19

Proxy caching in split TCP: dynamics, stability and tail asymptotics

François Baccelli
Ecole Normale Supérieure

Giovanna Carofiglio
Ecole Normale Supérieure and Politecnico di Torino

Serguei Foss
Heriot-Watt University

Abstract

The split of a multihop, point-to-point TCP connection consists in replacing a plain, end-to-end TCP connection by a cascade of TCP connections. In such a cascade, connection n feeds connection $n + 1$ through some proxy node n. This technique is used in a variety of contexts. In overlay networks, proxies are often peers of the underlying peer-to-peer network. split TCP is also already proposed and largely adopted in wireless networks at the wired/wireless interface to separate links with vastly different characteristics. In order to avoid losses in the proxies, a backpressure mechanism is often used in this context.

In this paper we develop a model for such a split TCP connection aimed at the analysis of throughput dynamics on both links as well as of buffer occupancy in the proxy. The two main variants of split TCP are considered: that with backpressure and that without. The study consists of two parts: the first part is purely experimental and is based on *ns2* simulations. It allows us to identify complex interaction phenomena between TCP flow rates and proxy buffer occupancy, which seem to have been ignored by previous work on split TCP. The second part of the paper is of a mathematical nature. We establish the basic equations that govern the evolution of such a cascade and prove some of the experimental observations made in the first part. In particular, we give the conditions for system stability and we show the possibility of heavy tail asymptotics for proxy buffer occupancy and delays in the stationary regime.

From Semantics to Computer Science Essays in Honour of Gilles Khan, eds Yves Bertot, Gérard Huet, Jean-Jacques Lévy and Gordon Plotkin. Published by Cambridge University Press.

19.1 Introduction

The panorama of access network technologies has been changing at an incredibly fast rate over the past few years, whilst almost any substantial change intervened at transport layer, where TCP has become a "standard de facto".

However, the increasing user demand for high quality services spurs development of performance-enhancing techniques to implement on top of the pre-existing IP infrastructure.

Particularly powerful for content delivery and media streaming in peer-to-peer systems, *overlay networks* have emerged as an attractive solution for throughput improvement without any change of the underlying network architecture.

One of the key features of overlay networks is the *split-connection* mechanism, that yields a considerable throughput improvement for a TCP connection when split in shorter segments on the route between the sender and the receiver host. Intermediate nodes act as proxies: incoming packets are locally acknowledged on each segment (LACKs), then stored and forwarded on the next TCP connection.

In the context of overlay networks, split TCP is addressed in [5] and [19], where, in addition, a backpressure mechanism is proposed to limit the sending rate to the forwarding rate in presence of saturated proxy buffer, thus preventing buffer overflows. In other contexts, split TCP has been shown to be particularly effective when the sender-to-receiver route includes network segments with very different characteristics, like wired and wireless links, that usually cause problems to TCP. In fact, the "split connection" approach was initially proposed in the context of wireless networks where a significant throughput degradation has been observed for TCP. The poor TCP performance in wireless networks is to ascribe to the congestion control that wrongly attributes to congestion losses due to link failures (consequence of mobility or channel errors), or is related to high propagation delays that slacken the growth of the congestion window. In the seminal work of [9] and [10] a new implementation of TCP was proposed, Indirect TCP (I-TCP), which handles the problem of wired-wireless link interaction and introduces the concept of split TCP. Two TCP connections in tandem replace the single TCP connection: the first running on the wired side, the second one running over the wireless link and characterized by different parameters to cope better with larger delays and channel losses. The same approach has been drawn on in [18] where the split TCP scheme is adapted to mobile ad hoc networks to cope with the additional issue of a dynamic placement of the proxy. The aim of the "split-connection" approach is to operate a clear separation between flow control and congestion control functionalities over two different network environments. Similar issues have been studied in *satellite networks* ([20], [16]) where long propagation delays cause TCP throughput degradation by lengthening the slow start duration and slowing the linear growth of the congestion window in the congestion avoidance phase. Such throughput limitations are aggravated by frequent losses related to channel errors or temporary link disconnections. In this context, proxies with specific capabilities, called performance enhancing proxies (PEP), have been introduced to perform a transport layer connection split oblivious to end systems (cf. [25]). Among all the approaches that attempt to isolate issues pertaining to different media, the split connection approach is the only one that does not require any modification of standard TCP implementations, and for that reason it has been the subject of an in-depth study in the literature. The diffusion and implementation of split-connection techniques is documented by a recent measurement study ([28]) where the authors detect, through the use of inference/detection methods, the deployment of split TCP in all commercial networks they consider. They also investigate the throughput improvement provided by split TCP with respect to standard TCP implementation, that can be up to 65%

The majority of related work on split TCP are either measurement or simulation studies targeted to the throughput evaluation along a chain of TCP connections. (e.g.[11]). There are only a few *analytical attempts* in the literature which study split TCP's dynamics.

In [29] the authors study a particular class of split-connection approaches in wired/wireless networks, that adopts a standard version of TCP on wired segment and an "ad hoc" lightweight transport protocol for the wireless hop. In [27] an estimate of the expected throughput is provided for a cascade of standard TCP connections based on the well known square root formula, thus neglecting the dependencies between the two connections. Similar models based on the square root formula for TCP throughput estimation are presented in [15], [19] and [26], where the authors make the assumption that the buffer in the proxy never empties nor fills.

In this work, we make the following analytical contributions: we establish the equations for throughput dynamics jointly with that of buffer occupancy in the proxy. We then determine the stability conditions by exploiting some intrinsic monotonicity and continuity properties of the system. Finally, we focus on the study of buffer occupancy in the proxy and end-to-end delays to derive tail asymptotics. The framework allows us to consider both the case with an infinite buffer at the proxy and that of a limited buffer size, where a backpressure algorithm is needed to limit the sender rate and avoid losses in the proxy.

The paper consist of two parts: the first part (Sections 19.2 and 19.3) exploits some simulation results to make some basic observations on the system dynamics in different scenarios. We identify there the complex interaction that exists between TCP flow rates and proxy buffer occupancy. To the best of our knowledge, this problem is addressed here for the first time and finds an analytical explanation in the second mathematical part of the paper, where we emphasize the role of the buffer size on the total throughput gain for the split. The second part (Section 19.4) contains further mathematical results such as the equations governing the overall dynamics, the stability condition for the case without backpressure. We also compute the tail asymptotics for proxy buffer occupancy and delays in the stationary regime and show that they are surprisingly heavy-tailed under certain natural statistical assumptions of the literature. Finally, Section 19.5 is devoted to discussions on future studies and concludes the paper.

19.2 Simulation scenarios

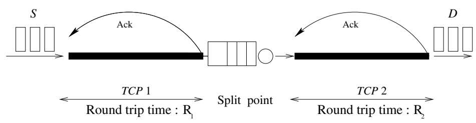

Fig. 19.1. split TCP network scheme.

19.2.1 Network scheme

We consider a saturated traffic source S that sends packets to the destination D through two long-lived TCP-Reno connections $TCP1$, $TCP2$ in cascade as in Figure 19.1. Due to the fact that S is saturated, $TCP1$ always has packets to send.

A layer-4 proxy is placed in the middle and forwards packets from the first link to the second one, sending local acknowledgments to S (LACKs). It prevents the loss of packets that cannot be immediately forwarded on the second link by storing them temporarily in a buffer. When the buffer approaches its maximal capacity, a backpressure mechanism limits the sender rate. The flow control is accomplished through the advertised window indication present on the acknowledgements sent back to the sender S. The transmission window of S is then regulated according to the minimum between the current congestion window and the advertised receiver window. Therefore, as the buffer occupancy approaches the buffer capacity, the backpressure algorithm timely starts working and prevents buffer overflows. In the *ns2* simulator, the backpressure algorithm is implemented by means of ack notifications to the sender of the instantaneous available space in the proxy buffer (RFC compliant).

19.2.2 Assumptions and notation

In the following we introduce the notation that will be used throughout the paper and the assumptions shared by the simulation setting and, thereinafter, by the model.

- The TCP connections are assumed to be in congestion avoidance phase, thus neglecting the initial slow start.
- $X(t)$, $Y(t)$ respectively denote $TCP1$, $TCP2$ rates at time t.
- The proxy buffer has size B. We will generally assume a limited buffer size, though the mathematical part also considers the ideal case of $B = \infty$.
- The local round trip times, R_1, R_2 (of $TCP1$, $TCP2$, respectively) are assumed to be constant, equal to twice the local propagation delay.
- Losses are modelled by two kinds of Poisson processes:
  - homogeneous Poisson processes with constant intensities λ_0, μ_0, which will be referred to as the **rate independent** (RI) case.
  - inhomogeneous Poisson processes with (stochastic) intensities $\lambda_1 X(t)$, $\mu_1 Y(t)$, proportional to the rates $X(t)$ and $Y(t)$, a case that will be referred to as the **rate dependent** (RD) case;

Concerning the loss process assumptions, the RI case corresponds to the physical layer loss pattern of wireless links (fast/slow fading) or some DSL links (multi-users interference), whereas the RD case fits well with the cases where there is a PER (packet error rate) due to congestion (cf. [7]). For example, the self congestion that arises in slow access links. In addition, these models allow one to consider at the same time both transmission error losses and congestion losses. Some interesting models are hybridations of the above simple cases. Here is a typical example to be used in what follows: $TCP1$ is a fast-wired link with a RD loss process and $TCP2$ is a slow DSL or wireless link with RI losses.

19.2.3 Scenarios

The following three scenarios focus on a cascade of two TCP connections. They correspond to different network settings, and contexts of application of split TCP.

- *The slow sender fast forwarder case (SF).*
  When the first TCP connection is "slower" than the second one, i.e. it has a smaller capacity and/or longer RTT and/or higher loss rate, we are in what we call the "SF" scenario. It can be the case in overlay networks, where the traffic of the first TCP connection is forwarded on a faster TCP connection.
- *The fast sender slow forwarder case (FS).*
  We call the "FS" case, the scenario where $TCP1$ is "faster" than $TCP2$. It is the case in hybrid wired/wireless scenarios where the faster reliable wired connection forwards its traffic to a slower lossy wireless connection. In the case of a wired/satellite configurations, in addition, the wireless part is characterized by higher propagation delays. The example where $TCP1$ is a fast link with a RD loss process and $TCP2$ is a slow one with RI losses will be referred to as the FS-RD/RI example.
- *The symmetric case.*
  In the wired/wireless cascade, the two TCP connections are strongly asymmetric, as the two media are notably different. In overlay networks, instead, it can happen for a long TCP connection to be split in smaller symmetric segments. In this setting, the two links have about the same characteristics.

19.2.4 Performance metrics

The majority of related work on split TCP are experimental evaluations of the throughput improvement achieved by splitting connection techniques. In addition to the throughput metric, our work is focused on the analysis of buffer occupancy in the proxy and on packet delays.

19.3 Simulation results

In this section we present a set of *ns2* simulations to illustrate the temporal evolution of congestion windows of $TCP1$ and $TCP2$ as well as the proxy buffer occupancy in the three cases mentioned above.

19.3.1 Rates

Let us start with a rather "symmetric" case, where the links have the same rate, $C_1 = C_2 = 100\,\text{Mbps}$, similar propagation delays, $R_1 = 100\,\text{ms}$, $R_2 = 90\,\text{ms}$, and where losses are generated according to homogeneous (RI) Poisson processes with intensity $\lambda = \mu = 0.3\,\text{losses/s}$. The proxy buffer can store $B = 15\,\text{pkts}$ and we assume a constant packet size equal to 1500 bytes. Figure 19.2 shows the *ns2* simulation of the congestion window patterns in Congestion Avoidance phase, together with the buffer occupancy $Q(t)$ in the proxy.

Looking at the buffer dynamics in Figure 19.2, we remark that:

Observation 19.1 The rate of $TCP1$ and $TCP2$ interact through the buffer occupancy. One can distinguish three operational phases:

(PH1) as long as the buffer is neither empty nor full, *TCP* rates X, Y follow the *AIMD* rule, i.e. they linearly increase until a jump occurs and halves the rate;

(PH2) the rate of $TCP2$, Y exhibits a nonlinear growth when the buffer is empty;

(PH3) the rate of $TCP1$, X exhibits a nonlinear growth when the buffer approaches saturation.

In Figures 19.3 and 19.4, we plot the congestion windows and the proxy buffer occupancy in the *SF* and *FS* cases, respectively. In the *SF* case we maintain the same links capacities ($C_1 = C_2 = 100\,\text{Mbps}$), the round trip times are $R_1 = 90\,\text{ms}$, $R_2 = 30\,\text{ms}$, and loss intensities are $\lambda = \mu = 0.4\,\text{losses/s}$ (still under the assumption of RI losses). The *FS* case is characterized by $R_1 = 40\,\text{ms}$, $R_2 = 80\,\text{ms}$, and $\lambda_0 = 0.4\,\text{losses/s}$, $\mu_0 = 0.2\,\text{losses/s}$. In both cases, the proxy buffer size is $B = 20\,\text{pkts}$. We observe that in both cases, the dynamics of one of the three possible phases can be neglected w.r.t. the others. In the *FS* scenario the buffer

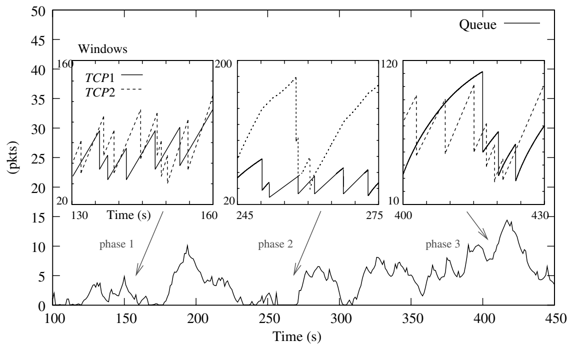

Fig. 19.2. Congestion windows and buffer occupancy in the symmetric case.

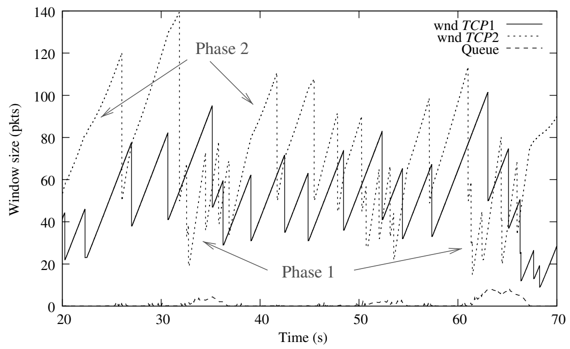

Fig. 19.3. Congestion windows and buffer occupancy in the SF case.

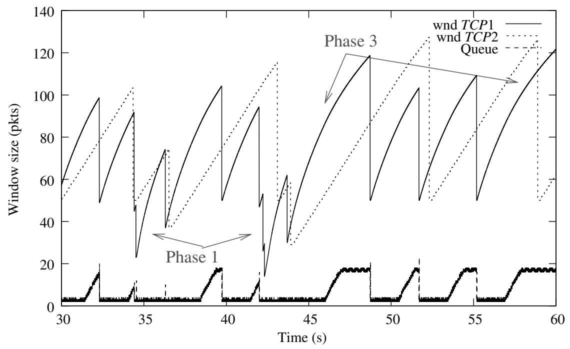

Fig. 19.4. Congestion windows and buffer occupancy in the FS case.

Table 19.1. *Impact of buffer size on stationary throughput averages.*

B (pkt)	Mean throughput (pkt/s)	Throughput improvement
1	780	—
5	1415	81%
10	1840	136%
20	2285	193%
50	2954	278%
100	3340	328 %
1000	3340	328 %

hardly ever empties, therefore the duration of phase 2 is negligible when compared to the other phases, whereas in the SF it is the opposite scenario: the buffer is rarely close to saturation, hence phase 3 is almost never visible. Therefore, we will consider later SF/FS cases with large buffers where only phases 1–2 and 1–3 are taken into account, respectively.

The nonlinear behavior of X or Y bears evidence of the fact that the two TCP connections interact. In particular, the window of $TCP2$ (and therefore its rate) evolves not only according to the windows dynamics of TCP, but also according to the availability of packets in the proxy. Similarly, the window of $TCP1$ (and therefore its rate) evolves not only according to the AIMD rule, but also according to the availability of space in the receiver buffer, advertised by the proxy to the source S via the backpressure mechanism.

In observation 19.1, we have already remarked the role of the buffer content on the interaction between $TCP1$ and $TCP2$. In Table 19.1 we report the mean values of total throughput in steady state for different values of the buffer size in a SF case. Here $C_1 = C_2 = 100\,\text{Mbps}$, $R_1 = 40\,\text{ms}$, $R_2 = 20\,\text{ms}$, $\lambda_0 = \mu_0 = 0.4\,\text{losses/s}$. Compared to the extreme case of $B = 1$ pkt, we remark a large improvement of total throughput as buffer size increases. At the other extreme, when B is large enough to never saturate, the total thoughput reaches its maximum value, which corresponds to the throughput of $TCP1$ in isolation (the last is computed via the mean throughput formula in Section 19.4.2). An important consideration follows:

Observation 19.2 In the RI case, the end-to-end throughput of split TCP increases with the proxy buffer size.

A large buffer size is beneficial to the total throughput, in that it makes it less common to use the flow control of the backpressure mechanism and it reduces the probability for the buffer to empty, which limits the forwarding rate.

On the TCP rates we observe that:

Observation 19.3 In the finite buffer backpressure case, the long term average of $TCP1$ coincides with that of $TCP2$ and is strictly smaller than that of each connection in isolation.

This is in contrast with the throughput estimation provided in [15], [26] and [27], through the application of the square root formula. These works rely on the assumption that the two connections are independent and evaluate the overall throughput as the minimum of the throughputs of each connection taken in isolation. As shown in the last table, such a minimum rule is in fact the best case and a significant throughput degradation can be observed w.r.t. this rule in the presence of small buffer size.

The simulation results presented so far share the assumption of RI losses, though all considerations still hold for the RD case. We report in Figure 19.5 a symmetric RD scenario with the following parameters: $C_1 = C_2 = 100\,\text{Mbps}$, $R_1 = R_2 = 60\,\text{ms}$, $\lambda_1 = \mu_1 = 0.03$, where all three phases can be observed.

19.3.2 Buffer occupancy

We present here the statistical analysis of the buffer occupancy in steady state and when $B = \infty$. In order to guarantee the existence of a steady

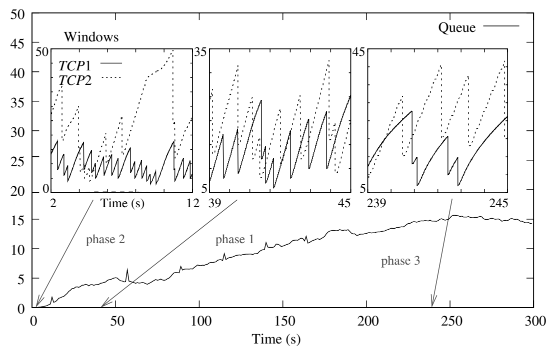

Fig. 19.5. Congestion windows and buffer occupancy in the RD loss case.

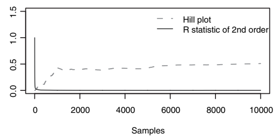

Fig. 19.6. Hill plot and R statistic for the queue tail.

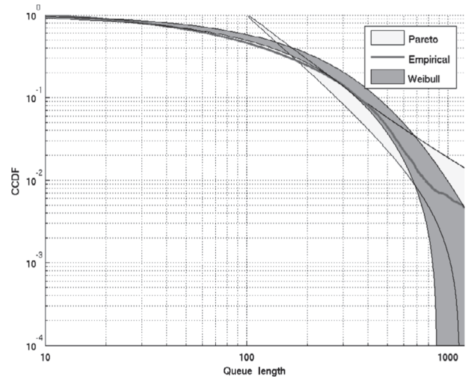

Fig. 19.7. α-Confidence intervals with $\alpha = 0.95$.

state, we consider the SF case (we will give in due time the exact conditions for such a steady state to exist). To infer some preliminary information on the queue distribution, we made a fit with the *R* software (free clone of *S*-plus) of the stationary queue distribution from the samples extracted via *ns2* simulation.

19.3.2.1 RI case

In the RI case, the *R* software suggests a Weibull distribution with shape parameter 0.5; the last parameter was obtained through a maximum likelihood estimator. We can then conjecture that:

Observation 19.4 In the RI case with $B = \infty$, when there is a stationary regime, the stationary buffer occupancy exhibits a heavy tail.

The presence of possible heaviness in the queue tail motivated a further inspection of the moments of Q by means of statistical methods suitable for heavy-tailed distributions.

For this purpose, two statistics were employed: the Hill plot and the R statistic (see definitions an further details of the analysis in [6]).

In Figure 19.6 we plot the parameter γ of the Hill estimator in the above SF scenario. It rapidly converges to a value between 0 and 1, which indicates a Pareto-like distribution with infinite second moment. In contrast with this result, the same figure also shows the R statistic for the second moment of the tail distribution, computed on the same set of samples; the fact that it becomes zero supports the thesis of the finiteness of the second moment. The discrepancy between these two results can be explained by looking at α-confidence intervals for the Pareto distribution or for the above-mentioned Weibull distribution in this example. We observe in Figure 19.7 that the 0.95-confidence intervals of a Pareto distribution with $\alpha = 0.5$ and a Weibull distribution with shape parameter 0.5 largely overlap in a way which compromises the inference of the tail distribution. In conclusion, we showed that these statistical methods aimed at the identification of the shape of these heavy tails provide discordant answers.

19.3.2.2 RD case

In the RD case both the Hill plot and the R statistic agree that the distribution of the buffer occupancy has all finite moments. We report in Figure 19.8, the Hill plot and the R statistic for the second moment of the distribution in a SF scenario with such RD losses. The results of statistical test in the RD case allows us to observe that:

Observation 19.5 In the RD case, the buffer occupancy exhibits a light tail decay.

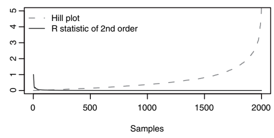

Fig. 19.8. Hill plot for the RD case.

19.4 Mathematical analysis

In this part, we first establish the differential equations which govern the joint evolution of the windows and the rates of the connections and of the buffer content. We then give formal proofs for some of the experimental observations made in the previous sections and add further results. More precisely:

- We address the stability issue, in case of infinite buffers and no backpressure (which is the only case where the stability question is of interest).
- We analyze the tail asymptotics for the buffer occupancy as well as for end-to-end delays.
- We prove that in the finite buffer, backpressured case, the stationary throughput is strictly less than the minimum of the stationary throughput of each connection in isolation.

19.4.1 The single connection model

Let us now briefly revisit the models of [8] and [7] for a single TCP connection. In the congestion avoidance phase, the evolution of the TCP congestion window is described through the following “hybrid” differential equation:

$$\mathrm{d}W(t) = \frac{\mathrm{d}t}{R} - \frac{W(t)}{2} N(\mathrm{d}t) \tag{19.1}$$

which states that the window increase between two loss events is linear with slope $1/R$, where R denotes the round trip time and that loss events produce jumps of the congestion windows which is cut by half. In this stochastic differential equation, $N(t)$ represents the loss point process. We assume here that this point process has a stochastic intensity (w.r.t. the natural history of $W(t)$ – see [4] for the definition of stochastic intensity), which is either constant or proportional to $W(t)$ depending on the case we consider (RI/RD).

Linked to the congestion window we define the instantaneous rate or throughput (here the two words are interchangeable), as $X(t) = W(t)/R$, a generally accepted assumption in the literature, which can be seen as an avatar of Little’s law. The rationale for the linear increase is as follows: TCP stipulates that in the congestion avoidance phase, the window is increased of 1 unit every W ack. In an infinitesimal interval of length $\mathrm{d}t$, the number of acks that arrive is $X(t)\mathrm{d}t$. Hence the window increases of $X(t)\mathrm{d}t/W(t) = \mathrm{d}t/R$. The rationale for the halving of the window in case of a loss is just the multiplicative decrease rule. Systems evolving according to eq. (19.1) were also studied in [22] and [1].

19.4.2 Mean throughput in steady state

The stationary distribution of the rate of each TCP connection *in isolation* has an analytical expression. In particular, the mean throughput of $TCP1$ (resp. $TCP2$), in isolation is given by $\bar{X} = 2\alpha/\lambda$ [resp. $\bar{Y} = 2\beta/\mu$], where $\alpha = 1/R_1^2$, $\beta = 1/R_2^2$. This result follows from the fact that $X(t) - \alpha t + \frac{\lambda}{2}\int_0^t X(u)du$ is a martingale (cfr.[6]). Thanks to the PASTA property, the mean value $\mathrm{E}_N^0[X(0-)]$ of the stationary rate just before a loss is equal to stationary rate in continuous time $\bar{X}$. Using this and the fact that the packet loss probability is $p = \lambda/\bar{X}$, we obtain the square root formula

$$\bar{X} = \sqrt{\frac{2\alpha}{p}}, \quad \left(\text{resp. } \bar{Y} = \sqrt{\frac{2\beta}{q}}\right). \tag{19.2}$$

In [8], the following square root formula is derived for the RD case:

$$\bar{X} = \Phi\sqrt{\frac{\alpha}{\lambda}}, \quad \left(\text{resp. } \bar{Y} = \Phi\sqrt{\frac{\beta}{\mu}}\right), \tag{19.3}$$

where $\Phi = \sqrt{\frac{2}{\pi}\frac{\sum_{i=0}^{\infty}(\prod_{j=1}^{i}(1-4^j))^{-1}}{\sum_{i=0}^{\infty}(\prod_{j=1}^{i}(1-4^j))^{-1}2^i}} \approx 1.309$. The mean throughput formulas in the RI and RD case have also appeared respectively in [2] and [24].

19.4.3 The split connection model

Let now introduce the stochastic equations for the split connection model.

With respect to the case of a single TCP connection, as observed in observation 19.1 there are three operational phases.

- Phase 1 or the *free phase*, where the buffer is neither empty nor full, and $X(t)$ and $Y(t)$ evolve independently;
- Phase 2 or the *starvation phase*, when the buffer is empty and Y is limited by the input rate X.
- Phase 3 or the *backpressure phase*, when the buffer has reached its storage capacity B and X is forced by the backpressure algorithm to slow down to the rate Y at which the buffer is drained off.

In the free phase, the AIMD rule gives:

$$\text{on } \{0 < Q(t) < B\} \quad \begin{cases} \mathrm{d}X(t) = \alpha \mathrm{d}t - \frac{X(t)}{2}M(\mathrm{d}t) \\ \mathrm{d}Y(t) = \beta \mathrm{d}t - \frac{Y(t)}{2}N(\mathrm{d}t). \end{cases} \tag{19.4}$$

In the starvation phase, as long as the buffer is empty (which requires that $X(t) \leq Y(t)$), we have:

$$\text{on} \quad \{Q(t) = 0\} \quad \begin{cases} \mathrm{d}X(t) = \alpha \mathrm{d}t - \frac{X(t)}{2} M(\mathrm{d}t) \\ \mathrm{d}Y(t) = \beta \frac{X(t)}{Y(t)} \mathrm{d}t - \frac{Y(t)}{2} N(\mathrm{d}t). \end{cases} \tag{19.5}$$

where $M(t)$, $N(t)$ represent the loss processes on X and Y. The rationale for a linear increase of $Y(t)$ proportional to $\frac{X(t)}{Y(t)} < 1$ is that when the buffer is empty, since $X(t) < Y(t)$, the rate at which packets are injected in $TCP2$ and hence the rate at which $TCP2$ acks arrive is $X(t)$. Hence the window of $TCP2$, W_2, increases of $X(t)dt/W_2(t) = dt\frac{X(t)}{R_2 Y(t)}$ in the interval $(t, t+dt)$ and the rate of $TCP2$ thus increases of $\beta dt \frac{X(t)}{Y(t)}$ during this interval. The ratio $X(t)/Y(t)$ can be interpreted as the *utilization factor* of the congestion window $W_2(t)$: in contrast with what happens in the free phase, where $TCP2$ is “independent” of $TCP1$ and where the window W_2 is fully utilized (draining packets from the buffer), in the starvation phase, the number of packets transmitted by $TCP2$ depends on $X(t)$, which brings the utilization factor below 1 and leads to a nonlinear evolution as observed in observation 19.1.

In the backpressure phase, which lasts until the buffer is saturated (this requires that $X(t) \geq Y(t)$), we have

$$\text{on } \{Q(t) = B\} \quad \begin{cases} \mathrm{d}X(t) = \alpha \frac{Y(t)}{X(t)} \mathrm{d}t - \frac{X(t)}{2} M(\mathrm{d}t) \\ \mathrm{d}Y(t) = \beta \mathrm{d}t - \frac{Y(t)}{2} N(\mathrm{d}t). \end{cases} \tag{19.6}$$

The rationale for this should be clear: acks of $TCP1$ now come back at a rate of $Y(t)$. Hence the congestion window, $W_1(t)$ of $TCP1$ grows at the rate $Y(t)/W_1(t)$.

The evolution of the buffer occupancy in the proxy within phase 1 ($Q(t) > 0$) is given by

$$Q(t) = Q(0) + \int_0^t (X(u) - Y(u))\mathrm{d}u. \tag{19.7}$$

Note that the queue at the proxy can be seen as a fluid queue with a fluid input rate $X(t)$ and a fluid output rate $Y(t)$ at time t. Hence we can also write:

$$\begin{aligned} Q(t) = \max\Bigg(& \sup_{0 \leq u \leq t} \int_u^t (X(v) - Y(v))\mathrm{d}v, \\ & Q(0) + \int_0^t (X(u) - Y(u))\mathrm{d}u \Bigg). \end{aligned} \tag{19.8}$$

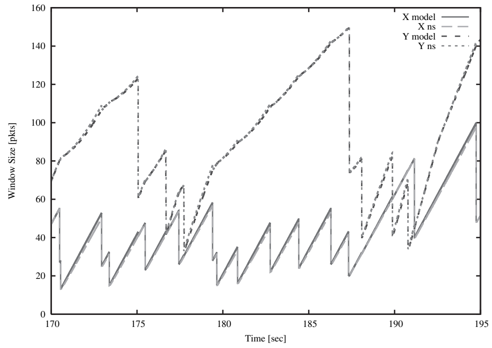

Fig. 19.9. Congestion windows: comparison between the model and *ns2*.

Figure 19.9 shows the perfect agreement between the evolution of congestion windows as predicted by equations (19.4)–(19.7) or provided by *ns2* simulations.

19.4.4 Stability in the infinite buffer, RI case

This subsection is focused on the case $B = \infty$, where only phase 1 and 2 exist. A natural question within this infinite buffer context is that of *system stability*, which is understood here as the finiteness of the stationary buffer occupancy at the proxy. In the stable case, this infinite buffer model is a good approximation of the SF case, whenever B is not too small (we have seen in Section 19.3 that in this case, phase 3 is almost never visited and that we can focus on phase 1/phase 2 dynamics). In the RI case, the stability proofs rely on some monotonicity properties which we introduce in the following paragraph.

19.4.4.1 Monotonicity properties in the RI case

Let us first consider sample paths of $X(t)$ and $Y(t)$ in case $B = \infty$, where the dynamics is governed by (19.4) or (19.5) depending on $Q(t)$.

We denote by Y^f ("free Y") some fictitious process which evolves according to the dynamics of phase 1 only. In the RI case, we can actually choose to make a *coupling* of Y^f and Y by building these two processes from the *same realization* of the Poisson point process N.

We can then state three main properties.

- If we consider two processes, $Y^f(t), \widehat{Y}^f(t)$ based on the same realization of N, but departing from different initial conditions, $Y^f(0) \leq \widehat{Y}^f(0)$, then, $Y^f(t) \leq \widehat{Y}^f(t), \forall t \geq 0$.

- If we consider the process $Y^f_v(t), t \geq v$ which starts from 0 at time v, then $Y^f_{v_1}(t) \geq Y^f_{v_2}(t)$, for all $v_1 < v_2 \leq t$.
- If $Y^f(0) = Y(0)$, then $Y(t) \leq Y^f(t)$, for all $t \geq 0$.

The proofs of the first two properties should be clear. The last one follows from the fact that in phase 2, $X(t) \leq Y(t)$, so that at any continuity point, the slope of $Y(t)$ is always less than or equal to the slope of $Y^f(t)$. Since both processes have the same discontinuity points (thanks to the coupling), the result immediately follows.

Consider now the case B finite with backpressure. The triple $(X(t), Y(t), Q(t))$ forms a continuous time Markov process. Thanks to the fact that Q is bounded from below by 0 and from above by B, one can show that this Markov process admits a unique stationary distribution, and that, starting from any initial value, this Markov process converges to the stationary one in the total variation norm.

We again compare the processes $(X(t), Y(t))$ in the split TCP system and the free processes $(X^f(t), Y^f(t))$. We build these processes on the same realizations of the point processes M and N. Assume the initial conditions to be the same: $(X(0), Y(0)) = (X^f(0), Y^f(0))$. When we are in phase 1, the two processes have exactly the same dynamics, so that if at the beginning of the phase, $(X(.), Y(.)) \leq (X^f(.), Y^f(.))$ coordinatewise, then this holds true at the end of the phase too. In phase 2 (resp. 3), the slope of Y (resp. X) is strictly less than that of Y^f (resp. X^f) and both are halved at the same epochs, whereas X and X^f (resp. Y and Y^f) have exactly the same dynamics. Hence, if $(X(.), Y(.)) \leq (X^f(.), Y^f(.))$ at the beginning of the phase, then $(X(.), Y(.)) < (X^f(.), Y^f(.))$ at the end. This leads to the following confirmation of observation 19.3:

Lemma 19.6 *In the RI case with $B < \infty$, when backpressure is used, the stationary rate of split TCP is* strictly *less than the minimum of that of $TCP1$ and $TCP2$ in isolation.*

19.4.4.2 Queue bounds

The triple $(X(t), Y(t), Q(t))$ forms a Markov process. Interestingly enough, the direct stability analysis of this Markov via Liapunov functions is not an easy task. In particular, we were unable to make use of the classical fluid limit techniques for Markov chains here, primarily because of the multiple phases. This is the reason why we use backward construction techniques to prove stability. This will be done by introducing two simple bounds on the queue size.

The proposed backward construction (see e.g. [4] Chapter 1 for classical instances of such constructions) consists in building the queue size $Q_t(0)$ at time 0 when departing from an appropriate initial condition at time $t < 0$. The initial condition that we select consists of a queue size $Q(t) = 0$, a rate for $TCP1$ which is the stationary rate $\widetilde{X}(t)$ of $TCP1$ at time t in isolation, and a rate for $TCP2$ which is the stationary rate $\widetilde{Y}^f(t)$ of $TCP2$ at time t in isolation. From (19.8), we have

$$Q_t(s) = \sup_{t \le u \le s} \int_u^s (\widetilde{X}(v) - Y_t(v))\mathrm{d}v, \tag{19.9}$$

for all $s \ge t$, where $Y_t(v)$ denotes the rate of $TCP2$ in the split TCP system and at time v under the above assumptions.

The stability issue can then be stated in the following terms: does $Q_t(0)$ have an almost surely (a.s.) finite limsup when t tends to $-\infty$? This is enough to ensure that the Markov chain $(X(t), Y(t), Q(t))$ is neither transient nor null recurrent.

We are now in a position to define the lower bound queue. From monotonicity property 3 and from (19.9), we get

$$Q_t(0) \ge L_t = \sup_{t \le u \le 0} \int_u^0 (\widetilde{X}(v) - \widetilde{Y}^f(v))\mathrm{d}v, \tag{19.10}$$

where $\widetilde{Y}^f(.)$ is the stationary free process for $TCP2$. In [6], we prove that the stochastic process $(\widetilde{X}(t), \widetilde{Y}^f(t))$, which describes the fluid input and the fluid drain in this queue, forms a stationary and geometrically ergodic Harris chain. In particular we show there that we can apply the splitting technique of Athreya and Ney (cfr.[3]) for such chains and that there exist renewal cycles for this process related to its return times to the compact set $C = [0, x] \times [0, \frac{2\beta}{\alpha}x]$, where x is an arbitrary positive real number. In what follows, we will denote by T the length of such a renewal cycle. We now define the upper bound queue. Let $\tau(t)$ denote the beginning of the last busy period of $Q_t(s)$ before time 0 (0 if $Q_t(0) = 0$ and t if $Q_t(s) > 0$ for all $t < s \le 0$). We have

$$\begin{aligned} Q_t(0) &= \int_{\tau(t)}^0 (\widetilde{X}(v) - Y_t(v))\mathrm{d}v \le \int_{\tau(t)}^0 (\widetilde{X}(v) - Y^f_{\tau(t)}(v))\mathrm{d}v \\ &\le U_t = \sup_{t \le u \le 0} \int_u^0 (\widetilde{X}(v) - Y^f_t(v))\mathrm{d}v, \end{aligned} \tag{19.11}$$

where the first inequality follows from the fact that the dynamics on $(\tau(t), 0)$ is that of the free phase and from the fact that $Y^f_{\tau(t)}(.)$ is the minimal value for the free $TCP2$ process (monotonicity property 1).

19.4.4.3 Stability

Lemma 19.7 *If $\rho < 1$, where $\rho = \alpha\mu_0/\beta\lambda_0$, then the RI system is stable. If $\rho > 1$, then it is unstable.*

Proof We prove first that if $\rho > 1$, then the system is not stable. The equation for L_t is that of a classical fluid queue with *stationary and jointly ergodic* arrival and service processes. The joint ergodicity follows from the fact that the couple $(\widetilde{X}, \widetilde{Y}^f)$ forms a Harris recurrent and geometrically ergodic Markov process (see [6]). We can hence apply classical results on fluid queues stating that under the above stationarity and ergodicity properties, if $\mathrm{E}[\widetilde{X}(0)] > \mathrm{E}[\widetilde{Y}^f(0)]$, then L_t tends a.s. to ∞, which in turn implies that we cannot have $\limsup_{t\to\infty} Q_t(0)$ a.s. finite. Hence $\rho > 1$ implies instability.

We now prove that if $\rho < 1$, then the system is stable.
Assume that $\limsup Q_t(0) = \infty$ with a positive probability. Then $\limsup U_t = \infty$ with a positive probability too. As $\widetilde{X}$ and Y^f_t are locally integrable for all t, this together with the second monotonicity property of the last subsection imply that there exists a sequence t_n tending to $-\infty$ and such that a.s.

$$\int_{t_n}^{0} (\widetilde{X}(v) - Y^f_{t_n}(v))\mathrm{d}v \to_{n\to\infty} \infty. \qquad (19.12)$$

Let us show that this is not possible under the assumption $\rho < 1$. Let θ_t denote the product shift of the point processes M and N (this shift is ergodic). The pointwise ergodic theorem implies that

$$\frac{1}{t}\int_{-t}^{0} \widetilde{X}(v)\mathrm{d}v = \frac{1}{t}\int_{-t}^{0} \widetilde{X}(0) \circ \theta_v \mathrm{d}v \to_{t\to\infty} \mathrm{E}[\widetilde{X}(0)], \qquad (19.13)$$

where the last limit is in the a.s. sense. We show now that the following a.s. limit also holds:

$$\frac{1}{t}\int_{-t}^{0} Y^f_{-t}(v)\mathrm{d}v \to_{t\to\infty} \mathrm{E}[\widetilde{Y}^f(0)]. \qquad (19.14)$$

This will conclude the proof since equations (19.13)–(19.14) and the assumption

$$\mathrm{E}[\widetilde{X}(0)] < \mathrm{E}[\widetilde{Y}^f(0)]$$

imply that a.s. $\lim_{t\to\infty} \int_{-t}^0 (\widetilde{X}(v) - Y_{-t}^f(v))\mathrm{d}v = -\infty$, which contradicts (19.12).

Let us now prove (19.14). From the monotonicity properties, the function $\varphi_t = \int_{-t}^0 Y_{-t}^f(v)\mathrm{d}v$ is super-additive: $\varphi_{t+s} \geq \varphi_t \circ \theta_{-s} + \varphi_s$. Thanks to the sub-additive ergodic theorem, this together with the fact that φ_t is integrable imply that a.s. $\exists \lim_{t\to\infty} \frac{1}{t}\int_{-t}^0 Y_{-t}^f(v)\mathrm{d}v = K$, for some constant K which may be finite or infinite. The fact that K is finite follows from the bound $0 < Y_{-t}^f(v) \leq \widetilde{Y}^f(v)$ and from the pointwise ergodic theorem applied to the stationary and ergodic process $\{\widetilde{Y}^f(v)\}$. Since K is finite, the last limit holds both a.s. and in L^1 [21]. Using again super-additivity of the forward process, we get by the same arguments:

$$K = \lim_t \frac{1}{t}\int_{-t}^0 Y_{-t}^f(v)\mathrm{d}v = \lim_t \mathrm{E}\left(\frac{1}{t}\int_0^t Y_0^f(v)\mathrm{d}v\right).$$

But from the fact that $Y_0^f(v)$, $v \geq 0$ is a geometrically ergodic Markov chain,

$$\exists \lim_{t\to\infty} \frac{1}{t}\int_0^t Y_0^f(v)\mathrm{d}v = \mathrm{E}[\widetilde{Y}^f(0)] \quad \text{a.s.}$$

Hence K=E$[Y^{\tilde{f}}(0)]$, which concludes the proof of (19.14). □

19.4.5 Stability in the infinite buffer, RD case

In the RD case, we use the same backward construction as above to prove the exact analogue of Lemma 19.7 (in the RD case, ρ is equal to $\alpha\mu_1/\beta\lambda_1$). We only sketch the main ideas of the proof. To get an upper bound on $Q_t(0)$, we consider the following optimization problem: what is the infimum over all $y > 0$ of the integral $\int_u^0 Y_{u,y}^f(v)\mathrm{d}v$ where $Y_{u,y}^f(v)$ denotes the value of the free process of $TCP2$ at time $v \geq u$ when starting from an initial value of y at time u? Let us first show that the last optimization problem admits an a.s. unique solution $y^*(u)$, and that this solution is a.s. finite.

For defining such an infimum, we need the following construction which builds the stochastic processes $Y_{u,y}(v)$, $v \geq u$ from a two dimensional homogeneous Poisson point process $\mathcal{N}$ of intensity μ on

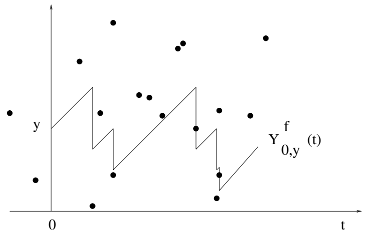

Fig. 19.10. Coupling of the RD processes.

the positive half plane ($t \in \mathrm{R}, y \in \mathrm{R}_+$). We start with $Y^f_{u,y}(u) = y$ and then have a linear growth of slope β until we find a point of $\mathcal{N}$ below the curve $Y^f_{u,y}(.)$. There we halve the value of $Y^f_{u,y}$ at that time and we proceed again using the same rule of a linear growth until the next point below the curve (see Fig. 19.10). It is easy to see that the stochastic intensity of the losses is exactly $\mu Y^f_{u,y}(t)$ at time t, which is precisely what we want. With this construction, all $Y^f_{u,y}(t)$ are defined as deterministic functions of $\mathcal{N}$ and the infimum over all y of $\int_u^{u+t} Y^f_{u,y}(v)\mathrm{d}v$ for fixed $t > 0$ and u is well defined in view of the fact that the function $y \to \int_u^{u+t} Y^f_{u,y}(v)\mathrm{d}v$ is piecewise continuous, with a finite number of discontinuities in any compact, is increasing between discontinuities and tends to ∞ when y tends to ∞. Denote by $Y^*_u(v)$ the function $Y^f_{u,y^*(u)}(v)$. Hence

$$Q_t(0) \leq U_t = \sup_{t \leq u \leq 0} \int_u^0 (\widetilde{X}(v) - Y^*_u(v))\mathrm{d}v. \tag{19.15}$$

The arguments to prove stability when $\rho < 1$ are then similar to those in the RI case: when t tends to ∞, $\frac{1}{t}\int_{-t}^0 \widetilde{X}(v)\mathrm{d}v$ tends to $\mathrm{E}[\widetilde{X}(0)]$ a.s. from the pointwise ergodic theorem. The function $\varphi_t = \int_{-t}^0 Y^*_{-t}(v))\mathrm{d}v$ is super-additive. We then use the sub-additive ergodic theorem to prove that $\frac{1}{t}\varphi_t$ tends to a constant K a.s. and the pointwise ergodic theorem again to show that this constant is necessarily $\mathrm{E}[\widetilde{Y}^f(0)]$.

The proof of the last property relies on the following two ingredients: (a) for all y, with probability 1, there exists a positive random variable $\epsilon(y) > 0$ such that the functions

$$y \to g_t(y) = \frac{1}{t}\int_0^t Y^f_{0,y}(v)\mathrm{d}v$$

are t-uniformly continuous; (b) let $y^o(t)$ be the initial condition that minimizes $\int_0^t Y^f_{0,y}(v)\mathrm{d}v$; the liminf of the function $y^o(t)$ as t tends to ∞

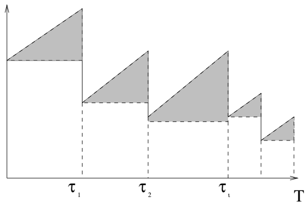

Fig. 19.11. Decomposition of the integral of X in a sum of trapezes.

is 0. From (b), we deduce that there exists a subsequence t_n such that $y^o(t_n)$ converges to 0 a.s. It is easy to see that $g_{t_n}(y^o(t_n))$ converges to K as n tends to infinity. But $g_{t_n}(0)$ converges to $\mathrm{E}\widetilde{Y}^f(0)$ due to the ergodicity of the Harris chain $\widetilde{Y}^f(.)$. This together with the continuity property (a) allow one to conclude that

$$K = \lim_{n\to\infty} g_{t_n}(y^o(t_n)) = \lim_{n\to\infty} g_{t_n}(0) = \mathrm{E}\widetilde{Y}^f(0) \quad \text{a.s.}$$

19.4.6 Tail asymptotics in the RI case

Here, we take up the observation 19.4 and confirm it with the following theoretical result.

Lemma 19.8 *In the RI case, the queue distribution is heavier than a Weibull distribution of shape parameter $k = 0.5$.*

This unexpected result suggests that the buffer occupancy in split TCP is not negligible. It also explains the quite important fluctuations observed on end-to-end delays within this context (Fig. 19.11).

We give the proof of this result in [6]. Let us summarize here the main steps of the proof. It relies on the lower bound of equation (19.10), and it is based on the fact that the fluid input process and the fluid draining process of this lower bound queue are jointly stationary and ergodic and have renewal cycles (see [6]). We denote by T the length of such a renewal cycle and we define

$$\Delta = \int_0^T \widetilde{X}(t) - \widetilde{Y}^f(t)\mathrm{d}t = I_x - I_y.$$

We first study the asymptotics for $\mathbf{P}(\Delta > x)$ as $x \to \infty$. We show that this is lower-bounded by random variables with a Weibull distribution with shape parameter 1/2. Hence Veraverbeke's theorem ([13]) can be used to show that Q is heavier than a Weibull distribution with shape parameter 1/2. We now provide an intuitive explanation of the result on the tail of Δ. By looking at I_x, we observe that each trapeze area has a triangular lower bound, so that

$$\mathbf{P}\left(\sum_{0}^{N_T} Trap_i > q\right) \geq \mathbf{P}\left(\sum_{0}^{N_T} \alpha\frac{\tau_i^2}{2} > q\right),$$

where N_T denotes the number of losses in the cycle. All triangular areas are i.i.d and heavy tailed: as τ_i are i.i.d exponentially distributed, each summand has a tail distribution

$$\mathbf{P}\left(\alpha\frac{\tau^2}{2} > x\right) = \mathbf{P}\left(\tau > \sqrt{\frac{2x}{\alpha}}\right) = \mathrm{e}^{-\mu\sqrt{\frac{2x}{\alpha}}},$$

which is Weibull with shape parameter $k = 0.5$. Thanks to the properties of subexponential distributions, the result applies to the integral of $\widetilde{X}$, and then propagates to Q. It is worth noting that such a result is not affected in practice by a limited congestion window (max congestion window). Indeed, if we decompose the integral of $\widetilde{X}$ as above, it is easy to see that the heavy tailness yet arises due to the presence in the sum of some terms like $\tau_i^2/2$ linked to the periods when the window is not clipped to the maximum value.

The communication literature contains many instances of heavy-tailed queues. The most famous example is probably that of a FIFO queue subject to a self-similar input process; it is proved in [23] that the stationary buffer content is then heavy tailed; it is also well known (see e.g. [12]) that such self-similar input processes arise in connection with HTTP traffic, when represented as the superposition of a large number of ON/OFF sources with Pareto on and/or off periods. In this example, heavy-tailed queues arise as a corollary of long-range dependence, which in turn is a consequence of the heavy tailedness of file sizes and off periods. In contrast, the heavy tailedness of the proxy contents in our split TCP model arises as a direct consequence of the AIMD dynamics under the assumption of a RI loss process. Such rate independent loss processes are quite natural in wireless or DSL lines, for example. Note however that the heaviness of the tail is linked to the loss model considered. In the RD case, arguments similar to those used in the RI case let us conjecture that the queue distribution is light, as suggested by observation 19.5.

19.4.7 Phases duality and End delays

The structure of equations (19.5), (19.6) points out the *duality* between phase 2 and 3: we can obtain one equation from the other by exchanging the roles of $X(t)$ and $Y(t)$ (and their parameters). In phase 3, the analogue of $Q(t)$ in phase 2 is what we can call the *antibuffer*: $A(t) = B - Q(t)$, the amount of the proxy buffer space available at time t. Based on this duality between the SF and FS scenarios, we can use the analysis of the tail asymptotics of Q in the SF case to evaluate $A(t) = B - Q(t)$ in the dual FS case.

Let us now look at end-to-end delays. This is the sum of three terms: the two local propagations delays and the proxy buffer waiting and forwarding time. In the FS case and when B is large enough, the processing delay of a packet arriving at time t at the proxy is well approximated by the queue length at time t divided by the mean value of the stationary service rate of $TCP2$, i.e.

$$D(t) \approx \frac{R_1}{2} + \frac{R_2}{2} + \frac{Q(t)}{\mathrm{E}[\tilde{Y}]} = \frac{R_1}{2} + \frac{R_2}{2} + \frac{B}{\mathrm{E}[\tilde{Y}]} - \frac{A(t)}{\mathrm{E}[\tilde{Y}]}. \qquad (19.16)$$

The fluctuations of $D(t)$ are then determined by those of $A(t)$. Duality shows that the fluctuations of $A(t)$ in this FS scenario are similar to those of $Q(t)$ in the SF case, namely are heavy tailed.

19.5 Conclusions

The first contribution of the paper is the set of equations (19.4)–(19.7) which, to the best of our knowledge, provides the first mathematical attempt for describing the dynamics of the split TCP system. Previous models neglect the dependence between the two connections, by assuming that the buffer at the intermediate node never empties nor saturates ([15]), whilst we have shown that there exist two phases, (ph. 2, 3) where the interaction between the TCP rates X and Y and buffer occupancy Q cannot be ignored. These equations also allowed us to show that the prediction of the expected throughput in steady state as provided by the square-root formula for the slowest TCP connection in isolation ([19], [27], [15], [26]) is *not* valid unless buffers are very large. Finally, these equations allowed us to identify situations in which the proxy buffer content has either heavy tails or important fluctuations which imply in turn important fluctuations for end-to-end delays. We also expect these equations to open a new analytical track for answering the following list of questions which remain open up to the writing of the present paper.

(i) In the finite buffer backpressured case,

(a) Is the split TCP stationary rate increasing in B as suggested by observation 19.2.

(b) What is the value of the stationary rate of the connection?

(ii) In the infinite buffer stable case,

(a) Is the stationary proxy buffer contents light tailed in the RD case?

(b) What is the distribution, or the mean value of the stationary buffer content?

Acknowledgments

This work was funded in part by the Euro FGI Network of Excellence.

Bibliography

[1] E. Altman, K. Avrachenkov, C. Barakat and R. Nunez-Queija. State-dependent M/G/1 type queueing analysis for congestion control in data networks. *Computer Networks*, **39**(6), 789–808, 2002.

[2] E. Altman, K. Avrachenkov and C. Barakat. A stochastic model of TCP/IP with stationary random losses. *IEEE/ACM Transactions on Networking*, **13**; 356–369, 2005.

[3] K. B. Athreya and P. Ney. A new approach to the limit theory of recurrent Markov chains. *Transactions of the American Mathematical Society*, 1978.

[4] F. Baccelli and P. Bremaud. *Elements of Queueing Theory*, 2nd ed. Wiley, 2003.

[5] F. Baccelli, A. Chaintreau, Z. Liu, A. Riabov and S. Sahu. Scalability of reliable group communication using overlays. *Proceedings of the IEEE Infocom*, February 2004.

[6] F. Baccelli, G. Carofiglio and S. Foss. Proxy Caching in split TCP: Dynamics, Stability and Tail Asymptotics, *INRIA Report*, July 2007.

[7] F. Baccelli, K. B. Kim and D. De Vleeschauwer. Analysis of the competition between wired, DSL and wireless users in an access network. In *Proceedings of IEEE Infocom*, Miami, FL, USA, March 2005.

[8] F. Baccelli, D. McDonald and J. Reynier. A mean-field model for multiple TCP connections through a buffer implementing RED. *Performance Evaluation*, No. 11, 77–97, 2002.

[9] A. Bakre and B. R. Badrinath. I-TCP: Indirect TCP for Mobile Hosts. *Proceedings of the 15th ICDCS*, May 1995.

[10] A.V. Bakre and B. R. Badrinath, Implementation and performance evaluation of Indirect TCP. *IEEE Transactions on Computers*, **46**(3):260–278, 1997.

[11] H. Balakrishnan, V. Padmanabhan, S. Seshan and R. Katz. A comparison of mechanisms for improving TCP performance over wireless links. *IEEE/ACM Transactions on Networking*, **5**(6), 1997.

[12] F. Brichet, J. Roberts, A. Simonian and D. Veitch. Heavy traffic analysis of a storage model with long range dependent on/off sources. *Queueing Systems*, **23**, 1996.

[13] P. Embrechts, C. Kluppelberg and T. Mikosch *Modelling Extremal Events*. Springer, 1997.

[14] S. Foss and S. Zachary, The maximum on a random time interval of a random walk with long-tailed increments and negative drift. *Annals of Applied Probability*, **13**(1):37–53, 2003.

[15] R. B. Jain and T. Ott. Design and Implementation of split TCP in the Linux Kernel. *Proceedings of the Globecom*, San Francisco, November 2006.

[16] M. Karaliopoulos, R. Tafazolli and B. Evans. Modeling split-TCP latency and buffering requirements in GEO satellite networks. *Proceeding of the IEEE Wireless Communications and Networking Conference*, vol. 3, pp. 1509–1514, March 2005.

[17] U. Keller and G. Biersack. A congestion control model for multicast overlay networks and its performance. *Proceedings of the Networked Group Communication*, Boston, MA, 2002.

[18] S. Kopparty, S. V. Krishnamurthy, M. Faloutsos and S. Tripathi. split TCP for mobile ad hoc networks. *Proceedings of IEEE Globecom '02*, Taipei, November 2002.

[19] G. Kwon and J. Byers. ROMA: reliable overlay multicast with loosely coupled TCP connections. *Proceedings of the IEEE Infocom*, February 2004.

[20] M. Luglio, M. Y. Sanadidi, M. Gerla and J. Stepanek. On-board satellite "split TCP" proxy. *IEEE JSAC*, **22**(2), 2004.

[21] T. Liggett, An improved subadditive ergodic theorem. *Annals of Probability*, **13**:1279–1285, 1985.

[22] V. Misra, W.-B. Gong and D. Towsley, Stochastic differential equation modeling and analysis of TCP-windowsize behavior. In *Performance'99*, Istanbul (Turkey), October 1999

[23] I. Norros, A storage model with self-similar input, *Queueing Systems*, **16**:387–396, (1994).

[24] T. Ott, J. Kemperman and M. Mathis, The stationary behavior of ideal TCP congestion avoidance. *Research Report*, 1996, available from T. Ott web page.

[25] http://www.ietf.org/rfc/rfc3135.txt.

[26] P. Rizk, C. Kiddle and R. Simmonds. Improving gridFTP performance with split TCP connections. *Proceeding of IEEE First International Conference on e-Science and Grid-Computing*, 2005.

[27] A. Sundararaj and D. Duchamp. Analytical characterization of the throughput of a split TCP Connection. *Technical Report*, Stevens Institute of Technology, 2003.

[28] W. Wei, C. Zhang, H. Zang, J. Kurose and D. Towsley. Inference and evaluation of Split-Connection Approaches in Cellular Data Networks. *Proceedins of the ACM PAM*, 2006.

[29] F. Xie, J. Hammond and D. L. Noneaker. Steady state analysis of a split-connection scheme for internet access through a wireless terminal. *IEEE/ACM Transactions on Networking*, **12**(3):2004.