HTTP Turbulence

François Baccelli, Augustin Chaintreau, Danny de Vleeschauwer, David R. Mcdonald

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

HTTP Turbulence

François Baccelli — Augustin Chaintreau — Danny De Vleeschauwer — David McDonald

N° 5205

Mai 2004

THÈME 1

ISSN 0249-6399 ISRN INRIA/RR--5212--FR+ENG

HTTP Turbulence

François Baccelli* , Augustin Chaintreau† ,
Danny De Vleeschauwer‡ , David McDonald§

Thème 1 — Réseaux et systèmes
Projet TREC

Rapport de recherche n° 5205 — Mai 2004 — 54 pages

Abstract: In this paper, we consider a set of HTTP flows using TCP over a common drop-tail link to download files. After each download, a flow waits for a random think time before requesting the download of another file, whose size is also random. When a flow is active its throughput is increasing with time according to the additive increase rule, but if it suffers losses created when the total transmission rate of the flows exceeds the link rate, its transmission rate is decreased. The throughput obtained by a flow, and the consecutive time to download one file are then given as the consequence of the interaction of all the flows through their total transmission rate and the link's behavior.

We study the mean-field model obtained by letting the number of flows go to infinity. This mean-field limit may have two stable regimes : one without congestion in the link, in which the density of transmission rate can be explicitly described, the other one with periodic congestion epochs, where the inter-congestion time can be characterized as the solution of a fixed point equation, that we compute numerically, leading to a density of transmission rate given by as the solution of a Fredholm equation. It is shown that for certain values of the parameters (more precisely when the link capacity per user is not significantly larger than the load per user), each of these two stable regimes can be reached depending on the initial condition. This phenomenon can be seen as an analogue of turbulence in fluid dynamics: for some initial conditions, the transfers progress in a fluid and interaction-less way; for others, the connections interact and slow down because of the resulting fluctuations, which in turn perpetuates interaction forever, in spite of the fact that the load per user is less than the capacity per user. We prove that this phenomenon is present in the Tahoe case and both the numerical method that we develop and simulations suggest that it is also be present in the Reno case. It translates into a bi-stability phenomenon for the finite population model within this range of parameters.

* INRIA-ENS, 45 rue d'Ulm 75005, Paris, France, francois.baccelli@ens.fr
† INRIA-ENS, 45 rue d'Ulm 75005, Paris, France, Augustin.Chaintreau@ens.fr
‡ ALCATEL-NSG, Antwerp, Belgium, danny.de_vleeschauwer@alcatel.be
§ INRIA and University of Ottawa, Canada, dmdsg@mathstat.uottawa.ca

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex (France)
Téléphone : +33 1 39 63 55 11 —Télécopie : +33 1 39 63 53 30

This research was supported in part by the "Opération Stratégique Conjointe" Alcatel-INRIA entitled "End to End Analysis of IP Traffic".

Key-words: Mean field, HTTP connection, TCP, congestion control, flow control, additive increase–multiplicative decrease algorithm, IP traffic, synchronization. AMS 1980 subject classifications: Primary 60K25; Secondary 60K20.

Turbulence HTTP

Résumé : Cet article étudie le partage de la bande passante d'un lien avec débordement par des flots HTTP. Dans ce modèle, après chaque téléchargement de fichier, un flot attend pendant une durée aléatoire avant de demander le téléchargement d'un nouveau fichier de taille elle aussi aléatoire. Lorsqu'un tel flot est actif, son débit évolue selon l'algorithme AIMD (additive increase, multiplicative decrease) de TCP: ce débit augmente linéairement tant qu'il n'y a pas de débordement et il est divisé par 2 (RENO) ou ramené à une valeur proche de 0 (TAHOE) lorsque des pertes sont détectées. Les débits obtenus par un flot et les durées de téléchargement des fichiers sont donc déterminés par l'interaction de tous les flots et leur compétition pour le partage de la bande passante de ce lien.

Nous étudions un modèle de champ moyen obtenu en faisant tendre le nombre des flots vers l'infini. Le modèle limite peut avoir deux régimes stables, l'un sans congestion dans lequel la distribution des débits peut être décrite de manière explicite, l'autre avec des congestions périodiques. Pour ce deuxième régime, la période peut être obtenue comme solution d'une équation de point fixe et la distribution des débits comme la solution d'une équation de Fredholm.

Nous montrons qu'il existe des cas où la capacité du lien par flot est supérieure à la charge par flot et où chacun de ces deux régimes peut cependant être atteint en fonction des conditions initiales. Ce phénomène peut être vu comme un analogue de la turbulence en dynamique des fluides: pour certaines conditions initiales, les transferts de fichiers progressent de manière fluide et sans interactions entre les flots. Pour d'autres, les flots entrent en interaction et sont ralentis par les fluctuations qui résultent de cette interaction et ce ralentissement favorise à son tour les interactions, qui persistent indéfiniment.

Nous prouvons que ce phénomène est présent dans le case de TCP TAHOE. Les méthodes numériques que nous développons dans l'article ainsi que les simulations suggèrent que ce phénomène est aussi présent dans le cas RENO.

Ces propriétés du modèle de champ moyen se traduisent par un phénomène de bi-stabilité pour les modèles avec population finie.

Ces travaux de recherche s'inscrivent dans le cadre de l'"Opération Stratégique Conjointe" Alcatel-INRIA intitulée "End to End Analysis of IP Traffic".

Mots-clés : Champ moyen, connexion HTTP, TCP, contrôle de congestion, contrôle de flux, algorithme de croissance additive et décroissance multiplicative, trafic IP, synchronisation. Classifications AMS 1980: principale 60K25; secondaire 60K20.

1 Introduction

Consider the following toy model: two users with the same round trip time R share a bufferless link of capacity C. Each user alternates between OFF periods and download periods. As an example, if $C = \infty$, each user has the following behavior: it remains silent during an OFF period that is of duration β^{-1} and then downloads a file of μ^{-1} packets, alternating between such OFF and download periods in a periodic way forever. Assume now that the file transfers of these users are controlled by a TCP-Reno like congestion control mechanism based on the additive increase multiplicative decrease (AIMD) rule for the transmission. To make the problem as simple as possible, disregard slow start and time out.

- In case the link capacity C is infinite, the additive increase rule of the congestion avoidance mechanism of TCP implies that each download takes $t = R\sqrt{2/\mu}$, so that each user has a periodic behavior of period $\beta^{-1} + t$ in which it downloads files at a long term average rate of $\rho = (\mu(\beta^{-1} + R\sqrt{2/\mu}))^{-1}$ packets per second.
- If $C > 2\sqrt{2/(\mu R^2)}$ (which is the sum of the peak rates of the two flows), then the two users never fill in the link capacity and TCP never lets them interact either.
- If $C < 2\sqrt{2/(\mu R^2)}$, then two things may happen depending on the initial phases of the two flows.
  - Assume first that the two flows are in phase, namely both simultaneously start a download at time 0. Then the capacity of the link is exceeded before the end of the two simultaneous transfers and (assuming full synchronization of the losses) both flows then experience a loss at some time $t' < t$ which results in a window being divided by two that we consider to be instantaneous at time t'. For certain values of the parameters (see Figure 1), the two flows then complete their download simultaneously at some epoch $t'' > t$ before the capacity of the link is again exceeded, so that the two flows actually reach a periodic regime of period $t'' + \beta^{-1}$ in which they download a long term average of $\rho' = (\mu(\beta^{-1} + t'')^{-1} < \rho$ packets per second.
  - Assume now that the two flows are out of phase, in such a way that, for instance, when one is downloading, then the other one is OFF (which is possible if $\beta^{-1} > t$). Then the situation seen by each flow is exactly as that in the case with $C = \infty$, so that the long term average download rate is ρ
- If $C < \sqrt{2/(\mu R^2)}$, then the two flows experience losses in any case.

The main conclusion from this toy example is that the average rate obtained by the flows depends on their relative phase.

The main question addressed in this paper is that of the possible persistence of this phase dependence phenomenon in a more realistic model where the number of users is large

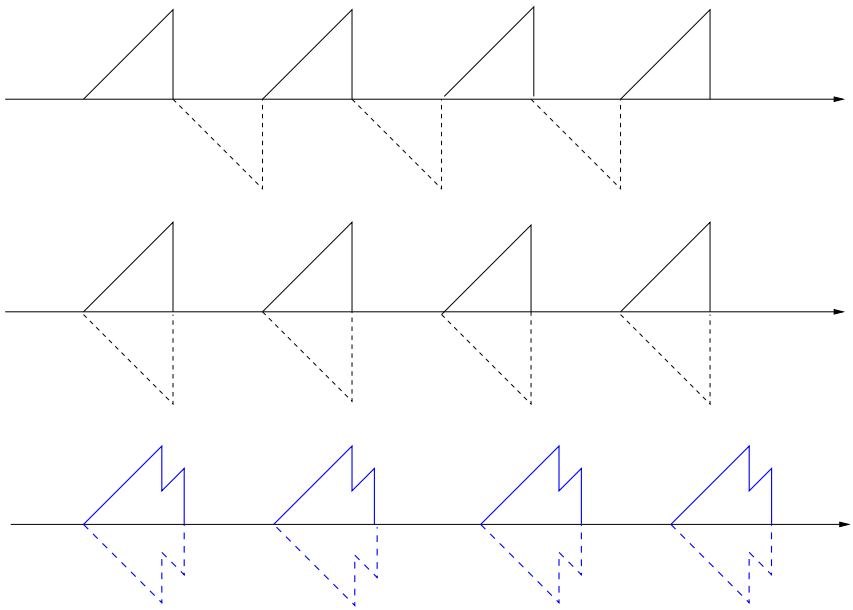

Figure 1: Flow 1 is the solid line and flow 2 the dotted line. Top figure: two out of phase flows with capacity above the peak rate of one flow; no interaction. Middle figure: two in phase flows with capacity larger than twice the peak rate of one flow: no interaction. Bottom figure: two in phase flows with capacity between the peak rate and twice the peak rate; the interaction regime is reached and this slows down the file transfers.

and where both the file sizes and the OFF periods are independent random variables, with given distribution functions on the positive real line. It could be expected that in contrast with the last deterministic model, the "mixing" operated by the randomness of the file sizes and the OFF periods leads to some form of dephasing of the various flows and hence to a throughput that is independent of the initial phase condition.

We will consider two cases: the Reno case based on the additive increase multiplicative decrease (AIMD) rule for of the transmission rate and the Tahoe case. The Reno case will be the default assumption throughout the paper.

Informed readers may be concerned that our model for interacting HTTP flows sharing a common link is oversimplified. It is well known that within the context of the Internet, it is appropriate to assume that the distribution of file sizes and OFF periods have heavy tails (e.g. Pareto file sizes and Weibull or lognormal OFF periods, as for example in [10]). However, in most of the mathematical derivations of the present paper, we will assume heavy tails because we are unable to solve the associated mathematics at this stage (nor can the rest of the scientific community to the best of our knowledge). We will rather concentrate

on the version of the problem where both file sizes and OFF periods are exponential random variables and where all files and OFF times are mutually independent.

Why study a model based on statistical assumptions that are clearly inappropriate? The rationale is as follows: the exponential case is tractable and allows one to identify and prove the presence of phenomena that are also observed by simulation in the heavy tailed case. So the mathematical study based on the exponential case will be important step in the direction of the understanding of the interaction of HTTP flows with the more realistic statistics.

In the mathematical analysis, we assume the existence of a stationary deterministic mean-field limit when the number of flows goes to infinity. In this deterministic limit there are two possible stable regimes. If the file sizes are small enough the link is able to carry all the traffic without congestion. The average transmission rate stabilizes at a value calculated below giving an overall utilization of the link which is less than one. In the other stable regime there is a series of congestion epochs where the buffer overflows and the active flows experience losses and cut their transmission rate in two. The main aim of this paper is to investigate these two regimes, and in particular the conditions under which they appear and the stationary distributions they lead to.

Section 3 gives a necessary condition for the existence of stationary regimes with congestion epochs. This necessary condition is based on the rate conservation principle which allows one to pose a fixed point problem for the rate of congestion epochs. The numerical aspects associated with this fixed point equation are discussed in detail in this section: the functions that are used in this fixed point equation are obtained as the solutions of Fredholm integral equations of the second kind, which are derived from a regenerative analysis of the rate of a tagged flow. This leads to an efficient way of calculating the possible values of the period of the mean-field model.

Section 4 focuses on a necessary and sufficient condition for the existence of stationary regimes with congestion epochs. For this, we first study the interaction-less regime, for which we establish a partial differential equation. We give both an explicit solution of this PDE and an efficient numerical way to solve it via yet another Fredholm equation of the second type, which has a natural regenerative interpretation. We then establish an invariant equation describing, for a given inter-congestion period of the mean-field process, the stationary distribution of rates at a congestion epoch. The existence of a probability measure solution of this invariant measure equation that satisfies certain conditions described in the paper is a necessary and sufficient condition for the existence of such a periodic congestion regime. The associated integral equation is again a Fredholm integral equation of the second kind.

One of the key observations is made in Section 4 : within this setting, it is possible to have multiple stationary mean-field regimes depending on the initial conditions: for certain values of the parameters, there exist both a "fluid regime" where flows do not interact at all and a "turbulent regime" where the fact that flows interact once implies a slow down of the whole system that propagates interaction forever.

Section 5 extends the approach to a model with a simple representation of slow start. Section 7 gathers simulation results on the bi-stability phenomenon and on the case with heavy tailed file sizes and OFF-times. We show by simulation and analysis that the phe-

nomena that are identified in the exponential model are also present in the heavy tailed case. Section 6 focuses on the comparison of our results with those of earlier models of the literature. In particular, we compare this model to the processor sharing Engset model. Finally Section 8 gathers some packet level simulations based on ns2 [1]. The aim of this section is to confirm the presence of the results identified via the fluid mathematical models and the fluid simulation of the present paper in a packet level simulation.

2 Related Work

Modeling TCP through the fairness it achieves (or equivalently the utility functions that it optimizes) has been a very active area of research since the work of Kelly in [15]. A general extension of this framework to dynamic traffic with a large number of flows is described in [10]. In [18] this framework is used to study the performance of networks with dynamic traffic (in [18] files to be transmitted arrive according to a Poisson process), with several types of fairness assumptions. In [12] the results proven in this previous paper are extended to a Poisson arrival process of sessions, each associated with a file download having a general distribution. [12] contains a proof that if the network can be modeled by a processor sharing queue - or equivalently if instantaneous fair sharing can be assumed in the network - then the mean throughput only depends on the average requested size per session. Comparison with simulations is provided but as the authors themselves remarked, this result might be challenged in real networks either for very small flows, that do not last long enough to benefit from their possible fair share in the network capacity, or for close to critical load where the discrimination between flows and the unequal sharing due to TCP are more frequent.

At the same time, a few papers focused on TCP bandwidth sharing for dynamic traffic when taking into account the AIMD rule. In [14], one of the first models developed on dynamic traffic, a version of the Engset model is proposed and shown to be insensitive w.r.t the file size distribution. TCP is modeled as a constant transfer rate calculated from the study of TCP sharing for a fixed number of persistent flows that are exactly in phase (increasing their window and decreasing it by the same amount at the same time). This model is extended by Kherani and Kumar under exponential assumptions in [16] where the inter-congestion period and the increase of the total rate is now dynamically changing with the traffic. In this model the flows contributing to the traffic are all in phase (they all react together at the same time and in the same way) ; the analytical result cannot be explicitly given in the general case but only in the low load, large file case where TCP bandwidth can be approximated by a completely fair allocation.

Our work extends these two papers; we are not assuming that the flows are in phase or that they share the bandwidth equally. We study the asymptotics of a model with N ON/OFF flows sharing a link according to an AIMD rules, when N tends to infinity. Our goal is to provide - in the exponential case that we can entirely characterized via fixed point equation - observations on the efficiency of the TCP sharing, with no assumption of the fairness achieved, and for any load on the network.

3 A Necessary Condition for the Existence of a Regime with Periodic Congestion

3.1 Model

We suppose N HTTP flows share a link which has *no* buffer or rather a small buffer that cushions collisions. The link rate is CN packets per second so the link drops packets at random when the transmission rates of the flows exceed the link rate. We assume each HTTP flow is silent for an exponential time with a mean $1/\beta$. After the silence period the flow transmits a file where the distribution of file sizes is exponential with a mean $1/\mu$. The default option is that each flow implements TCP Reno so the transmission rate increases at rate $1/R^2$ during the transmission of a file where R is the round trip time of packets. When the file has been transmitted the transmission rate is reset to zero.

The interaction between flows is via the sum of their rates. As long as this sum, which we refer to as the aggregate rate, is less than NC, then there is no interaction between the flows. When the aggregate rate reaches the link capacity CN, an event that we call a congestion epoch occurs. For the sake of tractability, we assume that all losses taking place before the flows react take place instantaneously. This reaction consists in the fact that Reno may cut the rate given to each of the N flows independently with a probability p. The parameter p, which is the proportion of flows that experience a loss at such a congestion epoch, is called the synchronization rate of the model (this parameter is evaluated from queueing theory by Baccelli and Hong in [6]). After this reaction, the aggregate rate is again less than C and a new interaction-less phase starts. In the TCP-Tahoe case, the rate of flows that experienced a loss is reset to 0.

3.2 Rate Conservation

Define $X(t)$ to be the transmission rate of a tagged flow participating in the steady state. Assume that there exists a stationary regime for $X(t)$, namely that it is a stationary stochastic process defined on a probability space $\{\Omega, \mathcal{F}, P\}$. The distribution of $X(t)$ is therefore the distribution of all the transmission rates in the steady state. $X(t)$ increases linearly at rate $1/R^2$ when it is active; i.e. with mean rate $\mathbb{P}(X(0) > 0)/R^2$. This increase is counteracted by negative jumps when a file finishes and the transmission rate drops to zero. It is also counteracted by a reduction by one half when a packet is lost at a congestion epoch.

The following point processes will be useful:

- T, the point process of congestion epochs, with inter-arrival times τ, with Palm expectation $\mathbb{E}_0^\tau$; let $\bar{\tau}$ denote the expectation of the inter-congestion times w.r.t. $\mathbb{P}_0^\tau$;
- D, the point process of file completions of the tagged flow, with intensity λ_δ and with Palm expectation $\mathbb{E}_0^\delta$.

When a file is completely downloaded, the throughput is reset to zero. Hence, with the introduced notation, the rate of decrease of the transmission rate due to file completions is

$\lambda_\delta \mathbb{E}_0^\delta(X(0^-))$. In addition to that, the mean rate at which the tagged flow suffers a packet loss is $p/\bar{\tau}$, and the tagged flow divides its transmission rate by 2 for each loss. Consequently the rate of decrease of the transmission rate due to packet loss is $\frac{p}{\bar{\tau}}\mathbb{E}_0^\tau[X(0-)/2]$. Since the utilization is exactly one when the congestion epoch begins it follows that $\mathbb{E}_0^\tau[X(0-)] = C$ so the rate of decrease of the transmission rate due to packet loss is $pC/(2\bar{\tau})$.

By the rate conservation principle (RCP, see e.g. [5], Chapter 1), the mean rate of increase equals the mean rate of decrease. So

$$\frac{\mathbb{P}(X(0) > 0)}{R^2} = \frac{pC}{2\bar{\tau}} + \lambda_\delta \mathbb{E}_0^\delta[X(0^-)]. \tag{1}$$

On the left hand side the unknown quantity is the steady state probability that a flow is active while on the right hand side we have λ_δ, the rate at which file completions occur and $\mathbb{E}_0^\tau[X(0^-)]$, the mean transmission rate observed when the file is completely downloaded.

In the Tahoe case, the RCP equation reads

$$\frac{\mathbb{P}(X(0) > 0)}{R^2} = \frac{pC}{\bar{\tau}} + \lambda_\delta \mathbb{E}_0^\delta[X(0^-)]. \tag{2}$$

In what follows, the RCP will be used as a way to determine the possible values of $\overline{\tau}$. As we shall see in §3.3 below the expressions that show up in the RCP equation, namely $\mathbb{P}(X(0) > 0)$ and $\mathbb{E}_0^\delta[X(0^-)]$ can be computed as a function of $\overline{\tau}$, so that this equation can be seen as a fixed point equation for $\overline{\tau}$.

3.3 The Fredholm Equations

In this section and in the rest of the paper, we let the parameter N tend to ∞ and we assume the existence of a stationary mean-field limit as $N \to \infty$ in the same spirit as in [7], [10] or [6]. In such a mean-field regime the inter-congestion times become deterministic and we have propagation of chaos; i.e. each flow becomes independent. We will concentrate on the case where the stationary regime of the mean-field limit has inter-congestion times are all equal some constant τ.

We will see below that when assuming τ known, all quantities in Equation (1) can be computed as the solutions of certain Fredholm integral equations, and that (1) can be used as fixed point for determining τ.

In this section, we assume τ to be given. We define a cycle to start at a congestion epoch where the tagged flow is idle. The cycle ends at the first congestion epoch when the flow is idle again. We use the following notation :

- Σ is the point process of congestion epochs where the tagged flow is idle, with inter-arrival times σ and with Palm expectation $\mathbb{E}_0^\sigma$.

The rationale for defining such cycles is that the sequence of successive cycles associated with the tagged flow is i.i.d. or in other words that the beginning of cycles are regeneration times for the tagged flow.

3.3.1 Expected number of files in a cycle

Define $f(t)$ to be the expected number of files that will be transmitted by the end of the current cycle given that the tagged flow is inactive at the current time t (where $0 \leq t < \tau$). Also define $g(z)$ to be the expected number of files that will be transmitted by the end of the current cycle given that the current transmission rate of the tagged flow is z packets per second and that the current time is immediately after a congestion epoch.

Our goal is to evaluate $f(0)$ but we find $f(t)$ for all $t \in [0, \tau[$. Since the silence period has an exponential distribution we can condition on the time when the flow has a new file to transmit. There are two possibilities. Either the file arrives before the next congestion epoch at some time r where $t \leq r \leq \tau$ or it does not. If it hasn't arrived, the current cycle ends and $f(t) = 0$.

If it does, for a time r where $t \leq r \leq \tau$, we condition on the size y of the arriving file. There are again two cases. Either the transmission of this file is completed before the next congestion or there is some remaining data to be transmitted after the next congestion epochs. We are in the first case if we can transmit y packets in $\tau - r$ time units given that the flow starts out with transmission rate zero. Since the transmission rate increases at rate $1/R^2$ it will take t' time units to transmit y packets if $y = (t'/2)(t'/R^2)$, ; i.e. if $t' = R\sqrt{2y}$. Consequently y packets can be transmitted before the next congestion epoch only if $y \leq (\tau - r)^2/(2R^2)$. In this case we add one to the number of files transmitted during the current cycle plus a renewal term. We can summarize this first case by

$$\int_t^\tau \beta e^{-\beta(r-t)} \left(\int_0^{\frac{(\tau-r)^2}{2R^2}} \mu e^{-\mu y} dy (1 + f(r + R\sqrt{2y}) \right) dr.$$

In the second case the y packets cannot be transmitted before the next congestion epoch. In this case, which occurs with probability $\exp(-\mu(\tau - r)^2/(2R^2))$, we do not add one to the number of files transmitted, but only the expected number of files transmitted after the next congestion epochs. It depends on the throughput seen after congestion : by the congestion epoch the transmission rate of the tagged flow is $(\tau - r)/R^2$. There is probability p that the tagged flow suffers a packet loss which reduces the transmission rate to $(\tau - r)/(2R^2)$.

We can summarize the expected number of files that will be transmitted by the end of the current cycle given we are in this second case as

$$\int_t^\tau \beta e^{-\beta(r-t)} e^{-\mu \frac{(\tau-r)^2}{2R^2}} \left(pg(\tfrac{\tau-r}{2R^2}) + (1-p) g(\tfrac{\tau-r}{R^2}) \right) dr.$$

We conclude that $f(t)$ is given by :

$$\int_t^\tau \beta e^{-\beta(r-t)} \left\{ \int_0^{\frac{(\tau-r)^2}{2R^2}} \mu e^{-\mu y} (1 + f(r + R\sqrt{2y})) dy \right.$$
$$\left. + e^{-\mu \frac{(\tau-r)^2}{2R^2}} \left(pg(\frac{\tau - r}{2R^2}) + (1-p) g(\frac{\tau - r}{R^2}) \right) \right\} dr. \quad (3)$$

We now turn to $g(z)$. The residual number of packets to transmit from the current file Y has an exponential distribution. Again there are two cases. Either the current file can be transmitted before the next congestion epoch or it can't.

In the first case Y is exponentially distributed between 0 and $z\tau + \tau^2/(2R^2)$ since this is the maximum amount that can be transmitted in τ time units. After Y units are transmitted we add one to the total number of files transmitted in the current cycle plus a renewal term representing the expected number of files we will transmit in the remaining time of the current cycle. The time t when the first transmission was completed satisfies $y = tz + t^2/(2R^2)$; i.e. $t = R\sqrt{R^2z^2 + 2y} - R^2z$. At this time, the tagged flow is becoming idle. Consequently this first case contributes to the value of $g(z)$ by :

$$\int_0^{z\tau+\frac{\tau^2}{2R^2}} \mu e^{-\mu y}(1 + f(R\sqrt{R^2z^2 + 2y} - R^2z))dy.$$

The second case occurs if $Y > z\tau + (\tau^2/(2R^2))$, and thus with a probability equal to $\exp(-\mu(z\tau + \tau^2/(2R^2)))$. When the next congestion epoch begins the tagged flows has the remaining of the file (exponentially distributed) to transmit, and a transmission rate of $z + \tau/R^2$. With a probability p the tagged flow suffers a packet loss in this congestion epochs ; in this case, its transmission rate is reduced to $(z + \tau/R^2)/2$. We summarize this case as follows:

$$e^{-\mu(z\tau+\frac{\tau^2}{2R^2})}\left(pg(\frac{z+\frac{\tau}{R^2}}{2}) + (1-p)g(z+\frac{\tau}{R^2})\right).$$

As a conclusion, g(z) can be written :

$$\int_0^{z\tau+\frac{\tau^2}{2R^2}} \mu e^{-\mu y}(1 + f(R\sqrt{R^2z^2 + 2y} - R^2z))dy + e^{-\mu(z\tau+\frac{\tau^2}{2R^2})}\left(pg(\frac{z+\frac{\tau}{R^2}}{2}) + (1-p)g(z+\frac{\tau}{R^2})\right). \quad (4)$$

Equations (3) and (4) constitute an integral equation of the Fredholm type for the pair (f, g).

3.3.2 Expected "on" period during a cycle

Let $h(t)$ denote the expected cumulative time that the flow is active in the remaining time of the current cycle given that the tagged flow is inactive at the current time t with $0 \leq t < \tau$. Let $i(z)$ be the expected cumulative time that the flow is active in the remaining time of the current cycle given that the current time is immediately after a congestion epoch, and that the tagged flow is active with a current transmission rate of z. With the notation defined above, we have $\mathbb{E}_0^\sigma[\int_0^\sigma 1_{X(t)>0}dt] = h(0)$. Arguments similar to those given above lead to the following Fredholm equations for the (h, i) pair that are given in (5).

3.3.3 Expected duration of a cycle

Let $j(t)$ denote the expected residual time before the end of the current cycle given that the tagged flow is inactive at the current time t with $0 \leq t < \tau$. Let $k(z)$ be the expected residual time before the end of the current cycle given that the current time is immediately after a congestion epoch, and that the tagged flow is active with a current transmission rate of z.

We have $\mathbb{E}_0^\sigma[\sigma] = j(0)$ and again (j, k) can be seen as the solution of a Fredholm equation (see (6))

3.3.4 Expected jumps down due to completions of files during a cycle

Finally, let $l(t)$ denote the expected cumulative throughput reductions due to file completions from now to the end of the cycle given that the tagged flow is inactive at the current time t with $0 \leq t < \tau$.

And let $m(z)$ be the expected cumulative jumps down due to file completions from now to the end of the cycle given that the current time is immediately after a congestion epoch, and that the tagged flow is active with a current transmission rate of z.

We have $\mathbb{E}_0^\sigma[\int_0^\sigma X(t-)D(dt)] = l(0)$, where D, already introduced, denotes the point process of file completions.

As well as the other functions introduced, (l, m) are linked by a Fredholm Equation that is described on (7).

3.3.5 Expression of the three unknowns of the fixed point equation

For given τ, from the numerical solution of the set of Fredholm equations, one can efficiently determine

- $\mathbb{E}_0^\sigma[K_B] := f(0)$, the mean number of births during a cycle (which is also the mean number of file completions during a cycle;
- $\mathbb{E}_0^\sigma[\int_0^\sigma 1_{X(t)>0} dt] = h(0)$, the mean cumulative ON time over a cycle;
- $\mathbb{E}_0^\sigma[\sigma] = j(0)$, the mean duration of a cycle and
- $\mathbb{E}_0^\sigma[\int_0^\sigma X(t-)D(dt)] = l(0)$, the mean cumulative throughput reductions due to file completions over a cycle.

From this we can deduce the following representations for the 3 unknowns of (1):

$$\begin{aligned} \lambda_\delta &= \frac{\mathbb{E}_0^\sigma[K_B]}{\mathbb{E}_0^\sigma[\sigma]} = \frac{f(0)}{j(0)} \\ \mathbb{E}_0^\delta[X(0-)] &= \frac{\mathbb{E}_0^\sigma[\int_0^\sigma X(t-)D(dt)]}{\mathbb{E}_0^\sigma[K_B]} = \frac{l(0)}{f(0)} \\ \mathbb{P}(X(0) > 0) &= \frac{\mathbb{E}_0^\sigma[\int_0^\sigma 1_{X(t)>0} dt]}{\mathbb{E}_0^\sigma[\sigma]} = \frac{h(0)}{j(0)}. \end{aligned}$$

Notice that the product $\lambda_\delta \mathbb{E}_0^\delta[X(0-)]$ which is used in (1) is equal to $\frac{l(0)}{j(0)}$ so that the (f, g) pair is actually not required for solving this fixed point equation.

$$
\begin{aligned}
h(t) &= \int_t^\tau \beta e^{-\beta(r-t)} dr \left\{ \int_0^{\frac{(\tau-r)^2}{2R^2}} \mu e^{-\mu y} (R\sqrt{2y} + h(r + R\sqrt{2y}))\, dy \right. \\
&\quad \left. + e^{-\mu \frac{(\tau-r)^2}{2R^2}} \left(\tau - r + pi(\frac{\tau - r}{2R^2}) + (1-p) i(\frac{\tau - r}{R^2}) \right) \right\} \\
i(z) &= \int_0^{z\tau + \frac{\tau^2}{2R^2}} \mu e^{-\mu y} (R\sqrt{R^2 z^2 + 2y} - R^2 z + h(R\sqrt{R^2 z^2 + 2y} - R^2 z))\, dy \\
&\quad + e^{-\mu(z\tau + \frac{\tau^2}{2R^2})} \left(\tau + pi(\frac{z + \frac{\tau}{R^2}}{2}) + (1-p) i(z + \frac{\tau}{R^2}) \right). \qquad (5)
\end{aligned}
$$

$$
\begin{aligned}
j(t) &= (\tau - t) e^{-\beta(\tau - t)} + \int_t^\tau \beta e^{-\beta(r-t)} dr \left\{ \int_0^{\frac{(\tau-r)^2}{2R^2}} \mu e^{-\mu y} ((r-t) + R\sqrt{2y} + j(r + R\sqrt{2y}))\, dy \right. \\
&\quad \left. + e^{-\mu \frac{(\tau-r)^2}{2R^2}} \left(\tau - t + pk(\frac{\tau - r}{2R^2}) + (1-p) k(\frac{\tau - r}{R^2}) \right) \right\}. \\
k(z) &= \int_0^{z\tau + \frac{\tau^2}{2R^2}} \mu e^{-\mu y} (R\sqrt{R^2 z^2 + 2y} - R^2 z + j(R\sqrt{R^2 z^2 + 2y} - R^2 z))\, dy \\
&\quad + e^{-\mu(z\tau + \frac{\tau^2}{2R^2})} \left(\tau + pk(\frac{z + \frac{\tau}{R^2}}{2}) + (1-p) k(z + \frac{\tau}{R^2}) \right). \qquad (6)
\end{aligned}
$$

$$
\begin{aligned}
l(t) &= \int_t^\tau \beta e^{-\beta(r-t)} dr \left\{ \int_0^{\frac{(\tau-r)^2}{2R^2}} \mu e^{-\mu y} (\frac{\sqrt{2y}}{R} + l(r + R\sqrt{2y}))\, dy \right. \\
&\quad \left. + e^{-\mu \frac{(\tau-r)^2}{2R^2}} \left(pm(\frac{\tau - r}{2R^2}) + (1-p) m(\frac{\tau - r}{R^2}) \right) \right\}. \\
m(z) &= \int_0^{z\tau + \frac{\tau^2}{2R^2}} \mu e^{-\mu y} (z + \frac{R\sqrt{R^2 z^2 + 2y} - R^2 z}{R^2} + l(R\sqrt{R^2 z^2 + 2y} - R^2 z)) dy \\
&\quad + e^{-\mu(z\tau + \frac{\tau^2}{2R^2})} \left(pm(\frac{z + \frac{\tau}{R^2}}{2}) + (1-p) m(z + \frac{\tau}{R^2}) \right). \qquad (7)
\end{aligned}
$$

3.4 The Tahoe Case

Given τ, the rate of the tagged flow is again a regenerative process with the same cycle structure as in the Reno case, namely starting with a congestion period when the rate of the tagged flow is 0 and ending at the next congestion is again 0. Using the same notation

as in the Reno case, we now get

$$\begin{aligned} f(t) &= \int_t^\tau \beta e^{-\beta(r-t)} \\ &\cdot \left\{ \int_0^{\frac{(\tau-r)^2}{2R^2}} \mu e^{-\mu y}(1+f(r+R\sqrt{2y})dy \right. \\ &\left. + e^{-\mu\frac{(\tau-r)^2}{2R^2}} \left(pg(0)+(1-p)g(\frac{(\tau-r)}{(R^2)}) \right) \right\} dr \end{aligned} \tag{8}$$

and

$$\begin{aligned} g(z) = & \int_0^{z\tau+\frac{\tau^2}{2R^2}} \mu e^{-\mu y}(1+f(R\sqrt{R^2z^2+2y}-R^2z))dy \\ & + e^{-\mu(z\tau+\frac{\tau^2}{2R^2})} \left(pg(0)+(1-p)g(z+\frac{\tau}{R^2}) \right). \end{aligned} \tag{9}$$

The other equations can easily be derived by similar arguments and are omitted.

3.5 Numerical Evaluation of the Fixed Point

In this section we present the method that we developed to numerically study the fixed point equation satisfied by τ. The main result is a common linear equation describing the integral equations for the pairs $(f,g),(h,i),(j,k),(l,m)$.

Each of the pairs of functions $(f,g),(h,i),(j,k),(l,m)$ satisfies a Fredholm equation of the second type where all equations share some common terms. It is shown in Appendix 10.5 that the general form of these equations is as follows: we look for a functions A, defined on $[0;\tau]$ and a function B defined on $[0;+\infty[$ such that they verify Equation (10) where $\kappa=\mu/(R^2)$ and where the functions U and V are given in the following table for all 4 cases:

$$\begin{aligned} A(t) &= \int_t^\tau \beta e^{-\beta(r-t)} \left(U(r) + \int_r^\tau \kappa(s-r)e^{-\kappa\frac{(s-r)^2}{2}} A(s)ds \right. \\ &\left. + e^{-\kappa\frac{(\tau-r)^2}{2}} \left(pB(\frac{\tau-r}{2}) + (1-p)B(\tau-r) \right) \right) dr \ . \\ B(r) &= V(r) + \int_0^\tau \kappa(r+s)e^{-\kappa\frac{s^2+2sr}{2}} A(s)ds \\ &+ e^{-\kappa\frac{\tau^2+2\tau r}{2}} \left(pB(\frac{\tau+r}{2}) + (1-p)B(\tau+r) \right) \ . \end{aligned} \tag{10}$$

$A(t)$	$B(r)$	$U(r)$	$V(r)$
$f(t)$	$g\left(\frac{r}{R^2}\right)-1$	1	0
$h(t)$	$i\left(\frac{r}{R^2}\right)$	$a_\tau(r)$	$b_\tau(r)$
$j(t)$	$k\left(\frac{r}{R^2}\right)$	$a_\tau(r)$	$\frac{1}{\beta}+b_\tau(r)$
$l(t)$	$m\left(\frac{r}{R^2}\right)-\frac{r}{R^2}$	$\frac{a_\tau(r)+\frac{p}{2}c_\tau(r)}{R^2}$	$\frac{b_\tau(r)+\frac{p}{2}d_\tau(r)}{R^2}$

with the functions $a_\tau, b_\tau, c_\tau, d_\tau$ defined as:

$$a_\tau(r) = \int_r^\tau e^{-\kappa\frac{(s-r)^2}{2}} ds \; ; \; b_\tau(r) = \int_0^\tau e^{-\kappa\frac{s^2+2sr}{2}} ds \; ;$$

$$c_\tau(r) = (\tau - r)e^{-\kappa\frac{(\tau-r)^2}{2}} \; ; \; d_\tau(r) = (r+\tau)e^{-\kappa\frac{\tau^2+2\tau r}{2}}.$$

Let $(\Gamma(t), \tilde{\Gamma}(r))$ be the solution (A, B) of Equation (10) for $(U, V) = (1, 0)$, let $(\Theta(t), \tilde{\Theta}(r))$ denote the solution for $(U, V) = (a_\tau, b_\tau)$, and let $(\Delta(t), \tilde{\Delta}(r))$ be the solution of this equation for $(U, V) = (c_\tau, d_\tau)$. According to the last table, we have :

$$\Gamma(t) = f(t) \; ; \; \Theta(t) = h(t) \text{ and, as Equation (10) is linear,}$$

$$\frac{1}{\beta}\Gamma(t) + \Theta(t) = j(t) \text{ and } \frac{\Theta(t) - \frac{p}{2}\Delta(t)}{R^2} = l(t).$$

We numerically solve Equation (10) in the following way. First, we set $B(r) = 0$ for $x > K\tau$. This is motivated by the fact that for physical reasons $B(r)$ has to decrease as r increases, a fact that can be proved mathematically, but we omit the proof here. Second, we discretize the functions $A(t)$ and $B(r)$ uniformly with a density of M samples per interval of length τ. So, the function $A(t)$ is approximated by a vector of M samples and $B(r)$ by a vector of KM samples. We stack both vectors and hence obtain a vector of dimension $(K + 1)M$. Approximating the integrals in (10) by weighted sums of the samples of the functions, Equation (10) reduces to a matrix equation. Solving this matrix equation involves the inversion of a $(K + 1)M \times (K + 1)M$ matrix.

The numerical error introduced in this procedure can be controlled by the choice of the parameters K and M. In each of the examples shown in this paper we made sure that K and M were large enough for the numerical errors to be small enough. With the parameters (β and κ) used in this paper and for ranges of τ-values considered in this paper K=8 and M=40 turned out to give accurate enough values for the fixed points. A detailed study of how the numerical error decreases as the parameters K and M increase is beyond the scope of this paper.

3.6 Determination of τ

As shown above, τ satisfies the following equation :

$$\frac{pC}{2\tau} + \frac{l(0)}{j(0)} = \frac{1}{R^2}\frac{h(0)}{j(0)} \tag{11}$$

or equivalently

$$\frac{pC}{2\tau} + \frac{\frac{1}{R^2}(\Theta(0) - \frac{p}{2}\Delta(0))}{\frac{1}{\beta}\Gamma(0) + \Theta(0)} = \frac{1}{R^2}\frac{\Theta(0)}{\frac{1}{\beta}\Gamma(0) + \Theta(0)}.$$

The most convenient form of this fixed point equation is the following one:

$$C = \left(\frac{\Delta(0)}{\frac{1}{\beta}\Gamma(0) + \Theta(0)} \right) \frac{\tau}{R^2}. \tag{12}$$

This form is valid both for the Reno and the Tahoe cases, for appropriate definitions of Θ and Γ. In Figure 2, we have computed the right-hand side of Equation (12), which does not depend on C, as a function of τ for a fixed setting of the parameters $1/\beta = 2s, 1/\mu = 2000$ Pkts, $R = 100ms$, $p = 0.8$ On this plot we can see that if the link capacity is large enough

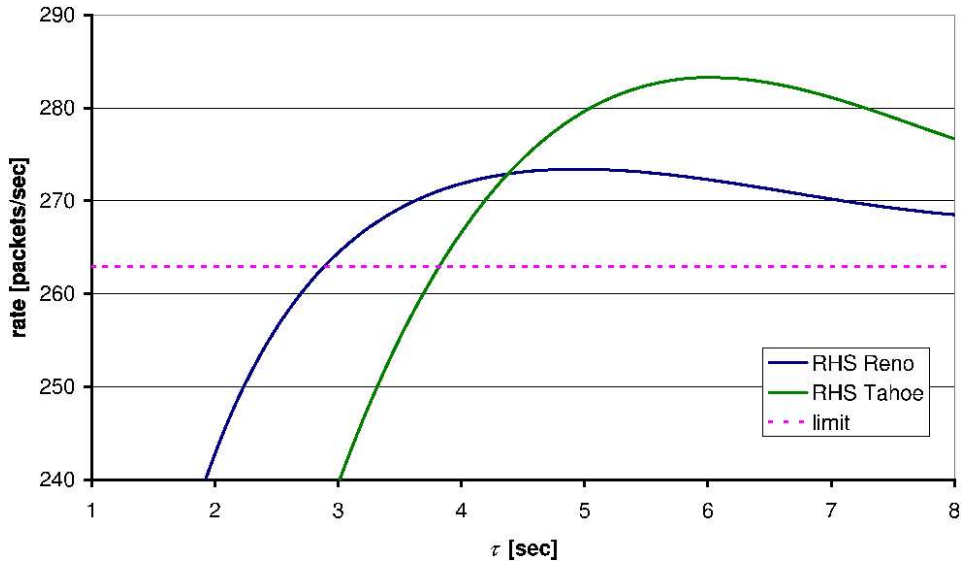

Figure 2: The RHS of (12) as a function of τ; the fixed points are the intersections of this RHS with the horizontal line C, in the Reno and the Tahoe cases.

there is no value of τ making this function vanish (here for $C = 290$ Pkts/s.). In this case, the only possible stable regime is congestion-less. For smaller values of the capacity, we observe either two fixed points (e.g. for C=270 Pkts/s.) or one (e.g. for C=250 Pkts/s.). In the case with two solutions, we have several candidates for a stable regime, with different periods. In the next section we will present a method helping to distinguish between solutions that may be the inter-congestion time of a stable regime and other solutions. From Figure 2 we can conclude more:

- for all C-values above 273.4 Pkts/s. (283.3 in the Tahoe case), there are no intersections;

- for $263 < C < 273.5$ Pkts/s. ($263 < C < 283.3$ in the Tahoe case), there are two intersections and
- for $C < 263$ Pkts/s., there is only one intersection.

4 A Necessary and Sufficient Condition for the Existence of a Congestion Periodic Regime with a Given Period

We start with a detailed study of the interaction-less regime (this is the free regime, i.e. the regime when $C = \infty$), which will be an essential ingredient of the analysis of the congestion regime which may occur when $C < \infty$ as we shall see in §4.2 below.

4.1 The Free Regime

4.1.1 The free regime regenerative rate process

In the case without congestion, each flow increases its transmission rate linearly at rate $1/R^2$ and can transmit a file of size y packets in time t where $y = t^2/(2R^2)$; i.e. in time $t = R\sqrt{2y}$. The density of the transmission time of a file is

$$\mu \frac{t}{R^2} e^{-\frac{\mu t^2}{2R^2}}$$

(as easily seen by the change of variable $t \to v = t^2/2R^2$) and the mean file transmission time is therefore

$$T_{\texttt{ON}} = R \int_0^\infty \mu \exp(-\mu y) \sqrt{2y} dy = R\sqrt{\frac{\pi}{2\mu}}. \tag{13}$$

A tagged flow alternates between periods composed of a silence period of exponential duration with parameter β and a active period of mean duration $T_{\texttt{ON}}$, distributed according to the above density.

The rate $X(t)$ of the tagged flow at time t is a regenerative process that stays equal to 0 during OFF periods and increases linearly with time during activity periods. This stochastic process regenerates after the completion of one OFF and one ON period. The point process of regeneration epochs of a tagged flow will be denoted by S.

During each ON period a flow transmits on average $1/\mu$ packets. Consequently the average transmission rate per flow is

$$\rho = (1/\mu)/(1/\beta + T_{\texttt{ON}})). \tag{14}$$

The proportion ν of flows which are idle is $(1/\beta)/(1/\beta + T_{\texttt{ON}}))$. Notice that the transmission rate equals $\nu\beta/\mu$. This is intuitively obvious since $\nu\beta$ is the rate at which new flows come on-line and each new flow must transmit on average $1/\mu$ packets.

Hence when the regime without congestion occurs, the average transmission rate per flow ρ is less than C; i.e. $\nu\beta/\mu < C$ and

$$\rho = \frac{\nu\beta}{\mu} = \left(\mu\left(1/\beta + R\sqrt{\frac{\pi}{2\mu}}\right)\right)^{-1} < C. \tag{15}$$

4.1.2 The free regime PDE

Let $\nu(t)$ be the proportion of idle flows at time t. Let $s(z,t)$ be the density of the transmission rates of active flows in the mean-field regime (we consider first the case with a density for the sake of clear exposition). Consequently,

$$\int_0^\infty s(z,t)dz = 1 - \nu(t). \tag{16}$$

From the partial differential evolution equation introduced by Baccelli et al. in [7] we can see that the density function verifies the PDE:

$$\frac{\partial s}{\partial t}(z,t) + \frac{1}{R^2}\frac{\partial s}{\partial z}(z,t) = -\mu z s(z,t). \tag{17}$$

Multiplied by dz, the second term on the left hand side represents the rate of change of the proportion of transmission rates in $[z, z+dz]$ due to the linear increase in the transmission rate. The right hand side represents the rate at which files complete transmission since $s(z,t)dz$ is the proportion of flows with transmission rates in the interval $[z, z+dz]$ and flows with transmission rates in this interval complete transmission at a rate μz.

The rate at which flows become active is $\beta\nu(t)$ hence in time dt the area $\beta\nu(t)dt$ is added under the graph of $s(z,t)$ between 0 and dt/R^2 because this area is cleared out by the additive increase in the transmission rates. The area under the graph of $s(z,t)$ between 0 and dt/R^2 is $s(0,t)dt/R^2$ to first order. Hence,

$$s(0,t)/R^2 = \beta\nu(t). \tag{18}$$

4.1.3 The Fredholm equation for solving the free regime PDE

We define the Laplace transforms

$$\widehat{s}_z(u) = \int_0^\infty e^{-ut} s(z,t)dt, \tag{19}$$

$$\widehat{\nu}(u) = \int_0^\infty e^{-ut} \nu(t)dt. \tag{20}$$

Taking the Laplace of (17) w.r.t. t, we get

$$u\widehat{s}_z(u) - s(z,0) + \frac{1}{R^2}\frac{\partial}{\partial z}\widehat{s}_z(u) = -\mu z \widehat{s}_z(u)$$

or equivalently

$$\frac{\partial}{\partial z}\widehat{s}_z(u) = -R^2(u+\mu z)\widehat{s}_z(u) + R^2 s(z,0).$$

The solution of this ordinary differential equation is

$$\begin{aligned}\widehat{s}_z(u) &= e^{-R^2(uz+\mu z^2/2)}\Big(\widehat{s}_0(u) + R^2\int_0^z e^{R^2(ux+\mu x^2/2)}s(x,0)dx\Big)\\ &= R^2\int_0^z e^{-R^2u(z-x)}e^{-R^2\mu(\frac{z^2}{2}-\frac{x^2}{2})}s(x,0)dx + R^2e^{-R^2(uz+\mu z^2/2)}\beta\widehat{\nu}(u), \quad (21)\end{aligned}$$

where we used the fact that $\widehat{s}_0(u) = \beta R^2\widehat{\nu}(u)$, which follows from (18).

We now remark that

$$R^2\left(\frac{z^2}{2}-\frac{x^2}{2}\right) = zR^2(z-x) - \frac{(R^2(z-x))^2}{2R^2},$$

and we introduce $t = R^2(z-x)$, so that $x = z - \frac{t}{R^2}$. The previous equation becomes :

$$\begin{aligned}\widehat{s}_z(u) &= R^2\int_0^z e^{-R^2u(z-x)}e^{-\mu(zR^2(z-x)-\frac{(R^2(z-x))^2}{2R^2})}s(x,0)dx + R^2e^{-R^2(uz+\mu z^2/2)}\beta\widehat{\nu}(u),\\ &= \int_0^\infty e^{-ut}e^{-\mu(zt-\frac{t^2}{2R^2})}s(z-\frac{t}{R^2},0)dt + R^2\beta e^{-R^2\mu\frac{z^2}{2}}\widehat{\nu}(u)e^{-uzR^2}. \quad (22)\end{aligned}$$

By immediate Laplace inversion, we can then write :

$$s(z,t) = s(z-\frac{t}{R^2},0)\; e^{-\mu\left(tz-\frac{t^2}{2R^2}\right)} + R^2\beta\; e^{-\mu R^2\frac{z^2}{2}}\nu(t-zR^2). \quad (23)$$

Using now (16), one finally gets the following Fredholm equation for $s(z,t)$:

$$s(z,t) = s(z-\frac{t}{R^2},0)\; e^{-\mu\left(tz-\frac{t^2}{2R^2}\right)} + e^{-\mu R^2\frac{z^2}{2}}R^2\beta\left(1-\int_0^\infty s(x,t-zR^2)dx\right) \quad (24)$$

which turns out to be quite handy for numerical exploitation as we shall see below.

Equation (24) is easy to interpret when considering the two cases: for the rate to be z at time t, either the transfer of the file transmitted at time 0 is not yet completed at time t, which requires that the rate was $z - t/R^2 \geq 0$ at time 0, or it is completed, which requires that the flow was inactive at time $t - zR^2 > 0$ and there was a transition from inactive to active at that time. In fact it is clear that (24) can be generalized to describe the evaluation of a measure $S(dz,t)$ representing the distribution of transmission rates at time t starting from an arbitrary measure $S(dz,0)$:

$$S(dz,t) = R^2\beta\left(1-\int_{x=0}^\infty S(dx,t-zR^2)\right)\; e^{-\mu R^2\frac{z^2}{2}}dz + S(dz-\frac{t}{R^2},0)\; e^{-\mu(zt-\frac{t^2}{2R^2})} \quad (25)$$

4.1.4 Further properties of the solution of the PDE

Let

$$\alpha(t) = \int_0^\infty zs(z,t)dz. \tag{26}$$

The function $\alpha(t)$ represents the aggregate rate (sum of the transmission rates at time t where the sum is over all flows). Let

$$\widehat{\alpha}(u) \quad = \quad \int_0^\infty e^{-ut}\alpha(t)dt. \tag{27}$$

The two following lemmas are proved in Appendix 10.

Lemma 1 The solution of the free regime PDE is such that

$$\widehat{\alpha}(u) \quad = \quad \frac{\nu(0)\frac{\beta}{\beta+u}\widehat{I}(u) + \widehat{J}(u)}{1-\mu\frac{\beta}{\beta+u}\widehat{I}(u)}, \tag{28}$$

where

$$\widehat{I}(u) = R^2\int_0^\infty xe^{-R^2ux-R^2\mu x^2/2}dx \tag{29}$$

and $\widehat{J}(u)$ is given by

$$R^2\int_{z=0}^\infty e^{R^2uz+\frac{R^2\mu z^2}{2}}s(z,0)\int_{x=z}^\infty xe^{-R^2ux-\frac{R^2\mu x^2}{2}}dxdz. \tag{30}$$

The limiting behavior of the solution of the free regime PDE is given by the following expressions:

Lemma 2 The stationary distribution of the rates is:

$$\nu(\infty) \quad = \quad \frac{\frac{1}{\beta}}{\frac{1}{\beta}+R\sqrt{\frac{\pi}{2\mu}}} \tag{31}$$

$$s(z,\infty) \quad = \quad \frac{R^2e^{-R^2\mu z^2/2}}{\frac{1}{\beta}+R\sqrt{\frac{\pi}{2\mu}}}. \tag{32}$$

The stationary aggregate rate is:

$$\alpha(\infty) \quad = \quad \frac{1}{\mu}\frac{1}{\frac{1}{\beta}+R\sqrt{\frac{\pi}{2\mu}}} = \rho. \tag{33}$$

As we shall see in Figure 9, for certain initial conditions, the aggregate rate function $\alpha(t)$ may have a "bump", namely a maximal value that is significantly larger that ρ (see also Figure 19 for the same phenomenon under other statistical assumptions).

In Appendix 10, we also give an interpretation of the transforms of Lemma 1 in terms of renewal theory (§10.3) and a closed form expression for the solution of the PDE in the time domain (§10.4).

4.2 The Interaction Regime(s)

4.2.1 The invariant measure equation

Assume there exists a periodic regime of period τ. Then τ should be a solution of (1). In addition the couple $(\nu_0, S_0(dz))$ that gives the proportion of OFF sources and the distribution of rates just after congestion epochs should be invariant w.r.t. the shift that moves from a congestion epoch to the next.

First τ and $(\nu_0, S_0(dz))$ should be such that the aggregate rate function α_0 obtained when taking $S(dz, 0) = S_0(dz)$ is such that $\alpha_0(\tau) = C$ and $\alpha_0(t) < C$ for all $0 < t < \tau$.

In addition, given that at congestion epochs, a proportion p of the windows are halved, the $(\nu_0, S_0(dz))$ should satisfy the integral equation (which will be referred to as the invariant measure equation)

$$S_0(dz) = (1-p)S(dz, \tau) + pS(d2z, \tau), \tag{34}$$

where $S(dz, t)$ is the solution of (25) with the initial condition $S(dz, 0)$ taken equal to $S_0(dz)$.

When using the integral representation of $\alpha_0(.)$ given in (50) one gets that the last integral equation for $S_0(.)$ can also be seen as a Fredholm type integral equation of the second kind.

In the Tahoe case the transmission rates of active sources has a measure which must have a point mass at zero at congestion epochs; the invariant measure equation then reads

$$S_0(dz) = (1-p)S(dz, \tau) + p\delta_0(dz) \int_0^\infty S(dv, \tau). \tag{35}$$

A few remarks are in order before addressing numerical issues:

- The existence of a couple $(\nu_0, S_0(dz)))$ solution of (34) and such that the $\alpha_0(\tau) = C$ and $\alpha_0(t) < C$ for all $t < \tau$ is necessary and sufficient for the existence of a congestion periodic regime of period τ. Using this, it is easy for instance to check that in the region where the RCP equation has two fixed points, the rightmost fixed point is spurious. This immediately follows from the fact that the condition $\alpha_0(t) < C$ for all $t < \tau$ is not satisfied for this other fixed point (see Figure 4).

- The more general problem of finding all possible periodic regimes can be stated as follows: find all pairs made of a real number $0 < \tau < \infty$ and of a couple $(\nu_0, S_0(dz)))$ such that (34) (or (35) in the Tahoe case) holds and such that $\alpha_0(\tau) = C$ and $\alpha_0(t) < C$ for all $t < \tau$.

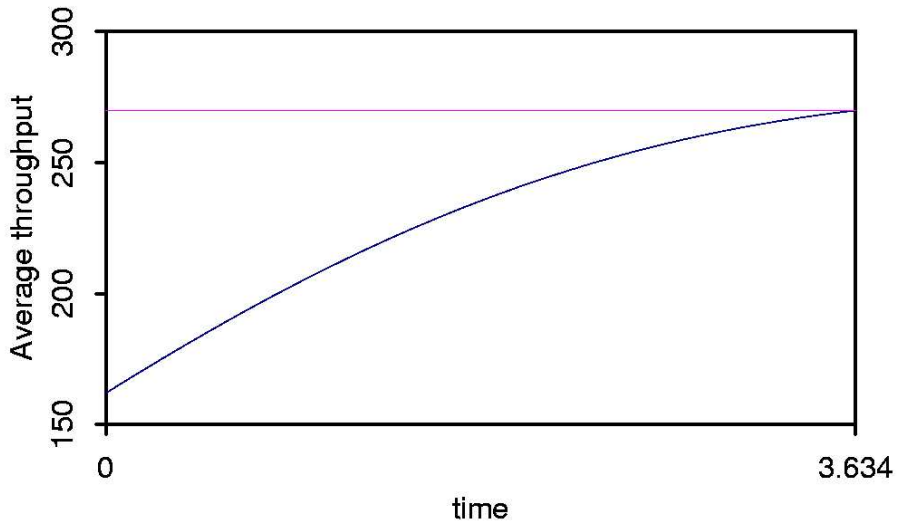

Figure 3: Evolution of the aggregate rate with time: non spurious case.

- Of course, other stationary regimes are possible like e.g. periodic regimes where the aggregate rate has a period that consists of $k > 1$ congestions, or even non periodic regimes (although we did not find such regimes by simulation).
- Injecting the couple $(\nu_0, S_0(dz)))$ as an initial condition into Equation (24) determines the proportion of active flows and the throughput distribution of active flows $S(dz, t)$ for all $0 \leq t < \tau$. The mean stationary throughput obtained from this function averaged over continuous time is given by the following cycle mean:

$$\frac{1}{\tau} \int_{t=0}^{\tau} \int_{z=0}^{\infty} zS(dz, t)dt. \tag{36}$$

4.2.2 Numerical solution

We have chosen a numerical procedure to find an approximation for $s(z, t)$ based on Equations (24) and (34). The alternative that consists in using the explicit expressions obtained

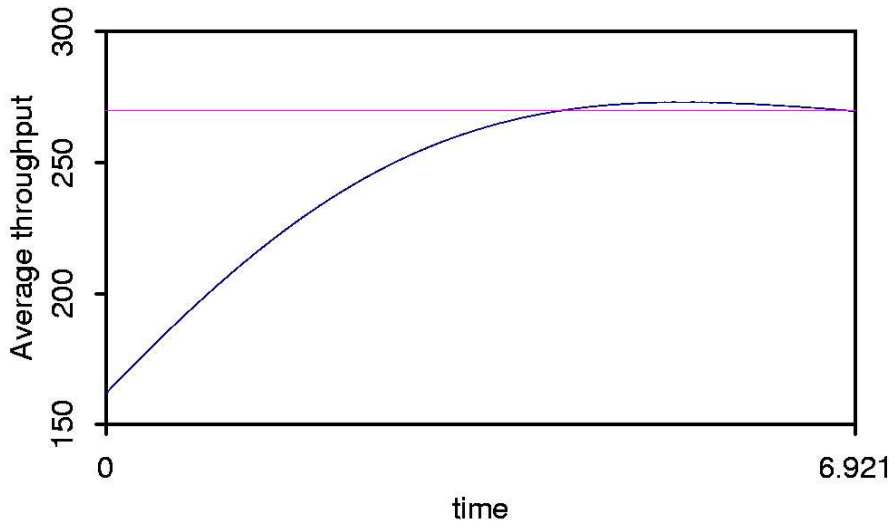

Figure 4: Evolution of the aggregate rate with time: case with a spurious solution to the RCP Equation

in the appendix for the $s(z,t)$ (Equation (48) combined with the expression for α given by (50)), would involve the numerical inversion of certain Laplace transforms (like $\hat{\xi}$ and $\hat{\eta}$ defined there) and this is not practical. Therefore we opted for the following method. We discretize the function $s(z,t)$ with $L+1$ samples over its time domain (an interval of length τ) and with a density of L samples per interval of length $\frac{\tau}{R^2}$ over its space domain (i.e. the z variable). We use $L+1$ samples in the time domain as there is a crucial difference between the time instant just before a congestion epoch (the L-th sample) and the time instant just after (the 0-th sample). We truncate the $s(z,t)$ function in the z direction by putting $s(z,t) = 0$ for $z > K\frac{\tau}{R^2}$. This truncation is motivated by the solution of the interaction-less system where this function decays like the tail of a Gaussian distribution.

The discretized version of Equations (24) and (34) define a matrix equation. Notice that in this case (in contrast to the case of solving for $A(t)$ and $B(r)$ in §3.5) there are L^2K unknowns and the matrices involved may become very large. Therefore, we used Equation

(24) and (34) as a recursive rule to calculate an approximation for $s(z,t)$. The larger L and K are chosen the better the approximation (but more computations are needed). For the examples considered in this paper K=5 and L=200 turned out to be adequate values.

4.2.3 The multiple stationary regime region

In this section, we give both numerical and simulation evidence showing that the condition that the load factor

$$\rho = (1/\mu)/(1/\beta + T_{\texttt{ON}}))$$

is less than C (namely the capacity per user is more than the mean load per user) is not sufficient for having an interaction-less mean-field regime for all initial conditions. The numerical part is based on the solution of the set of Fredholm equations of the last subsections. The simulation is based on the N2N code [3], a discrete event simulator which computes the AIMD sharing for a finite number of ON/OFF flows, interacting through the sum of their rates, as described in Section 3.1.

We also show that there exist values of the parameters such that depending on the initial condition describing the rates of the various flows, one may enter either into an interaction-less stationary regime or into a stationary congestion regime.

In the case considered here $1/\mu = 2000$ Pkts, $1/\beta = 2$ s., $p = 0.8$ and $R = 0.1$ s. The load factor ρ is then around 263 Pkts/s. We take $C = 270$ Pkts/s.

- When the initial condition is chosen according to the stationary law given in (31)–(32), then $\alpha(t) = \rho$ for all t and no congestions occur at all since $\rho < C$.
- As already shown in Section 3.6, the rate conservation principle gives two values of τ solution of the fixed point equation (1), the smallest of which is $\tau \sim 3.7s$. Using the solution of the invariant measure equation of Section 4.2.1, we find that for this value of τ, there exists a probability measure satisfying the integral equation (34) and satisfying the key condition that the associated α function first reaches C at time τ (see Figure 3). The p.d.f of this distribution as obtained numerically is depicted in Figure 5 for Reno.

  The existence of such a regime is confirmed by the N2N simulation of 1 Million HTTP users with the above characteristics and sharing a link of capacity 270 Pkts/sec (see Figure 7). Moreover, the steady state distributions found by simulation match quite precisely those obtained numerically.

In other words, depending on the initial phases of the flows, one either enters into a congestion-less regime or into a periodic regime with infinitely many congestions. The first case occurs when the initial conditions are chosen independently for all flows, and each flow is in the stationary regime it would reach if there were no interaction at all. The second case occurs if the flows are more in phase: here all start inactive at time 0.

Here are a few remarks of interest:

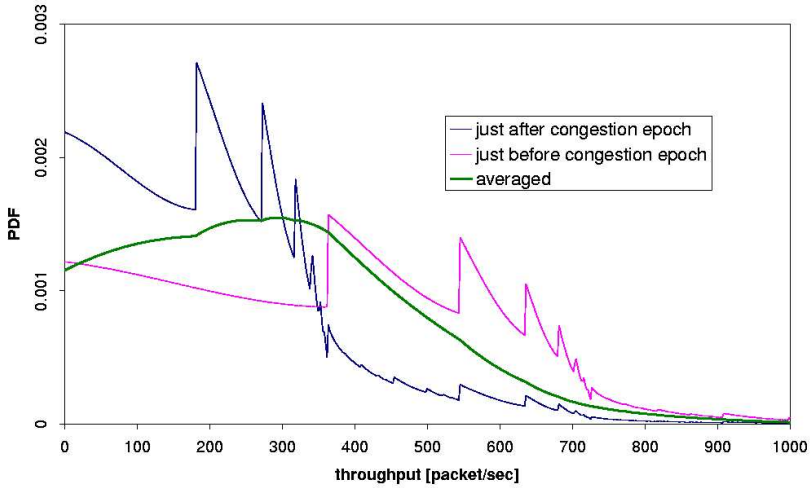

Figure 5: Numerical solution of the invariant measure equation. Distributions obtained for Reno. $1/\mu = 2000$ Pkts, $1/\beta = 2$ s., $p = 0.8$ and $R = 0.1$ s. and $C = 270$ Pkts/s. In red, steady state probability distribution function of the rate just after a congestion epoch; in green, continuous time stationary rate distribution.

- The same period and periodic regime are reached when the initial condition is that with all flows initially active and with null rate;
- The largest value of C for which we observe these two possible stationary regimes is approximately 273.5 Pkts/sec as shown independently by the N2N simulator and the fixed point method;
- the second solution of the RCP happens to be spurious. There exists a probability solution of (34) but as easily seen on Figure 4, the associated α function crosses the C level before this value of τ.
- Similar results hold for Tahoe. The associated distributions are plotted in Figure 8.

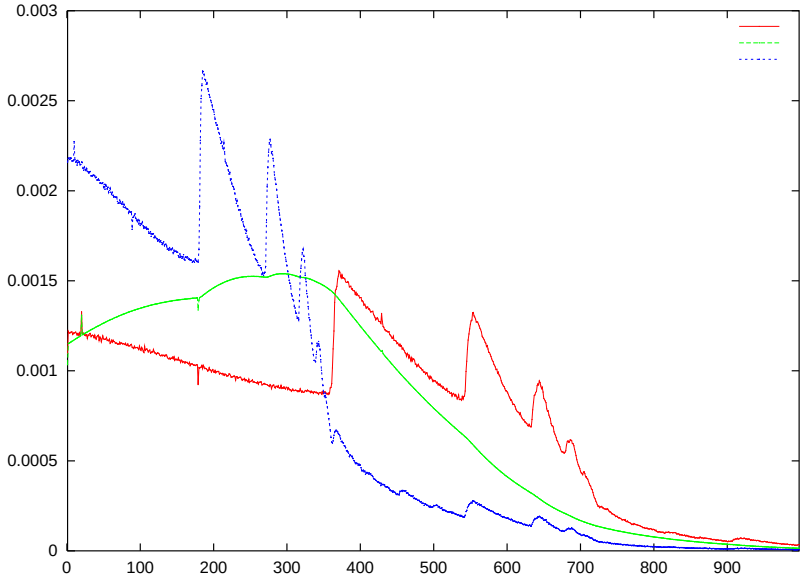

Figure 6: N2N simulation of 1 Million HTTP flows. Distributions obtained for Reno in the same case as in Figure 5.

4.2.4 Dependence of bi-stability region w.r.t. the parameters

Let C_T be the maximum C for which there is an interaction regime, ρ be given as in (14) and define the over-provisioning ratio (for guaranteeing the absence of interaction) to be $\omega = C_T/\rho$. Here are a few data on this ratio in the exponential case with $p = 0.8$ and $\frac{1}{\mu} = 2000$ Pkts.

- $1/\beta = 2$ s., $R = 0.1$ s.: $\omega = 1.04$;
- $1/\beta = 4$ s., $R = 0.1$ s.: $\omega = 1.06$;
- $1/\beta = 8$ s., $R = 0.1$ s.: $\omega = 1.09$;
- $1/\beta = 2$ s., $R = 0.05$ s.: $\omega = 1.06$;
- $1/\beta = 8$ s., $R = 0.05$ s.: $\omega = 1.12$;
- $1/\beta = 2$ s., $R = 0.025$ s.: $\omega = 1.09$;
- $1/\beta = 8$ s., $R = 0.025$ s.: $\omega = 1.15$.

The region is larger for small RTTs and for short think times.

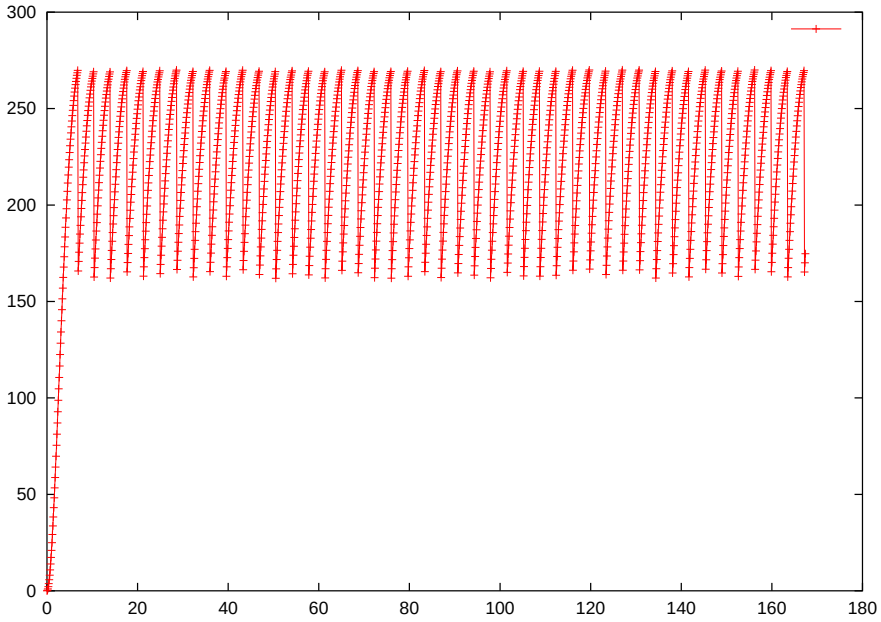

Figure 7: Evolution of aggregate rate when all flows are initially active and with null rate for $C = 270$ Pkts/s.

4.2.5 Proof of the existence of congestion regimes with load less than capacity

Let us consider the Tahoe case with an initial condition consisting of all sources active and with 0 rate. The functions $\alpha(t)$ (the aggregate rate defined in (26)) and $\gamma(t) = 1 - \nu(t)$ (the proportion of active flows) associated with this initial condition play a key role in the construction of this section. They are depicted in Figure 9 in the case $1/\mu = 2000$, $1/\beta = 2$ and $R = 0.1$.

Let

- M denote the maximum of $\alpha(t)$ over all $t > 0$;
- θ denote argmax of $\alpha(t)$;
- m denote the minimum of $\alpha(t)$ over all $t > \tau$;
- γ denote the minimum of $\gamma(t)$ over all $t > 0$.

In the particular case of Figure 9, we have $M = 301.8$, $\theta = 5.5$, $m = 258.1$ and $\gamma = 0.723$.

Let

$$\tilde{C} = p\gamma M + (1 - p\gamma)m. \tag{37}$$

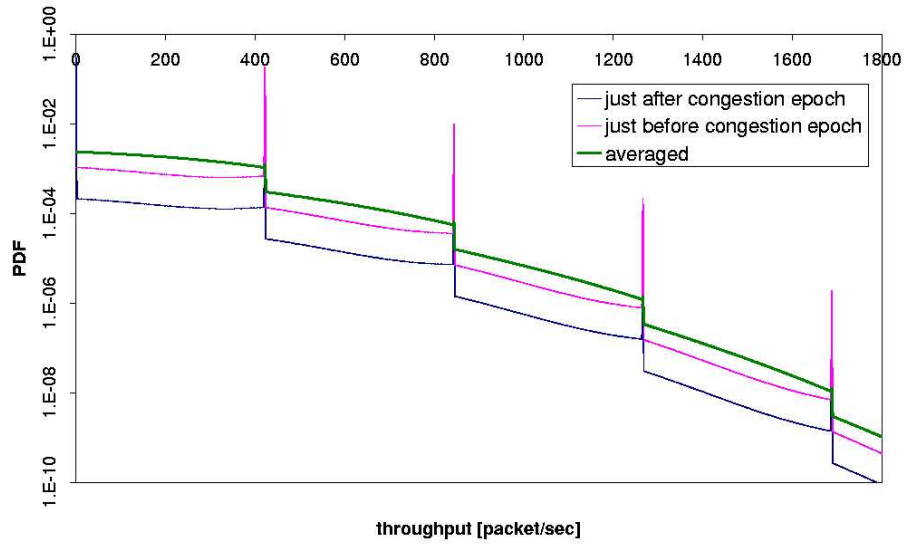

Figure 8: Example of distributions obtained numerically for Tahoe in log scale. C=270 Pkts/s., R=0.1 s., $1/\mu$=2000 Pkts, $1/\beta$=2 s., p=0.8. The RCP fixed point for Tahoe is τ=4.222s. Notice the Dirac-impulses at multiples of $\frac{\tau}{R^2}$=422.2 Pkts/s. They stem from sources that experience a loss at some congestion epoch and are put to a rate of 0 there. After one congestion epoch such a source has a rate of $\frac{\tau}{R^2}$.

Lemma 3 For the above initial condition, if $\tilde{C} > \rho$, then the Tahoe version of the model experiences an infinite number of congestion epochs for all C in the interval $\rho \leq C \leq \tilde{C}$.

Proof Assume that $k \geq 0$ congestion epochs (we include time 0 in the set of congestion epochs) took place at $T_0 = 0, T_1, \ldots, T_k$. At time T_k, flows can be partitioned into classes according to the index of their last congestion time (this is a partition of flows because we consider a flow that never experienced congestion on the $(0, T_k]$ interval as being in class 0). Let p_i^k be the proportion of flows that are of class i at time T_k.

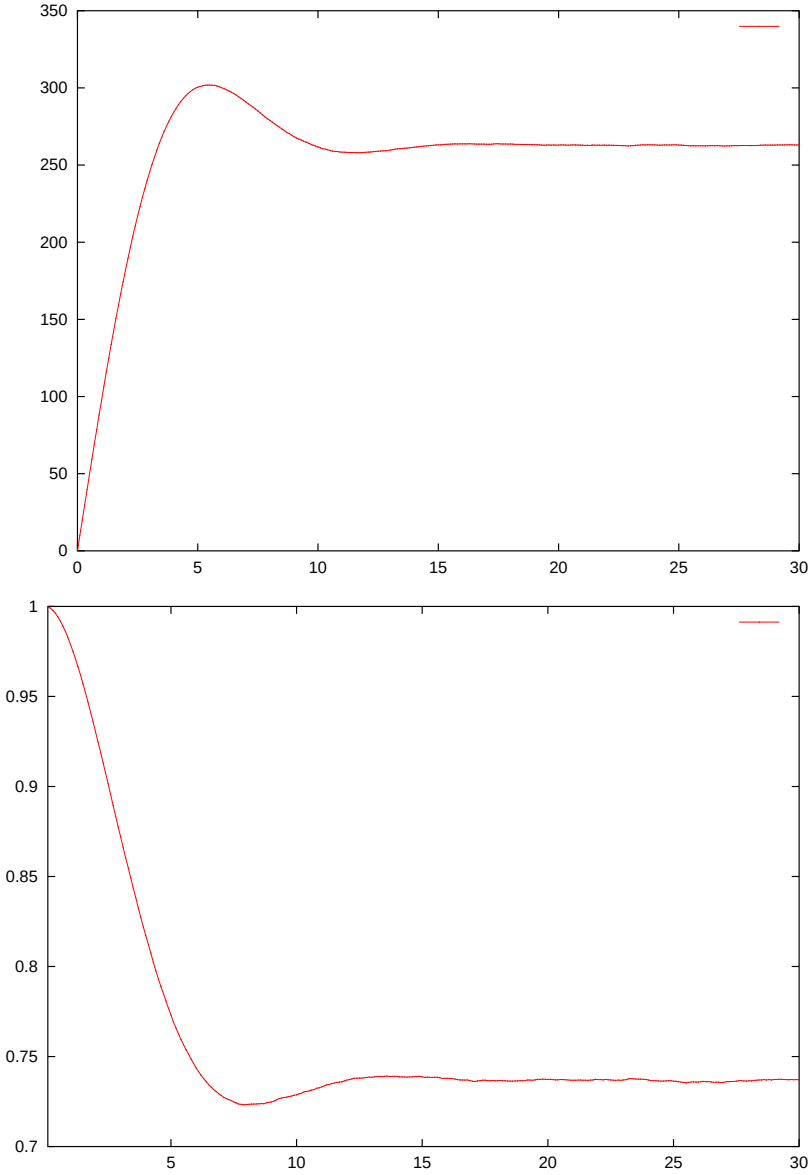

Figure 9: The α function with its bump (top) and the γ function (bottom).

Assume that there is no congestion after time T_k. Then for all $t \geq 0$, the aggregate rate of the Tahoe model at time $T_k + t$ is given by the following expression:

$$a(t) = p_k^k \alpha(t) + p_{k-1}^k \alpha(T_k - T_{k-1} + t) + \ldots + p_0^k \alpha(T_k - T_0 + t).$$

Similarly, if $k > 0$, the proportion of active flows at time T_k^- is

$$g(T_k) = p_{k-1}^{k-1} \gamma(T_k - T_{k-1}) + \ldots + p_0^{k-1} \gamma(T_k - T_0),$$

so that

$$g(T_k) \geq p_{k-1}^{k-1} \gamma + \ldots + p_0^{k-1} \gamma = \gamma.$$

This in turn implies that $p_k^k \geq g(T_k) p \geq p\gamma$ for all $k \geq 0$.

When evaluating a at θ, we get:

$$\begin{aligned} a(\theta) &= p_k^k \alpha(\theta) + p_{k-1}^k \alpha(T_k - T_{k-1} + \theta) + \ldots \\ &\qquad + p_0^k \alpha(T_k - T_0 + \theta) \\ &\geq p_k^k M + (p_{k-1}^k + \ldots + p_0^k) m \\ &= p_k^k M + (1 - p_k^k) m \\ &\geq p\gamma M + (1 - p\gamma) m = \tilde{C} > C \end{aligned}$$

which contradicts the fact that the continuous function $a(t)$ remains smaller than C for all t. Hence, there is at least one more congestion period. □

So in our example, when $p = 0.8$, we are sure that Tahoe exhibits infinitely many congestions as soon as $C \leq \tilde{C} = 283.38$. Notice that this is only a sufficient condition for congestion, namely $\tilde{C} > C_T$ in general.

From our numerical and simulation estimates, it seems that the bi-stability region for Tahoe is larger than for Reno (see Figure 2).

Of course, under the assumption of the last lemma, if the initial condition for the flows is that of the steady state of the interaction-less regime, then one remains in this regime forever.

We have no analogue of Lemma 3 in the Reno case at this stage. The fact that Reno could have a turbulent regime when the load per user is less than the capacity per user is hence only backed by simulation and numerical evidence at this stage.

4.3 Properties of the Stationary Rate

We first study the stationary continuous time mean throughput obtained by one flow. Figure 10 plots this mean throughput in function of the mean file size μ^{-1} in the Reno case and when when $C = 270$ Pkts/sec, $R = 0.1$ s., $p = 0.8$ and $\beta^{-1} = 2$ s.

We observe that a sharp decrease of the mean performance of about 15% takes place at a value of the mean file size that is significantly smaller than that obtained by a mean load analysis. This sharp decrease is that due to the jump from the congestion-less to the congestion stationary regimes described above.

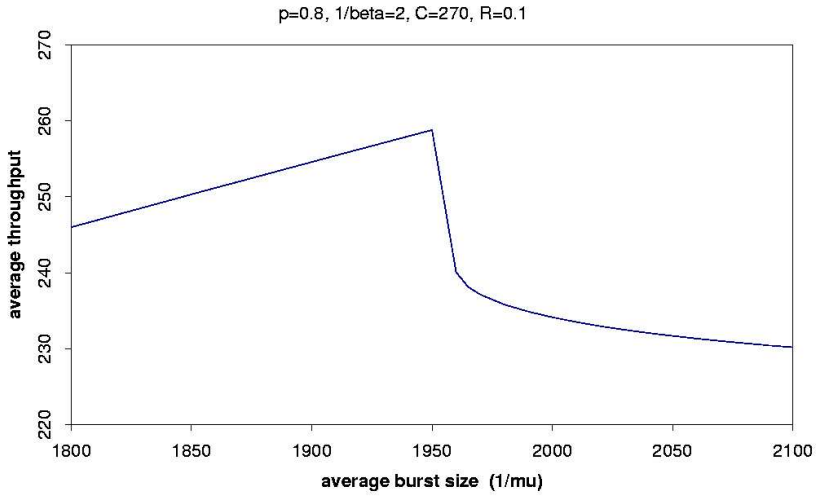

Figure 10: Average throughput as a function of the mean burst size $1/\mu$

We now study more detailed properties of the stationary throughput. Figure 11 gives the stationary rate pdfs obtained by simulation and numerically in the case C=250 Pkts/sec, p=.4, $1/\mu$=2200 Pkts, $1/\beta$=2 s., R=.1 s. The fractal and intricate structure of the pdf of the rate at congestion epochs should not come as a surprise (similar shapes were obtained for long lived sessions by Chaintreau and De Vleeschauwer in [9]). Compared to the case of Figure 5 the irregularities of the pdf are enhanced by the smaller value of p. This irregular structure is not a simulation artifact: the same structure is clearly observed via the numerical method too as shown by Figure 13.

The continuous time has a somewhat more regular rate pdf.

5 Extension of the Approach to the Slow Start

5.1 Mathematical Analysis

A thorough analytical or simulation treatment of slow start for dynamic flows is beyond the scope of this paper. We limit ourselves to a discussion of the simplest way to represent slow

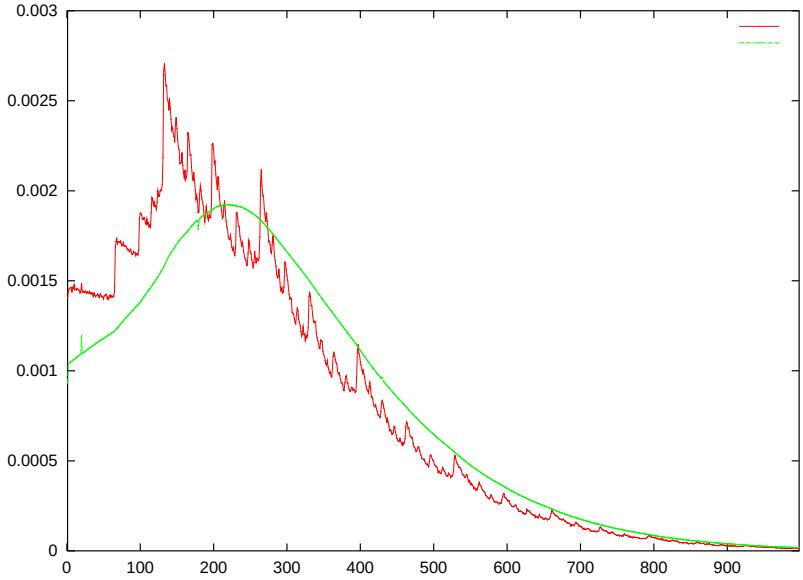

Figure 11: N2N Simulation: in red, steady state probability distribution function of the rate just after a congestion epoch. In green, continuous time stationary rate distribution. Both are obtained by simulation. Reno case, C=250 Pkts/s., p=.4, $1/\mu$=2200 Pkts, $1/\beta$=2 s.

start within this framework; i.e. an instantaneous jump of some random size at the birth of a flow. The rationale for this is that the associated exponential growth phase is quite quick compared to the congestion avoidance phase and that it can hence in a first approximation be seen as a jump from 0 to some random value H that may be either obtained from measurements or estimated as e.g. a proportion of the max window size.

The RCP equation of Section 3.2 then becomes

$$\frac{pC}{2\tau} + \lambda_\delta \mathbb{E}^\delta(X(0-)) = \frac{P(X(0) > 0)}{R^2} + \lambda_\delta \mathbb{E}^B(H). \tag{38}$$

It is also easy to extend the integral equations of Section 3.3 from knowledge of the distribution $\eta(z)$ of H. The regenerative cycles admit the very same definition as in the case without slow start, whereas the integral equation giving the expression of $f(t)$ (3) should be rewritten as indicated in (39). We limit ourselves to the expression for f as the others are

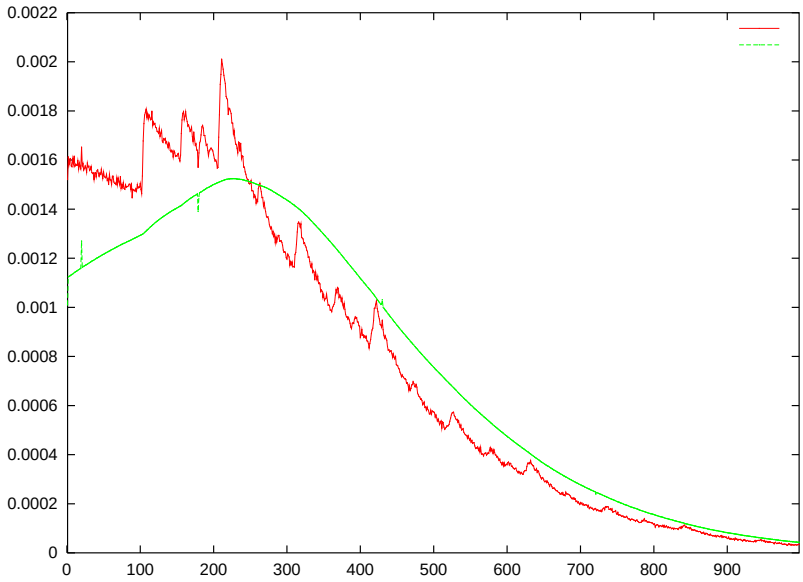

Figure 12: N2N simulation of the same case as in Figure 11 but with C=270 Pkts/s.

obtained in the same way; in particular the equations for g, i, k and m are unchanged.

$$\begin{aligned} f(t) &= \int_t^\tau \beta e^{-\beta(r-t)} dr \left\{ \int_0^\infty \eta(w) dw \right. \\ &\int_0^{w(\tau-r)+\frac{(\tau-r)^2}{2R^2}} \mu e^{-\mu y} dy (1 + f(r + R\sqrt{R^2w^2 + 2y} - R^2 w) \\ &\quad + e^{-\mu\left(w(\tau-r)+\frac{(\tau-r)^2}{2R^2}\right)} \\ &\left. \left(pg(w/2 + \frac{\tau-r}{2R^2}) + (1-p) g(w + \frac{\tau-r}{R^2}) \right) \right\}. \end{aligned} \tag{39}$$

So the fixed point equation based on the RCP can be extended almost directly to the case with this simplified representation of slow start.

Finally, the partial differential equation of Section 4.1 for the interaction-less process should be replaced by

$$\frac{\partial s}{\partial t}(z,t) + \frac{1}{R^2}\frac{\partial s}{\partial z}(z,t) = \nu(t)\beta\eta(z) - \mu z s(z,t). \tag{40}$$

It the density $\eta(z)$ converges weakly to a Dirac measure at zero then the slow-start case reduces to the congestion avoidance case studied in Section 4.1. In particular for any $z > 0$,

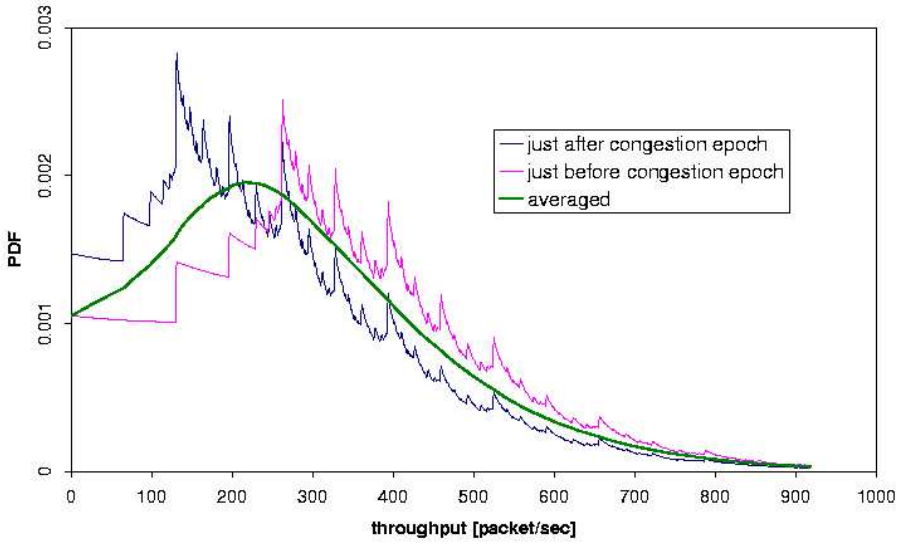

Figure 13: The same case as in Figure 11 obtained by the numerical method.

the solution $s_{SS}(z,t)$ to (40) converges pointwise to $s_{FR}(z,t)$, the solution of (17) and (18). To see this, suppose $\eta(z)$ has support on $[0, h/R^2]$. Equation (40) on $(z,t) \in [h/R^2, \infty] \times [0,\infty]$ is then the same as (17) on the same interval. The solution to (17) is determined by the boundary values $s_{SS}(h/R^2, t), t \geq 0$ and $s(z,0), z \geq h/R^2$. Now integrate (40) over $z \in [0, h/R^2]$. To first order it follows that $s_{SS}(h/R^2, t) = \nu(t)\beta R^2 = s_{FR}(0,t)$. By the continuity of $s_{FR}(z,t)$ it follows the solution to (40) is arbitrarily close, as $h \to 0$, to the solution of (17) and (18) on $(z,t) \in [h/R^2, \infty] \times [0,\infty]$.

The stationary aggregate rate associated with the solution of (40) is:

$$\alpha(\infty) \quad = \quad \frac{1}{\mu} \frac{1}{\frac{1}{\beta} + \int_{z=0}^{\infty} \int_{u=0}^{\infty} \eta(z) \mu e^{-\mu u} \left(\sqrt{z^2 R^4 + 2uR^2} - zR^2 \right) du dz} = \rho. \qquad (41)$$

By arguments similar to the ones of §4.1 one gets that the solution of (40) satisfies the Fredholm integral equation

$$\begin{aligned} s(z,t) &= R^2\beta \int_{v=0}^{z} \left(1 - \int_{x=0}^{\infty} s(x, t - R^2(z-v))dx\right) e^{-\mu R^2\left(\frac{z^2}{2} - \frac{v^2}{2}\right)} \eta(v)dv \\ &+ s(z - \frac{t}{R^2}, 0)\ e^{-\mu\left(tz - \frac{t^2}{2R^2}\right)}. \end{aligned} \tag{42}$$

The invariant measure equation keeps the same form as (34) but with $s_0(.,.)$ now obtained from the last equation rather than from (24). The same machinery can then be used, in particular for the necessary and sufficient condition for the existence of a periodic regime of period τ, which is the direct analogue of what was done above in the case without slow start.

The numerical methods used for solving the RCP and the invariant measure equation and the simulation methodology are direct extensions of those used in the case without slow start.

Consider the case where H is deterministic and equal to $C/2$ (see §5.2 below).

Consider, for instance, the case where the parameters are still $C = 270$ Pkts, $1/\beta = 2$ s., p=0.8, R=0.1 s. and $1/\mu$=2000 Pkts. Both the N2N simulator and the RCP equation (38) give and a period of $\tau = 1.89$ s. The numerical solution of (42) leads to an aggregate rate function $\alpha(.)$ that satisfies the required property of first hitting $C = 270$ Pkts at $\tau = 1.89$ sec, so that this solution of the RCP is non-spurious.

5.2 HTTP 1.1 Example

We propose to focus on HTTP 1.1 where the files successively downloaded by a flow use the same TCP connection. This assumes of course that the successive downloads of this user are made from the same server and that the Keepalive Timer (usually 15 s.) does not expire (for the last point, see [2]).

We then refer to IETF RFC 2581 [4] to state the following concerning TCP:

- When the TCP connection is idle for more than one retransmission timeout (RTO, roughly a few RTTs), CWND is reduced to IW (initial window), which we will assume to correspond to decreasing the rate to 0.
- SSTHRESH is however kept to save information on the previous value of the congestion window. We propose here to take SSTHRESH= $C/(2(1-\nu))$, where ν denotes the stationary probability that a flow is idle at a congestion epoch. The rationale for this is as follows: when the last loss occurred (a loss always occurs for each flow in the finite population model), the proportion of active flows was $1-\nu$ and the average rate was per flow was C; hence due to symmetry, each active flow had an average of $C/(1-\nu)$; so it indeed makes sense to take SSTHRESH= $C/(2(1-\nu))$.

Hence in our slow start model, the rate of a flow jumps to $C/(2(1-\nu))$ at the beginning of each file transfer, and a congestion avoidance phase then starts until file completion. This is

one model among many other possibilities, which has engineering meaning under the above assumptions (all flows access the same server, HTTP 1.1 is used, and the Keepalive Timer is large) and provided CWNDMAX is large and the exponential phase of the slow start is fast enough to be neglected.

Of course ν is unknown. To cope with this, in a first step, we solve the model of §5.1 with $H = C/2$. This determines τ_1 and ν_1. In a second step, we solve the model again with $H = C/2(1-\nu_1)$ and so on until convergence. When applying this procedure to the example of the last section, τ_1=1.89 s. and ν_1=0.226 at the first step and τ_n=1.73 s. and ν_n=0.225 for all $n \geq 2$. The regime associated with the last values is such that the α function first reaches $C = 270$ Pkts at τ=1.73 s. The joint trajectories of two flows (among thousands) controlled by these dynamics as obtained by simulation are plotted for illustration in Figure 14.

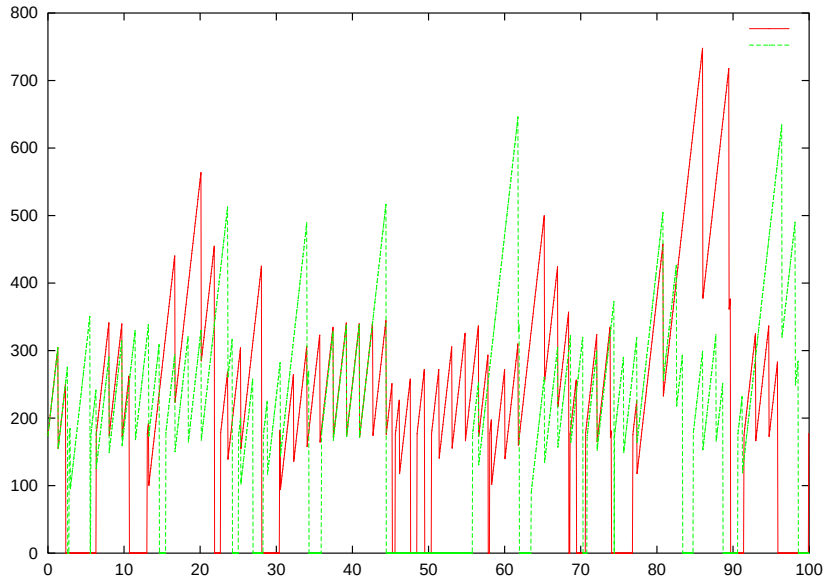

Figure 14: Two HTTP 1.1 flows with slow start C=270 Pkts/s., p=.8, $1/\mu$=2000 Pkts, $1/\beta$=2 s.

The stationary distribution of the rate is exemplified in Figure 15.

The basic observation is the same as in the case without slow start: in cases where the load per user is less than the capacity per user, one can get a turbulent mean-field limit with infinitely many congestions for appropriate initial conditions. Here is an example of such a turbulent regime: C=364 Pkts/s., p=.8, $1/\mu$=2000 Pkts, $1/\beta$=2 s. One gets a period of τ=5.568 s. and a load per user of 356.618 Pkts/s. Here, the load per user is defined using the same ideas as above: when the transfer of a file starts, the rate jumps from 0

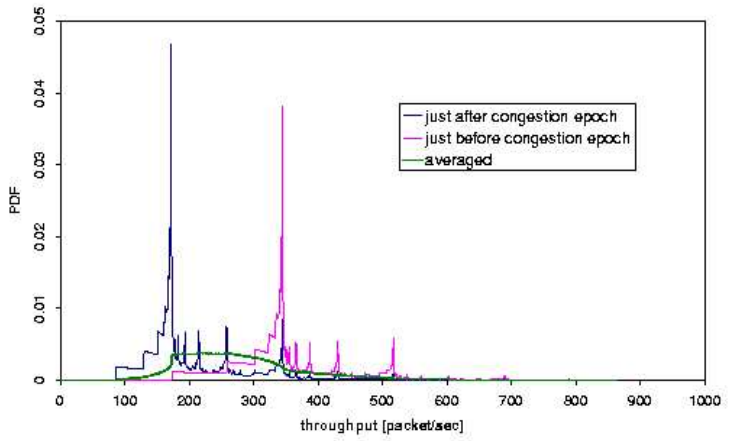

Figure 15: Stationary pdf of the rate of one flow with slow start; C=270 Pkts/s., p=.8, $1/\mu$=2000 Pkts, $1/\beta$=2 s.

to $H = C/(2(1-\nu))$ and then evolves according to the congestion avoidance AIMD rules. In this last expression, ν is the continuous time probability that a flow is active in the interaction-less regime. Notice that determining ν requires the solution of a fixed point equation (as this probability depends on H which itself depends on ν).

6 Comparisons

6.1 Comparison to the Long Lived Flow AIMD Model

The issue addressed here is that of the limiting behavior of the ON-OFF model considered in this paper when the size of the files tends to infinity and the comparison to the results of [6] and [9] on the long lived (or persistent) flow case. We recall that in the long lived flow case [6], the inter-congestion time is $\tau_{LL} = R^2Cp/2$ and the mean throughput of a flow is $C(1-p/4)$. In Figure 16, we plot (together with another curve) the mean throughput obtained by one flow in function of the mean file size $1/\mu$. The rightmost part of the AIMD curve has an horizontal asymptote of app. $C(1-p/4)$ which suggests that the limiting behavior in question is indeed that of the long lived flow case.

Consider the fixed point equation (11). Multiplying by R^2 we get:

$$\frac{R^2Cp}{2\tau} + R^2 l(0)/j(0) = h(0)/j(0),$$

where the second term in the LHS can be interpreted as the jumps down at the end of the downloads and the RHS as the fraction of flows that are "on". We see that when letting $1/\mu$ go to infinity, we have the following consistent limits: τ tends to τ_{LL}, the contribution of the second term in the LHS (the jumps down) tends to 0 and the RHS tends to 1 from below.

The $s(z,0)$ function that we obtain is also consistent with what was found in the case of long lived flows: $s(z,0)$ has spikes at $q = (p/2)C$ and all the other values predicted in [9] on long lived flows.

6.2 Comparison to the Processor Sharing Engset Model

Another interesting issue concerning non persistent flows is the comparison of the bandwidth sharing that results from the AIMD induced dynamics of the present paper to that of the processor sharing (PS) approximations proposed in the literature (see Section 2). The closest large population PS model would be the Engset model with N users, where N is large. In this model the active sessions generate $1/\mu$ packets which are queued at a single server processor sharing node serviced with rate CN packets per second. Once served, these sessions move to an infinite server think time node where they stay for a duration of $1/\beta$ seconds. In steady state these sources are independent and the proportion in the think state will be θ. The rate at which new sessions are created is therefore $\theta N\beta$. This must be matched by the rate at which sessions finish. This rate would be the same if we served the sessions in a FIFO manner and the rate for this is $1/((1/\mu)/NC)(1-\pi_N(0)) = \mu NC(1-\pi_N(0))$, where $\pi_N(0)$ denotes the steady state probability that the PS queue is empty. It is easy to check from the product form of the finite population Engset model that when N tends to infinity, there are two basic regimes:

- If $\beta < \mu C$, then $\pi_N(0)$ tends to 1 when N tends to infinity, so that the PS queue is always empty in the large population asymptotic model. So in this case, the mean rate obtained by each flow is $x = \beta/\mu$;
- If $\beta > \mu C$, then $\pi_N(0)$ tends to 0 and consequently $\theta = \mu C/\beta$. In this case, the intensity of the arrival point process in the PS queue grows to infinity like $N\mu C$ whereas the steady state queue size grows to infinity like $N(1-\theta)$. Hence Little's law allows us to determine the mean waiting time W of a tagged file transfer in the PS queue via the formula

  $$W = (1-\theta)\frac{1}{\mu C}.$$

  Consequently, the mean rate obtained by each flow is

  $$x = \frac{1}{\mu}\frac{1}{1/\beta + W} = C.$$

Our conclusion is that the rate in the large population Engset PS model is given by the formula:

$$x = \beta \min\left(\frac{1}{\mu}, \frac{C}{\beta}\right). \tag{43}$$

Figure 16 below compares this to the expressions obtained from our AIMD model, with and without slow start. In the case without slow start, the rate in the increasing part of the curve of the AIMD model (i.e. the part where no congestion occurs) is obtained from (14). As one can check, the match is not so good unless the load is small. Notice that there is actually no reason for these models to be close because in the processor sharing formula there is no dependence on the RTT.

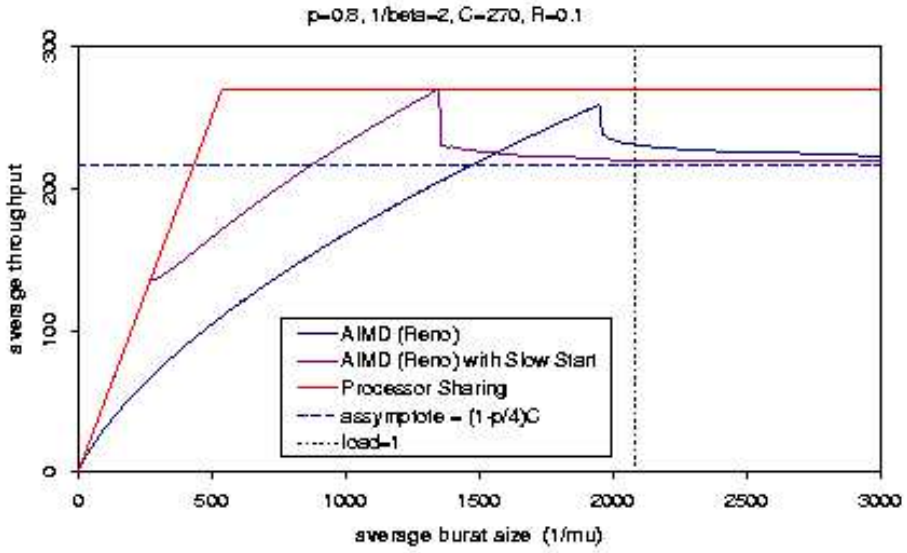

Figure 16: The average rate as predicted by the PS and the AIMD models.

The rightmost part of the PS curve postulates full bandwidth sharing whereas the AIMD dynamics does not. The rightmost part of the AIMD curve has an horizontal asymptote of app. $C(1 - p/4)$ (that is here $.8 \times C$) as predicted.

The qualitative properties found in the present study have no analogues in these processor sharing models: there are no multiple stationary regimes depending on the initial condition: above, we looked at the steady state of the Engset model and then let N (population) go to infinity. That is we let first time go to infinity (to get steady state) and we then let N to to infinity. Had we started the Engest model in some transient state (e.g. all users thinking, rather than in steady state), the steady state obtained when letting first N go to infinity and then letting time go to infinity is the same as the one obtained above as is easily seen by a direct analysis of the transient mean-field Engset model.

Notice that these multiple regimes appear in the vicinity of critical load, which is precisely a region where processor sharing is not expected to provide an accurate model for TCP bandwidth sharing anyway.

7 Simulation

The simulation results of this section are based on the N2N simulation tool [3].

7.1 Meta-stability

The fact that the mean-field limit has two stationary regimes for some values of the parameters translates into the existence of two meta-stable regimes for any finite stochastic system with the same mean parameters, with rare oscillations from one stable regime to the other. This phenomenon (see e.g. [13] for another example pertaining to protocols) is depicted in Figures 17 and 18 which feature both the Tahoe case with $1/\mu = 2000$ Pkts, $1/\beta = 2$ s. and $R = 0.1$ s.

In Figure 17, the number of sources is rather small (1000) and we wish to choose the capacity between ρ and the critical value C_T above which the mean-field system has only one uncongested mode. In practice we estimate C_T by the height of the bump in the empirical aggregate rate function (which estimates $\alpha(t)$, given at (26) and we estimate ρ by the long run average rate per flow. The two modes are clearly visible in the trajectories in Figure 17. The fluctuations are high enough to make the system move frequently enough from one mode to the other.

In Figure 18, the number of sources is larger (10000) so that fluctuations are more limited which implies more difficult transitions from one mode to the other. In this case, we let the system start with all sources active and with 0 rate. We chose $C = 278$, a region where both modes are possible in the limit. As one can check, the rare event that allows the system to move from the congested to the uncongested mode only happens after appr. 5400 s.

7.2 Heavy Tailed Case

The setting is the same as that of the previous sections with lognormal distribution functions for the file size and the OFF-time. The scenario is the following: TCP Reno, with RTT $R = 30$ ms. and with synchronization rate $p = 0.8$; the file size and the OFF-period follow

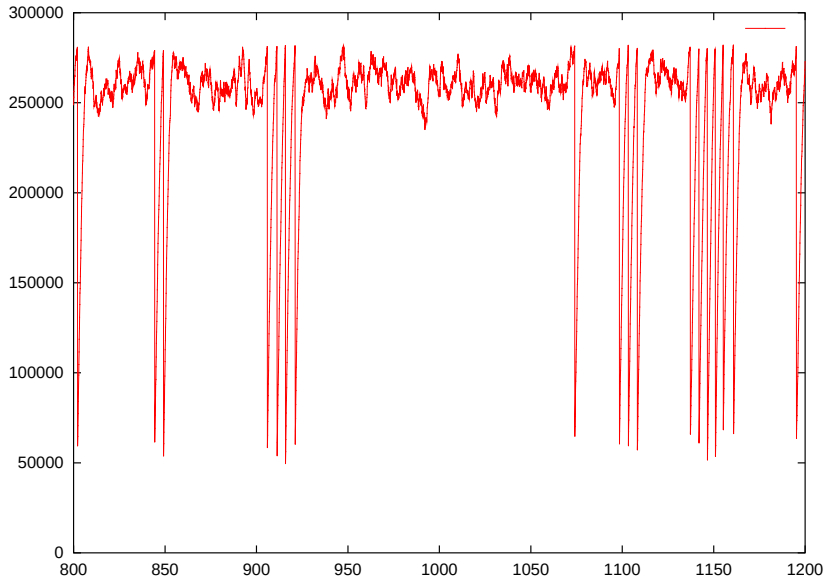

Figure 17: Bi-stability: 1000 Tahoe flows with $C = 282$, $p = .8$.

lognormal distributions: the file size has mean value 2000 Pkts and standard deviation 8669 Pkts, and the OFF-period has a mean value of 2 sec and a standard deviation of 8.7 s. Variance is much higher than in the exponential case.

Simulations (or direct calculations) show that the mean load per source is appr. $\rho = 620$ Pkts/s. Figure 19 gives the aggregate rate when $C = \infty$ for the initial condition with all sources active and with null rate. We observe the same phenomenon as in the exponential case, with a first maximum at 717 Pkts/sec, significantly larger than the horizontal asymptote at ρ, though with a shape that is different from that in the exponential case.

Our simulation suggests that as in the exponential case, congestion regimes show up for values of C larger than ρ. Here, such regimes are possible for all C between a threshold that seems to be located between 670 and 680.

8 Packet Level Simulation

8.1 The Bump

Simulations using ns2 [1] were conducted in which 400 ON/OFF flows access a shared link with capacity C through distinct links with capacity 20Mbps each. The throughput for individual flow never attain this high value even without losses because file tranfers finish

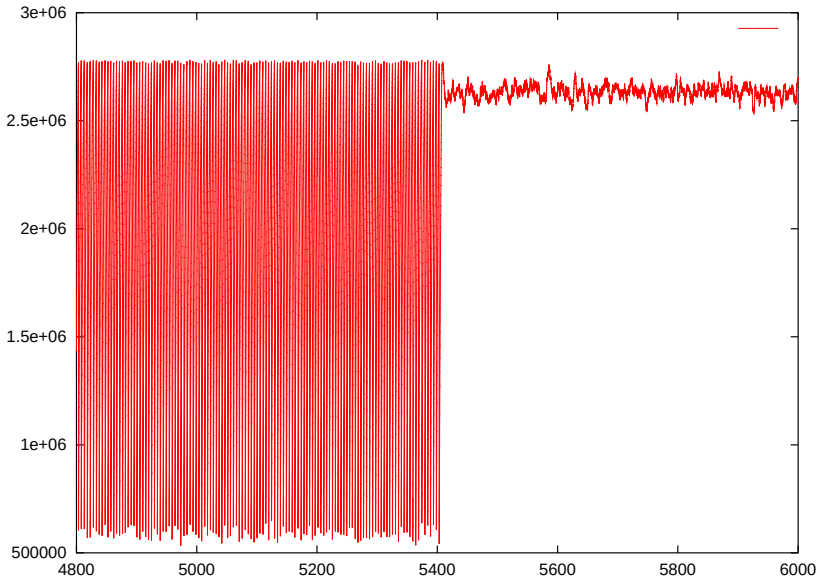

Figure 18: Bi-stability: 10000 Tahoe flows with C =278, $p = .8$.

before this rate is attained. Their common RTT is 61ms, and the shared buffer is 300 packets. The ON and OFF periods are exponentially distributed with mean 2 s. and 2000 pkts respectively. The slow start was switched off for each flow, replicating conditions close to those of the mathematical model.

Figure 20 gives the empirical aggregate rate function, and its associated bump, when C is chosen to be large (2.5 Gbps) to put the network in its free regime. We obtain a long run average throughput, or estimated load, of 3.31 Mbps per flow.

8.2 The Fluid and Turbulent Regimes

Next, we choose the capacity to be 5% more than the estimated load of all flows (C=1.39Gbps), that is less than the height of the bump. We observe that the traces of the ns2 packet simulator, Figure 21, exhibit the same bi-stability between the turbulent and the fluid regimes, as was observed in N2N simulation, Figure 18. Notice the transition to the interactionless regime at time t=1000s. In a later stage of the simulation, shown in Figure 22, the interactions regime is observed.

We characterize the interactionless regime in Figure 21 by the spectral analysis shown in Figure 23 (top), computed from the trace between time 1010s and 1350s. The same analysis is made for the interaction regime in Figure 23 (bottom), comptured from the traces between time 2600s and 2780s. Notice that the peak in the periodogram at frequency

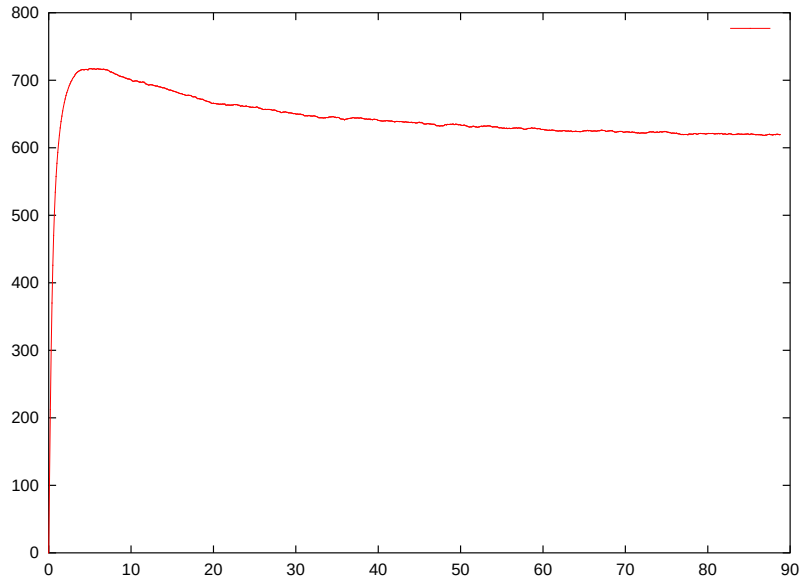

Figure 19: The mean-field aggregate rate of Reno when $C = \infty$ and all flows are initially active and with 0 rate.

0.27 corresponds to a period of 2.32s, which is roughly the observed intercongestion period (seen Figure 24

9 Conclusion

The main achievement of the present paper is an interaction model for TCP controlled dynamic flows that is based on the AIMD dynamics of TCP rather than on the frequently made assumption that TCP bandwidth sharing is well described by the processor sharing discipline. Thanks to a mathematical model based on the mean-field limit, some unexpected qualitative results are found. In particular the system may enter into a congestion regime for loads that are significantly smaller than the link capacity. Also multiple stationary regimes may be reached depending on the initial phases of the ON-OFF flows. This property that can be seen as an analogue of turbulence. These phenomena, which translate into a bi-stability property for systems with finite population, are absent in the PS model.

Another interesting property is the fractal nature of the p.d.f of the stationary rates as already observed in the long-lived flow case by Chaintreau and De Vleeschauwer in [9]: the randomness and the mixing of the ON-OFF structure seems to be compatible with a complex self-similar structure for the rates. Even for the (rather unrealistic) exponential

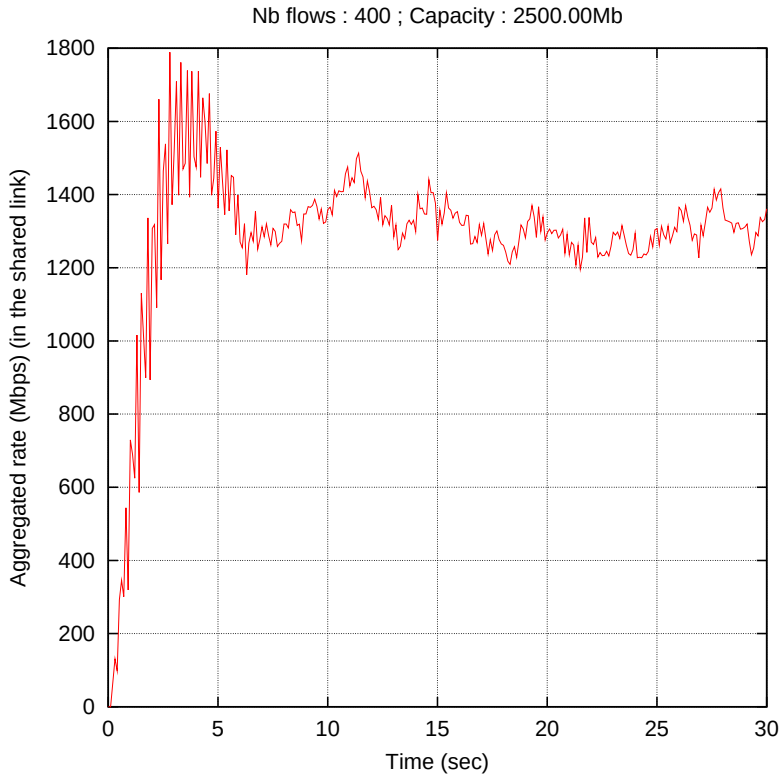

Figure 20: ns2 aggregated traffic when all flows start active at time 0

model analyzed here, several important theoretical questions have to be solved to complete the present study. These include the proof of the mean-field limit (which should be feasible along the lines of what was already done for the long lived flow case) and the mathematical confirmation of the numerical findings presented in Section 4 in the Reno case.

The main step after that is of course to extend the approach to non exponential file sizes and particularly to heavy tailed distributions. Other interesting extensions along the lines of what is already known for the long lived flow case would address the multiple link case and the non-linear dynamics induced by a large tail-drop buffer. Finally, it should be possible to mix this HTTP traffic model with the model for long lived flows to give a single interactive dynamical system.

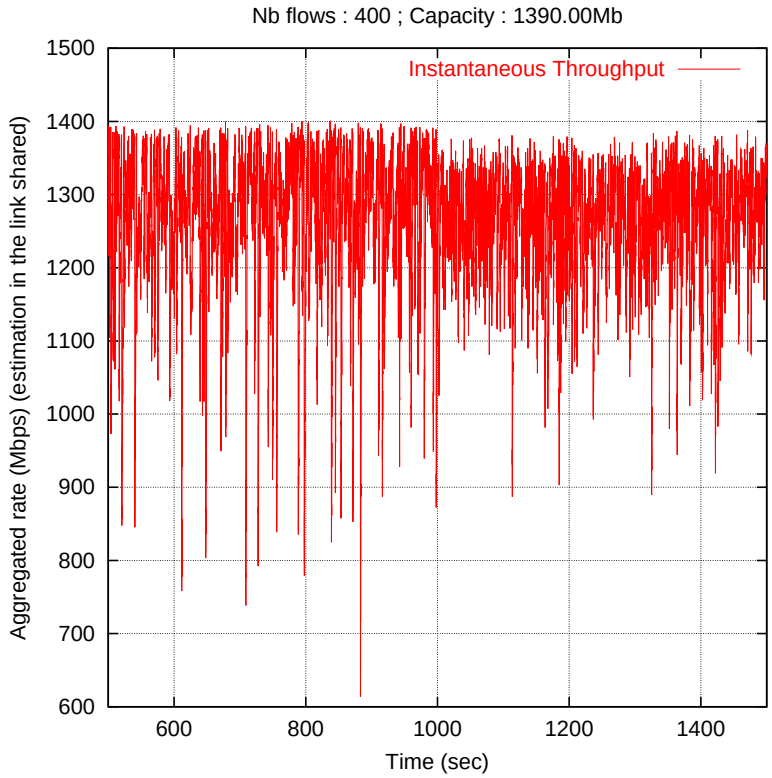

Figure 21: ns2 aggregated traffic: alternance between the two regimes.

Acknowledgments. The authors thank Dohy Hong for his valuable contributions to this paper. The analogy of the phenomena described in this paper with turbulence was brought to our attention by Bruce Hajek.

10 Appendix

10.1 Proof of Lemma 1

By arguments similar to those in the proof of (18),

$$\frac{d}{dt}\nu(t) = -\beta\nu(t) + \mu\alpha(t), \tag{44}$$

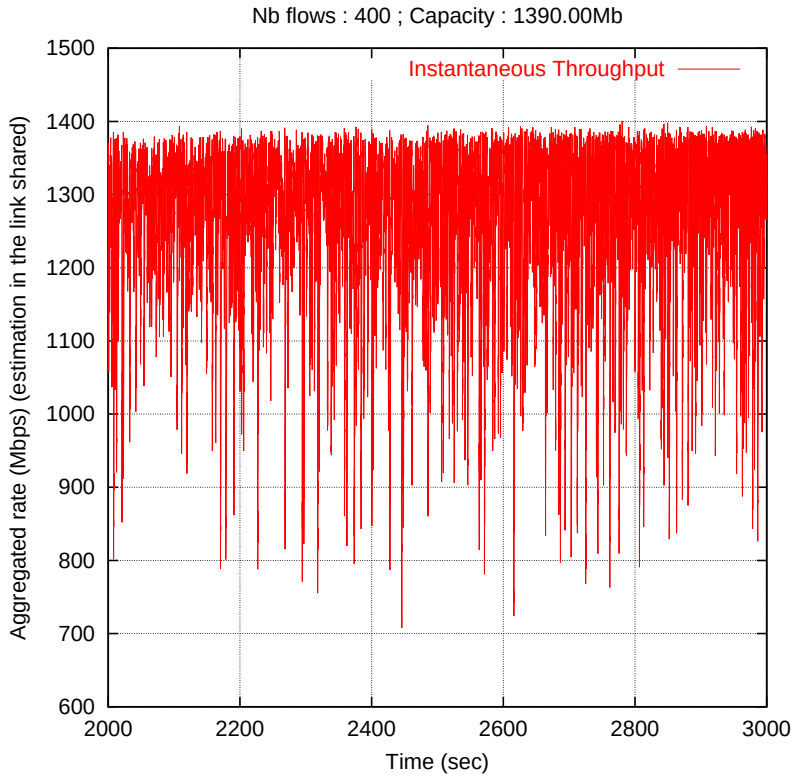

Figure 22: ns2 aggregated traffic: interaction regime.

so that

$$\widehat{\nu}(u) \quad = \quad \frac{1}{\beta+u}\left(\nu(0)+\mu\widehat{\alpha}(u)\right). \tag{45}$$

This and (18) give

$$\widehat{s}_0(u) \quad = \quad \frac{\beta R^2}{\beta+u}\left(\nu(0)+\mu\widehat{\alpha}(u)\right).$$

When using the last expression in (21), we get

$$\begin{aligned}\widehat{s}_z(u) = \quad & e^{-R^2(uz+\mu z^2/2)}\left(\tfrac{\beta R^2}{\beta+u}\left(\nu(0)+\mu\widehat{\alpha}(u)\right)\right.\\ & \left.+R^2\int_0^z e^{-R^2u(z-x)}e^{-R^2\mu(\frac{z^2}{2}-\frac{x^2}{2})}s(x,0)dx\right).\end{aligned} \tag{46}$$

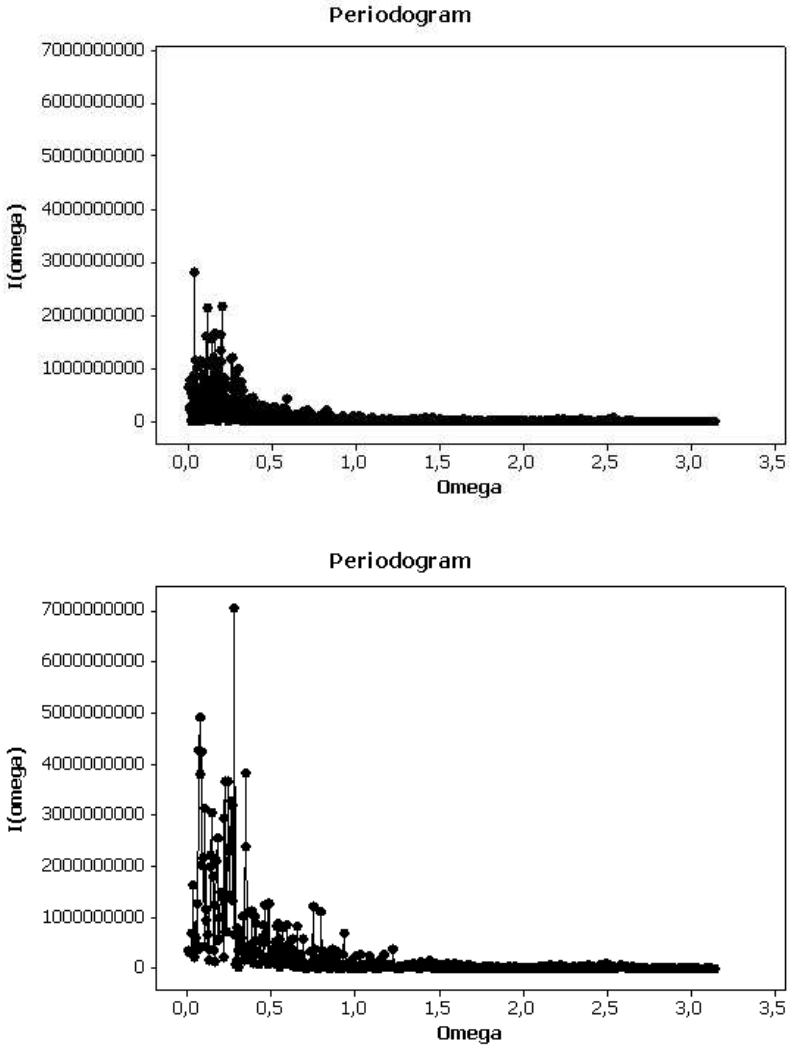

Figure 23: Periodograms : interactionless regime (top) and interaction regime (bottom).

When multiplying both sides of (46) by z and integrating w.r.t. z, we get

$$\widehat{\alpha}(u) = \frac{\beta}{\beta + u}\widehat{I}(u)\left(\nu(0) + \mu\widehat{\alpha}(u)\right) + \widehat{J}(u), \tag{47}$$

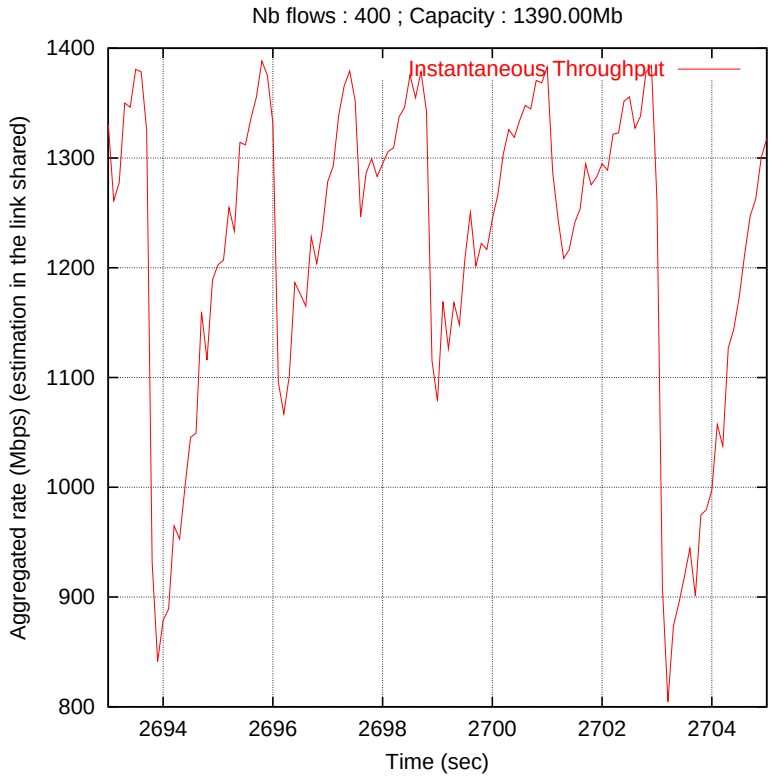

Figure 24: A zoom of aggregated traffic in the congestion regime.

where $\widehat{I}(u)$ and $\widehat{J}(u)$ are the integrals defined in §4.1.3. The expression in (28) follows immediately.

10.2 Proof of Lemma 2

Tauberian theorems (see e.g. [11]) applied to the expressions obtained above for $\widehat{\nu}(u)$ and $\widehat{s}_z(u)$ and $\widehat{\alpha}(u)$ lead to the stationary distribution (31) and (32) for the rates (note that this is a truncated Gaussian) and to (33) for the aggregate rate.

10.3 Regenerative Interpretation for the Results of Lemma 1

We now give the interpretation of the expressions in Lemma 1 in terms of the regenerative process $X(t)$.

- The function $\mu\widehat{I}(u)$ is the Laplace transform of the function

$$\mu I(t) = \mu\frac{t}{R^2}e^{-\frac{\mu t^2}{2R^2}},$$

which is the density of the random duration of an ON period (a file transfer);

- The function $\mu\widehat{I}(u)\frac{\beta}{\beta+u}$ is the Laplace transform of the random variable equal to the sum of one OFF and one ON period, namely the duration of a regeneration cycle;

- The function

$$\widehat{\xi}(u) = \frac{1}{1-\mu\frac{\beta}{\beta+u}\widehat{I}(u)}$$

is the Laplace transform of the density $\xi(t)$ of the renewal measure (see [11], Vol 2, pp. 184 and following) of the renewal point process S associated with the regenerative process $X(t)$; $\xi(t)$ can be interpreted as the density with respect to the Lebesgue measure of the expected number of points in the $(0,t]$ interval of the Palm version of the renewal process, or intuitively as the probability density that the Palm version of the renewal point process hits t;

- The function $\widehat{I}(u)$ is the Laplace transform of the product of t/R^2 and of the probability that a flow that starts active and with 0 rate at time 0 and is continuously active until time t: $I(t) = \frac{t}{R^2}e^{-\frac{\mu t^2}{2R^2}}$. From this and a simple renewal theory argument, one sees that the function

$$\widehat{\psi}(u) = \frac{\frac{\beta}{\beta+u}\widehat{I}(u)}{1-\mu\frac{\beta}{\beta+u}\widehat{I}(u)}$$

is simply the Laplace transform of the function $\psi(t)$ that gives the expected value of $X(t)$ when the tagged flow starts inactive at time 0 and has an arbitrary number of OFF and on periods in between;

- The function $\widehat{J}(u)$ is the Laplace transform of the function

$$J(t) = \int_0^\infty s(z,0)\left(z+\frac{t}{R^2}\right)e^{-\mu\left(zt+\frac{t^2}{2R^2}\right)}dz,$$

in which we recognize the expected value of $X(t)$ when the tagged flow starts active and with rate sampled according to $s(.,0)$ at time 0 and remains continuously active between time 0 and t; also notice that an integral of the form

$$\int_{z=0}^\infty s(z,0)dz\int_{v=0}^t\left(z+\frac{v}{R^2}\right)e^{-\mu\left(vz+\frac{v^2}{2R^2}\right)}dv$$
$$\int_{u=0}^{t-v}\mu\beta e^{-\beta(t-v-u)}e^{-\mu\frac{(t-v)^2}{2R^2}}\frac{t-v}{R^2}du$$

can also be seen as

$$\int_{z=0}^{\infty} s(z,0)dz \int_{v=0}^{t} \mu\left(z+\frac{v}{R^2}\right) e^{-\mu\left(vz+\frac{v^2}{2R^2}\right)} dv$$
$$\int_{u=0}^{t-v} \beta e^{-\beta(t-v-u)} e^{-\mu\frac{(t-v)^2}{2R^2}} \frac{t-v}{R^2} du,$$

where $\mu\left(z+\frac{v}{R^2}\right) e^{-\mu\left(vz+\frac{v^2}{2R^2}\right)}$ is the density of the random duration of an ON period that starts with initial rate z. So this integral is the expected value of $X(t)$ on the event that the initial ON period ends and there is then one OFF period and one uninterrupted ON period until time t; so, a renewal argument allows us to conclude that the term

$$\widehat{\zeta}(u) = \frac{\widehat{J}(u)}{1-\mu\frac{\beta}{\beta+u}\widehat{I}(u)}$$

is the Laplace transform of the function $\zeta(t)$ that gives the expected valued of $X(t)$ when the tagged flow starts active at time 0 with a rate sampled according to $s(.,0)$ and has an arbitrary number of OFF and ON periods in between.

10.4 Explicit Solution of the Solution of the PDE

By direct Laplace inversion of (21), one gets the following expressions for $\nu(t)$ and $s(z,t)$:

$$s(z,t) = \nu(0)\beta R^2\phi(z,t) + \mu\beta R^2 \int_0^t \alpha(u)\phi(z,t-u)du + e^{-\mu\left(tz-\frac{t^2}{2R^2}\right)} s(z-\frac{t}{R^2},0) \quad (48)$$

$$\nu(t) = \nu(0)e^{-\beta t} + \mu\int_0^t e^{-\beta(t-u)}\alpha(u)du, \quad (49)$$

where

$$\phi(z,t) = e^{-\frac{R^2\mu z^2}{2}} e^{-\beta(t-R^2 z)}$$

is the inverse Laplace transform of the function

$$\widehat{\phi}_z(u) = \frac{1}{\beta+u} e^{-R^2 zu - R^2\mu z^2/2}.$$

Similarly $\alpha(t)$ can be expressed as follows:

$$\alpha(t) = \nu(0)R^2\psi(t) + R^2\int_0^t e^{-\mu v^2/(2R^2)}\xi(t-v)\int_0^\infty e^{-\mu zv}(z+\frac{v^2}{R^2})s(z,0)dz\,dv \quad (50)$$

with ψ and ξ defined in §10.3.

10.5 A Common Linear Equation for $(f,g),(h,i),(j,k),(l,m)$

We start from the equation given for (f,g) and described for other functions in (5)-(7). They can be put in the following general form (where the γ and ζ functions have no connection to those defined in the earlier sections of the paper) :

$$\begin{aligned}\gamma(t) &= \eta_0(t) + \int_t^\tau \beta e^{-\beta(r-t)}dr \left\{ \int_0^{\frac{(\tau-r)^2}{2R^2}} \mu e^{-\mu y}dy(\eta_1(t,r,y) + \gamma(r+R\sqrt{2y}) \right. \\ &\left. + e^{-\mu\frac{(\tau-r)^2}{2R^2}} \left(\eta_2(t,r) + p\zeta(\frac{\tau-r}{2R^2}) + (1-p)\zeta(\frac{\tau-r}{R^2}) \right) \right\}. \qquad (51)\end{aligned}$$

$$\begin{aligned}\zeta(z) &= \int_0^{z\tau+\frac{\tau^2}{2R^2}} \mu e^{-\mu y}(\eta_3(z,y) + \gamma(R\sqrt{R^2z^2+2y} - R^2z))dy \\ &+ e^{-\mu(z\tau+\frac{\tau^2}{2R^2})} \left(\eta_4(z) + p\zeta(\frac{z+\tau/R^2}{2}) + (1-p)\zeta(z+\tau/R^2) \right). \qquad (52)\end{aligned}$$

for the following values:

$\gamma(t)$	$\zeta(r)$	$\eta_0(t)$	$\eta_1(t,r,y)$	$\eta_2(t,r)$	$\eta_3(z,y)$	$\eta_4(z)$
$f(t)$	$g(z)$	0	1	0	1	0
$h(t)$	$i(z)$	0	$R\sqrt{(}y)$	$\tau-r$	$R\sqrt{R^2z^2+2y}-R^2z$	τ
$j(t)$	$k(z)$	$(t-\tau)e^{\beta(t-\tau)}$	$r-t+R\sqrt{2y}$	$\tau-t$	$R\sqrt{R^2z^2+2y}-R^2z$	τ
$l(t)$	$m(z)$	0	$\frac{1}{R}\sqrt{2y}$	0	$\frac{1}{R}\sqrt{R^2z+2y}$	0

For the first equation defining γ, we introduce the following change of variable :

- For the integral in y, corresponding to the case where a file download is started and completed before τ, let s be the time of completion of this file:

$$s = r + R\sqrt{2y} \text{ such that } y = \frac{(s-r)^2}{2R^2} \text{ and hence } dy = \frac{s-r}{R^2}ds.$$

For the second equation, defining ζ, we introduce the change of variable :

- First we study ζ as a function of $r = R^2z$, which corresponds to the time to obtain a rate equal to z, if the flow start from rate zero, in the case it had no congestion.
- For the integral in y, corresponding to the case where the remaining file download is completed before τ, let s be the time of completion of this file:

$$s = R\sqrt{R^2z^2+2y} - R^2z = \sqrt{r^2+2R^2y} - r \text{ such that}$$

$$y = \frac{s^2+2sr}{2R^2} \text{ and hence } \quad dy = \frac{s+r}{R^2}ds.$$

Doing so, we obtain the following version of the general equation:

$$\begin{aligned}\gamma(t) &= \eta_0(t) + \int_t^\tau \beta e^{-\beta(r-t)} dr \left\{ \int_r^\tau \mu \frac{s-r}{R^2} e^{-\mu\frac{(s-r)^2}{2R^2}} (\eta_1(t,r,s) + \gamma(s)) ds \right. \\ &\left. + e^{-\mu\frac{(\tau-r)^2}{2R^2}} \left(\eta_2(t,r) + p\zeta(\frac{\tau-r}{2}) + (1-p)\zeta(\tau-r) \right) \right\}. \qquad (53)\end{aligned}$$

$$\begin{aligned}\zeta(r) &= \int_0^\tau \mu \frac{s+r}{R^2} e^{-\mu\frac{s^2+2sr}{2R^2}} (\eta_3(r,s) + \gamma(s)) ds \\ &+ e^{-\mu(\frac{2r\tau+\tau^2}{2R^2})} \left(\eta_4(r) + p\zeta(\frac{r+\tau}{2}) + (1-p)\zeta(r+\tau) \right). \qquad (54)\end{aligned}$$

with

$\gamma(t)$	$\zeta(r)$	$\eta_0(t)$	$\eta_1(t,r,s)$	$\eta_2(t,r)$	$\eta_3(r,s)$	$\eta_4(r)$
$f(t)$	$g(z)$	0	1	0	1	0
$h(t)$	$i(z)$	0	$s-r$	$\tau-r$	s	τ
$j(t)$	$k(z)$	$(t-\tau)e^{\beta(t-\tau)}$	$s-t$	$\tau-t$	s	τ
$l(t)$	$m(z)$	0	$\frac{1}{R^2}(s-r)$	0	$r+\frac{1}{R^2}s$	0

To save space, we now denote $\frac{\mu}{R^2}$ by κ. An integration by parts to simplify the integrals associated with η_1 and η_4 gives:

$$\int_r^\tau \kappa(s-r) e^{-\kappa\frac{(s-r)^2}{2}} ds \eta_1(t,r,s) = \eta_1(t,r,r) - e^{-\kappa\frac{(\tau-r)^2}{2}} \eta_1(t,r,\tau) \int_r^\tau e^{-\kappa\frac{(s-r)^2}{2}} \frac{\partial \eta_1(t,r,s)}{\partial s} ds.$$

$$\text{and} \int_0^\tau \kappa(s+r) e^{-\kappa\frac{s^2+2sr}{2}} \eta_3(r,s) ds = -e^{-\kappa\frac{\tau^2+2\tau r}{2R^2}} \eta_3(r,\tau) + \int_0^\tau e^{-\kappa\frac{s^2+2sr}{2}} \frac{\partial \eta_3(r,s)}{\partial s} ds$$

So we can write :

$$\begin{aligned}\gamma(t) &= \eta_0(t) + \int_t^\tau \beta e^{-\beta(r-t)} dr \left\{ \tilde{\eta}_1(t,r) + \int_r^\tau \kappa(s-r) e^{-\kappa\frac{(s-r)^2}{2}} \gamma(s) ds \right. \\ &\left. + e^{-\kappa\frac{(\tau-r)^2}{2}} \left(\tilde{\eta}_2(t,r) + p\zeta(\frac{\tau-r}{2}) + (1-p)\zeta(\tau-r) \right) \right\}. \qquad (55)\end{aligned}$$

$$\begin{aligned}\zeta(r) &= \tilde{\eta}_3(r) + \int_0^\tau \kappa(s+r) e^{-\kappa\frac{s^2+2sr}{2}} \gamma(s) ds \\ &+ e^{-\kappa(\frac{\tau^2+2r\tau}{2})} \left(\tilde{\eta}_4(r) + p\zeta(\frac{r+\tau}{2}) + (1-p)\zeta(r+\tau) \right). \qquad (56)\end{aligned}$$

where the coefficient are :

$$\tilde{\eta}_1(t,r) = \eta_1(t,r,r) + \int_r^\tau e^{-\kappa\frac{(s-r)^2}{2}} \frac{\partial \eta_1(t,r,s)}{\partial s} ds \; ; \; \tilde{\eta}_2(t,r) = \eta_2(t,r) - \eta_1(t,r,\tau)$$

$$\tilde{\eta}_3(t,r) = \eta_3(r,0) + \int_0^\tau e^{-\kappa\frac{s^2+2rs}{2}} \frac{\partial \eta_3(t,r,s)}{\partial s} ds \; ; \; \tilde{\eta}_4(t,r) = \eta_4(r) - \eta_3(r,\tau)$$

$\gamma(t)$	$\zeta(r)$	$\eta_0(t)$	$\tilde{\eta}_1(t,r)$	$\tilde{\eta}_2(t,r)$	$\tilde{\eta}_3(r,s)$	$\tilde{\eta}_4(r)$
$f(t)$	$g(r/R^2)$	0	1	-1	1	-1
$h(t)$	$i(r/R^2)$	0	$a_\tau(r)$	0	$b_\tau(r)$	0
$j(t)$	$k(r/R^2)$	$(t-\tau)e^{\beta(t-\tau)}$	$r-t+a_\tau(r)$	0	$b_\tau(r)$	0
$l(t)$	$m(r/R^2)$	0	$\frac{1}{R^2}a_\tau(r)$	$-\frac{\tau-r}{R^2}$	$\frac{r+b_\tau(r)}{R^2}$	$-\frac{r+\tau}{R^2}$

with

$$a_\tau(r) = \int_r^\tau e^{-\kappa\frac{(s-r)^2}{2}}ds \text{ and } b_\tau(r) = \int_0^\tau e^{-\kappa\frac{s^2-2rs}{2}}ds$$

Equation (10) can easily be deduced from this for functions $(f,g),(h,i)$ and (l,m). For the functions (j,k), we should perform another integration by part :

$$\int_r^\tau \beta e^{-\beta(r-t)}(t-\tau)dr = -(t-\tau)e^{-\beta(t-\tau)} + \int_r^\tau e^{-\beta(r-t)}dr.$$

The first term in the RHS compensates exactly $\eta_0(t)$ for this function, while the second term of the RHS can be seen as $1/\beta$ to add inside the integral w.r.t. the variable r.

References

[1] `http://www.isi.edu/nsnam`

[2] `http://lists.w3.org/Archives/Public/ietf-http-wg-old/2000SepDec/0078.html`

[3] `http://www.n2nsoft.com`

[4] `http://www.faqs.org/rfcs/rfc2581.html`

[5] F. Baccelli, P. Brémaud (2002), *Elements of Queueing Theory*, Springer Verlag, second edition.

[6] Baccelli, F., Hong, D. (2002) AIMD, Fairness and Fractal Scaling of TCP Traffic. *in Proc. of INFOCOM, New York, June.*

[7] Baccelli, F., McDonald, D. R., Reynier, J. (2002). A mean-field model for multiple TCP connections through a buffer implementing RED. *Performance Evaluation Vol. 11, (2002) pp. 77-97.* Elsevier Science.

[8] Barakat, C., Thiran, P., Iannaccone, C., Diot, C., Owezarski, P. (2002) A flow-based model for Internet backbone traffic. *Internet Measurement Workshop 2002.*

[9] Chaintreau, A., De Vleeschauwer, D. (2002) A closed form formula for long-lived TCP connections throughput. *Performance Evaluation 49(1/4): 57-76 (2000).*

[10] Chang, C.-S., Liu, Z. (2002) A Bandwidth Sharing Theory for a Large Number of HTTP-like Connections. *In Proceedings of the IEEE Infocom 2002 Conference, New York City , June 2002.*

[11] Feller, W. (1971) *An Introduction to Probability Theory and its Applications*, Wiley, second edition.

[12] Fred, S., Bonald, T., Proutiere, A., Régnié, G., Roberts, J. (2001) Statistical bandwidth sharing: a study of congestion at flow level. *in Proceedings of ACM SIGCOMM 2001: pp111-122*

[13] Gibbens, R. J., Hunt, P. J. and Kelly, F. P. (2002) Bistability in Communication Networks, *in Disorder in Physical Systems.*

[14] Heyman, D., Lakshman, T., Neidhardt, A. (1997) A new method for analyzing feedback-based protocols with applications to engineering web traffic over the Internet. *in ACM Sigmetrics, 1997, pp. 24–38.*

[15] Kelly, F. (1997) Charging and Rate Control for Elastic Traffic. *European Transactions on Telecommunications, vol. 8, pp. 33–37, 1997.*

[16] Kherani, A.A., Kumar, A. (2000) Performance Analysis of TCP with Nonpersistent Sessions. *Workshop on Modeling of Flow and Congestion Control, INRIA, Ecole Normale Supérieure, Paris, September 4-6, 2000.*

[17] Massoulié, L., Roberts, J. (1999) Bandwidth sharing: objectives and algorithms *in Proceedings of IEEE INFOCOM, Vol. 3. New York, NY, pp. 1395–1403.*

[18] Roberts, J., Massoulié, L. (1998) Bandwidth sharing and admission control for elastic traffic. *ITC Specialist Seminar, Yokohama, October.*

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay 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 Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

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