ANALYSIS OF A FLUID APPROXIMATION TO FLOW CONTROL DYNAMICS*

Jean-Chrysostome Bolot
INRIA
B. P. 93
06902 Sophia-Antipolis Cedex
France

A. Udaya Shankar†
Department of Computer Science
University of Maryland
College Park, MD 20742
USA

Abstract

We consider a flow control mechanism that dynamically regulates the rate of data flow into a network based on feedback information about the network state. Such mechanisms have been introduced recently in a variety of networks including the Internet, and have been advocated for future high-speed networks. We first model the flow control mechanism by a discrete-space stochastic process and define appropriate performance measures for transient and steady-state regimes. However, the model does not appear to be analytically tractable and we study it through simulation. We then simplify it to a continuous-space deterministic (or fluid) model for which we can easily derive closed-form solutions. We find the analytical results for the fluid model to agree well with the simulation results obtained using the discrete-space model. Both models explicitly consider delay of the feedback information, thus making them relevant for high-speed networks.

1 Introduction

In a computer network, packets generated by a source node are delivered to their destination by routing them via a sequence of intermediate nodes. The traffic flowing through an intermediate node depends upon the number of source-destination pairs that are routed through that node and the rates at which these sources introduce packets into the network. If the source rates are increased without constraint, queues of packets waiting to be routed build up at bottleneck nodes, leading to high delay. Eventually, the buffering capacity of these nodes is exceeded and packets are dropped, resulting in low throughput as well. Flow control mechanisms attempt to maximize throughput and minimize delays and losses by imposing constraints on the flow of data into the network.

Consider a source-destination pair whose packets are routed via intermediate nodes. Let μ_i be the service rate offered by intermediate node i to packets of this source-destination pair. For two nodes i and j, μ_i and μ_j can differ for several reasons: they may be connected to links with different bandwidths; the total traffic through the nodes may differ because they support different source-destination pairs; their hardware may differ; etc. If the network is in steady-state, the ideal flow control policy is to limit the source rate to the bottleneck rate $\mu = \min_i(\mu_i)$ [12], or equivalently the interpacket gap to $1/\mu$; a higher rate would result in packet delay or loss, and a lower rate in underutilization.

In reality, however, the bottleneck rate changes with time because connections (i.e., source-destination pairs) are being set up and terminated, and because their sources do not maintain constant data rates. Recently, flow control mechanisms have been proposed where the source rate is dynamically regulated based on feedback from the intermediate nodes [9, 15, 37, 18, 25, 31, 13, 38, 19]. These mechanisms attempt to adapt the source rate to the bottleneck rate in minimum time and with minimum packet delay or loss.

A variety of feedback schemes are possible. In the binary feedback scheme, which is used in DECNET [28] and is being advocated for ISO TP4 [23] and high-speed networks [38], each data packet received by the destination has a bit indicating whether or not the packet encountered an intermediate node with average queue size greater than 1; this information is sent back to the source in acknowledgement packets. In the Internet TCP, the source measures the times between sending a packet and receiving its acknowledgement and uses them to maintain a running roundtrip time estimate [15]. Other schemes rely on feedback information that includes average link utilization [10, 31], interarrival times between successive acknowledgement packets [18], etc.

There is also a variety of ways to regulate the source rate based on the feedback information. In *rate-based* mechanisms, the source rate is regulated by adjusting interpacket gaps. Examples are found in the transport protocols NETBLT [9] and VMTP [7], although it is not clear in

*This work was supported in part by National Science Foundation Grant No. NCR 89-04590 and by RADC and DARPA under contract number F30602-90-C-0010 to UMIACS at the University of Maryland. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency, RADC, or the U.S. Government. The first author was also supported by a University of Maryland Graduate School Dissertation Fellowship and a ACM Samuel N. Alexander Fellowship Award.

†Also with the Institute for Advanced Computer Studies.

these protocols exactly how interpacket gaps are adjusted or what the feedback would be in a wide-area network. Keshav [19] describes a scheme in which the source adjusts its rate to keep the bottleneck queue size around some fixed value. Other rate-based schemes have been proposed in [8, 39, 10, 18, 31, 37, 13, 38]. However, most protocols currently in operation [33, 28, 15] use a *window-based* scheme, rather than the rate-based scheme. Here, a limit referred to as the *window size* is placed on the number of packets that can be outstanding at the source, but no constraint is placed on the rate at which these packets can be sent. For example, in the binary feedback scheme, the source decreases the window size by a multiplicative factor if the feedback congestion bit is set to 1; otherwise, it increases the window size linearly. Other window-based schemes have been proposed in [25, 14]. The window size is intended to be a bound on the number of data and acknowledgement packets of this connection in transit in the network. Note that because packets can be lost, a good roundtrip time estimate is an essential part of window-based schemes.[1] Obtaining a good roundtrip time estimate in a network when losses are frequent is not trivial [16, 15, 32], and maybe as difficult as the original flow control problem.

Performance evaluation of flow control schemes

We are interested in evaluating the performance of dynamic flow control mechanisms. Performance measures of interest to the user include the average throughput and response time. Performance measures of interest to the network designer include the maximum bottleneck queue size and the time to adapt to variations in the bottleneck rate.

For performance evaluation, it is natural to model a flow control mechanism by a queueing system. Much is known about the steady-state behavior of queueing systems, when the service and job arrival rates are stationary (i.e., have distributions that do not change with time) and certain product-form conditions hold [1, 20]. These results have been used extensively to study fixed-window flow control mechanisms in loss-less networks [30, 33].

However, to model flow control mechanisms that dynamically regulate the source rate based on feedback information, we need queueing models where the service times and job arrivals are time-varying and state-dependent. Unfortunately, the analysis of such systems is extremely difficult and it has received limited attention. Only a few special cases have been solved [35]. More complex systems are often analyzed by numerically solving the Chapman-Kolmogorov equations describing the system state [36, 3], but this rapidly becomes unmanageable for realistic models.

These difficulties led to the development of approximate techniques. One such technique is the fluid approximation, where the basic idea is to replace discrete-space stochastic processes by corresponding continuous-space deterministic processes, which are related by means of ordinary differential equations [27, 20, 5]. An extension of this is the diffusion approximation [20, 26]. Another technique is the state approach technique [34]. Finally, it should be pointed out that due to the lack of analytical results, a number of flow control mechanisms have been evaluated with discrete-event simulation [28, 39, 11] or experimental approaches [15, 22, 32].

[1] Otherwise a packet can be retransmitted while its previous copy is still in transit.

Summary of the rest of the paper

In this paper, we consider a rate-based flow control mechanism with feedback similar to that of the binary feedback scheme. In Section 2, we present a queueing model that explicitly represents the delay in the feedback. In such a model, a high-speed network would be represented by a large ratio of feedback delay to packet transmission time. The model does not appear to be analytically tractable, and we study it via simulation.

In Section 3, we formulate a continuous-space deterministic model, or fluid model, of the connection. We obtain analytical closed-form solutions for the transient and steady-state behavior of this model, and for performance measures such as the period of the steady-state cycle, the average throughput, the maximum bottleneck queue size, etc. We compare the analytical results for the fluid model with the simulation results obtained using the discrete-space model, and find them to correspond closely. A preliminary version of the fluid model was presented in [2].

Section 4 concludes the paper.

2 A Discrete-space Stochastic Model

In this section, we present a discrete-space stochastic model of a connection with rate-based flow control. Each data packet sent by the source has a congestion bit initially equal to 0. When the packet arrives at the destination, the congestion bit is 1 if and only if the packet encountered an intermediate node with queue size greater than 1. The destination sends the information back to the source in acknowledgement packets. The source rate is increased linearly in time if the congestion bit is equal to 0, and decreased exponentially in time if the congestion bit is 1.

We assume that there is exactly one bottleneck node in the sequence of nodes from the source to the destination and back to the source. We explicitly model the bottleneck node by a queue of service rate μ bits/s. Let $Q(t)$ denote the queue size at time t. We do not explicitly model the non bottleneck nodes. It has been shown that this does not introduce significant errors [29, 4]. We use a constant delay D_F to represent the queueing and propagation delays from the source to the bottleneck, and another constant delay D_R to represent the queueing and propagation delays from the bottleneck back to the source (see Figure 1).

Let S_i denote the size, expressed in bits, of data packet i. We assume that data packet sizes are independent and identically distributed. The source controls its sending rate by adjusting the gaps between the sending times of successive data packets. We assume that there are always data packets waiting to be sent. Let t_i denote the time at which data packet i is sent. Let $g_i = t_{i+1} - t_i$ denote the gap between the sending times of packets i and $i+1$. Let $X(t_i)$

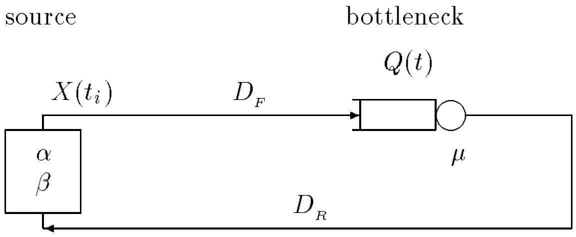

Figure 1: Discrete space stochastic model

denote the source rate at time t_i, expressed in bits/s. We define $X(t_i) = S_i/g_i$.[2] In our queueing model, the source rate is changed only at the t_i's, i.e., when packets are sent. Our model computes t_{i+1} to reflect the desired $X(t_i)$.

Consider the instant just before sending packet n. If the last congestion bit received by the source equals 0, we want the value of t_{n+1} (or equivalently the interpacket gap g_n) to be such that the sending rate is increased linearly over the time interval t_{n-1} to t_n. That is, we want $X(t_n) = X(t_{n-1}) + \alpha\ (t_n - t_{n-1})$, where α is a proportionality constant expressed in bit/s^2. This is equivalent to

$$\frac{S_n}{g_n} = \frac{S_{n-1}}{g_{n-1}} + \alpha\ (t_n - t_{n-1})$$

This implies that $t_{n+1} = t_n + g_n$, where

$$g_n = \frac{S_n g_{n-1}}{S_{n-1} + \alpha\ g_{n-1}^2} \quad (= \frac{S_n (t_n - t_{n-1})}{S_{n-1} + \alpha (t_n - t_{n-1})^2})$$

If the last congestion bit received by the source just before sending packet n equals 1, we want the value of t_{n+1} (or equivalently the interpacket gap g_n) to be such that the sending rate is decreased exponentially over the time interval t_{n-1} to t_n. That is, we want $X(t_n) = X(t_{n-1}) \exp(-(t_n - t_{n-1})/\beta)$, where β is a time constant expressed in seconds. This is equivalent to

$$\frac{S_n}{g_n} = \frac{S_{n-1}}{g_{n-1}} \exp(-g_{n-1}/\beta)$$

This implies that $t_{n+1} = t_n + g_n$, where

$$g_n = \frac{S_n}{S_{n-1}} g_{n-1} \exp(g_{n-1}/\beta)$$

Observe that if several congestion bits are received between t_{n-1} and t_n, then only the last one is used. A more accurate model would be to update $X(t_n)$ whenever a congestion bit is received. In fact, both models yield the same value of $X(t_n)$ as a function of $X(t_{n-1})$, if all the congestion bits received between t_{n-1} and t_n are identical (all zeroes or all ones). A little thought shows that it is very rare that congestion bits of different values are received between t_{n-1} and t_n.

In the above queueing system, the source rate at time t_i is a random variable that depends on the values of the bottleneck queue size in the interval $[t_{i-1} - D_R ... t_i - D_R]$. Thus, $X(t_i)$ is a function of the sample path properties of the process $\{Q(t), t \geq 0\}$ in the interval $[t_{i-1} - D_R ... t_i - D_R]$. This makes the analysis of $X(t_i)$ intractable. In [26], Mukherjee shows that even a simple diffusion approximation to the above queueing system is analytically intractable. Therefore, we study the behavior of our model with simulations. It turns out to have an initial transient regime followed by a steady-state regime in which it has an "almost cyclic" behavior. We define the following performance measures for a simulation run, where T is the length of the simulation run and N is the number of packets sent in the run:

Q_{max}= $\max\{Q(t) : 0 \leq t \leq T\}$. The maximum value of $Q(t)$. This indicates the buffer size required at the bottleneck to avoid packet loss.

$\overline{Q} = \frac{1}{T} \int_0^T Q(t)\ dt$. The time average of $Q(t)$.

$\overline{X} = \frac{1}{T} \sum_{i=1}^{i=N} S_i$. The time average of the source rate.

T_{TR}. The transient duration of $X(t_n)$. This is obtained by the truncation method. [3]

Note: In the deterministic case when all data packets have the same size, $Q(t)$ and $X(t_n)$ are deterministic functions of time. The above model is then analytically tractable, but the derivation of performance measures is still very cumbersome. It requires solving sets of inequalities to obtain indices, then using these indices in other equations (e.g., find index k which is the lowest integer satisfying $g_k < 1/\mu$; then $Q(t)$ reaches the value Q_{max} for the first time when packet n arrives at the bottleneck, where n is the lowest index greater than k such that $g_n > 1/\mu$).

2.1 Simulations

We now present simulations of the above queueing system. They were obtained using a modified version of the REAL simulator [17]. In all cases, we assume that the data packet sizes $\{S_i\}$ are independent and identically distributed with a uniform distribution over an interval $[s_{min} \ldots s_{max}]$. We characterize the distribution by its mean *DPSIZE* = 8000 bits and by $c = (s_{max} - s_{min})/(s_{max} + s_{min})$. The value of c can range from 1 to 0, with 0 indicating that all data packets have the same size of 8000 bits. c is equal to the coefficient of variation of the uniform distribution, divided by $\sqrt{3}$.

We consider two different settings for the parameters μ, D_F and D_R in our model. The first setting characterizes an 'ordinary' network, i.e., a network similar to networks currently in operation, and is as follows: $\mu = 120$ kbit/s, $D_F = D_R = 330$ ms. These values are consistent with measures obtained experimentally from Internet connections spanning the continental United States [32]. The second setting characterizes a high-speed network spanning the continental United States [25], and is as follows: $\mu = 40$ Mbit/s,

[2] There are other ways of defining $X(t_i)$, for example S_{i-1}/g_{i-1}, or $(S_i + S_{i-1})(1/g_{i-1} + 1/g_i)$, etc.

[3] In this method, the duration of the transient regime in a series of values $\{X(t_i): i \geq 1\}$ is equal to the smallest t_j such that $X(t_j)$ is neither the maximum nor the minimum of the series $\{X(t_i): i \geq j\}$ [21]

$D_F = D_R = 20$ ms. Throughout the rest of the paper, we let $D = D_F + D_R$.

The values of D_F and D_R are different for the ordinary and the high-speed networks. This is because D_F, for example, includes both propagation and transmission delays between the source and the bottleneck. Propagation delays are identical for ordinary and high-speed networks covering similar geographical areas. Transmission delays, however, depend on the bandwidth of the links in a network, and are much lower in a high-speed network.

For convenience of presentation we eliminate the parameter α by assuming that the coefficients α and β satisfy the relation $\alpha\beta = \mu$ (As explained in Section 3, this corresponds to having the rate of increase of $X(t)$ equal to the rate of decrease of $X(t)$ around steady-state). We express β as $\beta = \hat{\beta}\ D$. $\hat{\beta}$ is a dimensionless number that is equal to the number of roundtrip delays necessary to decrease the source rate from some value x to the value x/e. Since $\alpha = \mu/\beta = \mu/\hat{\beta}\ D$, $\hat{\beta}$ is also equal to the number of roundtrip delays necessary to increase the source rate from some value x to the value $x + \mu$. For example, in the ordinary network, $\hat{\beta} = 2$ corresponds to $\beta = 1.32$ s and $\alpha = 90$ kbit/s^2. In the high-speed network, $\hat{\beta} = 2$ corresponds to $\beta = 80$ ms and $\alpha = 500$ Mbit/s^2. The smaller the value of $\hat{\beta}$, the faster the source rate increases and decreases.

Performance versus packet size variance

In this subsection, we examine the system for different values of c. Figure 2 shows the evolutions of $Q(t)$ and $X(t)$ for the cases $c = 0, 0.5$ and 0.9 in the range $0 \leq t \leq 30$ s for the ordinary network when $\hat{\beta} = 2$.

We observe that the bottleneck queue size and the source sending rate display an "almost periodic" behavior. To understand this behavior, consider the case $c = 0$, i.e., all data packets have fixed size. Consider the situation where the source starts sending data packets at time t_1 at a rate less than μ, and assume that the bottleneck queue is initially empty. Clearly, the first acknowledgement packet arrives at the source with a congestion bit equal to 0. Therefore, the source rate starts increasing linearly in time with rate α and eventually exceeds μ at some time t_2. After a delay D_F, i.e., at time $t_2 + D_F$, a queue starts building up at the bottleneck. The source detects the presence of the queue after a delay D_R, at which point it starts decreasing its rate. However, the queue keeps increasing as long as $X(t)$ remains higher than μ. The source rate keeps decreasing until the source detects that the bottleneck queue size is lower than 1, at which point it starts increasing again. When it reaches μ again, the state of the system is identical to that at time t_2. Therefore, the system has reached a steady-state limit cycle. When $c > 0$, the cycle is not perfect because the random variations in packet size introduce perturbations in the dynamics of $Q(t)$ and $X(t)$. However, the behavior remains almost cyclic.

Table 1 presents values of the performance measures for the ordinary network for different values of c when $\hat{\beta} = 2$.

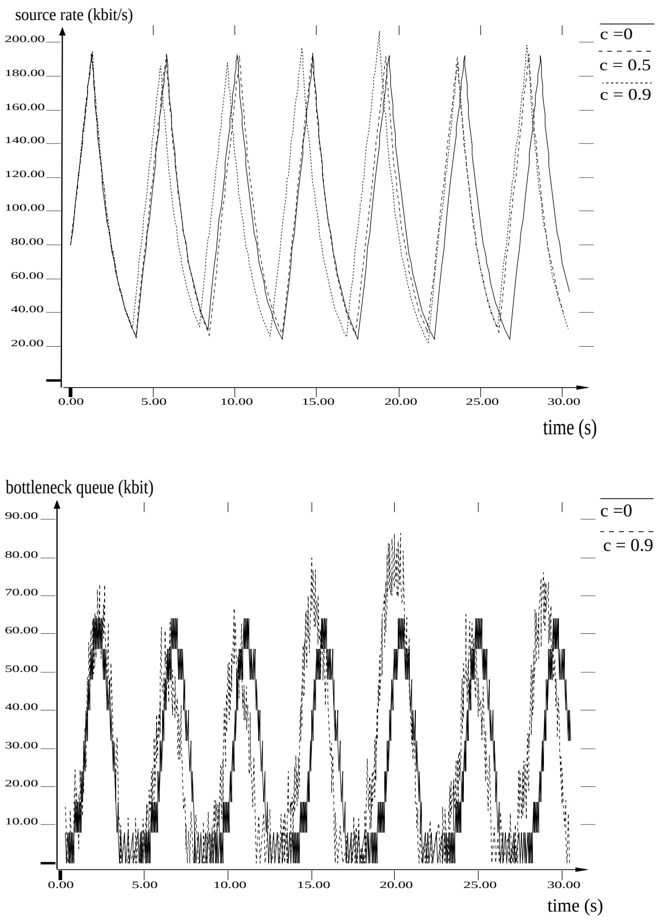

Figure 2: Ordinary network: Evolutions of $Q(t)$ and $X(t)$ for $\hat{\beta} = 2$ and $c = 0$, 0.5 and 0.9

c	Q_{max} (kbit)	$\overline{Q}$ (kbit)	$\overline{X}$ (kbit/s)	T_{TR} (s)
0	64	23.0	94.4	0.58
0.1	72	23.4	94.6	0.6
0.3	80	23.8	94.8	0.6
0.5	88	24.0	95.1	0.58
0.9	120	23.6	94.6	0.6

Table 1: Performance measures versus c for the ordinary network for $\hat{\beta} = 2$

Observe that the value of Q_{max} increases as c increases. This is to be expected because increasing c increases the randomness of the data packet arrivals at the bottleneck. More surprising is the observation that the values of $\overline{Q}$ and $\overline{X}$ appear to be almost independent of c, even as c approaches 1. We do not have an explanation for this behavior. We note, however, that this is not the behavior of a $GI/G/1$ system. There, the expected steady-state queue size increases with increasing randomness of customers arrivals. In our model, the interarrival times of data packets at the bottleneck are not independent. In particular, they depend on the state of the bottleneck at some point in the past, as reported by the feedback information.

Table 2 presents values of the performance measures for the high-speed network for different values of c when $\hat{\beta} = 2$.

c	Q_{max} (kbit)	$\overline{Q}$ (kbit)	$\overline{X}$ (Mbit/s)	T_{TR} (s)
0	728	240.0	31.9	0.024
0.1	744	238.4	32.1	0.024
0.3	792	237.2	32.1	0.024
0.5	808	236.5	32.0	0.023
0.9	840	235.2	31.7	0.024

Table 2: Performance measures versus c for the high-speed network for $\hat{\beta} = 2$

We observe a behavior similar to that of the ordinary network. Namely, Q_{max} increases as c increases, whereas $\overline{Q}$ and $\overline{X}$ appear to be independent of c. We note that the larger bandwidth-delay product of the high-speed network requires larger buffers at the bottleneck to avoid packet losses [24].

Fixed length packets

In this subsection, we consider the case where $c = 0$, i.e., all data packets have a fixed size of 8000 bits. We examine the variations of the performance measures versus $\hat{\beta}$ and $D = D_F + D_R$. Table 3 shows the values of the performance measures for different values of $\hat{\beta}$.

ordinary network

$\hat{\beta}$	Q_{max} (kbit)	$\overline{Q}$ (kbit)	$\overline{X}$ (kbit/s)	T_{TR} (s)
10	24	10.9	111.7	4.1
5	32	13.9	105.8	1.4
2	64	23.0	94.4	0.58
1	98	43.3	86.3	0.3

high-speed network

$\hat{\beta}$	Q_{max} (kbit)	$\overline{Q}$ (kbit)	$\overline{X}$ (Mbit/s)	T_{TR} (s)
10	160	51.1	37.4	0.28
5	320	98.7	35.5	0.11
2	728	240.0	31.9	0.024
1	1336	470.9	30.0	0.01

Table 3: Performance measures versus $\hat{\beta}$ for $c = 0$

We observe that the duration of the transient regime decreases as $\hat{\beta}$ decreases. Therefore, steady-state is reached quickly if $\hat{\beta}$ is small. However, a small value of $\hat{\beta}$ implies a large value of Q_{max}, and thus larger buffer requirements if we are to avoid buffer overflow at the bottleneck. In addition, the average steady-state sending rate $\overline{X}$ moves farther away from the ideal value μ as $\hat{\beta}$ decreases. Recall that $\mu = 120$ kbit/s for the ordinary network, and $\mu = 40$ Mbit/s for the high-speed network. Therefore, a tradeoff has to be found between a rapid relaxation to steady-state and the efficiency of the steady-state regime. This is consistent with the results reported in [28, 25].

Figure 3 shows the evolutions of the source rate and the bottleneck queue size, for the ordinary network when $\hat{\beta} = 2$ and $\hat{\beta} = 8$. The tradeoff between a short transient regime and an efficient steady-state regime is clearly visible.

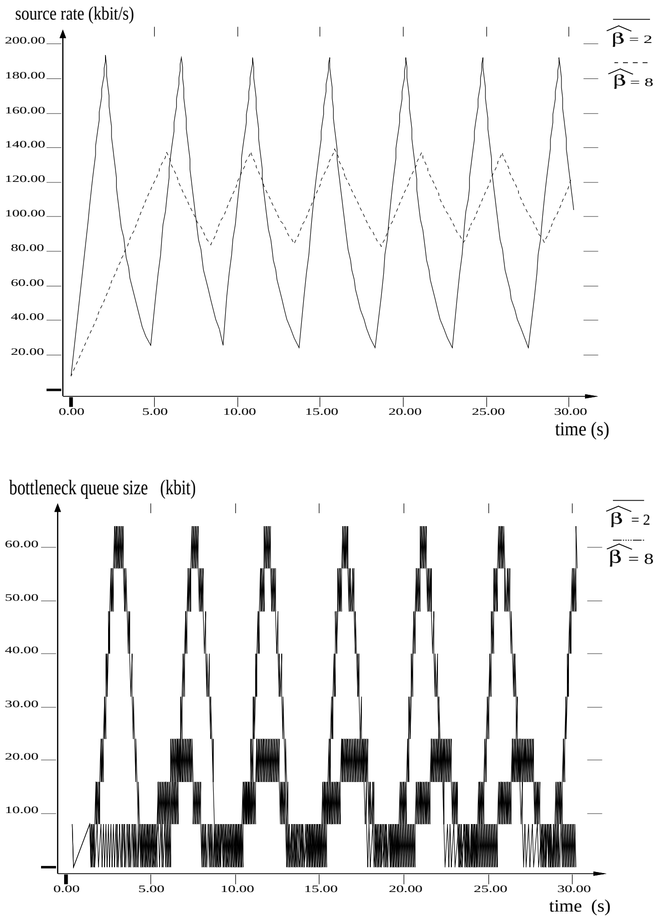

Figure 3: Ordinary network: Evolutions of $Q(t)$ and $X(t)$ for $\hat{\beta} = 2$ and $\hat{\beta} = 8$, for $c = 0$

We now examine the variations of the performance measures for different values of $D = D_F + D_R$. Table 4 shows the values of the performance measures for the ordinary network and the high-speed network for $\hat{\beta} = 2$.

ordinary network

D (ms)	Q_{max} (kbit)	$\overline{Q}$ (kbit)	$\overline{X}$ (kbit/s)	T_{TR} (s)
50	16	8.8	111.2	0.81
400	40	14.5	96.8	0.4
1000	90	43.2	87.2	0.22

high-speed network

D (ms)	Q_{max} (kbit)	$\overline{Q}$ (kbit)	$\overline{X}$ (Mbit/s)	T_{TR} (s)
2	16	8.8	39.1	0.2
20	200	64.7	34.6	0.05
60	1544	526.4	30.6	0.02

Table 4: Performance measures versus D for $\hat{\beta} = 2$, for $c = 0$

As D decreases, the performance of the system improves, i.e., Q_{max} and $\overline{Q}$ decrease while $\overline{X}$ increases. This is expected since a decrease of the reverse propagation delay implies that the feedback information, when it is received at the source, more accurately reflects the state of the bottleneck queue. In addition, a decrease of the forward propagation delay implies that the rate adjustments done by the source are quickly reflected at the bottleneck.

3 A Continuous-space Deterministic (Fluid) Model

In this section, we consider a continuous-space deterministic (or fluid) model of our rate-based flow control mechanism. We model the bottleneck node by a queue of service rate μ. Let $q(t)$ denote the bottleneck queue size at time t, and let $x(t)$ denote the source rate at time t. In the fluid model, $x(t)$ and $q(t)$ are real-valued continuous functions of time. We let a constant delay d_F represent the queueing and propagation delays from the source to the bottleneck, and another constant delay d_R represent the queueing and propagation delays from the bottleneck back to the source (see Figure 4). The relation between these parameters and the parameters of the discrete-space model is described in Section 3.2.

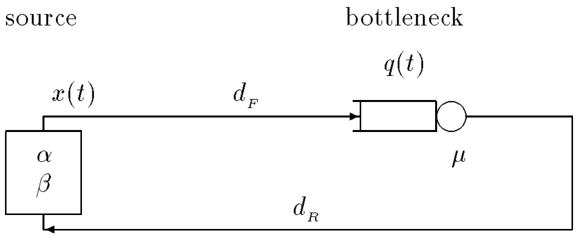

Figure 4: Fluid model

One advantage of having $x(t)$ and $q(t)$ be real-valued continuous functions of time is that the adjustments to $x(t)$ can be simply described by specifying $\dot{x}(t)$ (we use $\dot{x}(t)$ to denote $dx(t)/dt$). Thus, the behavior of the model can be conveniently described in terms of coupled differential equations involving $x(t)$, $\dot{x}(t)$, $q(t)$ and $\dot{q}(t)$.

We now derive the equations describing $\dot{q}(t)$ and $\dot{x}(t)$ in terms of $q(t)$ and $x(t)$. Observe that data sent by the source at time t arrives at the bottleneck at time $t+d_F$. Thus, the data arrival rate at the bottleneck at time t is $x(t-d_F)$. If $q(t)>0$, data departs from the bottleneck at rate μ. If $q(t)=0$, data departs at the same rate as it arrived, i.e., $x(t-d_F)$. This leads us to the following equation for $\dot{q}(t)$:

$$\dot{q}(t) = \begin{cases} 0 & \text{if } q(t)=0 \text{ and} \\ & \quad x(t-d_F)-\mu<0 \\ x(t-d_F)-\mu & \text{otherwise} \end{cases} \qquad (1)$$

The situation of 100% bottleneck utilization and no waiting delay corresponds to $x(t-d_F)=\mu$ and $q(t)=0$.

We now derive the equation for $\dot{x}(t)$. We assume that, at every time t, the source receives a congestion bit indicating whether or not the bottleneck was 100% utilized d_R seconds ago; i.e., whether $q(t-d_R)>0$ or $q(t-d_R)=0$. $x(t)$ is decreased exponentially with time constant β if $q(t-d_R)>0$, i.e., the bottleneck was 100% utilized at time $t-d_R$. $x(t)$ is increased linearly with rate α if $q(t-d_R)=0$. Thus, we get the following:

$$\dot{x}(t) = \begin{cases} \alpha & \text{if } q(t-d_R)=0 \\ -\frac{x(t)}{\beta} & \text{if } q(t-d_R)>0 \end{cases} \qquad (2)$$

3.1 Analytic solution

The evolution of $x(t)$ and $q(t)$, as governed by equations (1) and (2), is depicted pictorially in Figure 5. We now explain it.

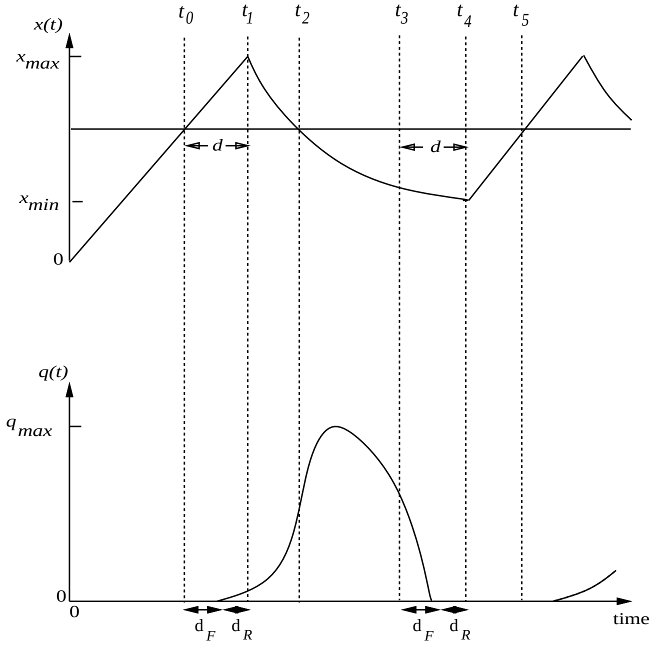

Figure 5: Analytic behavior of $x(t)$ and $q(t)$

Assume that the source starts sending data at time $t=0$, and that the bottleneck queue is initially empty. The source rate increases at rate α and reaches the value μ at time $t_0=\mu/\alpha$. Only now is $x(t)$ large enough to create a queue buildup at the bottleneck. Because of the delay d_F, the queue at the bottleneck starts building up at time t_0+d_F. The source detects the existence of the queue at the bottleneck only at time t_0+d, where the round trip delay $d=d_F+d_R$. Thus, $x(t)$ increases linearly from $t=0$ to $t=t_1=t_0+d$. We let $x_{max}=x(t_1)=\mu+d\alpha$.

At time t_1, $x(t)$ starts decreasing. However, the queue keeps increasing since $x(t)$ remains higher than μ for some time. Let t_2 be the time when $x(t)$ again equals μ. Then $q(t)$ reaches its maximum at time t_2+d_F, after which it starts decreasing. We define t_3 as the time at which $x(t)$ has been lower than μ long enough to drain out the extra packets introduced between t_0 and t_2 (t_3 is evaluated below). $q(t)$ becomes zero at time $t_3 + d_{_F}$. Meanwhile, $x(t)$ keeps decreasing until the source detects, at time $t_4 = t_3{+}d$, that the queue size at the bottleneck is down to 0, at which point it starts increasing again. We let $x_{min} = x(t_4)$.

Let t_5 be the time at which $x(t_5) = \mu$. Observe that $x(t_5) = x(t_0)$ and $q(t_5) = q(t_0)$, i.e., the state of the system at time t_5 is identical to that at time t_0. Therefore, the steady-state behavior of both $x(t)$ and $q(t)$ is cyclic with period $\tau = t_5 - t_0$. The transient behavior is defined by its length t_0.

We now derive closed-form expressions for the values of x_{max}, q_{max}, the period τ, and the time instants t_2 and t_3. Recall that $t_0 = \mu/\alpha$, $t_1 = t_0 + d$, $t_4 = t_3 + d$, and $t_5 = t_0 + \tau$.

We first consider x_{max}. Since $x(t)$ increases linearly at rate α from $t = 0$ to $t = t_1$, it follows that

$$\boxed{x_{max} = \mu + \alpha d}$$

The maximum value q_{max} of the bottleneck queue is given by $q_{max} = \int_{t_0}^{t_2}(x(t) - \mu)dt$. Since $x(t) = (\mu + \alpha d)\exp(-\frac{1}{\beta}(t - t_1))$ for $t_1 \leq t \leq t_2$, it follows that

$$q_{max} = \alpha\frac{d^2}{2} + \int_0^{t_2 - t_1}((\mu + \alpha d)e^{-\frac{1}{\beta}t} - \mu)dt$$

The value of $t_2 - t_1$ is computed by observing that $x(t_2) = \mu$, i.e., $\exp(-\frac{1}{\beta}(t_2 - t_1)) = \mu/(\mu + \alpha d)$. Therefore

$$t_2 - t_1 = -\beta\ln(\frac{\mu}{\mu + \alpha d})$$

and

$$\boxed{q_{max} = \alpha\frac{d^2}{2} + \alpha\beta d + \mu\beta\ln(\frac{\mu}{\mu + \alpha d})}$$

We now compute the period $\tau = t_5 - t_0$. Observe that $\tau = (t_1 - t_0) + (t_3 - t_1) + (t_4 - t_3) + (t_5 - t_4) = d + (t_3 - t_1) + d + (\mu - x_{min})/\alpha = 2d + (t_3 - t_1) + (\mu - x_{min})/\alpha$. Since $x_{min} = (\mu + \alpha d)\exp(-\frac{1}{\beta}(d + t_3 - t_1))$, the problem reduces to that of finding $t_3 - t_1$. However, t_3 is defined such that

$$\int_{t_0}^{t_3}(x(t) - \mu)dt = \alpha\frac{d^2}{2} + \int_0^{t_3 - t_1}((\mu + \alpha d)e^{-\frac{1}{\beta}t} - \mu)dt = 0$$

After some algebraic manipulations, we obtain

$$\boxed{t_3 - t_1 = \beta Root(\frac{\mu}{\mu + \alpha d}, \frac{\alpha d^2}{2\beta(\mu + \alpha d)})} \qquad (3)$$

where $Root(a, b)$ denotes the unique real solution of $1 - e^{-x} = ax - b$, with $0 < a < \infty$ and $0 < b$. We have from above

$$\boxed{\tau = 2d + (t_3 - t_1) + \frac{\mu - x_{min}}{\alpha}}$$

This concludes the computation of τ.

To summarize, the transient behavior of the system is defined by :

$$\boxed{\begin{array}{ll} x(t) = \alpha t & 0 \leq t \leq t_0 \\ q(t) = 0 & \end{array}}$$

and the steady-state behavior is defined by :

$$\boxed{\begin{array}{ll} x(t) = & \left\{\begin{array}{l} (\mu + \alpha d)\exp(-\frac{1}{\beta}(t - t_1 - nT)) \\ \qquad \text{for } t_1 + nT \leq t \leq t_4 + nT \\ \alpha(t - t_4 - nT) + x_{min} \\ \qquad \text{for } t_4 + nT \leq t \leq t_1 + (n+1)T \end{array}\right. \\ q(t) = & \int_{t_0}^{t}(x(t - d_{_F}) - \mu)dt \end{array}}$$

where $t_4 = t_3 + d$ is obtained from equation (3). Note that the behavior from t_0 to t_1 is included in the steady-state behavior with $n = -1$.

Using the above results, we can obtain the following performance measures (although the expressions for $\overline{q}$ and $\overline{x}$ would be messy):

- q_{max}: The maximum value of $q(t)$.
- $\overline{q}$: The time average of $q(t)$: $\overline{q} = \frac{1}{\tau}\int_{t_0}^{t_0+\tau} q(t)\, dt$.
- $\overline{x}$: The time average of $x(t)$: $\overline{x} = \frac{1}{\tau}\int_{t_0}^{t_0+\tau} x(t)\, dt$.
- $t_{_{TR}}$: The transient duration of $x(t)$: $t_{_{TR}} = t_0$.

First order approximation

We now consider a first-order approximation to our original model. Specifically, we consider a system model in which $x(t)$ is adjusted as follows

$$\dot{x}(t) = \left\{\begin{array}{ll} \alpha & \text{if } q(t - d_{_R}) = 0 \\ -\frac{\mu}{\beta} & \text{if } q(t - d_{_R}) > 0 \end{array}\right.$$

This amounts to approximating the exponentially decreasing function $(\mu + \alpha d)\exp(-\frac{1}{\beta}(t - t_1))$ by the linearly decreasing function $(\mu + \alpha d) - \frac{\mu}{\beta}(t - t_1)$. Thus, we expect the approximation to be accurate when $\alpha d \ll \mu$. We obtain:

$$\boxed{\begin{array}{lcl} q_{max} & = & \alpha\frac{d^2}{2}(1 + \frac{\alpha\beta}{\mu}) \\ \overline{q} & = & \alpha\frac{d^2}{12\tau}\frac{\alpha\beta}{\mu}(1 + \frac{\alpha\beta}{\mu})(1 + \frac{\mu}{2\alpha\beta} + \sqrt{1 + \frac{\mu}{\alpha\beta}}) \\ \overline{x} & = & \mu - \frac{\alpha d}{2}(-1 + \frac{\mu}{\alpha\beta} + \sqrt{1 + \frac{\mu}{\alpha\beta}}) \\ t_{_{TR}} & = & \mu/\alpha \\ x_{max} & = & \mu + \alpha d \\ \tau & = & d(1 + \frac{\alpha\beta}{\mu})(1 + \frac{\mu}{\alpha\beta} + \sqrt{1 + \frac{\mu}{\alpha\beta}}) \end{array}}$$

Setting $\alpha\beta = \mu$

For convenience of presentation, we eliminate the parameter α by assuming, throughout the rest of the paper, that coefficients α and β satisfy the relation $\alpha\beta = \mu$. This means that, near the fixed point $x(t) = \mu$, $q(t) = 0$, the rate at which $x(t)$ increases is equal to the rate at which it decreases. We can then simplify the results derived above. For convenience of reference, we now present the values of the performance measures obtained with the first order approximation.

$$
\begin{array}{lcl}
q_{max} & = & \frac{\mu d^2}{\beta} \\
\overline{q} & = & \frac{\mu d^2}{12\beta}(2+\sqrt{2}) \\
\overline{x} & = & \mu - \frac{\mu d}{\beta\sqrt{2}} \\
t_{TR} & = & \beta \\
x_{max} & = & \mu + \frac{\mu d}{\beta} \\
\tau & = & 2d(2+\sqrt{2})
\end{array}
\qquad (4)
$$

The duration of the transient regime is proportional to β. Therefore, steady-state is reached quickly if β is small. However, a small value of β implies large values of q_{max} and $\overline{q}$. In addition, equation (4) indicates that the difference between the average source rate $\overline{x}$ and the ideal value μ increases as β decreases. Therefore, a tradeoff has to be found between a rapid relaxation to steady-state and the efficiency of the steady-state behavior. This is consistent with the simulation results reported in Section 2.

In a real network, d varies with time. Therefore, keeping q_{max} at a given value requires β to vary with time such that β is proportional to d^2. Similarly, keeping $\overline{x}$ at a given value requires β to vary with time as d. Which objective is more important will dictate how β should vary with time. But in any case, these results indicate the need to adapt the value of β to variations in the roundtrip delay d.

3.2 Comparison with the discrete-space model

In this subsection, we compare the analytical results for the fluid model with the simulation results obtained in Section 2.1. To do this, we need to relate the parameters of the fluid model to the parameters of the discrete-space model.

We first relate $Q(t)$ and $X(t)$ to $q(t)$ and $x(t)$. Consider the discrete-space model with $c = 0$, i.e., all data packets have the same size *DPSIZE*. Assume that packets arrive at the bottleneck with a constant bit rate X_0, starting from time 0. The number of packets received at the bottleneck at time t is then equal to $\lceil X_0 t/\mathit{DPSIZE} \rceil$. If $X_0 > \mu$, the number of packets that have left the bottleneck at time t is equal to $\lfloor \mu t/\mathit{DPSIZE} \rfloor$. Therefore, $Q(t) = \mathit{DPSIZE}\,(\lceil X_0 t/\mathit{DPSIZE} \rceil - \lfloor \mu t/\mathit{DPSIZE} \rfloor)$. Consider now the corresponding situation in the fluid model, where the source rate is constant at $x_0 > \mu$. Then equation (1) in Section 3 implies that $q(t) = x_0 t - \mu t$.

If we equate X_0 to x_0, then it is easily shown that $q(t) \leq Q(t) \leq q(t) + 2\ \mathit{DPSIZE}$. The relationship between $Q(t)$ and $q(t)$ in more complicated situations where source rates depend on feedback information, packet sizes are randomly distributed, etc., is not clear. However, we take the above relation as an indication that $x(t)$ and $q(t)$ can be used as approximations to $X(t)$ and $Q(t)$, respectively.

We now relate D_F and D_R to d_F and d_R. Consider the situation where the source receives a packet with a congestion bit that brings new information. We want the delay from the source sending this data packet to the source reacting to the new feedback information to be the same in both the discrete-space and the fluid models. In the discrete-space model, this delay is the roundtrip delay D, plus the queueing delay at the bottleneck R, plus the time A between the source receiving the congestion bit and the next scheduled data packet send time. We can approximate A by $\frac{1}{2}\mathit{DPSIZE}/\mu$. In the fluid model, this delay is the roundtrip delay d plus the queueing delay at the bottleneck r. There is no delay corresponding to A since the source adjusts its rate as soon as it receives new feedback information. Assuming that $q(t)$ is a good approximation to $Q(t)$, we have $R = r + \mathit{DPSIZE}/\mu$. Therefore, we want $d = D + \frac{3}{2}\mathit{DPSIZE}/\mu$. We choose to let $d_F = D_F$ and $d_R = D_R + \frac{3}{2}\mathit{DPSIZE}/\mu$.

Examples

Using the above relations, we can compare the evolutions of the source rate and the bottleneck queue size in the fluid and in the discrete-space models. We first consider the case of the ordinary network. There, the parameters in the discrete-space model are as follows: $\mu = 120$ kbit/s, and $D_F = D_R = 330$ ms. Therefore, in the fluid model, $d_F = 330$ ms and $d_R = 430$ ms. Figure 6 shows the behavior of the discrete-space model for $c = 0$ and $c = 0.9$, as well as the behavior of the fluid model.

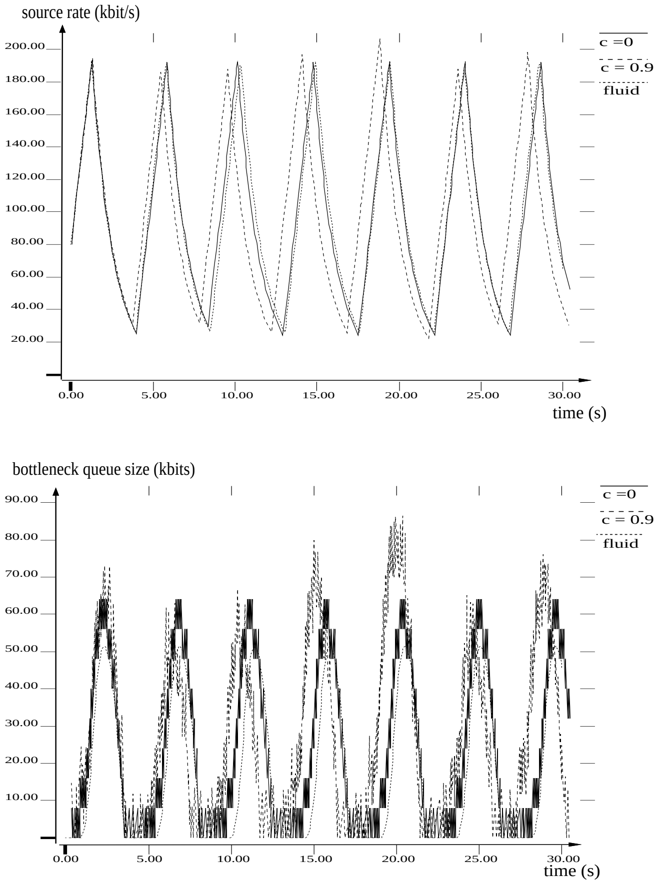

Figure 6: Ordinary network: Evolutions of the bottleneck queue size and the source rate for the fluid model and for the discrete-space model with $c = 0$ and $c = 0.9$

As expected, we observe that the fluid model is a very good approximation for the deterministic discrete-space model, i.e., when $c = 0$. More interesting is the observation that it also provides a reasonable approximation to the stochastic discrete-space model, even for large values of c. This is also the case for the high-speed network. There, $\mu = 40$ Mbit/s and $D_F = D_R = 20$ ms. Therefore, $d_F = 20$ ms and $d_R = 20.3$ ms. Figure 7 shows the behavior of the discrete-space model for $c = 0$ and $c = 0.9$, as well as the behavior of the fluid model.

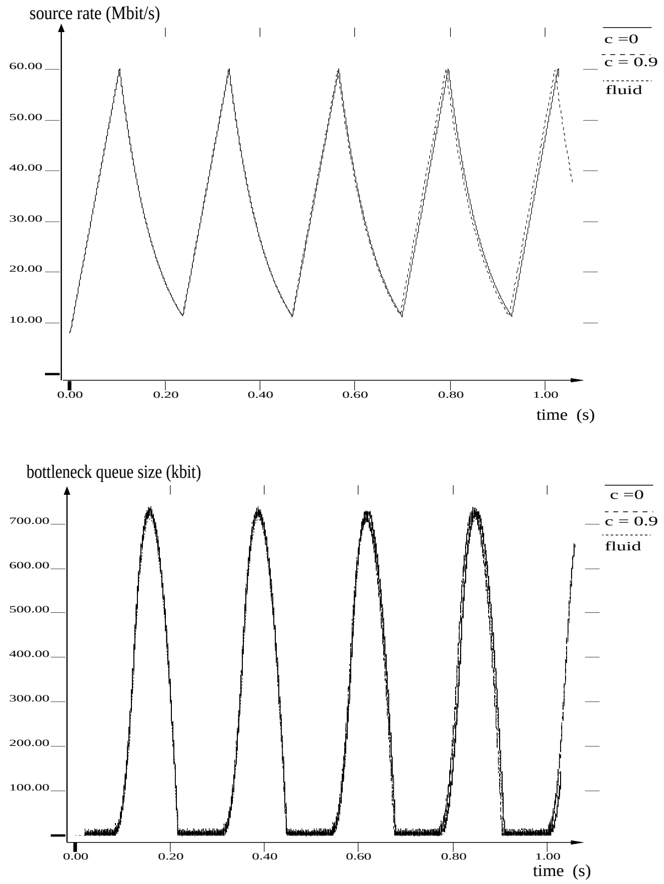

Figure 7: High-speed network: Evolutions of the bottleneck queue size and the source rate for the fluid model and for the discrete-space model with $c = 0$ and $c = 0.9$

In Table 5, for the high-speed network, we compare the performance measures obtained by the fluid first order approximation with the performance measures obtained by simulation of the discrete-space model with $c = 0$.

We verify that the analytical results obtained with the first order approximation are accurate when $\beta \gg d$, i.e., $\beta \gg 0.04$ s. In particular, the values in Table 5 are consistent with the fact that Q_{max} is inversely proportional to β and $\overline{X} - \mu$ decreases as β increases.

4 Conclusion

It appears that fluid modeling holds great promise in analyzing network congestion and flow control problems. It is very encouraging that the behavior (and thus design choices) indicated by the relatively simple fluid model in this paper seems to tie in very well with the results obtained by simulation. For example, our model brings out the tradeoff between good steady-state behavior and rapid adaptability to changing network conditions such as bottleneck rate and round trip delay changes. It also shows that it is important to take the current roundtrip delay into account when adjusting the sending rate at the source. This is consistent with the results reported in the literature [28, 23, 25, 26, 38, 19].

In our model, rate adjustments are made based on feedback information that reflects an instantaneous value of the bottleneck queue size. Some schemes attempt to eliminate the effects of short-lived transient variations in the bottleneck queue size by considering an average of recent values of the queue size. We are extending our model to consider the effects of such averaging techniques in the feedback mechanism.

In this paper, we analyzed the performance of a single connection with rate-based flow control. Fluid approximations have been found convenient for studying the time-dependent behavior of networks with more general topologies [6]. We are extending our model to consider multiple connections sharing one or more bottleneck nodes. Preliminary results indicate that the source sending rates and the bottleneck queue sizes follow a rich panoply of dynamical behaviors, with variations depending on round trip delays and rate control schemes.

	simulation				first order approximation			
β (s)	Q_{max} (kbit)	$\overline{Q}$ (kbit)	$\overline{X}$ (Mbit/s)	τ (s)	q_{max}	$\overline{q}$	$\overline{x}$	τ
0.8	88	27.7	38.6	0.27	81.2	23.1	38.6	0.27
0.4	160	51.1	37.4	0.26	162.4	46.2	37.2	0.27
0.2	320	98.7	35.5	0.25	324.8	92.4	34.3	0.27
0.08	728	240.0	31.9	0.23	812	231	25.8	0.27
0.04	1336	470.9	30.0	0.21	1624	462	11.5	0.27

Table 5: High-speed network: Performance measures versus β

References

[1] F. Baskett, K. M. Chandy, R. Muntz, F. Palacios, "Open, closed and mixed networks of queues with different classes of customers", *JACM*, vol. 22, no. 2, pp. 248-260, April 1975.

[2] J-C. Bolot, A. U. Shankar, "Dynamical behavior of rate-based flow control mechanisms", *Computer Communication Review*, vol. 20, no. 2, pp. 35-49, April 1990.

[3] J-C. Bolot, A. U. Shankar, B. Plateau, "Performance analysis of transport protocols over congestive channels", *Perf. Eval.*, vol. 11, no. 1, pp. 45-65, Apr. 1990.

[4] O. J. Boxma, "Sojourn times in cyclic queues - the influence of the slowest server", *Proc. 2nd Intl. Workshop on Applied Math. and Perf. Models of Comp. and Comm. Systs.*, Rome, Italy, 1987, pp. 13-24.

[5] D. Browning, "A unified model for flow control in packet switched networks", *Proc. 18th Conf. Info. Sciences and Systems*, Princeton, NJ, 1988, pp. 1044-1050.

[6] H. Chen, A. Mandelbaum, "Discrete flow networks: bottleneck analysis and fluid approximations", *Math. of Ops. Res.*, vol. 16, no. 2, pp. 408-446, May 1991.

[7] D. R. Cheriton, "VMTP: a transport protocol for the next generation of communication systems", *Proc. ACM SIGCOMM '86*, Stowe, Vermont, pp. 406-415, Aug. 1986.

[8] I. Cidon, I. Gopal, "Control mechanisms for high-speed networks", *Proc. ICC '90*, Atlanta, Georgia, April 1990.

[9] D. D. Clark, M. L. Lambert, L. Zhang, "NETBLT: a high throughput transport protocol", *Proc. ACM SIGCOMM '87*, Stowe, Vermont, pp. 353-359, 1987.

[10] D. Comer, R. Yavatkar, "A rate-based congestion avoidance and control scheme for packet switched networks", *Proc. 10th ICDCS*, pp. 390-397, Paris, May 1990.

[11] A. Demers, S. Keshav, S. Shenker, "Analysis and simulation of a fair queueing algorithm", *Proc. ACM SIGCOMM '89*, Austin, Texas, pp. 1-12, September 1989.

[12] M. Gerla, L. Kleinrock, "Flow control: A comparative survey", *IEEE Trans. Comm.*, vol. COM-28, April 1980.

[13] Z. Haas, J. Winters, "Congestion control by adaptive admission", *Proc. INFOCOM '91*, Bal Harbour, FL, May 1991, pp. 6A.3.1-6A.3.10.

[14] E. Hahne, C. R. Kalmanek, S. P. Morgan, "Fairness and congestion control on a large ATM data network with dynamically adjustable windows", *Proc. 13th ITC*, Copenhagen, June 1991.

[15] V. Jacobson, "Congestion avoidance and control", *Proc. ACM SIGCOMM '88*, Stanford, CA, pp. 314-329, August 1988.

[16] P. Karn, C. Partridge, "Improving round-trip time estimates in reliable transport protocols", *Proc. ACM SIGCOMM '87*, Stowe, Vermont, pp. 2-7, 1987.

[17] S. Keshav, "The REAL network simulator", Technical Report UCB/CSD 88/472, University of California at Berkeley, December 1988.

[18] S. Keshav, A. K. Agrawala, S. Singh, "Design and analysis of a flow control algorithm for a network of rate allocating servers", in *Protocols for High-speed Networks II*, Elsevier/North-Holland, April 1991.

[19] S. Keshav, "A control-theoretic approach to flow control", *Proc. ACM SIGCOMM '91*, Zurich, Switzerland, September 1991.

[20] L. Kleinrock, *Queueing Systems. Volume 2: Computer Applications*, Wiley-Interscience, New York 1976.

[21] A. M. Law, "Statistical analysis of simulation output data", *Operations Research*, vol. 31, no. 6, pp. 983-1029, 1983.

[22] A. Mankin, "Random drop congestion control", *Proc. ACM SIGCOMM '90*, Philadelphia, Pa., pp. 1-7, September 1990.

[23] A. Mankin et *al.*, "Evaluation of Internet Performance", MITRE Report MTR-8900216, The Mitre Corporation, McLean, Virginia, April 1990.

[24] N. F. Maxemchuk, M. El Zarki, "Routing and flow control in high speed, wide area networks", *Proc. IEEE*, pp. 10-23, December 1989.

[25] D. Mitra, J. B. Seery, "Dynamic adaptive windows for high speed data networks: Theory and simulations", *Proc. ACM SIGCOMM '90*, Philadelphia, PA, pp. 30-40, September 1990.

[26] A. Mukherjee, J. C. Strikwerda, "Analysis of dynamic congestion control protocols: A Fokker-Plank approach", *Proc. ACM SIGCOMM '91*, Zurich, Switzerland, Sept. 1991.

[27] G. Newell, *Applications of Queueing Theory*, Chapman and Wall, London, 1971.

[28] K. K. Ramakrishnan, R. Jain, "A binary feedback scheme for congestion avoidance in computer networks with a connectionless network layer", *Proc. ACM SIGCOMM '88*, Stanford, California, pp. 303-313, August 1988.

[29] M. I. Reiman, "Some diffusion approximations with state-space collapse", in *Modeling and Performance Evaluation Methodology*, F. Baccelli and G. Fayolle, eds., Springer-Verlag, NY, 1984, pp. 209-240.

[30] M. Reiser, "A queueing network analysis of computer communication networks with window flow control", *IEEE Trans. on Comm.*, vol. COM-27, no. 10, pp. 1199-1209, Oct. 1979.

[31] J. Robinson, D. Friedman, M. Steenstrup, "Congestion control in BBN packet-switched networks", *Computer Communication Review*, vol. 20, no. 1, pp. 76-90, January 1990.

[32] D. Sanghi, M. Subramanian, A. U. Shankar, O. Gudmundsson, P. Jalote, "A TCP instrumentation and its use in evaluating roundtrip-time estimators", *Internetworking: Research and Experience*, vol. 1, no. 2, December 1990, pp. 77-99.

[33] M. Schwartz, "Performance analysis of the SNA virtual route pacing control", *IEEE Trans. on Comm.*, vol. COM-30, no. 1, pp. 172-184, Jan. 1982

[34] D. Tipper, M. K. Sundareshan, "Numerical methods for modeling computer networks under nonstationary conditions", *IEEE JSAC*, vol. 9, no. 8, December 1990, pp. 1682-1695.

[35] S. K. Tripathi, A. Duda, "Time-dependent analysis of queueing systems", *INFOR*, vol. 24, no. 3, 1986, pp. 199-220.

[36] H. van As, "Transient analysis of Markovian queueing systems and its application to congestion control modeling", *IEEE JSAC*, vol. 4, no. 6, pp. 891-904, Sept. 1986.

[37] L. Wong, M. Schwartz, "Access control in metropolitan area networks", *Proc. ICC '90*, pp. 190-195, Atlanta, April 1990.

[38] N. Yin, M. G. Hluchyj, "A dynamic rate control mechanism for integrated networks", *Proc. INFOCOM '91*, Bal Harbour, FL, May 1991, pp. 6.A.1.1-6A.1.10.

[39] L. Zhang, *A New Architecture for Packet Switching Network Protocols*, PhD Thesis, Massachussets Institute of Technology, July 1989.