Performance of TCP/IP: the case of two Controlled Sources

Omar AIT-HELLAL, Eitan ALTMAN, Driss ELOUADGHIRI
Mohamed ERRAMDANI, Nouffisa MIKOU
INRIA
2004, Route des Lucioles, B.P. 93
06902 Sophia-Antipolis Cedex
FRANCE

Keywords: TCP, modeling, fluid approximations, simulations.

Abstract

In order to control the flow in TCP/IP protocol, each TCP source uses a transmission window which indicates the number of packets that may be transmitted by the source but not yet acknowledged by the destination. We consider the Tahoe algorithm [7] developed by Van Jacobson for the adjustment of the window size and the computations of time-out values. This paper analyses the synchronization and coupling that occur between several controlled sources using this algorithm. We restrict to two controlled sources for simplicity of presentation.

1 Introduction

During a TCP connection, each source disposes of a transmission window: the size of this window indicates the number of packets that may be transmitted by the source but not yet acknowledged. This paper analysis the performance of the protocol in the case of several sources having different round-trip delays. Specifically, we compute the throughpouts, queue lengths and round trip times (both averages as well as the steady state oscillating behavior). We analyze the case where synchronization and coupling occur. between the controlled sources. Although our approach holds for more than several controlled sources, we restrict to two controlled sources for simplicity of presentation. We consider a single router and a single destination (Figure 1) and analyze this situation by using a fluid model.

Our analysis is based on a deterministic fluid model similar to the one used in [2] where quantities such as the queue size and window size are allowed to take continuous values. Such models have been found to be helpful in providing insights into the dynamic as well as stationary behavior of feedback control mechanisms (see e.g. [2, 4, 6]). Even though the deterministic approach quite simplifies the analysis, we are still faced with a system of coupled differential equations, which do not have an analytic solution. We therefore proceed by proposing some approximations based on simplifying assumptions. We then test our approach using simulations (using the REAL simulator [9]); our numerical results are shown to agree pritty well with the simulations when the sources are not too asymmetric (when the round trip delays are not too different). We point out where the simplifying assumptions fail to hold in the highly asymmetric case, and propose an improved, more complex model to handle that situation. More extensive analysis and other proposed models can be found in [5].

Our analysis is restricted to the Tahoe version of TCP/IP of Van Jacobson's algorithm [7]. However, the type of problems we encounter is typical to many other control protocols in telecommunications. In particular, similar systems of coupled differential equations arrise in the analysis of the control of ABR (Available Bit Rate) in ATM [1] and in other versions of TCP/IP (such as Reno and Vegas).

The structure of the paper is as follows. Section 2 describes Van Jacobson's algorithm as depicted in [7]. Section 3 describes the fluid model and presents the results of the analysis of the model. Section 4 presents the second model. Finally section 5 makes some concluding remarks.

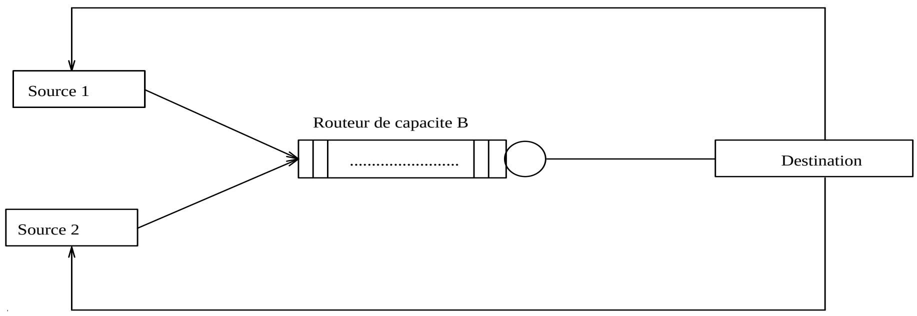

Figure 1: Model With Two Controlled Sources

2 The model, notation and description of TCP/IP

We use the following notation

- $W_i(s)$:= the window size of source i at time s,
- W^i_{max} := maximum window size of source i during a cycle,
- W^i_{th} := slow-start threshold of source i,
- τ_i := the total time between a transmission of a packet from source i till its acknowledgment arrives back to the source, without counting the time spent in the queue,

- $\tau_{i,1}$:= the time between the transmission of a packet from source i till it arrives to the queue,
- τ^2 := the time from the instant that a packet leaves the queue till it arrives to the destination. We have $\tau_i = 2(\tau_{1,i} + \tau^2)$.
- $\tilde{t}$:= The instant that the queue starts building up.
- $\tilde{W}_i$:= The window size of source i at time $\tilde{t}$
- thp^i_{in}:= The input rate of fluid originating from the controlled source
- thp^i_{out}:= The rate of fluid originating from the controlled source, at the output of the queue

Algorithm:

- At time 0, for i=1,2 $W_i = 1$
- The threshold of each source is given by: $W^i_{th} = W^i_{max}/2$
- When an acknowledgement is received by a source at time s:

$$W_i(s) = \begin{cases} W_i(s^-) + 1 & \text{if } W_i(s^-) < W^i_{th}(s^-) \text{ (Slow Start phase)} \\ W_i(s^-) + \frac{1}{[W_i(s^-)]} & \text{else (Congestion phase)} \end{cases}$$

  s^- is the recent instant when W_i is calculated and $[x]$ is the integer part of x

- When a loss occurs at time s_{loss}, we have: $W_i(s_{loss}) = 1$ and $W^i_{th}(s_{loss}) = W^i_{max}(s_{loss})/2$

Then a new cycle begins with a slow start phase with a threshold equal to $W^i_{th}(s_{loss})$.

We restrict our analysis to the case where the following simplifying assumptions hold:

- (A1) Full synchronization:

  When the queue is full, packets belonging to both sources are lost. When this assumption holds, then the window size of both sources will drop to one in a synchronized way, i.e. shortly after the queue is filled. This will result in a relatively simple cyclic behavior, where the window sizes of both sources have the same cycle duration.

  This "full synchronization" assumption will typically hold if the rate at which packets are sent (the throughput) both sources are considerably higher than the service rate, just before the queue is full. Note that there may be situations in which this assumption does not hold; after a packet from source 1 is lost, it could happen that service of a packet ends

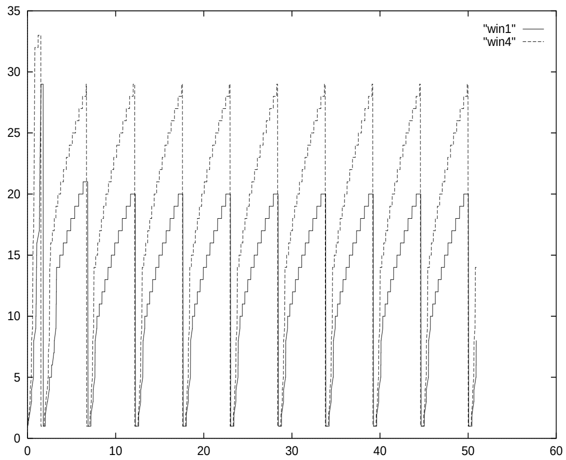

Figure 2: Two Controlled sources: a synchronized behavior

before a packet arrives from source 2, so the new arriving packet need not be lost. If this situation repeats sufficiently long, then the window of source 1 may drop to one, which will enable the other source to continue increasing its window without losing packets for some considerably long time. The full synchronization holds for example in the situation described in Figure 2, but not in that of Figure 3.

- (A2) One cycle (the duration of one cycle is the time between consecutive times that the queue fills up) contains one slow-start phase for each source. For the case of a single source, Lakshman and Madhow [8] have showed that this assumption is equivalent to having a small bandwidth-delay product (less than 3).
- (A3) The evolution of thp^i_{out} is as follows:
  - (A3i) For all $t < \tilde{t}$ (the queue is empty): $thp^i_{out}(t) = thp^i_{in}(t) = W_i(t)/T_i$. This is approximation is motivated by Little Law.
  - (A3ii) For all $t > \tilde{t}$: $thp^i_{out}(t) = \tilde{W}_i/T_i$. This approximation has the following interpretation. Once the queue starts filling up, the share of each source in the available bandwidth will remain constant. This assumption will be relaxed in Section 4.

3 The first model

3.1 Evolution of the Size of Window

To compute W_i, we must solve the differential equation:

$$\frac{dW_i}{dt} = \frac{dW_i}{dack_i}\frac{dack_i}{dt} \qquad \frac{dack_i}{dt} = thp^i_{out}$$

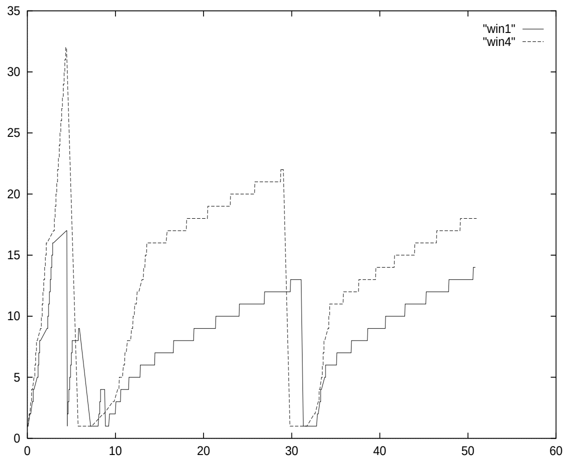

Figure 3: Two Controlled Sources: a non synchronized behavior

with

$$\frac{dW_i}{dack_i} = \begin{cases} 1 & \text{During the slow-start phase} \\ \\ \frac{1}{W_i} & \text{During the congestion avoidance phase.} \end{cases}$$

we shall distinguish between four cases.

1. $\tilde{W}_1 < W_{th}^1$ and $\tilde{W}_2 < W_{th}^2$

   - $W_1 < \tilde{W}_1$ $W(0) = 1$ and $W(\tilde{t}) = \tilde{W}_1$
     We have $\frac{dW_1}{dt} = \frac{W_1}{T_1}$
     Finally,
     $$W_1 = e^{t/T_1}$$
   - $W_1 > \tilde{W}_1$ and $W_1 < W_{th}^1$
     $W(\tilde{t}) = \tilde{W}_1$ and $W(t_1^2) = W_{th}^1$
     We have $\frac{dW_i}{dt} = \frac{\tilde{W}_i}{T_1}$
     Finally
     $$W_1(t) = \frac{\tilde{W}_1}{T_1}t + \tilde{W}_1(1 - Log(\tilde{W}_1))$$
   - $W_1 > W_{th}^1$
     $W_1(t_1^2) = W_{th}^1$ and $W_1(t_1^3) = W_{max}^1$
     We have $\frac{dW_1}{dt} = \frac{\tilde{W}_1}{T_1 W_1}$
     Finally
     $$W_1 = \sqrt{2\frac{\tilde{W}_1}{T_1}t + (W_{th}^1)^2 - 2(W_{th}^1 - \tilde{W}_1) - 2\tilde{W}_1 Log(\tilde{W}_1)}$$

We compute W_2 with same calculation for W_1:

2. $\tilde{W}_1 < W^1_{th}$ and $\tilde{W}_2 > W^2_{th}$
   The evolution of W_1 is the same as in case 1.
   For W_2, we have:

   - $W_2 < W^2_{th}$
   $$W_2 = e^{t/T_2}$$
   - $W_2 < \tilde{W}_2$ and $W_2 > W^2_{th}$
   $$W_2(t) = \frac{1}{T_2}t + W^2_{th} - Log(W^2_{th})$$
   - $W_2 > \tilde{W}_2$
   $$W_2 = \sqrt{2\frac{\tilde{W}_2}{T_2}t + (\tilde{W}_2)^2 - 2\tilde{W}_2(\tilde{W}_2 - W^2_{th}) - 2\tilde{W}_2 Log(W^2_{th})}$$

3. $\tilde{W}_1 > W^1_{th}$ and $\tilde{W}_2 < W^2_{th}$
   The evolution of W_1 is the same as of W_2 in case 2, and The evolution of W_2 is the same as in case 1,

4. $\tilde{W}_1 > W^1_{th}$ and $\tilde{W}_2 > W^2_{th}$
   The evolution of W_1 is the same that in case 3 and The evolution of W_2 is the same that in case 2.

3.2 Computation of $\tilde{W}_i$ and W^i_{th}

Lemma 3.1
$$\tilde{W}_i < W^i_{th} \Longleftrightarrow T_i < \frac{B}{\mu}$$

Proof: This follows from the fact $\dfrac{W^i_{max}}{\frac{B}{\mu} + T_i} = thp^i_{out} = \dfrac{\tilde{W}_i}{T_i}$ and $W^i_{th} = \dfrac{W^i_{max}}{2}$, ■

Lemma 3.2 *We define* $A_i (i = 1, 2) : A_i = \frac{2T_i}{\frac{B}{\mu}+T_i}$

- *in case of* $T_1 < \frac{B}{\mu}$ *and* $T_2 < \frac{B}{\mu}$, $\tilde{t}$ *is the zero of the function* F_1*:*
  $F_1(t) = \frac{t/T_1}{T_1} + \frac{e^{t/T_1}}{T_1} - \mu$

- *in case of* $T_1 < \frac{B}{\mu}$ *and* $T_2 > \frac{B}{\mu}$ $\tilde{W}_2$ *is the zero of function* F_2*:*

$$F_2(t) = \frac{T_1}{T_2} log(T_1(\mu - \frac{w}{T_2})) + (\frac{1}{A_2} - 1)w - log(\frac{w}{A_2})$$

- *in case of* $T_1 > \frac{B}{\mu}$ *and* $T_2 < \frac{B}{\mu}$ $\tilde{W}_1$ *is the zero of function* F_3*:*

$$F_3(w) = \frac{T_2}{T_1} log(T_2(\mu - \frac{w}{T_1})) + (\frac{1}{A_1} - 1)w - log(\frac{w}{A_1})$$

- *in case of* $T_1 > \frac{B}{\mu}$ *and* $T_2 > \frac{B}{\mu}$, $\tilde{W}_2$ *is the zero of the function* F_4*:*

$$F_4 = \frac{-T_2(1-A_2)w}{T_1 A_2} + \frac{T_2}{T_1} log(\frac{w}{A_2}) + \frac{T_1(1-A_1)}{A_1}(\mu - \frac{w}{T_2}) - log(\frac{T_1}{A_1}(\mu - \frac{w}{T_2}))$$

Proof: In the case 1, we use the following equalities:

$$\tilde{W}_1 = e^{\tilde{t}/T_1} \qquad \tilde{W}_2 = e^{\tilde{t}/T_2} \qquad \frac{\tilde{W}_1}{T_1} + \frac{\tilde{W}_2}{T_2} = \mu \tag{1}$$

In the case 2, we use:

$$\tilde{W}_1 = e^{\tilde{t}/T_1} \qquad \tilde{W}_2 = \frac{1}{T_2}\tilde{t} + W_{th}^2 - log(W_{th}^2) \qquad \frac{\tilde{W}_1}{T_1} + \frac{\tilde{W}_2}{T_2} = \mu, \tag{2}$$

In the case 3, we use:

$$\tilde{W}_2 = e^{\tilde{t}/T_2} \qquad \tilde{W}_1 = \frac{1}{T_1}\tilde{t} + W_{th}^1 - log(W_{th}^1) \qquad \frac{\tilde{W}_1}{T_1} + \frac{\tilde{W}_2}{T_2} = \mu, \tag{3}$$

In the case 4, we use:

$$\tilde{W}_1 = \frac{1}{T_1}\tilde{t} + W_{th}^1 - log(W_{th}^1) \qquad \tilde{W}_2 = \frac{1}{T_2}\tilde{t} + W_{th}^2 - log(W_{th}^2) \qquad \frac{\tilde{W}_1}{T_1} + \frac{\tilde{W}_2}{T_2} = \mu. \tag{4}$$

■

3.3 Computation of $\overline{thp_i}$

We define the following quantities:
$T_i^j = t_i^j - t_i^{j-1}$,$t_i^j$ refer to subsection 3.1 and $N_i^j = \int_{t_i^{j-1}}^{t_i^j} thp_{out}^i(t)dt$ we have the same four cases:

- $\tilde{W}_1 < W_{th}^1$ and $\tilde{W}_2 < W_{th}^2$
  Further more we obtain
  $T_1^1 = T_1 Log(\tilde{W}_1) \qquad T_1^2 = \frac{T_1}{\tilde{W}_1}(W_{th}^1 - \tilde{W}_1) \qquad T_1^3 = \frac{T_1}{2\tilde{W}_1}((W_{max}^1)^2 - (W_{th}^1)^2)$
  $T_2^1 = T_2 Log(\tilde{W}_2) \qquad T_2^2 = \frac{T_2}{\tilde{W}_2}(W_{th}^2 - \tilde{W}_2) \qquad T_2^3 = \frac{T_2}{2\tilde{W}_2}((W_{max}^2)^2 - (W_{th}^2)^2),$

  For N_i^j, we obtain
  $N_1^1 = \tilde{W}_1 - 1 \qquad N_1^2 = W_{th}^1 - \tilde{W}_1 \qquad N_1^3 = \frac{(W_{max}^1)^2 - (W_{th}^1)^2}{2}$
  $N_2^1 = \tilde{W}_2 - 1 \qquad N_2^2 = W_{th}^2 - \tilde{W}_2 \quad N_2^3 = \frac{(W_{max}^2)^2 - (W_{th}^2)^2}{2},$

- $\tilde{W}_1 < W_{th}^1$ and $\tilde{W}_2 > W_{th}^2$
  The computation of T_1^j and N_1^j is the same that in case 1, for source S2, we have:
  $T_2^1 = T_2 Log(W_{th}^2) \qquad T_2^2 = T_2(\tilde{W}_2 - W_{th}^2) \qquad T_2^3 = \frac{T_2}{2\tilde{W}_2}((W_{max}^2)^2 - (\tilde{W}_2)^2)$
  $N_2^1 = W_{th}^2 - 1 \qquad N_2^2 = \frac{1}{(T_2)^2}\frac{(T_2^2)^2}{2} + W_{th}^2 \frac{T_2^2}{T_2} \qquad N_2^3 = \frac{(W_{max}^2)^2 - (\tilde{W}_2)^2}{2},$

- $\tilde{W}_1 > W_{th}^1$ and $\tilde{W}_2 < W_{th}^2$
  The computation of T_1^j and N_1^j is the same that T_2^j and N_2^j in case 2, the computation of T_2^j and N_2^j is the same as in case 1

- $\tilde{W}_1 > W_{th}^1$ and $\tilde{W}_2 > W_{th}^2$
  The computation of T_1^j and N_1^j is the same as in case 3 and the computation of T_2^j and N_2^j is the same that in case 2

This finally yields to explicit expressions $\overline{thp_i}$: $\overline{thp_i} = \frac{N_i^1 + N_i^2 + N_i^3}{T_i^1 + T_i^2 + T_i^3}$

Remark 3.1 *We see easily that :*

- F_1 *is strictly increasing*
- F_2 *is strictly decreasing*
- F_3 *is strictly decreasing*
- F_4 *is strictly increasing*

Finally, the zero of all functions can be computed by using the dichotomy algorithm

3.4 Computation of $\overline{rtt_i}$

The performance parameters $\overline{rtt_i}$ (i=1.2) are given by:$\overline{rtt_i} = \frac{\sum_{j=1}^{3} N_i^j R_i^j}{\sum_{j=1}^{3} N_i^j}$ With

$$R_i^j = \begin{cases} Ti & \text{If the queue is empty} \\ \\ \frac{1}{t_i^j - t_i^{j-1}} \int_{t_i^{j-1}}^{t_i^j} [T_i + \frac{Q(t)}{\mu}] dt & \text{Otherwise.} \end{cases}$$

Lemma 3.3 *We distinguish between the following four cases*

in case of $T_1 < \frac{B}{\mu}$ *and* $T_2 < \frac{B}{\mu}$,
the source S_1 *enter to congestion before* $S_2 \Longleftrightarrow T_1 > T_2$

- *in case of* $T_1 < \frac{B}{\mu}$ *and* $T_2 > \frac{B}{\mu}$,
  the source S_2 *enter to congestion before* S_1
- *in case of* $T_1 > \frac{B}{\mu}$ *and* $T_2 < \frac{B}{\mu}$,
  the source S_1 *enter to congestion before* S_2

Proof: For proof, in case 1 we use the difference $T_1^2 - T_2^2$ and for the rest cases the proof is trivial ■

Lemma 3.4 *When the queue is not empty* R_i^j *is expressed as follow:*

$$\begin{aligned} R_i^j &= T_i + \frac{1}{\mu(t_i^j - t_i^{j-1})} \int_{t_i^{j-1}}^{t_i^j} [W_1 + W_2] dt \\ &- \frac{1}{\mu(t_i^j - t_i^{j-1})} thp_{out}^1 \int_{t_i^{j-1}}^{t_i^j} S_1(t) dt - \frac{1}{\mu(t_i^j - t_i^{j-1})} thp_{out}^2 \int_{t_i^{j-1}}^{t_i^j} S_2(t) dt, \end{aligned}$$

With

$$S_i(t) = \begin{cases} 3\tau_{i,1} + 2\tau^2 & \textit{during the slow-start phase} \\ 2\tau_{i,1} + 2\tau^2 & \textit{during the congestion avoidance phase} \end{cases}$$

Proof:

$$Q(t) = W_1(t) + W_2(t) - \tau_{1,1} thp_{in}^1(t) - (\tau_{1,1} + 2\tau^2) thp_{out}^1(t) - \tau_{2,1} thp_{in}^2(t) - (\tau_{2,1} + 2\tau^2) thp_{out}^2(t) \quad (5)$$

There is one relation between thp^i_{in} and thp^i_{out}, indeed:

$$\begin{aligned} thp^i_{in}(t) &= \frac{dW_i}{dt} + \frac{dack_i}{dt} \\ &= \frac{dack_i}{dt}(1 + \frac{dW_i}{dack_i}), \end{aligned}$$

During the congestion avoidance phase, W_i is sufficiently large, then $W_i^{-1} << 1$, thus we obtain the following expression:

$$thp^i_{in}(t) = \begin{cases} 2thp^i_{out}(t) & \text{during the slow-start phase} \\ thp^i_{out}(t) & \text{during the congestion avoidance phase} \end{cases} \tag{6}$$

■

Remark 3.2 *The computation of R_i^j depends on whether S_1 enters in congestion before S_2 or not, hence:*
if S_1 enters the congestion phase before S_2:

$$\begin{aligned} \int_{\tilde{t}}^{t_1^2} [W_1 + W_2]dt &= \int_{\tilde{t}}^{t_1^2} (\frac{\tilde{W}_1}{T_1} t + \tilde{W}_1(1 - log(\tilde{W}_1)))dt \\ &+ \int_{\tilde{t}}^{t_1^2} (\frac{\tilde{W}_2}{T_2} t + \tilde{W}_2(1 - log(\tilde{W}_2))), \end{aligned}$$

if S_2 enters the congestion phase before S_1:

$$\begin{aligned} \int_{\tilde{t}}^{t_1^2} [W_1 + W_2]dt &= \int_{\tilde{t}}^{t_1^2} (\frac{\tilde{W}_1}{T_1} t + \tilde{W}_1(1 - log(\tilde{W}_1)))dt \\ &+ \int_{\tilde{t}}^{t_2^2} (\frac{\tilde{W}_2}{T_2} t + \tilde{W}_2(1 - log(\tilde{W}_2))) \\ &+ \int_{t_2^2}^{t_1^2} (\sqrt{2\frac{\tilde{W}_2}{T_2} t + (W_{th}^2)^2 - 2(W_{th}^2 - \tilde{W}_2) - 2\tilde{W}_2 Log(\tilde{W}_2)}), \end{aligned}$$

To simplify the computations, we introduce an algorithmic approach and hence some functions will be developed.
The function $integlin(ns, a, b, nc)$ calculates the integral of W_i when W_i has a linear behavior.

- Computation of $integlin(ns, a, b, nc)$
  ns: the index of the source
  a: the lower boundary of the interval of integration

b: the upper boundary of the interval of integration
nc: the case number

If $nc = 1$ then

$$\alpha \stackrel{\text{def}}{=} \begin{cases} \tilde{W}_1/T_1 & \text{if ns=1} \\ \tilde{W}_2/T_2 & \text{if ns=2} \end{cases} \qquad \beta \stackrel{\text{def}}{=} \begin{cases} \tilde{W}_1(1 - Log(\tilde{W}_1) & \text{if ns=1} \\ \tilde{W}_2(1 - Log(\tilde{W}_2)) & \text{if ns=2} \end{cases}$$

if $nc = 2$ then

$$\alpha \stackrel{\text{def}}{=} \begin{cases} \tilde{W}_1/T_1 & \text{if ns=1} \\ 1/T_2 & \text{if ns=2} \end{cases} \qquad \beta \stackrel{\text{def}}{=} \begin{cases} \tilde{W}_1(1 - Log(\tilde{W}_1) & \text{if ns=1} \\ W^2_{th} - Log(W^2_{th}) & \text{if ns=2} \end{cases}$$

if $nc = 3$ then

$$\alpha \stackrel{\text{def}}{=} \begin{cases} 1/T_1 & \text{if ns=1} \\ \tilde{W}_2/T_2 & \text{if ns=2} \end{cases} \qquad \beta \stackrel{\text{def}}{=} \begin{cases} W^1_{th} - Log(W^1_{th}) & \text{if ns=1} \\ \tilde{W}_2(1 - Log(\tilde{W}_2) & \text{if ns=2} \end{cases}$$

if $nc = 4$ then

$$\alpha \stackrel{\text{def}}{=} \begin{cases} 1/T_1 & \text{if ns=1} \\ 1/T_2 & \text{if ns=2} \end{cases} \qquad \beta \stackrel{\text{def}}{=} \begin{cases} W^1_{th} - Log(W^1_{th}) & \text{if ns=1} \\ W^2_{th} - Log(W^2_{th}) & \text{if ns=2} \end{cases}$$

The value of the integral is $\frac{\alpha}{2}(b^2 - a^2) + \beta(b - a)$

- Computation of $t(ns, np, nc)$
  There are three periods where the behavior of W_i is changed. The instants $t(ns, np, nc)$ are the points of regeneration.

$$t(ns, np, nc) = \begin{cases} (Duration(ns, 1, nc)) & \text{if np=1} \\ (Duration(ns, 1, nc) + Duration(ns, 2, nc)) & \text{if np=2} \\ (Duration(ns, 1, nc) + Duration(ns, 2, nc) + Duration(ns, 3, nc) & \text{if np=3,} \end{cases}$$

- Computation of $Duration(ns, nc, np)$
  For each case meted, $Duration(ns, nc, np) = T^{np}_{ns}$ and T^{np}_{ns} are the parameters already computed.

- Computation of $integrac(ns, a, b, nc)$
  This function is used to calculate the integral of W_i when its behavior is sub-linear.

$$integrac(ns, a, b, nc) = \frac{T(ns)}{3\tilde{W}(ns)}\left(\left(\frac{2\tilde{W}(ns)}{T(ns)}b + gr(ns, nc)\right)^{\frac{3}{2}} - \left(\frac{2\tilde{W}(ns)}{T(ns)}a + gr(ns, nc)\right)^{\frac{3}{2}}\right)$$

if $ns = 1$ then

$$gr(ns, nc) = \begin{cases} (W^1_{th})^2 - 2(W^1_{th} - \tilde{W}_1) - 2\tilde{W}_1 Log(\tilde{W}_1) & \text{if nc=1 or 2} \\ (\tilde{W}_1)^2 - 2\tilde{W}_1(\tilde{W}_1 - W^1_{th}) - 2\tilde{W}_1 Log(W^1_{th}) & \text{if nc=3 or 4,} \end{cases}$$

if $ns = 2$ then

$$gr(ns, nc) = \begin{cases} (W^2_{th})^2 - 2(W^2_{th} - \tilde{W}_2) - 2\tilde{W}_2 Log(\tilde{W}_2) & \text{if nc=1 or 3} \\ (\tilde{W}_2)^2 - 2\tilde{W}_2(\tilde{W}_2 - W^2_{th}) - 2\tilde{W}_2 Log(W^2_{th}) & \text{if nc=2 or 4,} \end{cases}$$

For the computation of R_i^j we distinguish four cases. We present the first case (i.e. $\tilde{W}_1 < W_{th}^1$ and $\tilde{W}_2 < W_{th}^2$); the computation of R_i^j in the other cases is similar (and can be found in [5]).

- $\tilde{W}_1 < W_{th}^1$ $\tilde{W}_2 < W_{th}^2$

$$R_1^1 = T_1$$
$$R_2^1 = T_2$$

As it's indicated the computation of R_1^2 and R_2^2 depends on the difference $T_2^2 - T_1^2$, it's the fact that we will distinguish two cases.

  - $T_2^2 < T_1^2$

$$\begin{aligned} R_1^2 &= T_1 + \frac{1}{\mu[t(1,2,1) - t(1,1,1)]}[integlin(1, t(1,1,1), t(1,2,1), 1) \\ &+ integlin(2, t(1,1,1), t(2,2,1), 1) + integrac(2, t(2,2,1), t(1,2,1), 1)] \\ &- \frac{thp_{out}^1}{\mu} S(1, slow) - thp_{out}^2 \frac{1}{\mu[t(1,2,1) - t(1,1,1)]} \\ &* [S(2, slow)[t(2,2,1) - t(1,1,1)] + S(2, congestion)[t(1,2,1) - t(2,2,1)]]. \end{aligned}$$

$$\begin{aligned} R_2^2 &= T_2 + \frac{1}{\mu[t(2,2,1) - t(1,1,1)]}[integlin(1, t(2,1,1), t(2,2,1), 1) \\ &+ integlin(2, t(2,1,1), t(2,2,1), 1)] - \frac{thp_{out}^1}{\mu} S(1, slow) \\ &- \frac{thp_{out}^2}{\mu} S(2, slow), \end{aligned}$$

$$\begin{aligned} R_1^3 &= T_1 + \frac{1}{\mu[t(1,3,1) - t(1,2,1)]}[integrac(1, t(1,2,1), t(1,3,1), 1) \\ &+ integrac(2, t(1,2,1), t(1,3,1), 1)] - \frac{thp_{out}^1}{\mu} S(1, congestion) \\ &- \frac{thp_{out}^2}{\mu} S(2, congestion), \end{aligned}$$

$$\begin{aligned} R_2^3 &= T_2 + \frac{1}{\mu[t(2,3,1) - t(2,2,1)]}[integlin(1, t(2,2,1), t(1,2,1), 1) \\ &+ integrac(1, t(1,2,1), t(2,3,1), 1) + integrac(2, t(2,2,1), t(2,3,1), 1)] \\ &- \frac{thp_{out}^2}{\mu} S(2, congestion) - thp_{out}^1 \frac{1}{\mu[t(2,3,1) - t(2,2,1)]} \\ &+ [S(1, slow)[t(1,2,1) - t(2,2,1)] + S(1, congestion)[t(2,3,1) - t(1,2,1)]], \end{aligned}$$

We can compute R_i^j (i=1,2 and j=1,2,3) for the rest cases as above

3.5 Numerical Results and Analysis

The input parameters B, μ and τ^2 are the same for Table 1 and Table 2. We therefore omit these parameters from Table 2.

- B is expressed in packets.
- thp_{in}, μ and λ are expressed in packets/seconds.
- τ_1 and τ_2 are expressed in seconds.
- A packet size is 576 bytes.

	Parameters					$\overline{thp_1}$			$\overline{rtt_1}$		
case	B	μ	τ^2	thp_{in}^1	$\tau_{1,1}$	anal	simu	%err	anal	simu	%err
1	30	108.5	0.0200	217.01	0.0200	54.24	53.09	02	0.298	0.270	10
2	30	108.5	0.0200	217.01	0.0200	93.81	68.57	37	0.337	0.277	21
3	30	108.5	0.0622	217.01	0.0622	101.7	63.31	61	0.482	0.425	13
4	30	108.5	0.0750	217.01	0.0750	54.11	45.37	19	0.470	0.459	02

Table 1: Performances of source S1

	Parameters		$\overline{thp_2}$			$\overline{rtt_2}$		
case	thp_{in}^2	$\tau_{2,1}$	anal	simu	%err	anal	simu	%err
1	217.01	0.0200	54.25	53.22	02	0.298	0.277	08
2	217.01	0.0300	14.68	37.87	61	0.245	0.294	16
3	217.01	0.0625	06.78	35.21	81	0.299	0.421	29
4	217.01	0.0750	54.11	52.32	03	0.470	0.456	03

Table 2: Performances of source S2

Table 1 shows the performances of source S1 and Table 2 these of S2. For the equidistant cases (i.e. cases 1 and 4) we get small differences between the analytic and simulation results: at most 19% for the throughput and 10% fot the rtt. Large differences are obtained, however, for the nonsymetric cases. Thus, we conclude that there is one situation where the simulations validate our analytic model, namely the situation where the sources S1 and S2 are kept equidistant from the router. In what follows, we try to explain the large differences in the asymmetric case. For all the cases above we obtain behaviors respectively represented in Figures 4-7. As shown in Section 2, the behaviors are either synchronized or non-synchronized.

In analyzing the numerical results for the asymmetric case, we discovered that the values $\tilde{W}_i$ of the window size when the queue starts building up are very sensitive to the parameters. Indeed, very small differences between $\tau_{1,1}$ and $\tau_{2,1}$ translate into very large differences between $\tilde{W}_1$ and $\tilde{W}_2$.

Possible ways to try to improve the model for the asymmetric case would thus be to relax assumption (A3ii). Instead of assuming that the throughputs of the sources are constant after time $\tilde{t}$ and equal $\tilde{W}_i/T_i$, one could

- drop completely the assumption that the throughputs of the sources are constant after time $\tilde{t}$. One could still approximate the evolution of the throughputs by piecewize constant values;
- still assume that the throughputs of the sources are constant after time $\tilde{t}$, but have values other than $\tilde{W}_i/T_i$.

We pursued in our research an approach related to the first direction, i.e. approximate the evolution of the throughputs after the time that the queue sills by piecewize constant values. As will be seen in Section 4, we decomposed the time after $\tilde{t}$ into two parts appropriately chosen.

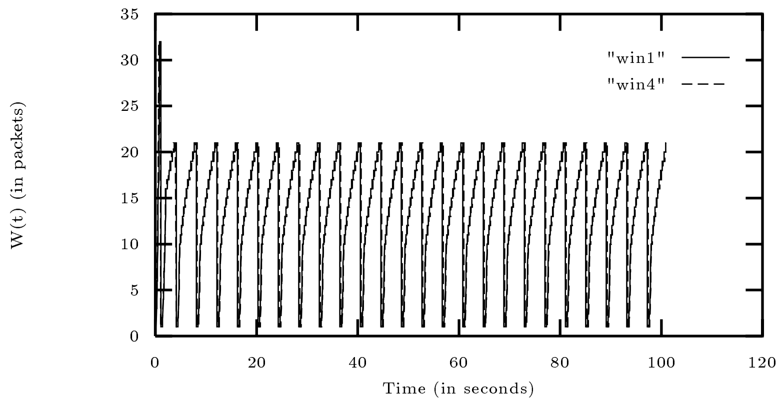

Figure 4: Figure of cse 1 with two TCP connections

4 The second model

We consider a modification of the first model where we analyze what happens when one source is in congestion and the other source in slow start. We consider the case that both $T_1 < B/\mu$ and

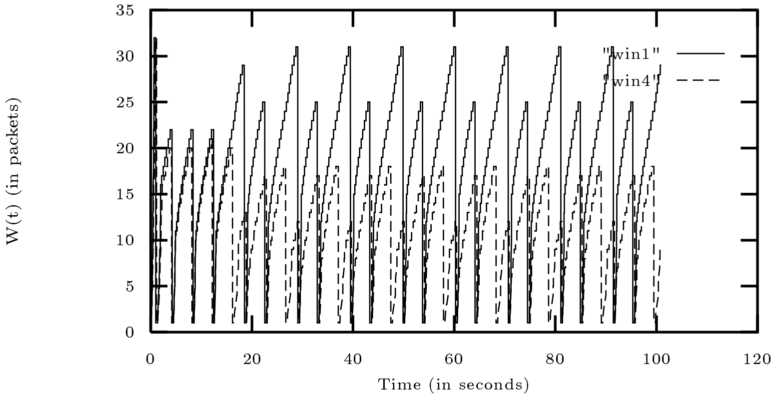

Figure 5: Figure of case 2 with two TCP connections

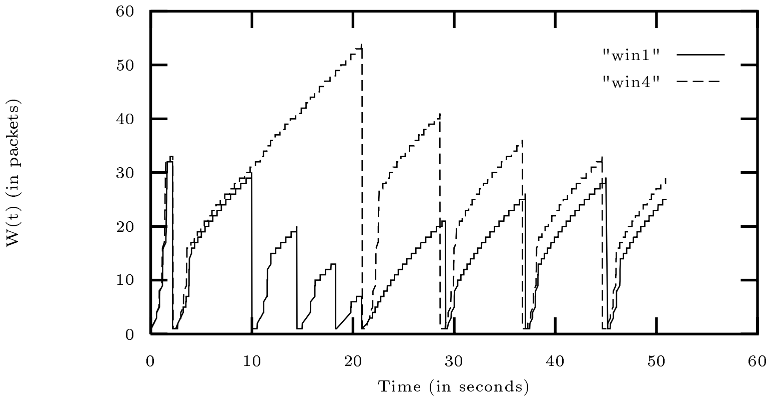

Figure 6: Figure of case 3 with two TCP connections

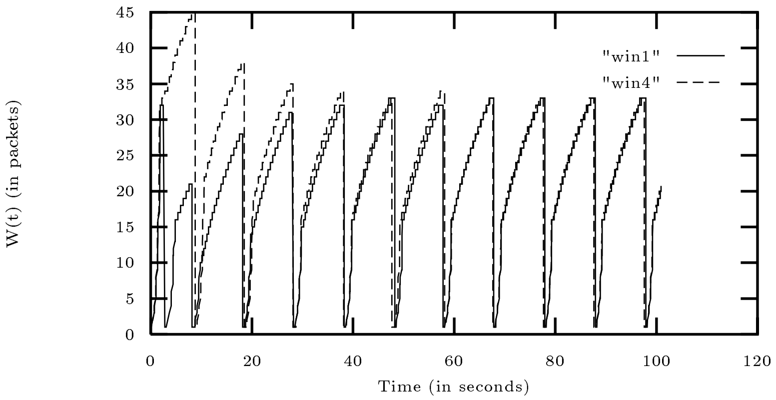

Figure 7: Figure of case 4 with two TCP connections

$T_2 < B/\mu$. We suppose that $T_1^2 > T_2^2$. $\exists$ $(t_s^1\ ,t_s^2)$; $(\tilde{t} \leq t_s^2 \leq t_s^1)$: During $[\tilde{t}, t_s^2]$, S_1 and S_2 are both in slow start and during $[t_s^2, t_s^1]$, S1 in slow start and S2 in congestion. In the previous section we assumed that during $[t_s^2, t_s^1]$, $thp_{out}^i = \tilde{W}_i/T_i$. We thus ignored the fact that the two sources are in different phases, and just assumed that the throughput of each one of them remains unchanged when the phase changes. In order to obtain a more realistic model, we make use of the following insight obtained in [2, 3]. If a controlled source shares a bottleneck node with an uncontrolled exogenous source with a constant rate of λ, then once the queue starts filling up, the bandwidth taken by the controlled source will be $\mu - \lambda/2$ if it is in the slow-start phase, and it will be $\mu - \lambda$ if it is in the congestion avoidance phase. Hence, a source in a slow-start phase is less agressive than an uncontrolled constant rate source, whereas a controlled source is more aggressive. To express this difference in "aggressivity" (or in priority) between the sources, we pretend that the source S2, which is in the congestion phase, and which has used a throughput of $\tilde{W}_2/T_i$ just before $\tilde{t}$, is an uncontrolled exogenous source, that tries to transmit at a rate of $\lambda = \tilde{W}_2/T_i$. When competing against S1, its throughput will thus decrease to $\lambda/2$, and S1 will use the remaining bandwidth. This gives us half the value of the throughput for source S2 as the one we assumed in the previous section. In reality, since source S2 is controlled, it is expected to have a throughput even smaller than $\lambda/2$, since a controlled source in congestion avoidance phase behaves even less aggressively than an uncontrolled source. We conclude that in $(t_s^2\ ,t_s^1)$,

$$\lambda = \frac{\tilde{W}_2}{T_2} \qquad thp_{out}^2 = \frac{\lambda}{2} \qquad thp_{out}^1 = \mu - \frac{\lambda}{2}.$$

After time t_s^1, S1 and S2 are in congestion phase, and we use an approximation based on the assumption that the bandwidth used by each source will remain constant. The throughput can then be approximated using Little's law by: $thp_{out}^i = \frac{\tilde{\tilde{W}}_i}{T_i + Q(t_s^i)/\mu}$, with

$$\tilde{\tilde{W}}_1 = W_1(t_s^1) \quad \tilde{\tilde{W}}_2 = W_2(t_s^1) \quad \tilde{\tilde{Q}} = Q(t_s^1).$$

The size of window, the throughput and the round trip delay can be derived using the same calculations as in section 3. The results that we obtained following this approach were still unsatisfactory for the asymmetric case, see [5].

5 Conclusion

In this paper, we study the synchronization and coupling that occur between two sources by utilizing a fluid approximation. The first model agrees with simulation in the case where the two sources are symmetric with respect to their distances to the buffer. This can model well two sources that are connected to a common LAN, and share the same destination. We proposed in the end of Section 3 some possible ways to improve the model for asymmetric distances, and studied one of these in Section 4.

References

[1] The ATM Forum Technical Committee, *Traffic Management Specification*, Version 4.0, 95-0013R8, April 1996.

[2] E. Altman, F. Boccara, J. Bolot, P. Nain, P. Brown, D. Collange, and C. Fenzy, "Analysis of the TCP/IP Flow Control Mechanism in High-Speed Wide-Area Networks", *IEEE Conf. on Decision and Control (CDC' 95)*, pp. 368-373, New Orleans, LA, USA, Dec. 1995.

[3] E. Altman, J. Bolot, D. Elouadghiri, M. Erramdani, P. Nain, P. Brown, D. Collange, Performance of TCP/IP over the French Research network: Measurements and Analysis, final report of INRIA-CNET contract No. 94-5B-012, 1995.

[4] J-C. Bolot, A. U. Shankar, "Analysis of a fluid approximation to flow control dynamics", *Proc. IEEE Infocom '92*, pp. 2398-2407, Florence, Italy, May 1992.

[5] M. Erramdani, *Evaluation de Performance de TCP/IP: Cas de deux Sources Contrôllées*, stage report, September 1995.

[6] K. W. Fendick, M. Rodrigues, A. Weiss, "Analysis of a rate-based control strategy with delayed feedback", *Proc. ACM Sigcomm '92*, pp. 136-147, Baltimore, MD, USA, Sept. 1992.

[7] V. Jacobson, "Congestion avoidance and control", *Proc. ACM Sigcomm '88*, Stanford, CA, Aug. 1988.

[8] T. V. Lakshman and U. Madhow, "Window-based congestion control for networks with bandwidth-delay products and random loss: a study of TCP/IP performance", *Proc. HPN '94 (High Performance Networking)*, pp. 133-147, Grenoble, France, Jun. 1994.

[9] S. Keshav, "REAL: A network simulator", *Department of Computer Science, UC Berkeley, Technical Report*, 88/472, 1988.