A Note on the TCP Fluid Model

Dohy Hong

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

A Note on the TCP Fluid Model

Dohy Hong

N° 4703

January 2003

THÈME 1

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--4703--FR+ENG

A Note on the TCP Fluid Model

Dohy Hong *

Thème 1 — Réseaux et systèmes
Projet TREC

Rapport de recherche n° 4703 — January 2003 — 11 pages

Abstract: In this paper, we revisit the fluid TCP model and propose an explicit solution of the non-linear differential equation. We show that as soon as the queue never empties, we have a very non-intuitive rate dynamic and that in partcular the ratio $w(t)/rtt(t)$ is a constant. We test the analytical solution by comparison with NS2 simulation results.

Key-words: TCP, AIMD, fluid model.

* INRIA-ENS, ENS, 45 rue d'Ulm 75005 Paris, France {Dohy.Hong@ens.fr}

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

Note sur le modèle fluide de TCP

Résumé : Nous revisitons ici le modèle fluide de TCP et proposons une solution explicite de l'équation différentielle non-linéaire associée. Nous montrons que tant que la file d'attente n'est pas vide, nous avons une évolution du débit contre-intuitive et qu'en particulier le ratio $w(t)/rtt(t)$ est une constante. Nous testons la solution analytique par comparaison avec des résultats de simulation NS2.

Mots-clés : TCP, AIMD, modèle fluide

1 Introduction

There are many papers studying the dynamic of TCP congestion control mechanism based on fluid approximation (e.g. cf. [6, 7, 11, 5, 3, 9, 1, 2]). Fluid models are extensively used to study in particular AQM type routers interacting with TCP sources.

The aim of this paper is to revisit this family of model to underline some basic properties of TCP that were introduced in [3].

2 Fluid model of DropTail queues

2.1 Single Flow Model

A single bottleneck link model is described by:

- **[network scenario]**: C the service capacity, B the buffer size, $R_{\min}$ the minimum RTT or the pure propagation delay without queueing delay;
- **[state variables]**: $R(t)$ the RTT at time t, $W(t)$ the window size, $Q(t)$ the queue size and $X(t)$ the packet send rate of the source.

TCP's AIMD dynamic can be described by fluid model as follows.
Before the congestion, i.e., if $X(t) < C$:

$$\dot{W}(t) = \frac{1}{R_{\min}} \quad (1)$$

$$R(t) = R_{\min} \quad (2)$$

$$Q(t) = 0 \quad (3)$$

$$X(t) = \frac{W(t)}{R_{\min}} \quad (4)$$

When the congestion starts, i.e., when $X(t) \geq C$, and when $Q(t) < B$:

$$\dot{W}(t) = \frac{1}{R(t)} \quad (5)$$

$$R(t) = R_{\min} + \frac{Q(t)}{C} \quad (6)$$

$$\dot{Q}(t) = X(t) - C \quad (7)$$

$$X(t) = \left(1 + \frac{1}{W(t)}\right) C \quad (8)$$

Equation (8) can be explained as follow: when the input rate to the bottleneck link is higher than its service capacity, the output of this link becomes constant equal to C, which also is the rate at which the source receives ACKs. Since the source increases its window size by $1/W(t)$ of every new ACK reception, its packet send rate is exactly: $(1 + 1/W(t)) \times$ acks_rate. This in particular implies that once the congestion starts, the send rate of the source decreases in time! This is neither obvious nor very intuitive.

Theorem 1. *The explicit solutions of equations* (5) - (8) *are:*

$$\begin{aligned} R(t) &= \sqrt{\frac{2}{C}t + R_{\min}^2} && (9)\\ W(t) &= CR(t) = \sqrt{2Ct + W(0)^2} && (10)\\ Q(t) &= W(t) - W(0) = \sqrt{2Ct + W(0)^2} - W(0) && (11)\\ X(t) &= C + \frac{1}{\sqrt{\frac{2}{C}t + R_{\min}^2}} && (12) \end{aligned}$$

In particular, the solution satisfies the equality:

$$\frac{W(t)}{R(t)} = C.$$

Proof We have from (5) and (6)-(8) that:

$$\dot{W}(t)R(t) = \dot{R}(t)W(t) = 1.$$

Therefore

$$(\dot{RW})(t) = 2$$

which implies that $R(t)W(t) = 2t + R_{\min}^2 C$. From this last equality, we get that:

$$\dot{R}(t) = \frac{2}{W(t)} - \frac{2t + R_{\min}^2 C\dot{W}(t)}{W(t)^2}.$$

Multiplying by $W(t)$, we get

$$\frac{\dot{W}(t)}{W(t)} = \frac{1}{2t + R_{\min}^2 C}$$

which admits as solution:

$$W(t) = \sqrt{2Ct + W(0)^2}$$

with $W(0) = R_{\min}^2 C^2$. From this, we find the other solutions. □

□

Remark 1. *In literature on TCP fluid modelling, Equation* (8) *is not generally used. Instead, many authors write:*

$$X(t) = \frac{W(t)}{R(t)}. \tag{13}$$

This is in a first view incompatible with the remark that $W(t)/R(t) = C$. *There are two reasons that explain why this last approximation* (13) *is not so bad apparently:*

- *the first one is the integer part effect of cwnd variable. As shown in Figure 1 (Top), when taking* $W(t)$ *without the integer part,* W/R *is almost as the packet send rate. Whereas taking the integer part (because only the integer part is used in packet send action by TCP), the quantity* W/R *is quite constant as predicted by our results;*

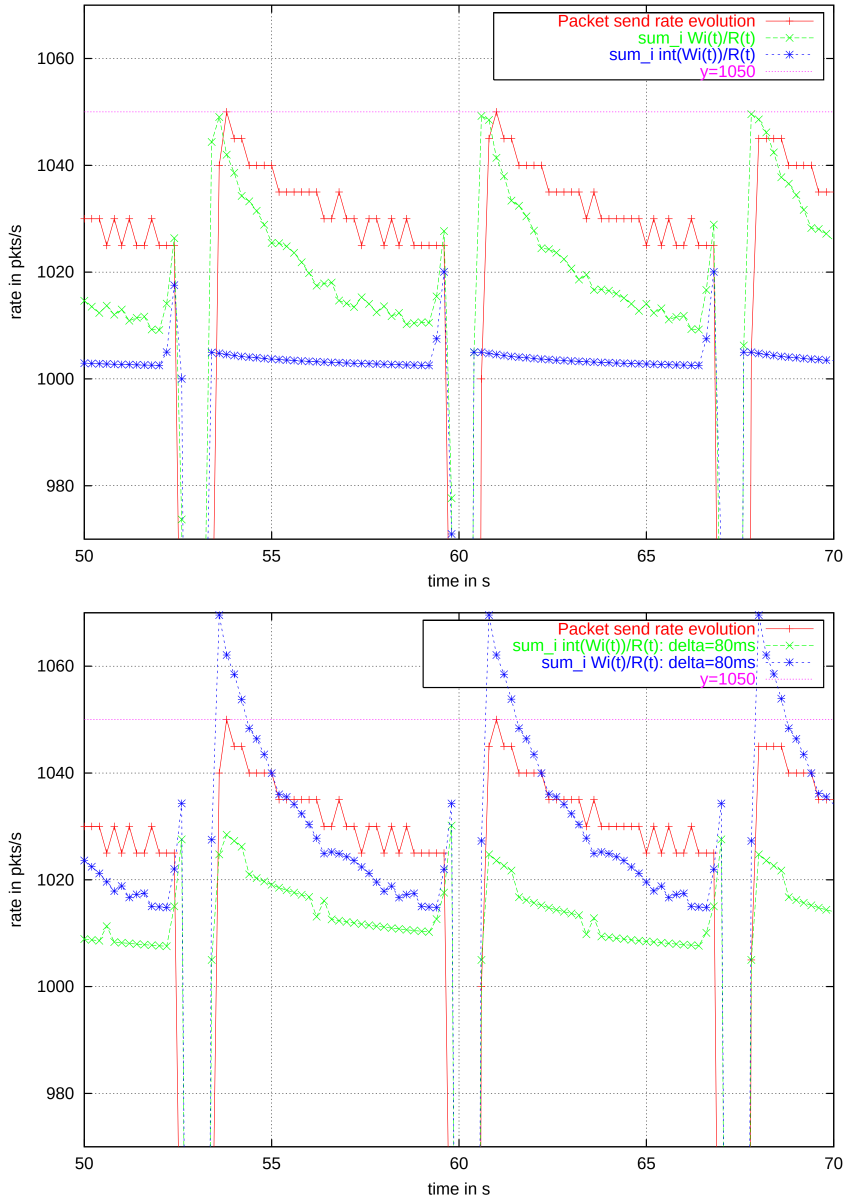

Figure 1: Top: Integer part effect on W. Bottom: Delay effect on W.

- *the second one is the pure transmission delay effect between the source and the bottleneck queue. To erase this effect we took by default a delay between source and bottleneck queue equal to zero (in particular in Figure 1 Top). Let's call this delay* δ. *Figure 1 (Bottom) shows that this has a similar effect as the first point. This is due to the fact that when the window size is increased, the impact on the queue size (therefore on RTT) is delayed by* δ. *Figure 1 (Bottom) shows also the combined effect of these two points.*

For the above two reasons, the approximation (13) *has the same monotonicity than* $X(t)$, *but it may also be very bad. In particular modelling interaction of TCP with AQM router, one* *should be aware that while the buffer is not empty, the approximation* (13) *may introduce an arbitrary deviation from the real case up to* $1/R(t)$.

2.2 Multiple Flows Model

For N parallel flows sharing a bandwidth C, Equations (5)-(8) become:

$$\dot{W}_i(t) = \frac{1}{R_i(t)} \tag{14}$$

$$R_i(t) = R_{i,\min} + \frac{Q(t)}{C} \tag{15}$$

$$\dot{Q}(t) = X(t) - C \tag{16}$$

$$X_i(t) = \left(1 + \frac{1}{W_i(t)}\right) Y_i(t), \tag{17}$$

where $Y_i(t)$ is the rate at which ACKs arrive to source i.

Theorem 2. *We have the equality:*

$$\sum_{i=1}^{N} \frac{W_i(t)}{R_i(t)} = C.$$

Proof We have:

$$\dot{W}_i(t) = \frac{Y_i(t)}{W_i(t)} \text{ and } \sum_{i=1}^{N} Y_i(t) = C.$$

Therefore,

$$\sum_{i=1}^{N} \dot{W}_i(t) W_i(t) = \sum_{i=1}^{N} \frac{W_i(t)}{R_i(t)} = \sum_{i=1}^{N} Y_i(t) = C.$$

□

□

We introduce the following notation:

- $w(t) = W(t)/N = \sum_{i=1}^{N} W_i/N$;
- $x(t) = X(t)/N = \sum_{i=1}^{N} X_i/N$;
- $c = C/N$;
- $q(t) = Q(t)/N$;
- $r(t) = \sum_{i=1}^{N} R_i(t)/N$;
- $r_{\min} = \sum_{i=1}^{N} R_{i,\min}/N$;

Theorem 3. *q is the solution of the equation:*

$$\dot{q}(t) = \sum_{i=1}^{N} \frac{c}{q(t) + cR_{i,\min}}$$

Proof From (16)-(17), we have:

$$\dot{Q}(t) = \dot{W} = \sum_{i=1}^{N} \frac{1}{R_i(t)} = \sum_{i=1}^{N} \frac{1}{Q(t)/C + R_{i,\min}} = \sum_{i=1}^{N} \frac{C}{Q(t) + CR_{i,\min}}$$

□

□

Now we consider the following approximation equations:

$$\textbf{(H0)}: \quad \dot{w}(t) = \frac{1}{r(t)} \tag{18}$$

$$r(t) = r_{\min} + \frac{q(t)}{c} \tag{19}$$

$$\dot{q}(t) = x(t) - c \tag{20}$$

$$x(t) = c + \dot{w}(t) \tag{21}$$

Among the equations (18)-(21), only (18) is an approximation. (18) becomes exact if $\sum_{i=1}^{N} R_i/N = N/\left(\sum_{i=1}^{N} 1/R_i\right)$. This is true in the homogeneous case (i.e., all $R_{i,\min}$ are equal) or is asymptotically true if $R_{i,\min} << q(t)/c$.

Theorem 4. *Under the assumption (H0), the solution is given by:*

$$r(t) = \sqrt{\frac{2}{c}t + r_{\min}^2} \tag{22}$$

$$w(t) = \sqrt{2ct + w(0)^2} \tag{23}$$

$$q(t) = \sqrt{2ct + w(0)^2} - w(0) \tag{24}$$

$$x(t) = c + \frac{1}{\sqrt{\frac{2}{c}t + r_{\min}^2}}. \tag{25}$$

Theorem 5. *While $q(t) > 0$, we have:*

$$\frac{w(t)}{r(t)} = C$$

except during the transient period, between packet losses time and the window decrease updates time (i.e. during Fast retransmit/Fast recovery phase).

Proof Before the first loss, we have.

$$\frac{w(t)}{r(t)} = C.$$

If $p(t)$ is the synchronization rate at time t (that is the window decrease at time t is $w(t)p(t)/2$), then just after the first window decrease updates, we have:

$$\frac{w(t^+)}{r(t^+)C} = \frac{w(t^-) - w(t^-)p/2}{r_{\min}C + w(t^-) - w(t^-)p/2 - W(0)} = 1.$$

□

□

It is well known that TCP does not utilize 100% of the bandwidth (e.g. cf. [8]). The following Lemma gives a simple buffer dimensionning rule.

Lemma 1. *For the link utilization to be 100 %, it is sufficient to have:*

- *[Under the DropTail policy]* $B \geq r_{\min} C$.
- *[Under any AQM policy]* $p(q) \leq \frac{2q}{q + r_{\min} C}$, *under the assumption that* $p(q)$ *(the packet drop probability when the queue size is* q*) is small enough to be approximately equal to the proportion of flows experiencing packet losses when the queue size is* q.

Proof The queue never empties if

$$q(t) - w(t)p(t)/2 = q(t)(1 - p(t)/2) - p(t) r_{\min} C/2 > 0,$$

which is equivalent to

$$q(t) > \frac{p}{2-p} r_{\min} C.$$

Under the DropTail policy, losses only occur when $q(t) = B$, so that the queue never empties if $B > r_{\min} C > \frac{p}{2-p} r_{\min} C$. □

□

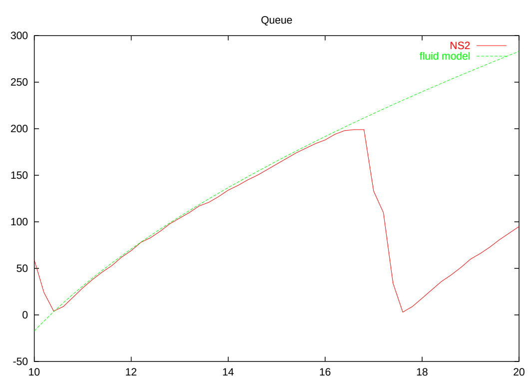

Figure 2: Queue size evolution.

3 Comparison with NS2 simulation

The default simulation scenario is: C=1000 pkts/s, B=200 pkts, N=10, RTTmin=200ms, Packet size = 1 Kbytes.

Figure 2 shows the square root evolution of the queue size in time (in seconds) during the congestion avoidance phase.

Figure 3 shows the accuracy of the condition of Lemma 1: in this case, the buffer never empties if B>200 pkts. With B=202, the buffer indeed never empties, whereas with B=198, the buffer empties just at congestion epoch.

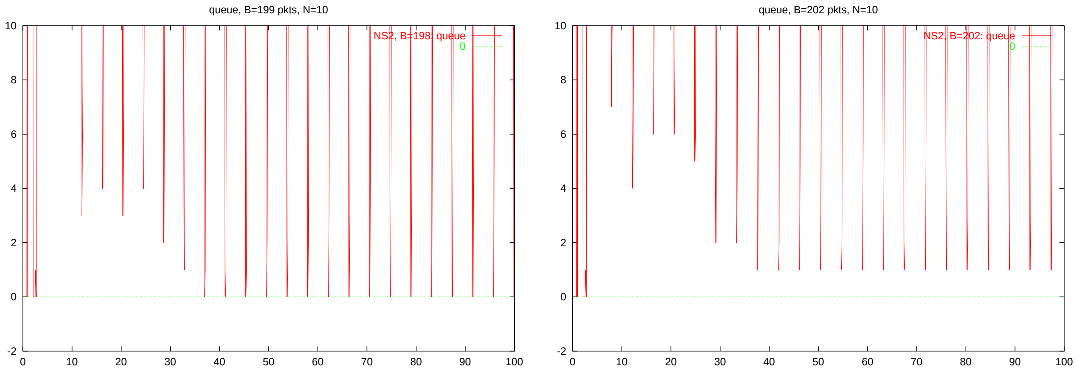

Figure 3: Queue emptiness tests. Left: B=198, queue empties. Right: B=202, queue never empties.

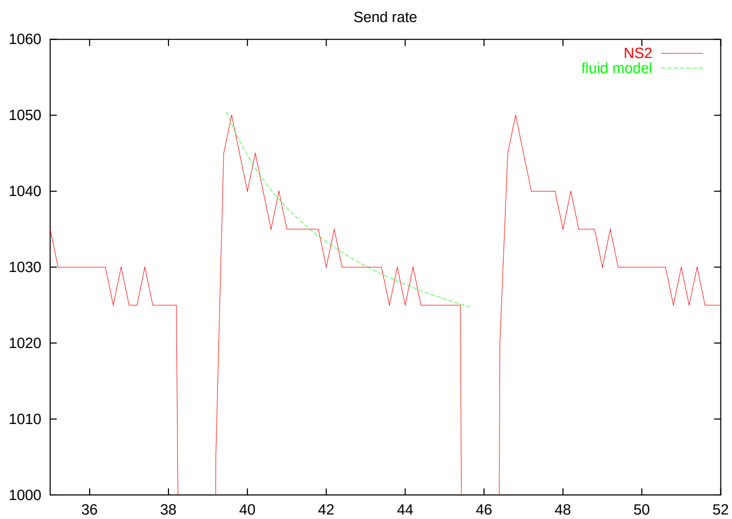

Figure 4: Send rate evolution.

Figure 4 shows that as predicted by the theory, the total send rate of the sources decreases !!! when the queue fills in with a pick rate of $C + 1/r_{\min} = 1050$ pkts/s as initial condition.

On Figure 5, the RTT is computed based on the queue size (and not on the srtt). Figure 5 shows that the property that W/RTT is constant is very well satisfied.

Figure 6 shows that the property that W/RTT is constant is quite well satisfied even for the model with different RTTs.

We see that our estimate is quite accurate to predict the results shown by NS2 simulations.

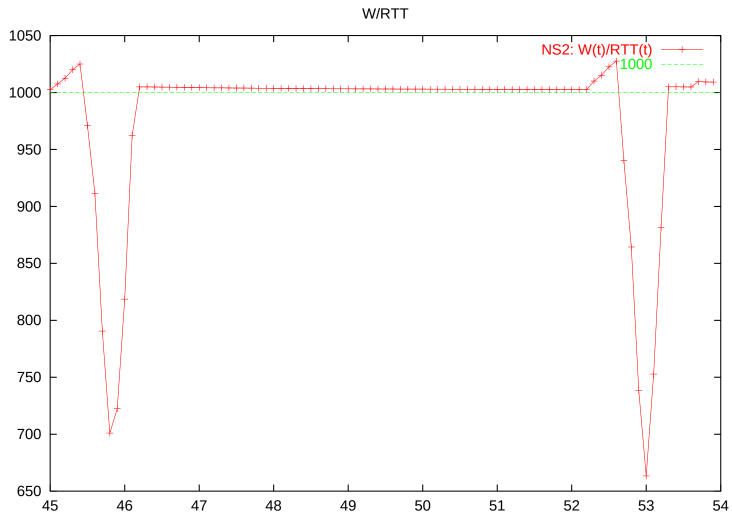

Figure 5: Evolution of W/RTT.

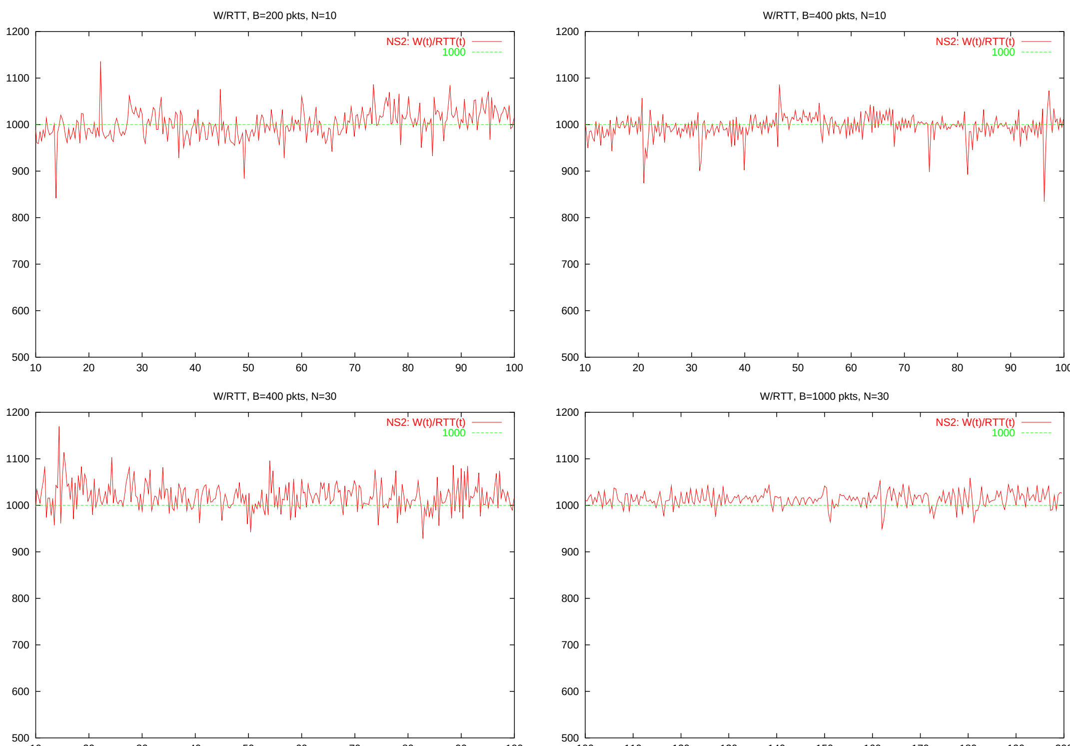

Figure 6: Heterogeneous RTTs: RTT=0.1+Unif[0,0.2]. N=10 (above. Left: B=200 pkts, Right: B=400), N=30 (below. Left: B=400, Right: B=1000).

References

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

[2] Baccelli, F. and Hong, D. (2003) Interaction of TCP flows as Billiards. *Proc. of INFOCOM*, *INRIA Report*, RR-4437, April 2002, INRIA Rocquencourt.

[3] Bonald, T. (1998) Comparison of TCP Reno and TCP Vegas via Fluid Approximation. *INRIA Report*, RR-3563, INRIA Sophia-Antipolis.

[4] Fall, K. and Floyd, S. (1995) Simulation-based Comparison s of Tahoe, Reno and SACK TCP. *Technical Report*, http//www-nrg.ee.lbl.gov/nrg-papers.html

[5] Hollot, C.V., Misra, V., Towley, D. and Gong, W.B. (2001) A Control Theoretic Analysis of RED. *Proc. of INFOCOM.*

[6] Kelly, F., Maulloo, A. and Tan, D. (1998) Rate control for communication networks: shadow price, proportional fairness and stability. *J. of the Operational Research Society*, 49, pp. 237-252.

[7] Kunniyyr, S. and Srikant, R. (2000) End-to-end congestion control schemes: utility functions, random losses and ECN marks. *Proc. of INFOCOM.*

[8] Lakshman, T.V. and Madhow, U. (1997) The performance of TCP/IP for networks with high bandwidth-delay products and random loss. *IEEE/ACM Trans. on Networking*, 5-3, pp. 336-350.

[9] Low, S.H., Paganini, F., Wang J., Adlakha S.A., Doyle J.C. (2002) Dynamics of TCP/AQM and a Scalable Control. *Proc. of INFOCOM*, June.

[10] Mathis, M., Semke, J., Mahdavi, J. and Ott, T. (1997) The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm. *Computer Communication Review*, 27(3), July.

[11] Misra, V., Gong, W. and Towsley, D. (2000) Fluid-based Analysis of a Network of AQM Routers Supporting TCP Flows with an Application to RED. *Proc. of SIGCOM*, Stockholm, September.

[12] G. Wright, R. Stevens, *TCP Illustrated Vol. 2*, Addison Wesley, 1995.

[13] Zhang, L., Shenker, S. and Clark, D. (1991) Observations on the dynamics of a congestion control algorithm: The effects of two-way traffic. *proc. of SIGCOMM.*

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