Bounds on the Throughput of Congestion Controllers in the Presence of Feedback Delay

Sanjay Shakkottai, R. Srikant and Sean Meyn

Abstract— **We consider decentralized congestion control algorithms for low-loss operation of the Internet using the ECN bit. There has been much analysis of such algorithms, but with a few exceptions, these typically ignore the effect of feedback delays in the network on stability. We study a single node with many flows passing through it, with each flow (possibly) having a different round trip delay. Using a fluid model for the flows, we show that even with delays, the total data rate at the router is bounded; and this bound shows that the (peak) total rate grows linearly with increase in system size, i.e., the fraction of over-provisioning required is constant with respect to N, the number of flows in the system. Further, for typical user data rates and delays seen in the Internet today, the bound is very close to the data rate at the router *without* delays. Earlier results by Johari and Tan have given conditions for a linearized model of the network to be (locally) stable. We show that even when the linearized model is not stable, the nonlinear model is upper bounded, i.e., the total rate at the bottleneck link is upper bounded, and the upper bound is close to the equilibrium rate for TCP.**

I. INTRODUCTION

This paper concerns decentralized end-to-end algorithms for Internet congestion control. The goal is to study the performance of algorithms designed for maintaining low-loss and low-delay in the network.

In recent years there has been much research in this area, mainly concerning algorithms based on ECN-marking. These control algorithms are designed on the premise that each user has a utility function, which the user is trying to maximize, while the network is simultaneously trying to maintain some sort of fairness amongst various users. In the algorithms proposed, the network tries to achieve its goal by *marking* packets during congestion (see [18], [**?**]). The notion of fairness (from the network's point of view) which has been used is weighted *proportional fairness* (see [6]). Through appropriate choice of the weights, other fairness criteria such as minimum potential delay fairness (see [15]) can be realized. The algorithms proposed have a *decentralized* implementation to achieve the network and user objectives simultaneously.

Our concern in this paper is performance of these systems in the presence of delay. It is known that in a network model without delays, a well-designed marking scheme will ensure that all of the user-rates converge [7], [**?**], [11]. However, if one includes delay in the model, it is possible for the user rates to oscillate around some equilibrium point. To operate the network with very low levels of loss, it then becomes necessary to provide enough capacity to account for these oscillations. In this paper, we derive an upper bound on these oscillations, and using these bounds we provide an estimate of the *over provisioning* (above the equilibrium rate) required for low-loss network operation.

Prior research in [5] addresses the effect of delays on stability through the analysis of a *linearized system*. In their work it is assumed that all users have the utility function $\log(x)$, and all have the same round trip delays for a general network topology. The network is shown to be locally stable if the marking function (i.e., the algorithm implemented at each router to mark packets), gain and delay satisfy suitable bounds (see Section V). In [14], [21], this result has been generalized for the case where the delays may be different.

We also refer the reader to a linearized analysis of congestion controllers, where the marking function is itself adapted as a function of the arrival rate at a router [12]. In this paper, we do not consider such adaptation, but assume that the marking function is fixed. We point out that round-trip delays are considered explicitly in [13], where dual solutions are obtained for the problem proposed in [6]. However, the results in[13] simply state that, for sufficiently small gains in the network control algorithms, asymptotic stability is achieved.

Related research in the ATM domain also considers rate control algorithms in the presence of feedback delay (for instance, [2]). However, those studies are primarily for single flows, whereas, we consider multiple flows with different delays. Further, the form of the controller is different as well. For instance, in [2] the authors study controllers which explicitly make use of the router queue-length.

We consider the following setting in the present paper.
(a) The network consists of a single node that accommodates many individual flows. There is a round trip delay, which may be different for each flow.
(b) All of our analysis, and our motivation for the algorithms considered, is based on a deterministic fluid model. We ignore stochastic effects in the network which could occur possibly due to randomness in marking or short uncontrolled flows (web mice). We also ignore discretization effects due to packet based implementation.
(c) We assume that the queueing delay at the router is negligible compared to the propagation delays. A simple technique to achieve this is to use a virtual queue at each link whose capacity is slightly smaller than the capacity of the real queue. Packets are then marked based on the occupancy of the virtual queue [8], [12].
In related papers we discuss these effects which could potentially cause differences between the stochastic model and the deterministic fluid model we study here, and show that the fluid models that we study in this paper are valid when there are large number of flows in the network [19], [20].

Regarding (c), we note that Kelly has argued that as the the cost of router hardware decreases and network bandwidth increases rapidly, we do expect that queue lengths at routers will

Research supported by NSF Grants ANI-9813710, NCR-9701525 and DARPA Grant F30602-00-2-0542.

This work was done at the University of Illinois at Urbana-Champaign. S.Shakkottai is currently with the Department of Electrical and Computer Engineering, The University of Texas at Austin, shakkott@ece.utexas.edu. R. Srikant and S. Meyn are with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, {rsrikant, s-meyn}@uiuc.edu.

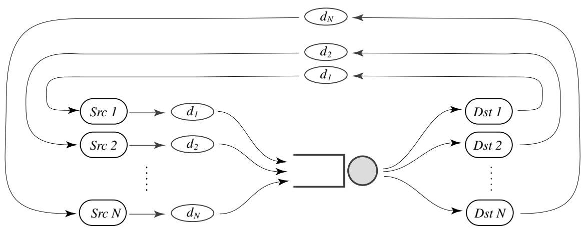

Fig. 1. Model of the system with delays. We have N connections, each with (possibly) different round trip delays.

be small in practice [8]. It has also been argued in [12] that one can design low-loss networks which achieve high utilization while simultaneously having small queues at the router. On the other hand, propagation delays over fibers will not decrease, as this is a fundamental property of nature.

The main contributions of this paper are the following:
(i) Using a fluid model for the flows, we show that even with delays, the total data rate at the router is bounded, and this bound shows that the (peak) total rate grows at most linearly with increase in system size, i.e., the number of flows. We assume throughout that the ratio of the link capacity to the number of flows is kept constant. In other words, the link capacity is is a linear function of the number of flows.
(ii) In simulations, we have considered the regime where the round trip delays are of the order of tens of milliseconds, and each user desires a throughput of more than 100 kbytes/sec. In this regime, numerical results in Section VI indicate that the overprovisioning required to operate a low-loss network is a small fraction of the link capacity. We have analytically substantiated this observation in Section IV, where we have shown that, if the bandwidth-delay product is large, then the amount of over-provisioning required becomes a small fraction of the link capacity.
(iii) In Section V, we will demonstrate that even if the linearized system is unstable, it is possible for the user rates to oscillate very close to the equilibrium rate. Further, in [5], a condition has been given for stability of the linearized network. In Section V, we also we give a condition for the non-linear congestion controllers (with the $\log(.)$ utility function and same round trip delays) which ensures that the sum of the user rates does not oscillate very much (i.e., the upper and lower bounds on the total rates are "close"). The condition yields a design rule for controller gains which coincides with the rule obtained in [5] via linearization.

II. Description of the Model

Consider a system consisting of N best-effort sources accessing a single-node network, where each source could experience a different, large propagation delay as shown in Figure 1. We denote the ith connection's (one-way) delay by d_i. We consider a fluid approximation for the system, where each user transmits data at a time-varying rate depending on feedback from the network.

At the router, the bottleneck link has a capacity of $N\hat{c}$. However, the system is designed so that in the delay-free case, the total arrival rate converges to $Nc < N\hat{c}$. The excess, $N(\hat{c}-c)$ is reserved for handling arrival rate oscillations due to delays and other perturbations; so that the router can maintain loss free operation in-spite of these effects. Our objective is to derive an upper bound on the total arrival rate, so that by choosing $\hat{c}$ appropriately, loss free operation can be maintained in the presence of delays. In the rest of this paper, we will refer to Nc as the *target arrival rate* of the link. Associated with the router is a marking function $p^{(N)}(.)$. This function is based on the total data rate accessing the router and determines the fraction of flow to be marked, and satisfies satisfies the following criteria:
(i) $0 \leq p^{(N)}(x) \leq 1$
(ii) $p^{(N)}(x)$ is an increasing function.
(iii) $p^{(N)}(x)$ is Lipschitz continuous.
The first property is obvious, as the marking function represents the fraction of flow marked. The second property is again clear: the larger the arrival rate is, the greater is the fraction marked. Finally, the last condition is a technical condition, which says that the function is "smooth". As an example, a possible rate-based marking function is of the form

$$p(x) = \frac{(x-\tilde{c})^{+}}{x} \tag{1}$$

In a deterministic fluid model, this has the interpretation of the fraction of fluid lost when the arrival rate exceeds a certain level, called the "virtual" capacity, $N\tilde{c}$ [10].

Next, we consider each user's behavior. Let us denote the transmission rate of the ith source at time t by $x_i(t)$. We first look at the case when the round trip delay for every flow is zero. Under this assumption, we now describe each user's rate adaptation mechanism (proposed in [6]) which will lead to a proportionally-fair allocation of rates. For each user $i = 1, \ldots, N$, the rate adaptation is as follows:

$$\begin{aligned} \dot{x}_i(t) &= \Delta_i - \beta\gamma_i(t) \\ \gamma_i(t) &= x_i(t)p^{(N)}\left(\sum_{j=1}^{N} x_j(t)\right) \end{aligned}$$

where Δ, β are positive constants which determine the rate at which flow increases or decreases it's transmission rate. We can interpret $\gamma_i(t)$ as the rate at which marked packets arrive to user i. Thus, in the absence of marks each user, each user additively increases the rate with parameter Δ_i and if marks are present, then, multiplicatively decreases the rate in proportion to $\gamma_i(t)$.

It has been shown in [6] that with the marking function described by (1), the above system of equations converge (as $t \to \infty$), and the equilibrium rates are proportionally fair.

We now consider the same system, *but with delays*. Then, we can describe the user rate adaptation by means of the following delay-differential equations. For each $i = 1, \ldots, N$,

$$\dot{x}_i(t) = \begin{cases} \Delta_i - \beta x_i(t-2d_i) \\ p^{(N)}\left(\sum_{j=1}^{N} x_j(t-d_i-d_j)\right) & \text{if } x_i(t) > 0 \\ \Delta_i & \text{if } x_i(t) = 0 \end{cases} \tag{2}$$

To understand the above expression, we note that the evolution of the transmission rate of user i at time t depends on the marking rate at the router at time $t - d_i$. However, the marking rate at the router at time $t - d_i$ depends on the transmission rate of every user $j = 1, \ldots, N$ at times $t - d_i - d_j$ respectively. In this case, we explicitly have a non-negativity constraint. Without this constraint, in the presence of delays, it is possible for the transmission rates to be negative, which is physically not possible.

The other two systems we consider approximate the behavior of TCP. We note that we do not consider window flow control mechanisms here. Instead, we focus on a rate-control model for the congestion control phase of TCP. The congestion control phase of TCP can be described by the following delay-differential equations. For each $i = 1, \ldots, N$,

$$\dot{x}_i(t) = \Delta_i - \beta x_i(t)\, x_i(t-2d_i)\, p^{(N)}\left(\sum_{j=1}^{N} x_j(t-d_i-d_j)\right) \quad (3)$$

with $\Delta_i = \frac{1}{4d_i^2}$, and $\beta \approx \frac{2}{3}$. The above delay-differential equation is an obvious extension of the model without delay that was derived in [10]. In [10], the value of β was shown to be approximately $\ln(2)$. Other empirical evidence (for example, [16]) suggests that $beta$ is $2/3$. Since $\ln(2)$ is approximately equal to $2/3$, we choose to use $2/3$ as the value for $beta$ in this paper.

Another model for TCP suggested in [7] is given by

$$\dot{x}_i(t) = \Delta_i\left(1 - p^{(N)}\left(\sum_{j=1}^{N} x_j(t-d_i-d_j)\right)\right) - \beta x_i(t)\, x_i(t-2d_i)\, p^{(N)}\left(\sum_{j=1}^{N} x_j(t-d_i-d_j)\right) \quad (4)$$

In the rest of this paper, we will refer to the above two models as TCP-like algorithms.

In [9], it has been shown that when $\{d_i\} = 0$, the above systems (which are now ordinary differential equations) converge, i.e., there is a unique equilibrium $x^* \in \mathcal{R}^N$ such that $x(t) \to x^*$ as $t \to \infty$, for any initial condition $x_0 \in \mathcal{R}^N$. In this paper, we are concerned with the boundedness of the delay-difference equations. We will show that even with delays, the total rate at the router is bounded, and this bound can be used as the over-provisioning required to account for delays.

As described earlier, the marking function $p^{(N)}(.)$ is the means by which the router decides how aggressively the flows back-off during congestion, and this function depends on the target arrival rate of the link Nc. This function is designed so that in the absence of delays, the total user rate in equilibrium[1] $\sum_{j=1}^{N} x_j(t)$ is equal to Nc, and $p'(Nc) > 0$. To illustrate the above, we consider the following example.

Example II.1: Consider the proportionally-fair system described by (2), with $\Delta_i = \Delta\ \forall i$. Hence, in a delay-free system, under equilibrium, the rates of individual users $x^*_{i,N} = c\ \forall, i, N$. The marking function we consider is that described in (1), i.e.,

$$p^{(N)}(\lambda) = \frac{(\lambda - N\tilde{c})^+}{\lambda}$$

for some $0 < \tilde{c} < c$. As we discussed earlier, in a deterministic fluid model, this has the interpretation of the fraction of fluid lost when the arrival rate exceeds the "virtual" capacity $N\tilde{c}$. It can also be shown ([10]) that this function is the marking function derived by taking the limit of the $M/M/1/B$ loss formula, when we scale the arrival rate, target arrival rate (Nc) and buffer size simultaneously. We note that for $\lambda = Nc$, $p^{(N)}(Nc) = \frac{c-\tilde{c}}{c}$, which is a constant independent of N.

From (2), in equilibrium, we have

$$p^{(N)}(Nc) = \frac{c-\tilde{c}}{c} = \frac{\Delta_i}{\beta c}$$

which follows from the fact that $x_i^* = c$. Therefore, by choosing $\tilde{c} = c - \frac{\Delta_i}{\beta}$, the required scaling of $p^{(N)}(.)$ is achieved. Finally, we comment that the scaling property obtained above,

$$p^{(N)}(Nx) = p(x) \quad x \geq 0 \quad (5)$$

is also satisfied for several other proposed rate based marking functions (see the marking functions for REM and virtual queue overflow in [8]).

[1] The equilibrium rates are the solution to the algebraic equation obtained from the delay-differential equation when we set the propagation delays as well as all derivatives to be equal to zero.

III. An Upper Bound on the Transmission Rates

In studies of the system described in the previous section, *assuming* that the delays are zero, [9] shows that the rates converge to an equilibrium value as $t \to \infty$. From simulations, we observe that with delays, the systems defined in the previous section need not converge in general. In this section, we derive an upper bound on how much the rates can diverge from the equilibrium rate obtained with zero delays.

We first consider a system of the form

$$\dot{x}_i(t) = \begin{cases} \Delta_i - \beta \frac{U'(x_i(t))^{-1}}{x_i(t)}\, x_i(t-2d_i)\, p^{(N)}\left(\sum_{j=1}^{N} x_j(t-d_i-d_j)\right) & \text{if } x_i(t) > 0 \\ \Delta_i \quad \text{if } x_i(t) = 0 \end{cases} \quad (6)$$

where $U(x)$ is a strictly concave, increasing function called the user utility function. Note that for the models considered earlier, the proportionally fair model (2) corresponds to $U(x) = \log(x)$, and (3), a model for TCP, corresponds to $U(x) = \frac{-1}{x}$.

It is clear that there is a trivial upper bound for the total rate (i.e., $\sum_{i=1}^{N} x_i(t)$) at the router. Consider any single flow, say flow 1. To construct the upper bound, we construct an upper bound of each user's flow, and take the sum of these bounds as the bound on the total rate at the router.

Let us consider user 1. Its transmission rate will become very large if there are no other flows in the system, thus allowing this flow to utilize the entire link bandwidth. Thus, a trivial bound can be obtained on a single flow rate by assuming that the congestion at the router is caused solely by this flow (i.e., $x_j(t) = 0$ for all $j \neq 1$). We observe that as the target arrival rate of the system is Nc, for a single flow to cause congestion, the rate should be $O(N)$ before marking will be significant enough to cause the source to back-off. Denote R as a bound on $x_1(t)$. We will illustrate the above procedure and derive the upper-bound R for the special case when the marking function is given by

$$p^{(N)}(x) = \frac{(x - N\tilde{c})^+}{x}, \quad x \geq 0.$$

We observe that for any marking to take place (and thus, for the flow to back off), we must at least have $R > N\tilde{c}$. This argument suggests that an upper bound on $\sum_i x_i(t)$ is at least NR, which is greater than $N^2\tilde{c} \sim O(N^2)$. The implication of this is that as we increase the capacity of the network, more and more over-provisioning is required. We will find that this implication is *false* for the model considered here.

We derive a much tighter bound in this section, and show that the bound is actually $O(N)$ for marking functions satisfying (5), i.e., the fraction of over-provisioning required is constant with respect to N. Further, numerical computation of the bound will show that this fraction is actually quite small. To this end, let

$$\begin{aligned}\Delta &= \frac{1}{N}\sum_{i=1}^{N}\Delta_i\\ D_i &= \max_j \frac{d_i+d_j}{2}\\ D &= \max_i D_i\\ d &= \min_i d_i\\ K &= \max_i 2\Delta_i d_i\end{aligned}$$

Now, we make the following assumption on the utility function.

Assumption III.1: Let

$$f(x) = U'(x)^{-1}\left(1-\frac{K}{x}\right)$$

We assume that $f(x)$ is convex and increasing for $x \geq K$. Note that $U(x)=\log(x)$ and $U(x)=\frac{-1}{x}$ satisfy this assumption.

Theorem III.1: Let $M>0$ satisfy

$$p^{(N)}\left(N(M-2\Delta D)\right)U'(M)^{-1}\left(1-\frac{K}{M}\right) = \frac{\Delta}{\beta} \quad (7)$$

Then, given $\epsilon>0$, there exists a finite $t^*(\epsilon)$ such that $\forall t \geq t^*$,

$$\sum_{i=1}^{N} x_i(t) \leq NM+\epsilon.$$

Proof: We first note that

$$\frac{dx_i}{dt} \leq \Delta_i, \qquad \forall i.$$

Now, suppose that $\sum_j x_j(t) > NM$ for some $t>0$. Then,

$$\begin{aligned}\sum_j x_j(t-d_i-d_j) &\geq \sum_j x_j(t) - \sum_j \Delta_j(d_i+d_j)\\ &\geq \sum_j x_j(t) - 2N\Delta D\\ &> N(M-2\Delta D).\end{aligned}$$

We also note that $x_i(t-2d_i) \geq x_i(t)-K$. Thus,

$$\dot{x}_i(t) \leq \Delta_i - \beta\frac{(U'(x_i))^{-1}}{x_i}(x_i-K)p^{(N)}(N(M-2\Delta D)).$$

Therefore,

$$\begin{aligned}\sum_i \dot{x}_i(t) &\leq \sum_i \Delta_i - \beta p^{(N)}(N(M-2\Delta D))\sum_i f(x_i)\\ &\leq N\Delta - \beta p^{(N)}(N(M-2\Delta D))Nf\left(\frac{1}{N}\sum_i x_i\right)\\ &\leq N\Delta - \beta p^{(N)}(N(M-2\Delta D))Nf(M).\end{aligned}$$

The last two lines in the above set of inequalities follow from the Assumption III.1. Thus, if

$$p^{(N)}\left(N(M-2\Delta D)\right)U'(M)^{-1}\left(1-\frac{K}{M}\right) > \frac{\Delta}{\beta},$$

then $\sum_i \dot{x}_i(t) < 0$ whenever $\sum_j x_j(t) > NM$. Further, it is easy to show that given any $\epsilon>0$, there exists $\delta>0$ such that whenever $\sum_j x_j(t) > NM+\epsilon$, $\sum_i \dot{x}_i(t) < -\delta$. Thus, the result follows. ⋄

We comment that from the proof, it is clear that if the initial condition is small enough, the bound works for all $t \geq 0$. As discussed earlier [7] suggests that (4) could be a suitable model for a rate-based TCP evolution. While this model doesn't fit into the class of equations (6), note that it is nearly identical to (3), except for the fact that the additive increase part is slightly different to account for the fact that when a mark is received, the rate is not increased. However, it is easy to verify that the proof of the previous theorem holds for (4) also and the upper bound for (4) is the same as the upper bound for (3).

Example III.1: The results in the previous section provide means of constructing bounds on the total rate at the router for various utility functions. As an example, we will now study a system with N flows, each with delay d, growth parameter Δ, and the logarithmic utility function. We will assume that the marking function is $p^{(N)}(\lambda) = \left(\frac{\lambda-N\tilde{c}}{\lambda}\right)^+$, which is the same as that used in the example in Section II. Thus, we have $\forall 1 \leq i \leq N$,

$$\dot{x}_i(t) = \begin{cases}\Delta - \beta x_i(t-2d) \\ p^{(N)}\left(\sum_{j=1}^{N} x_j(t-2d)\right), & x_i(t)>0\\ \Delta, & x_i(t)=0\end{cases}$$

For the above congestion controller, (7) becomes

$$(M-2\Delta d-\tilde{c}) = \frac{\Delta}{\beta}.$$

Note that this leads to an upper bound on the total rate which is $O(N)$, as opposed to that from the trivial $O(N^2)$ bound derived earlier.

IV. Discussion of the Upper Bound

Consider a congestion controller with $U(x)=-1/x$ which, as discussed earlier, can be used to model TCP. Suppose that the round-trip delays of all the flows are approximately the same, i.e., $d_i \approx d$ for all i. Then, using the fact that for a TCP-like controller, $\Delta_i = \frac{1}{4d_i^2}$, (7) becomes

$$p^{(N)}\left(\frac{N}{d}(Md-\frac{1}{2})\right)(Md)^2\left(1-\frac{1}{2Md}\right) = \frac{1}{4\beta}$$

Suppose the *bandwidth-delay product*, $cd \gg 0.5$. Then, as $M>c$, the above equation can be approximated by

$$p^{(N)}(NM)(Md)^2 \approx \frac{1}{4\beta} \quad (8)$$

Now, recall that (from steady-state analysis) $p^{(N)}(Nc)$ satisfies

$$\frac{1}{4d^2} - \beta c^2 p^{(N)}(Nc) = 0 \quad (9)$$

Comparing (8) and (9), we have $M \approx c$. The conclusion is that for such a system with large bandwidth-delay products, even if the system with delay does not converge, the actual rate at the router will be nearly equal to the target arrival rate c. It is well-known that, for controlled systems with time delays in the feedback path, the controller gains should be sufficiently small to ensure stability. *This approximate analysis suggests that the congestion avoidance mechanism of TCP [4] naturally chooses the additive increase rate (i.e., $\Delta = \frac{1}{4d^2}$), to prevent large oscillations even in the presence of delay.* This can be justified more formally for the marking function described in (1) as follows.

For the marking function (1), (7) becomes

$$M(M-\frac{1}{2d}-\tilde{c}) = c(c-\tilde{c}).$$

Thus,

$$M = \frac{c(c-\tilde{c})}{M-\frac{1}{2d}-\tilde{c}}.$$

Since $M>c$, the above equation equation implies that $M-c < 1/2d$. Therefore,

$$\frac{M-c}{c} < \frac{1}{2dc}$$

which goes to zero as $cd \to \infty$.

V. Instability and Boundedness

We derive a lower bound on the total rate of (2), for the special case where all the round trip delays $\{d_i\}$ are equal. We consider this special case to illustrate that even though a system may be unstable, it may be globally upper and lower bounded, with the upper and lowers bounds being close to the equilibrium rate.

In this section, we consider only the special case of the $\log(.)$ utility function. Let us fix any N and define

$$y(t) = \frac{1}{N}\sum_{i=1}^{N} x_i(t)$$

so that by (2), with d equal to the common delay value, we have

$$\dot{y}(t) = \Delta - \beta y(t-2d)\, p\,(y(t-2d)) \quad (10)$$

with the initial conditions given by $y(t) = r$ for $t \in [-2d, 0]$. We will assume that $0 \le r \le c$, i.e., the initial condition is a constant non-negative trajectory. The differential equation for $y(t)$ is obtained by ignoring the non-negativity constraint on $x_i(t)$. It can be shown that, for reasonable initial conditions, the trajectories will remain non-negative even without explicitly having the non-negativity constraint on each source's rate. However, for the sake of brevity, we omit the proof of this fact here and refer the interested reader to [19], [20]. The function $p(.)$ in the above differential equation is the *unscaled* marking function.

The derivation of the upper bound is much simpler and more intuitive here than in the earlier section because of the specific form of (10), and therefore, we present it below. We first reason that there is an M such that $y(t) < M$ for $t > 0$. To do this, we observe that $\dot{y}(t) \le \Delta$. Now, suppose $y(t) > M > c$. Then, since $\dot{y}(t) \le \Delta$, we have that $y(t-2d) > M - 2\Delta d$. Thus, choosing M such that

$$\Delta - \beta(M - 2\Delta d)p(M - 2\Delta d) \le 0$$

we have $\dot{y}(t) < 0$. Finally, as $p(.)$ is increasing, and $\Delta = \beta c p(c)$, it follows that an upper bound for $y(t), t \ge 0$ is given by

$$M = c + 2\Delta d$$

Let us now suppose that at time $t \ge 0$, $y(t) < l$ for some l. Then, it follows from (10) that

$$y(t-2d) < l + 2d\,(\beta M p(M) - \Delta)$$

Let $w(l) = l + 2d\,(\beta M p(M) - \Delta)$. Then,

$$\dot{y}(t) > \Delta - \beta w(l) p(w(l)).$$

Hence, choosing l so that it satisfies

$$w(l)p(w(l)) = \frac{\Delta}{\beta} \quad (11)$$

ensures that $\dot{y}(t) > 0$. Thus, if the trajectory of $y(t)$ ever hits l, then the derivative $\dot{y}(t)$ will be greater than zero, and therefore, $y(t)$ will continue to be greater than or equal to l. Hence, l is a lower bound on the total rate after some finite time. Now, we have $cp(c) = \frac{\Delta}{\beta}$. Thus, (11) can be equivalently represented by $w(l) = c$. Thus, the lower bound l is chosen such that

$$l = c - 2d(\beta M p(M) - \Delta) \quad (12)$$

Next, we provide some intuition on how the marking function, the gain and the round-trip delay are related. From the mean value theorem applied to the function $\beta x p(x)$, we have for some $y \in [c, M]$,

$$\beta M p(M) = \beta c p(c) + (M - c)\beta\,(y p'(y) + p(y))$$

Thus, subtracting Δ from both sides and using the fact that $\Delta = \beta c p(c)$, we have

$$\beta M p(M) - \Delta = (M - c)\beta\,(y p'(y) + p(y))$$

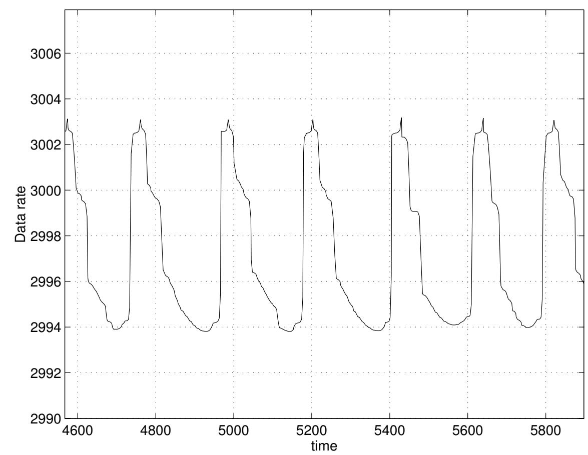

Fig. 2. Fluid simulation for a single connection with the $\log(x)$ utility function. The figure plots the data rate at the shared link versus time.

Thus, the lower bound l defined in (12) can be rewritten as

$$l = c - 2d(M - c)\beta\,(y p'(y) + p(y))$$

for some $y \in [c, M]$. If $(M-c)\beta\,(yp'(y) + p(y))\,2d$ is small *for all* $y \in [c, M]$, we clearly have $l \approx c$. In [5], the authors showed that the linearized system is stable if $\beta\,(yp'(y) + p(y))\,2d$ is small at $y = c$. From the discussion above, it indicates that a design rule for marking function is the following: *Choose* $p(.)$ *such that* $(yp'(y) + p(y))$ *is small for* y *in an interval about the target arrival rate*. This condition says that near the target capacity, the fraction of flow marked should be small. Further, the change in the fraction marked should be small as well, i.e., the marking function should not increase too "steeply". This intuitively makes sense because there would be large oscillations in the arrival rate if the fraction of flow marked changed a lot for small changes in the arrival rate. Depending on how close one would like the upper and lower bounds to be, one can obtain a bound on $\beta\,(yp'(y) + p(y))\,2d$.

On the other-hand, suppose a marking function is fixed. For example, if the marking function is that described by (1), then we have for all $y > \tilde{c}$, $(yp'(y) + p(y)) = 1$. Thus, in this case, we can see that if we want the trajectory to lie close to the equilibrium point, we need to have a small gain-delay (i.e., βd) product.

A. Illustration

In this section, we study a system which can be proven to be locally (and hence, globally) unstable, i.e., there *does not* exist a x^* such that the data rate $x(t) \to x^*$ as $t \to \infty$. However, we will show that the trajectory is bounded, with the bounds being close to the nominal throughput.

Consider a system of N users, with the $\log(.)$ utility function and all users sharing the same delay. Suppose the target arrival rate is $N * c$ bytes/unit-time. Since the delays are the same, and the initial conditions are the same, the evolution of the average rate is the same as that of a single connection, with target arrival rate per source being c bytes/unit-time. Thus, the evolution of this system can be given by

$$\dot{x}(t) = \Delta - \beta x(t-2d)p\,(x(t-2d))$$

with $p(\lambda) = \frac{(\lambda - \tilde{c})^+}{\lambda}$. Let us choose $\Delta = 1$, $\beta = 0.67$, $d = 2$ and $c = 3000$. Thus, for the equilibrium rate x^* to equal c, we

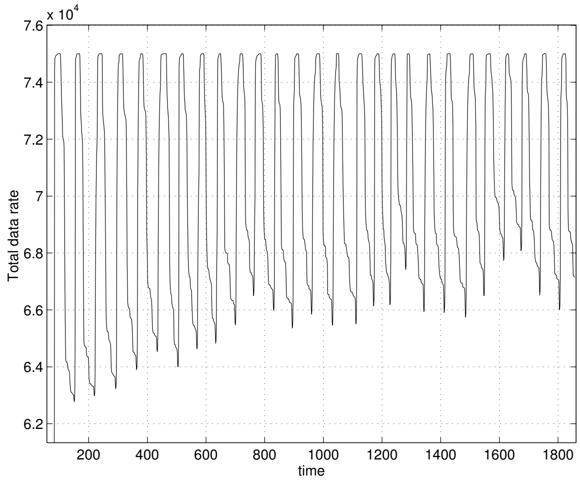

Fig. 3. Systems SYS-1: Fluid simulations for 25 flows, with the utility function being $\frac{-1}{x}$ and with different round trip delays. The figures plot the total rate at the shared link versus time. The marking function is $\left(\frac{\lambda - N\tilde{c}}{\lambda}\right)^{+}$.

need to choose $\tilde{c} = 2998.5$. We observe that in equilibrium, we have $x^* p'(x^*) + p(x^*) = 1$. From Theorem 5 in [5], this system is locally (and hence, globally) unstable if and only if

$$\beta(xp'(c) + p(c)) > \frac{\pi}{4d}$$

which is satisfied here. However, from the bounds derived in Section V, we get the upper bound on the rate as 3004 and the lower bound as 2989. This matches very well with the simulations in Figure 2, and the oscillations are small. *Thus, we can infer that even if the system is unstable, in some sense, it does not really matter as the rates are very close to the desired value.*

We note that our results are applicable only to rate-based marking functions and not to the queue-based marking functions presented in [17], [3]. An interesting topic for future research would be to check whether tight upper and lower bounds can be derived for queue-based marking functions when the local stability conditions are violated.

VI. Numerical and Simulation Results

In this section, we present simulation results for the fluid model and compare these with the bounds derived in the previous section.

We first a consider a network with 25 connections passing through it, with the total bandwidth at the shared link being 24Mbps, and the one-way delays of the individual connections ranging from 25msec to 125msec. Thus, scaling the one-way delay of the shortest round trip time to be 1, we have (one-way) delays ranging from 1 to 5, and the target arrival rate being $Nc = 25 * 3000 = 75000$ bytes/unit-time.

In Figure 3 (SYS-1), we have plotted the sum of the user rates at the shared link for the TCP-like system defined in (3). We use a threshold-based marking function of the form

$$p(\lambda) \quad = \quad \left(\frac{\lambda - \widetilde{C}}{\lambda}\right)^{+} \qquad (13)$$

where $\widetilde{C} = N\tilde{c}$. Table I presents the peak rate at the router observed from fluid simulations and the numerical upper bound computed from Section III. We first observe that the bounds

System	FS	UB	ER
SYS-1	75009	75036.5	75000
SYS-2	75215	78920.3	75215
SYS-3	75200	75538	75000

TABLE I

Comparison of fluid simulations and bounds for various systems. The first column indicates the system name, the second indicates the peak rate at the router after time t^* from fluid simulations, the third column indicates the numerical upper bound and the fourth column gives the equilibrium rate. Abbreviations key: Fluid Simulation (FS), Upper Bound (UB), Equilibrium Rate (ER).

are very close to the values observed in simulations. Further, for the system SYS-1, with no delays, the total rate would have converged to 75000 bytes/unit-time. As the upper bound from Table I is 75036, it indicates that very little over-provisioning is required at the router to handle delays.

Another marking function that we considered is

$$p^{(N)}(\lambda) \quad = \quad \left(\frac{\lambda}{Nc_1}\right)^{B}$$

This is motivated by an $M/M/1$ queue, in which the above expression is the probability that the queue-length exceeds some buffer size B when the mean virtual link capacity is Nc_1 and the mean arrival rate is λ. In Figure 4, we consider another system (SYS-2), where we use the above marking function, and the target capacity is approximately 75000 bytes/unit-time (to be exact, 75215 bytes/unit-time). This marking function is "smoother" than the earlier one, and the non-linear system converges. From Table I, the bound is given by 78920.3, which is within 5% of the converged value. As before, from the bound, this indicates that we do not need much over-provisioning. Fluid simulations indicate that the throughput of this particular system converges, and therefore, we do not need any over-provisioning due to the presence of feedback delays. However, the bounds suggests that we may need at most 5% over-provisioning, which suggests that the bound is not very conservative.

For the two systems we have considered, we provide results from packet based simulations. The main difference between the implementation used in this section and TCP is that time-outs and the slow-start phase are not implemented in our model. The reason is that, with appropriate design of the marking function [10], it is possible to nearly eliminate loss and thus, minimize the impact of time outs. In Table II and Table III, we present upper and lower bounds from packet-based simulations for various values of N, the number of flows, for SYS-1 and SYS-2 respectively. The upper and lower bounds are computed based on a 95% confidence interval. We have computed the sample mean μ and standard deviation σ of the total rate at the router. We have then denoted the upper bound as $(\mu + 3 \times \sigma)$ and the lower bound as $(\mu - 3 \times \sigma)$. We observe from the packet based simulations that as the number of flows get large, the packet simulations indicate that very little over-provisioning is required, which was our conclusion from the bounds as well as from the fluid simulations.

In Figure 5, we study a system of 25 users, with varying delays as before, but using the $\log(.)$ utility function (SYS-3). As

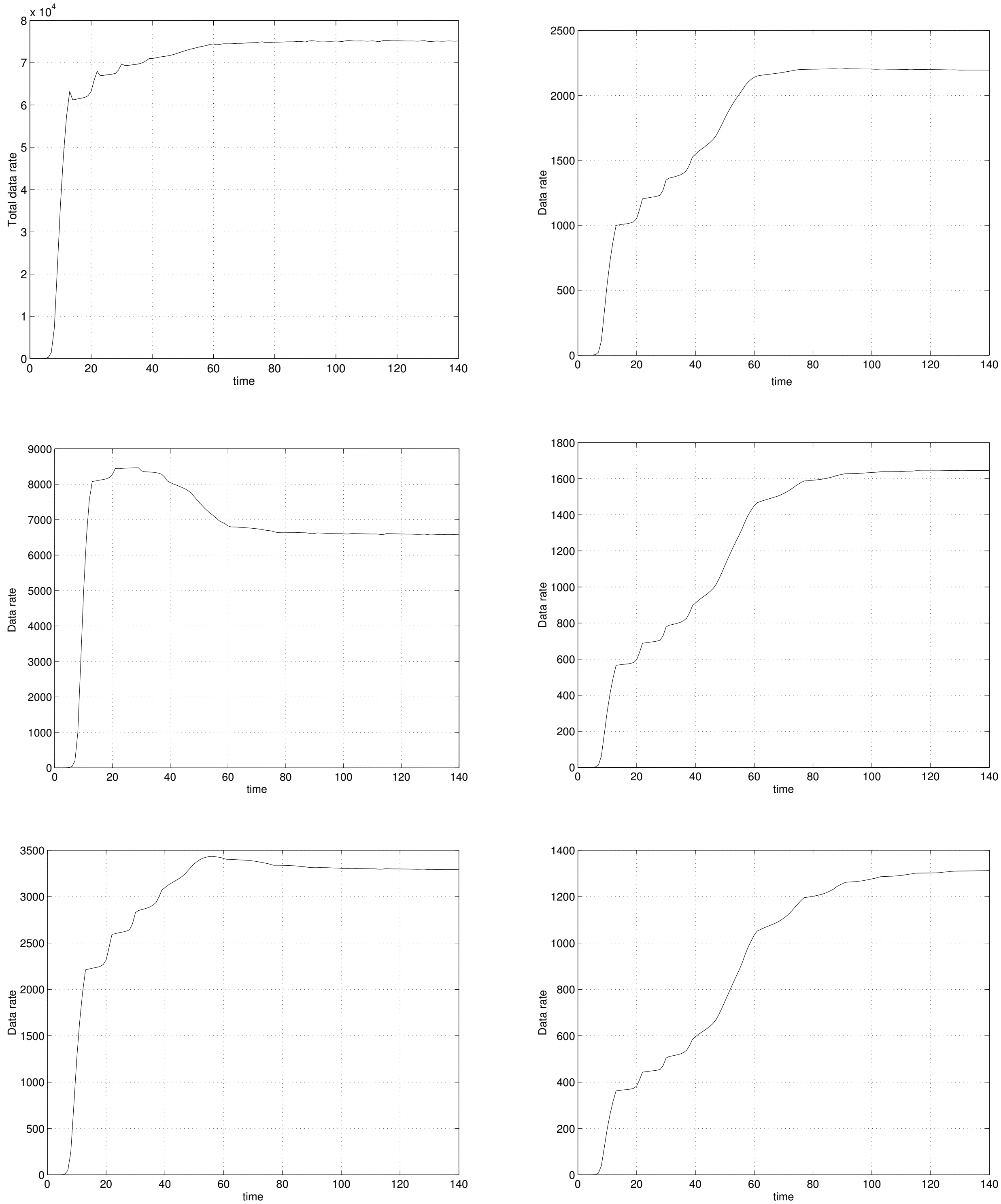

Fig. 4. System SYS-2: Fluid simulations for 25 flows with the utility function being $\frac{-1}{x}$ and with different round trip delays. The first plot is for the total rate at the shared link versus time. The others are the transmission rates of individual connections as a function of time. The marking function is $\left(\frac{\lambda}{Nc_1}\right)^B$.

in Table I, the bound exceeds the actual peak rate at the router by less than 3%. If there were no delays, the total rate would have converged to about 75000. Hence, the results again indicate that very little over-provisioning is required.

VII. Conclusion

In this paper, we studied some decentralized congestion control algorithms corresponding to those proposed for a low-loss Internet. We studied a single node with many flows passing

N	LB	UB	ER	% O-P required
25	1.76	2.24	2.0	12%
100	7.55	8.39	8.0	4.87%
500	38.79	40.74	40.0	1.85%
1000	78.04	80.96	80.0	1.2%
2500	195.76	201.72	200.0	0.86%

TABLE II

UPPER AND LOWER BOUNDS BASED ON PACKET SIMULATIONS FOR SYS-1. ABBREVIATIONS KEY: UPPER BOUND (UB), LOWER BOUND (LB), EQUILIBRIUM RATE (ER), OVER-PROVISIONING (O-P). THE UPPER BOUND, LOWER BOUND AND THE EQUILIBRIUM RATE ARE SCALED BY 37500. THUS, WITH 25 FLOWS, THE LOWER BOUND IS 1.76×37500.

N	LB	UB	ER	% O-P required
25	1.79	2.21	2.0	10.5%
100	7.54	8.38	8.0	4.75%
500	38.56	40.6	40.0	1.5%
1000	77.31	80.85	80.0	1.06%
2500	194.8	200.58	200.0	0.29%

TABLE III

UPPER AND LOWER BOUNDS BASED ON PACKET SIMULATIONS FOR SYS-2. ABBREVIATIONS KEY: UPPER BOUND (UB), LOWER BOUND (LB), EQUILIBRIUM RATE (ER), OVER-PROVISIONING (O-P). THE UPPER BOUND, LOWER BOUND AND THE EQUILIBRIUM RATE ARE SCALED BY 37500. THUS, WITH 25 FLOWS, THE LOWER BOUND IS 1.79×37500.

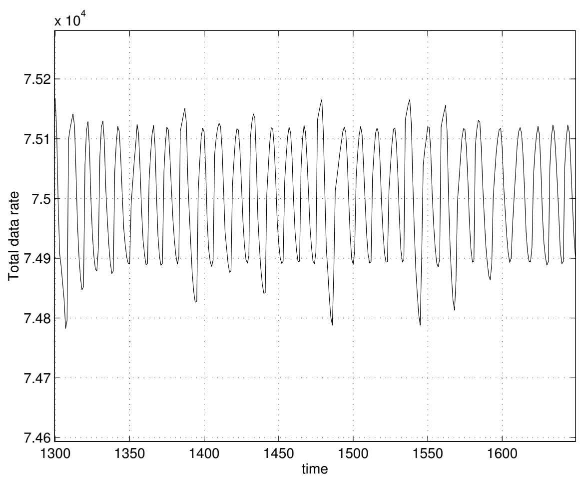

Fig. 5. System SYS-3: Fluid simulations for 25 connections, with different round trip delays. The figure plots the total rate at the shared link versus time. The utility function is $\log(x)$ with the marking function being $\left(\frac{\lambda - N\tilde{c}}{\lambda}\right)^{+}$.

through it, with each flow (possibly) having a different round trip delay. Using a fluid model for the flows, we showed that even with delays, the total data rate at the router is bounded. Further, for typical user data rates and delays seen in the Internet today, the bound is very close to the data rate at the router *without* delays. Our conclusions are two fold: first, the results indicate that in the regime of interest, very little over-provisioning is required at the router to have a low-loss, low-delay network. Second, for a special case of the $\log(.)$ utility function, with all connections having the same round trip delay, we have a lower bound on the total rate. This, in conjunction with the upper bound provides valuable intuition regarding link-utilization and the efficiency of the system with delays. Further the lower bound for this special case, along with the upper bound, leads to a design rule for the marking function which corroborates an earlier result in [5] for the linearized system.

REFERENCES

[1] D. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, MA, 1999.

[2] K. W. Fendick, M. A. Rodrigues and A. Weiss, "Analysis of a Rate-Based Control Strategy with Delayed Feedback," *Proc. ACM Sigcomm*, 1993.

[3] C. Hollot, V. Misra, D. Towsley and W. Gong, "A Control Theoretic Analysis of RED," *Proc. IEEE Infocomm 2001*, Anchorage, AK.

[4] Van Jacobson, "Congestion Avoidance and Control," *Proc. ACM Sigcomm*, Stanford, CA, August, 1988.

[5] R. Johari and D. Tan, "End-to-End Congestion Control for the Internet: Delays and Stability," *IEEE/ACM Transactions on Networking*, vol. 9, no. 6, pp. 818-832, December, 2001.

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

[7] F. P. Kelly, "Mathematical Modeling of the Internet," *Proc. of the Fourth International Congress on Industrial and Applied Mathematics*, July 1999, Edinburgh.

[8] F. P. Kelly, "Models for a Self-Managed Internet," *Philosophical Trans. of the Royal Society of London*, 2000.

[9] F. P. Kelly, A. Maulloo, and D. Tan, "Rate Control in Communication Networks: Shadow Prices, Proportional Fairness and Stability," *Journal of the Operational Research Society*, vol. 49, pp. 237–252, 1998.

[10] S. Kunniyur and R. Srikant, "End-to-end Congestion Control: Utility Functions, Random Losses and ECN marks," *Proc. IEEE Infocom 2000*, Tel Aviv, Israel.

[11] S. Kunniyur and R. Srikant, "A Time-Scale Decomposition Approach to Adaptive ECN Marking," *Proc. IEEE Infocom 2001*, Anchorage, AK.

[12] S. Kunniyur and R. Srikant, "Analysis and Design of an Adaptive Virtual Queue Algorithm for Active Queue Management," *Proc. ACM Sigcomm*, San Diego, CA, 2001.

[13] S. H. Low and D. E. Lapsley, "Optimization Flow Control, I: Basic Algorithm and Convergence," *IEEE/ACM Transactions on Networking*, 1999.

[14] L. Massoulie, "Stability of Distributed Congestion Control with Heterogeneous Feedback Delays," *IEEE Trans. Auto. Control*, vol. 47, no. 6, pp. 895-902, June, 2002.

[15] L. Massoulie and J. Roberts, "Bandwidth Sharing: Objectives and Algorithms," *Proc. IEEE Infocom 1999*, New York, NY.

[16] J. Padhye, V. Firoiu, D. Towsley and J. Kurose, "Modeling TCP Throughput: A Simple Model and its Empirical Validation," *Proceedings of ACM SIGCOMM*, 1998.

[17] F. Paganini, J. C. Doyle and S. H. Low, "Scalable Laws for Stable Network Congestion Control," *IEEE Conf. Decision and Control*, Orlando, FL, December 2001.

[18] K.K. Ramakrishnan and S. Floyd, "A Proposal to Add Explicit Congestion Notification (ECN) to IP," *RFC 2481*, January 1999.

[19] S. Shakkottai and R. Srikant, "Mean FDE Models for Internet Congestion Control Under a Many-Flows Regime," *Technical Report*, University of Illinois, 2001.

[20] S. Shakkottai and R. Srikant, "How Good are Deterministic Fluid Models for Internet Congestion Control?" *Proc. IEEE Infocom*, New York, NY, July, 2002.

[21] G. Vinnicombe. On the stability of end-to-end congestion control for the Internet, 2001. University of Cambridge Technical Report.