ON FLOW CONTROL IN COMPUTER NETWORKS*

Leonard Kleinrock
Computer Science Department
University of California
Los Angeles, California 90024

Summary

We discuss the effect that flow control procedures have on the throughput in a computer network. We introduce the notion of starvation and use it as the basis for the measure of the cost of any flow control procedure. For any given flow control procedure, we then apply a previously defined measure of power and show how to identify that operating point which yields maximum power.

Introduction

Among the remaining technical problems in computer network design, perhaps one of the most important is that of flow control (we distinguish technical problems from the "softer" social, political, legal and ecological problems). Although "The Flow Control Problem" does not have a universally accepted definition, it is enough for this paper to describe it as the problem of allocating network resources to the demands placed upon those resources by the user population. The user population includes any source of data which requires transmission through the computer network. The network resources include communications capacity, processing capacity (at the network switching nodes) and network storage capacity. The purpose of a flow control procedure is to throttle the flow of traffic entering (and leaving) the net in a way which protects the network and the data sources from each other while at the same time maintaining a smooth flow of data in an efficient fashion. It would be lovely if we knew how to accomplish all of this at the same time - unfortunately, we do not. To date, there is no clear procedure for allocating network resources to provide a large throughput at low delay in an equitable fashion among competing demands distributed across a network.[2] Occasionally we flood low speed output devices or "nickel-and-dime" high speed input devices with incessant interrupts, or overwhelm the network with too much traffic or starve the network with too little traffic, or mis-allocate network resources so that deadlocks and performance degradations occur.[5]

In this paper, we wish to make some definitions, introduce a useful cost measure, and draw some general conclusions about flow control procedures. It is not out intent to describe any particular flow control procedure, but rather to step back and identify some phenomenological issues and concepts of importance.

The Model

Our view of a network is simply a system to which is applied a certain input message traffic. The network accepts a portion of this traffic (the carried traffic) and delivers it to its destination. The average time it takes to deliver this traffic we denote by T. Thus, we have the following definitions:

λ = input rate applied to the network (msg/sec)

$\gamma = \gamma(\lambda)$ = traffic carried by the network = network throughput (msg/sec)

T = average network delay (msg)

Let us also define the network capacity as the maximum traffic that can be handled by the network:

$$\gamma_0 = \text{network capacity (msg/sec)}$$

We are interested in the function $\gamma(\lambda)$ which expresses the throughput achieved by a given flow control procedure which is subject to an input traffic intensity of λ. Under ideal conditions we would have the $\gamma(\lambda)$ profile shown in Figure 1. Here we see that all the input traffic is accepted until we saturate at a value equal to the network capacity. This corresponds to perfectly regulated traffic. However, due to random

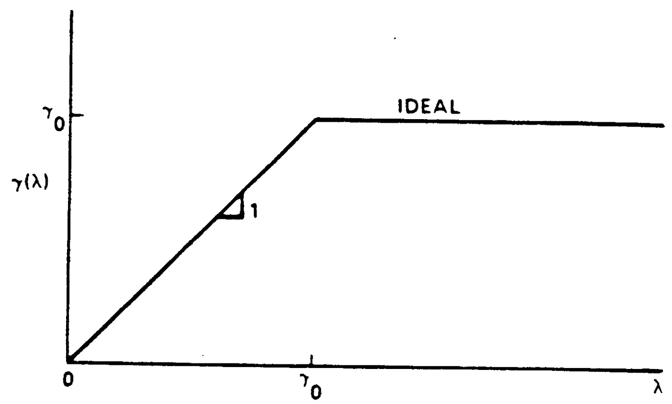

Fig. 1. An ideal throughput characteristic.

fluctuations in the input traffic, we do not achieve this ideal behavior. Indeed, we may encounter at least three kinds of $\gamma(\lambda)$ curves according to the type of flow control we exert, as shown in Figure 2. First we could

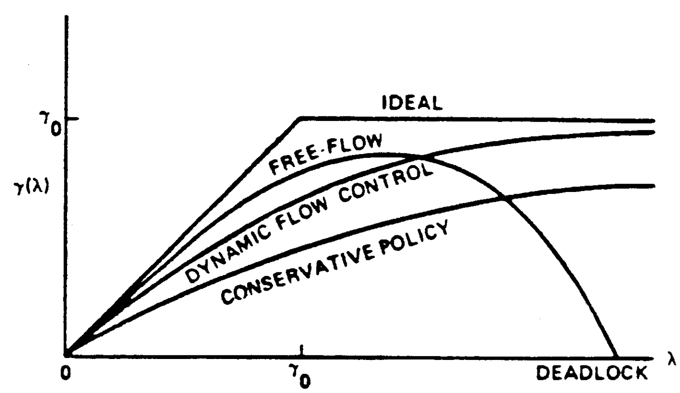

Fig. 2. The behavior of various flow control policies.

*This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract MDA 903-77-C-0272.

adopt a conservative policy and accept less traffic than the network can handle, thereby insuring no congestion (but lots of rejected input traffic). Second, we could relax completely and allow the traffic to flow freely; that is, we could put no limitations on the accepted traffic (no flow control). As we see, this policy often leads to serious degradations, and even to deadlocks if the congested traffic flow patterns in the network result in "resource-smashing" (poor or destructive use of resources, such as excessive retransmissions). The more sensible policy is to apply some form of dynamic flow control which will throttle the accepted traffic as a function of the network congestion, capacity, delay, etc.[3] In what follows, we usually assume the existence of some reasonable flow control policy.

Some Definitions

Consider the flow control function $\gamma(\lambda)$ shown in Fig. 3. In this figure we have extended the two asymptotes of the ideal curve. Let us examine the regions

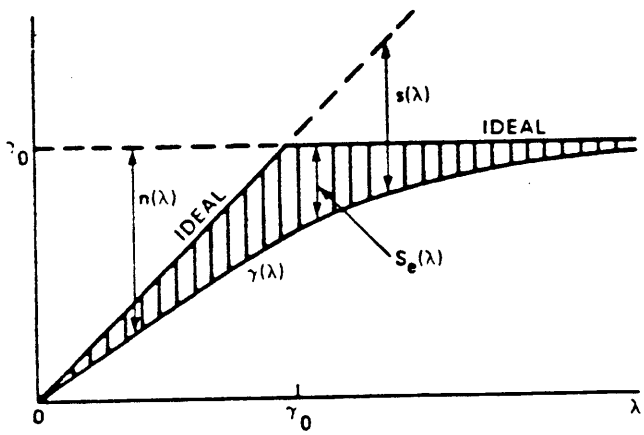

Fig. 3. Starvation

in this figure. First, we consider the ideal curve, $\gamma_{IDEAL}(\lambda)$ defined as

$$\gamma_{IDEAL}(\lambda) \triangleq \begin{cases} \lambda & \lambda \leq \gamma_0 \\ \gamma_0 & \lambda \geq \gamma_0 \end{cases} \tag{1}$$

Let us now define the function $f(\lambda)$ as

$$f(\lambda) \triangleq \gamma_0 - \lambda \tag{2}$$

Now, for $\lambda \leq \gamma_0$, we have

$$f(\lambda) = \gamma_0 - \gamma_{IDEAL}(\lambda) \tag{3}$$

In this region, we see that the network is carrying all of the applied traffic. However, the net is being "starved" in the sense that it is capable of carrying more traffic (in fact, it can carry $f(\lambda)$ more). Therefore, we refer to the region $\lambda < \gamma_0$ as the region of "network starvation." On the other hand, for $\lambda \geq \gamma_0$ we have

$$f(\lambda) = \gamma_{IDEAL}(\lambda) - \lambda \tag{4}$$

In this region we see that the network has reached its traffic-carrying capacity. However, a growing fraction of the applied traffic is being turned away (or, if you will, held back in an explosive queue). Thus let us refer to this region as the region of "source starvation" (i.e., the source wants more capacity). Since $f(\lambda)$ is negative for $\lambda > \gamma_0$, we measure the source starvation by $-f(\lambda)$. Thus, we define the "irreducible starvation" function $S_{IRR}(\lambda)$ as

$$S_{IRR}(\lambda) = |f(\lambda)| \triangleq \begin{cases} f(\lambda) & \lambda \leq \gamma_0 \\ -f(\lambda) & \lambda \geq \gamma_0 \end{cases} \tag{5}$$

This function is shown in Figure 4.

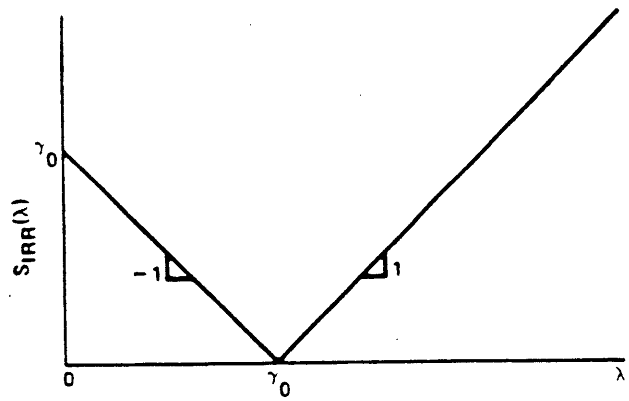

Fig. 4. The irreducible starvation.

Since we have been considering the ideal curve, we see that the irreducible starvation $S_{IRR}(\lambda)$ is the minimum we can ever have; that is, we must pay at least the "price" $S_{IRR}(\lambda)$.

Now, let us consider the "typical" flow control function $\gamma(\lambda)$ as in Figure 3. For this function, we may define the starvation function $S(\lambda)$ as

$$S(\lambda) \triangleq \begin{cases} \gamma_0 - \gamma(\lambda) & \lambda \leq \gamma_0 \\ \lambda - \gamma(\lambda) & \lambda \geq \gamma_0 \end{cases} \tag{6}$$

$S(\lambda)$ is sketched in Figure 5, in which we have also shown $S_{IRR}(\lambda)$.

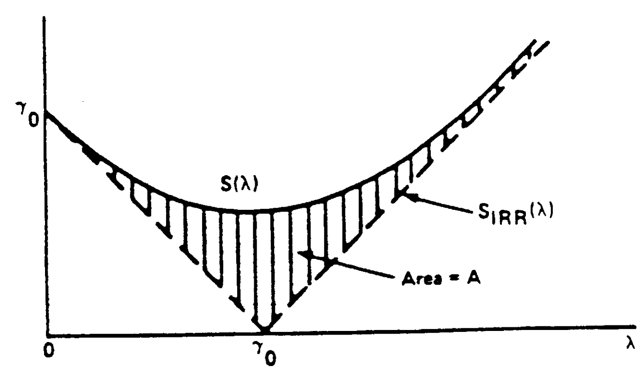

Fig. 5. A typical starvation function.

We may write

$$S(\lambda) = S_{IRR}(\lambda) + \gamma_{IDEAL}(\lambda) - \gamma(\lambda) \tag{7}$$

We further define the "excess starvation" function $S_e(\lambda)$ as

$$S_e(\lambda) \triangleq s(\lambda) - s_{IRR}(\lambda)$$

$$= \gamma_{IDEAL}(\lambda) - \gamma(\lambda) \qquad (8)$$

Thus, we see that $S_e(\lambda)$ is simply the vertical offset between γ_{IDEAL} and $\gamma(\lambda)$ as shown in Figure 3.

Note that the maximum of $S_e(\lambda)$ occurs at $\lambda = \gamma_0$ as shown in Figure 6. This is true in general so long as $\gamma(\lambda)$ is concave (i.e., with non-increasing slope).

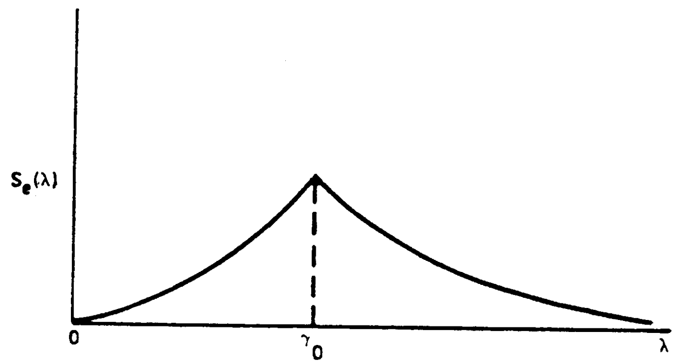

Fig. 6. The excess starvation.

We may further define the network starvation as

$$n(\lambda) \triangleq \gamma_0 - \gamma(\lambda) \qquad (9)$$

and the source starvation as

$$s(\lambda) \triangleq \lambda - \gamma(\lambda) \qquad (10)$$

as shown in Figs. 3 and 7. Clearly we have $n(\gamma_0) = s(\gamma_0)$ and we see in general why we refer to the region $\lambda < \gamma_0$ as the region of network starvation and to the region $\lambda > \gamma_0$ as the region of source starvation.

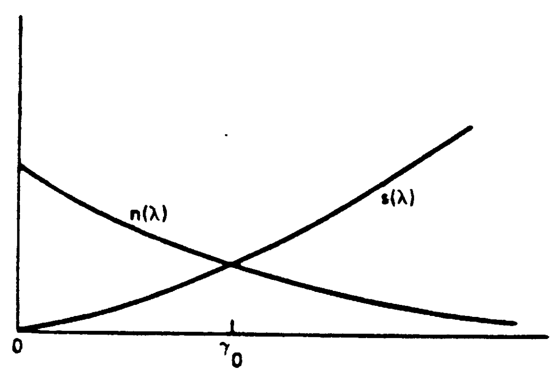

Fig. 7. Network and source starvation

Clearly, if we subtract out the irreducible net starvation, $\gamma_0 - \lambda$ $(\lambda \leq \gamma_0)$, from $n(\lambda)$ and subtract out the irreducible source starvation, $\lambda - \gamma_0$ $(\lambda \geq \gamma_0)$, from $s(\lambda)$ we obtain the excess net and source starvation functions $n_e(\lambda)$ and $s_e(\lambda)$, respectively. Thus

$$n_e(\lambda) \triangleq \begin{cases} \lambda - \gamma(\lambda) & \lambda \leq \gamma_0 \\ \gamma_0 - \gamma(\lambda) & \lambda \geq \gamma_0 \end{cases} \qquad (11)$$

$$s_e(\lambda) \triangleq \begin{cases} \lambda - \gamma(\lambda) & \lambda \leq \gamma \\ \gamma_0 - \gamma(\lambda) & \lambda \geq \gamma \end{cases} \qquad (12)$$

and so we have that

$$n_e(\lambda) = s_e(\lambda) = \gamma_{IDEAL}(\lambda) - \gamma(\lambda) \qquad (13)$$

and so

$$n_e(\lambda) = s_e(\lambda) = S_e(\lambda) \qquad (14)$$

(see Figure 6).

A Cost Measure

We now propose an extremely simple measure of the "cost" C of a flow control procedure, namely,

$$C \triangleq \int_0^\infty S_e(\lambda)\, d\lambda \qquad (15)$$

That is, we suggest that the cross-hatched area shown in Figure 3 be the measure of badness of a flow control procedure. If we study Figure 2, we see that

$$C_{FREE-FLOW} = \infty$$

$$C_{CONSERVATIVE} >> C_{DYNAMIC}$$

$$C_{IDEAL} = 0$$

and so the measure seems to have the correct flavor. (Actually, a better measure of cost would be the probability-weighted value of $S_e(\lambda)$, that is, the expected value of $S_e(\lambda)$; unfortunately, it is far more difficult to evaluate this expectation - we need some knowledge of the statistical structure - than it is to evaluate the expression given in Eq. (15).)

Let us now consider the cross-hatched area A shown in Figure 5. Clearly

$$A = \int_0^\infty [s(\lambda) - s_{IRR}(\lambda)]\, d\lambda$$

$$= \int_0^\infty S_e(\lambda)\, d\lambda \qquad (16)$$

and so A = C. Thus we can see that both areas are identical and both are intuitively pleasing measures of badness which are easy to evaluate.

Flow Control and Power

An interesting performance ratio was recently defined in [1] which they refer to as "power". It combines the throughput $\gamma(\lambda)$ and the average network delay T into the following single measure of power, P:

$$P \triangleq \frac{\gamma(\lambda)}{T} \qquad (17)$$

Clearly $T = T(\gamma(\lambda))$ which we choose to write as $T(\gamma)$. We inquire as to the conditions necessary for P to be maximum. We have

$$\frac{dP}{d\lambda} = \frac{T(\gamma)\, d\gamma(\lambda)/d\lambda - \gamma(\lambda)\, dT(\gamma)/d\lambda}{[T(\gamma)]^2}$$

Since $T(\gamma) \geq 0$, the condition for P_{MAX} is simply

$$T(\gamma)\,\frac{d\gamma(\lambda)}{d\lambda} = \gamma(\lambda)\,\frac{dT(\gamma)}{d\lambda}$$

or

$$\frac{dT(\gamma)}{T(\gamma)} = \frac{d\gamma(\lambda)}{\gamma(\lambda)} \tag{18}$$

For the rest of this section (unless stated otherwise) we shall assume that the average network delay may be modelled as a k-hop network with each hop modeled by an M/M/1 queue and with an instantaneous end-to-end acknowledgement. That is, each hop adds an amount

$$\frac{1}{\gamma_0 - \gamma(\lambda)}$$

seconds to the average delay and so the total average network delay $T(\gamma)$ is given by[4]

$$T(\gamma) = \frac{k}{\gamma_0 - \gamma(\lambda)} \tag{19}$$

Then

$$\frac{dT(\gamma)}{d\gamma(\lambda)} = \frac{k}{[\gamma_0 - \gamma(\lambda)]^2}$$

Using this in Eq. (18) gives

$$\frac{1}{\gamma_0 - \gamma(\lambda)} = \frac{1}{\gamma(\lambda)}$$

whose solution is

$$\gamma(\lambda) = \gamma_0/2 \tag{20}$$

Thus, the condition for P_{MAX} is that λ be selected at the value λ^* such that $\gamma^* \triangleq \gamma(\lambda^*) = \gamma_0/2$, and this implies that $T(\gamma^*) = 2k/\gamma_0 = 2T(0)$. That is, we balance throughput and delay such that we operate at half the maximum throughput and twice the minimum delay. This applies to any throughput function $\gamma(\lambda)$!

Note that for all $\gamma(\lambda)$ that $P = \gamma/T = \gamma(\gamma_0 - \gamma)/k$. For $\gamma_{IDEAL}(\lambda)$, we have $\gamma(\lambda) = \gamma_0/2$ at $\lambda^* = \gamma_0/2$. (See Fig. 8 for the ideal case.)

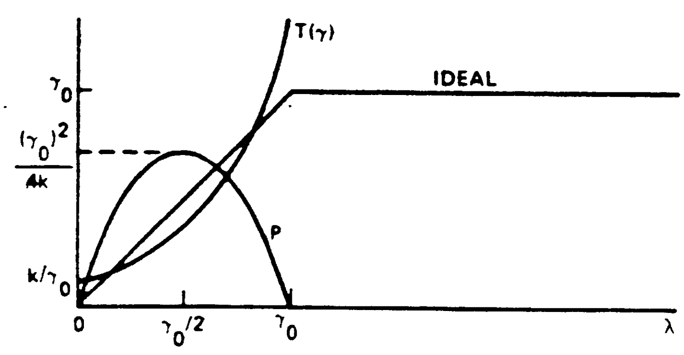

Fig. 8. Power, delay and throughput in the ideal case.

Moreover, as a special case, let us consider a dynamic flow control model using a window scheme.[2] This we model by claiming that $\overline{N}$, the average number of messages in the network, is restricted to a number w (the window size). Thus by Little's result[4]

$$\overline{N} = w = \gamma(\lambda)T(\gamma) \tag{21}$$

or

$$w = \frac{k\gamma(\lambda)}{\gamma_0 - \gamma(\lambda)}$$

which gives

$$\gamma(\lambda) = \frac{w}{k+w}\,\gamma_0 \tag{22}$$

For optimality, we know that $\gamma(\lambda) = \gamma_0/2$ which implies that $w = k$. This implies that the network should contain k messages, on the average, to give maximum power. That is, we should just keep the pipe full (i.e., one message per hop).

We see that for $\lambda < \gamma_0$, we have net starvation; for $\lambda > \gamma_0$, we have source starvation. Therefore, one is inclined to suggest that the "proper" operating range is in the vicinity $\lambda = \gamma_0$. Indeed, for the ideal curve shown in Figure 1, we see that $\lambda = \gamma_0$ is the proper operating region for deterministic flow.† The apparent contradiction between this last statement and the fact that maximum power is achieved at $\lambda = \gamma_0/2$ in Fig. 8 is reconciled when one replaces the M/M/1 assumption from Fig. 8 with the D/D/1 assumption in the last sentence. We note that the condition necessary such that $\lambda = \gamma_0$ be the value which maximizes P is simply $\gamma(\gamma_0) = \gamma_0/2$.

An additional observation regarding the power P may be made by studying the delay-throughput function in Fig. 9. From Eq. (17) we see that P is the inverse

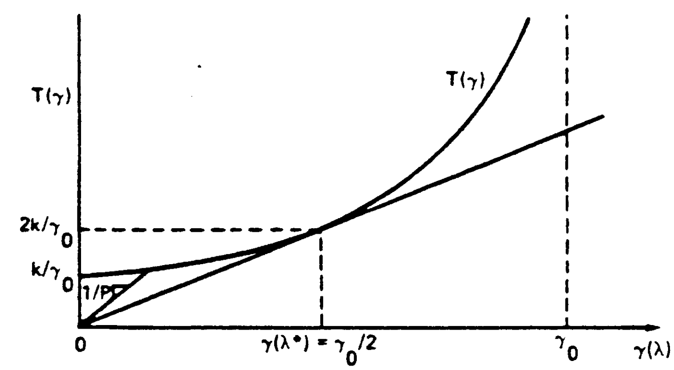

Fig. 9. Maximum power in the delay-throughput plane.

slope of the line joining the origin to the $T(\gamma)$ curve. For the assumption of k M/M/1 hops as given in Eq. (19), P_{MAX} occurs when $\gamma(\lambda^*) = \gamma_0/2$ and $T(\gamma^*) = 2T(0)$. If we select any other function for $T(\gamma)$ (instead of the form given in Eq. (19)), then P_{MAX} may be found from Fig. 9 as the inverse slope of the tangent to $T(\gamma)$ from the origin. To prove that last statement mathematically, we return to Eq. (18) and rewrite it as

$$\frac{dT(\gamma)}{d\gamma(\lambda)} = \frac{T(\gamma)}{\gamma(\lambda)}$$

† We note at this operating point ($\lambda = \gamma_0$) that $n(\lambda) = s(\lambda)$ and so the net starvation and source starvation are, in some sense, "matched"; for $\lambda < \gamma_0$, $n(\lambda) > s(\lambda)$, whereas for $\lambda > \gamma_0$, $n(\lambda) < s(\lambda)$.

which simply states that the condition on $\gamma(\lambda)$ for P_{MAX} is that the slope of $T(\gamma)$ be equal to $T(\gamma)/\gamma$ which is the condition we stated earlier (namely, that γ^* is that value of γ at which a ray out of the origin is tangent to $T(\gamma)$). This is true for any $T(\gamma)$ and any $\gamma(\lambda)$.

Moreover, we may observe that when the optimality condition given in Eq. (18) is satisfied, then the relative increase in delay (dT/T) is equal to the relative increase in throughput ($d\gamma/\gamma$). This last observation lends credibility to the usefulness of the notion of maximum power as follows: For $\gamma < \gamma^*$ we find that $dT/T < d\gamma/\gamma$ and therefore we should increase λ since we then gain more in relative throughput than we lose in relative delay. For $\gamma > \gamma^*$ we find that $dT/T > d\gamma/\gamma$ and therefore we should decrease λ thereby reducing the relative delay faster than we are reducing the relative throughput. Clearly, this forces us to converge to $\lambda = \lambda^*$, $\gamma = \gamma^*$ at which point the relative increase in delay exactly equals the relative increase in throughput.

Conclusions

Let us now summarize the main results of this paper:

(1) We have defined the concept of starvation and how it behaves in various operating regions. Specifically, we have the net starvation ($n(\lambda)$), the source starvation ($s(\lambda)$), the irreducible starvation ($S_{IRR}(\lambda)$), and the excess starvation ($S_e(\lambda)$). We have identified $S_{IRR}(\lambda)$ as the minimum price which must be paid, and have shown that the excess net starvation ($n_e(\lambda)$) and source starvation ($s_e(\lambda)$) are such that

$$n_e(\lambda) = s_e(\lambda) = S_e(\lambda).$$

(2) We have proposed a cost function C which is a useful measure of the poorness of a flow control scheme. C is equal to the area between the ideal and the actual throughput functions. Perhaps a more accurate cost function would be the difference between the ideal and the actual throughput weighted by the probability of achieving each throughput (this would give the average difference); however, such a measure would require an enormous amount of additional information (the probability measure) which is difficult to calculate, whereas the area is independent of such information.

(3) We used the power P as a means to determine where the "best" operating point should be (where best means maximum P). We found the following:

(a) For any delay function $T(\gamma)$ and any throughput function $\gamma(\lambda)$, the optimum point is the place where a ray out of the origin is tangent to the $T(\gamma)$ function. That is, we must operate at the "knee" of the $T(\gamma)$ curve; this is an intuitively pleasing result.

(b) As an example, if we assume a k-hop M/M/1 form for $T(\gamma)$ (see Eq. (19)), then for any $\gamma(\lambda)$ we have the following:

(i) The power is maximized when λ is selected such that $\gamma(\lambda) = \gamma_0/2$. That is, drive the system to achieve only half of the maximum throughput.

(ii) At this optimum point, we have a delay which is twice the minimum delay, i.e., $T(\gamma_0/2) = 2T(0)$.

(iii) At this optimum point, the window should be $w = k$; this assumes no delay for acknowledgements.

(4) The "balanced starvation point" $n(\lambda) = s(\lambda)$ always occurs at $\lambda = \gamma_0$. The power will be maximized there if, as in (3a) above, the tangent occurs at $\lambda = \gamma_0$ or if, as in (3b) above, $\gamma(\gamma_0) = \gamma_0/2$.

Although we have emphasized the application of flow control to a single network in this paper, it should be clear that these considerations are equally well applied to the far more complicated structure of many interconnected networks, a subject of considerable current interest.

References

[1] A. Giessler, A. Koenig, J. Haenle and E. Pade, "Deadlock-Free Packet Networks, " to appear in Computer Networks.

[2] R. E. Kahn and W. R. Crowther, "Flow Control in a Resource Sharing Computer Network," Proc. of the Second ACM-IEEE Symposium on Problems in the Optimization of Data Communications Systems, Palo Alto, California, October 1971, pp. 108-116.

[3] P. Kermani, "Switching and Flow Control Techniques in Computer Communication Networks," Ph.D. Dissertation, School of Engineering and Applied Science, University of California, Los Angeles, 1977.

[4] L. Kleinrock, Queueing Systems, Vol. 1: Theory, Wiley Interscience, New York, 1975.

[5] L. Kleinrock, Queueing Systems, Vol. 2: Computer Applications, Wiley Interscience, New York, 1976.

[6] H. Rudin, "An Introduction to Flow Control," Proc. of the Third International Conference on Computer Communication, Toronto, Ontario, Canada, August 1976, pp. 463-466.