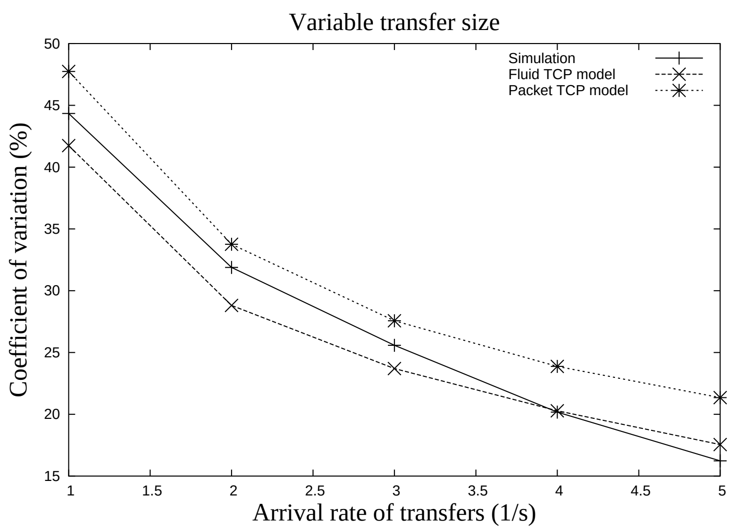

Figure 7: Variance of $R(t)$ versus arrival rate of transfers

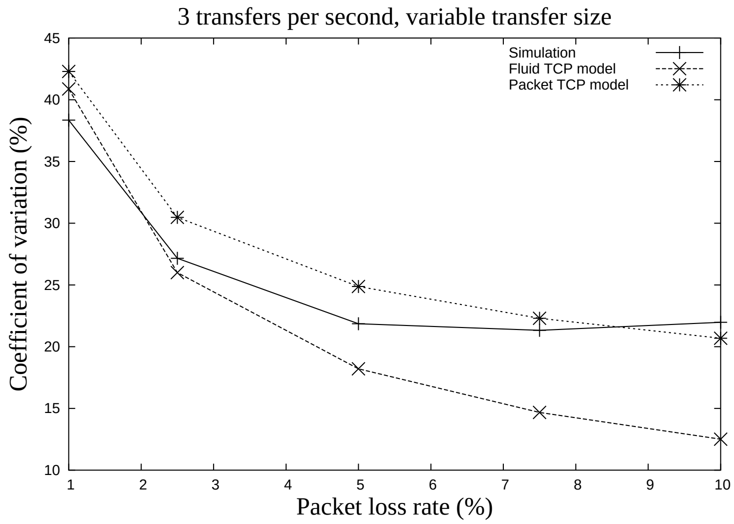

Figure 8: Variance of $R(t)$ versus packet loss rate

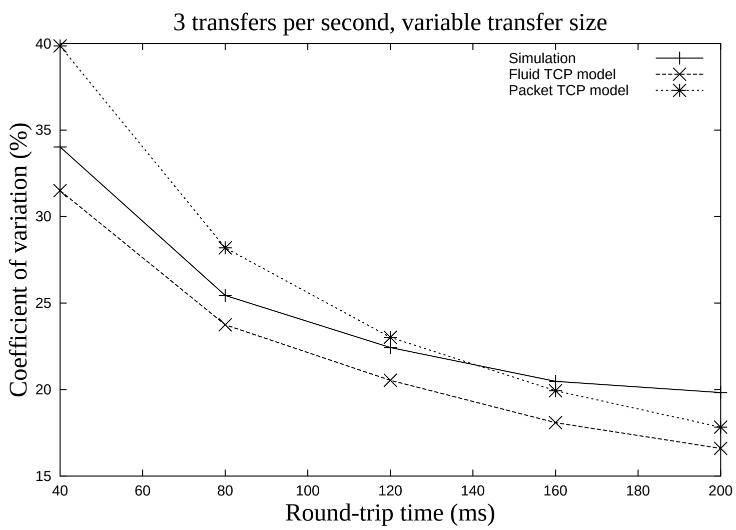

Figure 9: Variance of $R(t)$ versus round-trip time

to be corrected so that to eliminate the Timeout intervals out of $D - D_s$. Using $A(p, RTT)$, the time during which the TCP source is actually transmitting can be approximated by $(D - D_s)/A$ [6]. We substitute this into (11) to get the following corrected version of our fluid model,

$$V_R = \lambda \mathbf{E}\left[\ln\left(\sqrt{\frac{S_s + b}{b}}\right)\frac{S_s(S_s+2b)}{D_s} + \frac{4}{3}\frac{A(S-S_s)^2}{D-D_s}\right].$$

p and RTT can be approximated using the statistics on S_s and S. In fact, the term A is only a function of the packet loss probability p, since the round-trip time disappears when multiplying $\overline{X}$ by $Z(p)$ (Equations (16) and (17)).

For the computation of p, we use the expression of p_l (Equation (13)). $p_l(s)$ represents the probability that a TCP transfer of size s (packets) suffers from the loss of one of its packets. We also use the expression of $\mathbf{E}\left[S_s \,|S = s\right]$, the average volume of data transmitted during the initial slow start phase, given that the transfer has a size $S = s$. We take S_s equal to s when the flow does not suffer from losses. All TCP transfers are assumed to encounter the same packet loss probability p in the network, otherwise the computation of p will be difficult for the reasons we mentioned at the beginning.

Using a development similar to that in (14), one can show that

$$\mathbf{E}\left[S_s \,|S = s\right] = M\frac{p_l(s)}{p}.$$

All TCP packets have the same size M. By summing over all the values of s, we get the following expression for p:

$$p = M\frac{\mathbf{E}\left[p_l(S)\right]}{\mathbf{E}\left[S_s\right]}.$$

$\mathbf{E}\left[S_s\right]$ is the average volume of data transmitted during the initial slow start phase for an arbitrary TCP transfer. It can be easily approximated from the statistics on S and S_s. $\mathbf{E}\left[p_l(S)\right]$ is the probability that an arbitrary TCP transfer suffers from losses. Again, it can be approximated from the statistics on S and S_s. We do that by dividing the number of TCP flows that suffer from losses by the total number of flows monitored in the backbone.

To validate our corrected fluid model, we consider the third simulation scenario in Section 4.2 where we change the packet loss rate. We compare the coefficient of variation given by our fluid model to that given by the present corrected version of

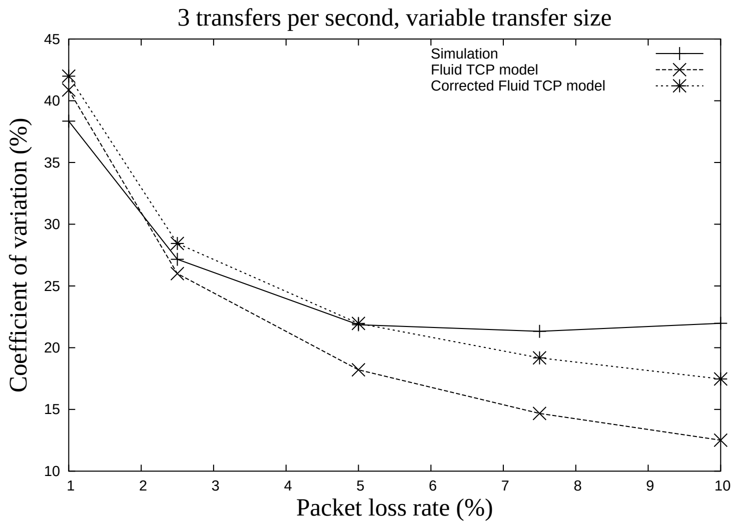

Figure 10: Variance of $R(t)$ versus packet loss rate

this model. The results are plotted in Figure 10. Clearly, the correction we introduce improves the performance of the model in case of high loss rates. In case of low loss rates, the original and the corrected versions give similar performances.

7 Conclusion

We proposed in this paper a model for the traffic on an IP backbone network. This model represents the traffic as the aggregate of TCP flows. The backbone link is supposed to be non-congested, which is the case in most tier 1 ISPs backbone networks. Using statistics on the sizes and durations of TCP flows, we computed expressions of the moments of the total link rate that results from the multiplexing of all TCP flows on this link.

Our main findings in this paper are close-form expressions for the moments of the aggregate of TCP traffic (average, variance) using only the joint moments of flow sizes and flow durations. In contrast to other modeling approaches in the literature, our expressions are independent of round-trip times and loss rates, which are difficult to measure in the core of the network. Our model accounts for the dynamics of TCP, and is valid for the different kinds of distributions of times between packet loss events.

It also accounts for Timeouts, which frequently occur at high loss rates. Based on our analysis, we provide a simple model for the shot of TCP that can be used to simulate and generate TCP traffic on uncongested backbone links (and evaluate the impact of a change e.g. in the traffic demands, or in the capacity of the access network, on the actual utilization of the link).

This model can be used for network management as it relies on parameters that are simple to compute on-line, namely S and D, the volume and the duration (respectively) of each TCP flow that compose the link traffic. Various router embedded tools can currently compute these parameters without affecting the performance of the router.

Our results can be extended in different directions. First, we are working on the validation of our results with real measurements on a major ISP backbone. In previous works, the validation of our model with simple forms for the shot on the same backbone has yielded very promising results. Second, we are working on the extension of our model to environments where the arrivals of flows are not really Poisson, and we are investigating whether a flow arrival process which has more correlation than a Poisson process has a big impact on the results. Finally, we are working on algorithms that implement the results of our model, and that could be used by an ISP to anticipate variations in its backbone traffic, using the history on the evolution of users' demand.

References

[1] M. Allman and A. Falk, "On the Effective Evaluation of TCP", *ACM Computer Communication Review*, vol. 29, no. 5, Oct. 1999.

[2] E. Altman, K. Avranchenkov, and C. Barakat, "A stochastic model for TCP/IP with stationary random losses", *ACM SIGCOMM*, Aug. 2000.

[3] E. Altman, K. Avrachenkov, C. Barakat, and R. N. Queija, "State-dependent M/G/1 Type Queueing Analysis for Congestion Control in Data Networks", *IEEE INFOCOM*, Apr. 2001.

[4] E. Altman, J. Bolot, P. Nain, D. Elouadghiri- M. Erramdani, P. Brown, and D. Collange, "Performance Modeling of TCP/IP in a Wide-Area Network", *IEEE Conference on Decision and Control*, Dec. 1995.

[5] F. Baccelli and P. Bremaud, "Elements of queueing theory: Palm-Martingale calculus and stochastic recurrences", *Springer-Verlag*, 1994.

[6] C. Barakat, "TCP modeling and validation", *IEEE Network*, vol. 15, no. 3, pp. 38-47, May 2001.

[7] C. Barakat, P. Thiran, G. Iannaccone, C. Diot, and P. Owezarski, "A flow-based model for Internet backbone traffic", *SPRINT ATL Technical Report TR02-ATL-020181*, Feb. 2002.

[8] S. Ben Fredj, T. Bonald, A. Proutiere, G. Regnie and J. Roberts, "Statistical Bandwidth Sharing: A Study of Congestion at Flow Level", *ACM SIGCOMM*, Aug. 2001.

[9] N. Cardwell, S. Savage and T. Anderson, "Modeling TCP Latency", *IEEE INFOCOM*, Mar. 2000.

[10] T.V. Lakshman and U. Madhow, “The performance of TCP/IP for networks with high bandwidth-delay products and random loss”, *IEEE/ACM Transactions on Networking*, vol. 5, no. 3, pp. 336-350, Jun. 1997.

[11] The LBNL Network Simulator, *ns*, http://www.isi.edu/nsnam/ns/

[12] V. Misra, W.-B. Gong, and D. Towsley, “Stochastic differential equation modeling and analysis of TCP-windowsize behaviour”, *Performance*, Oct. 1999.

[13] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, “Modeling TCP Throughput: a Simple Model and its Empirical Validation”, *ACM SIGCOMM*, Sep. 1998.

[14] K. Thompson, G.J. Miller, and R. Wilder, "Wide-Area Internet Traffic Patterns and Characteristics", *IEEE Network*, vol. 11, no. 6, pp. 10-23, Nov. 1997.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399