Reprinted from INTERNATIONAL CONFERENCE ON COMMUNICATIONS, June 1981

AN INVARIANT PROPERTY OF COMPUTER NETWORK POWER*

Richard Gail and Leonard Kleinrock

Computer Science Department
University of California
Los Angeles, California 90024

Abstract

Power has recently evolved as a potentially useful measure of computer network performance in that it suggests an appropriate operating point for simple nets. In this paper we elaborate on the observation that the average number in system is an invariant which defines the maximum power point, is a measurable quantity, and which can therefore serve as a flow control variable (for example, in a window flow control scheme). We show this invariance first for a series network of queues (Poisson input and constant service time) with arbitrary loads and then for a parallel configuration of M/G/1 systems (arbitrary service time distribution for each parallel branch) but with identical loads. We also show how these exact solutions may be extended as approximations to other queueing configurations.

1. Introduction

How does one select an "appropriate" operating point for a computer communications network? This seemingly simple question does not have a straightforward answer. Indeed, what is meant by the word "appropriate"? Furthermore, the modeling and analysis of computer networks draws upon techniques from many fields (e.g. queueing theory, optimization theory, networks and graphs, information theory), and one is faced with many different issues involving routing, topological design, flow control, etc. Various performance criteria such as delay, line cost, buffer size, throughput, and efficiency may be studied. Different physical networks (conventional wire, packet radio, satellite, local nets, etc.) lead to different models and emphasis. For example, recent work on multi-access protocols for one-hop satellite and packet radio nets emphasizes throughput (protocol capacity) as the major performance criterion of interest [1].

In this paper our interest is the tradeoff between throughput and delay inherent in the choice of a system operating point. One might expect that the more traffic (messages, packets) allowed into a network the higher the throughput (in congestion-prone systems, this need not be the case [2]) — but also the higher the delay. We choose to analyze this tradeoff using the notion of *power* introduced by Giessler et al [3] and subsequently studied by others [2, 4, 5, 6, 7]. When the network is operated at the maximum power point, we show below that the analysis leads to a value for the average number in system which is invariant under scaling of line capacities in some cases and invariant under distribution of message length (service time) in other cases. This value of the average number in system is easily calculated for the simple networks we consider, and may be implemented under a window flow control scheme [8].

The structure of this paper is as follows. After a description of the notation used (which follows that of [9] and [10]) we review the previous work on power. Next a series network with constant message length (e.g., a path in a virtual circuit network) is analyzed. Finally a parallel configuration is studied (e.g., a packet switch with numerous outgoing channels where the average number represents buffer size or window size). Several special cases for certain network parameters are considered and a general discussion follows. It is seen that the results obtained may be used to approximate known M/M/1 power results.

Now for the queueing theoretic notation. We first set

λ = total arrival rate of messages to the network.

We always consider the arrival process to be *Poisson*. Also in the simple networks considered here we assume messages are not blocked or lost (e.g. infinite buffer capacity, no noise on the lines and no possibility of collision as in packet radio networks), and thus throughput is synonymous with the traffic applied to the system. We next set

$T = T(\lambda)$ = average total time spent in the network by a message

which is the sum of

$W = W(\lambda)$ = average total waiting time (on queues) spent in the network by a message

and

$\bar{x}$ = average total service time of a message.

(That is, $\bar{x}$ is the total time a message spends in transmission on all channels in its journey through the net.) Thus $T = W + \bar{x}$. Finally we set

$\bar{N}$ = average total number of messages in the network.

We also will use Little's result [11] which relates several of the above quantities, i.e., $\bar{N} = \lambda T$.

These various network parameters are composed of corresponding quantities for each of the message channels in the network. We let $\bar{N}_i$, λ_i, T_i, W_i, $\bar{x}_i$ be the values for channel i. The service time for a message on the ith channel, $\bar{x}_i$, may be expressed as the average length of a message, $\bar{b}$, in bits, divided by the capacity of the ith channel, C_i, in bits per second. That is, $\bar{x}_i = \bar{b}/C_i$. Thus the variation in service time at the ith channel occurs due to the variation in message length. We also define the utilization (efficiency) for the ith channel as $\rho_i = \lambda_i \bar{x}_i$.

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

2. Review Of Power

A performance measure combining throughput and delay into a single function is the notion of *power* introduced in [3]. This is simply defined as

$$P(\lambda) = \frac{\lambda}{T(\lambda)}. \tag{2.1}$$

One might consider that an appropriate operating point for a network is to choose that value of λ which maximizes power. Kleinrock showed in [2] that for $P(\lambda)$ a differentiable function of λ, power is maximized for a value of λ where

$$\lambda \frac{dT}{d\lambda} = T(\lambda). \tag{2.2}$$

Thus the optimum power point defines the "knee" of the $T(\lambda)$ curve. That is, it occurs at that value of λ where a straight line through the origin in the $(\lambda, T(\lambda))$ plane is tangent to the $T(\lambda)$ curve. (In fact, $T(\lambda)$ itself need not be a convex function of λ in order for this last statement to hold; namely, in the case where more than one tangent line can be found, maximum power will occur for that tangent line which makes the smallest angle with the horizontal axis.)

Using equation (2.2), Kleinrock showed in [4] for any M/G/1 system, that

$$\bar{N}^* = 1. \tag{2.3}$$

Here we use the convention that the superscript $*$ indicates that the variable in question has been *maximized* with respect to power. Thus equation (2.3) indicates that if the input rate λ to an M/G/1 system is chosen so as to maximize power $P(\lambda)$, then the average number in system $\bar{N}^*$ at this point λ^* is equal to 1. This agrees with the reasoning that the proper operating point of the deterministic D/D/1 system is exactly when $\bar{N}^* = 1$. It also can be shown that the server utilization at maximum power for M/G/1 is

$$\rho^* = \frac{1}{1+\sqrt{(1+\nu^2)/2}} \tag{2.4}$$

where ν is the coefficient of variation of the service time distribution (recall that ν is simply the ratio of the standard deviation to the mean service time).

Using the independence assumption [12] a message path in a network can be modeled as a series of independent M/M/1 queues (with $\bar{b} \triangleq 1/\mu$). Kleinrock showed in [2] that for such a series with M channels and *equal* channel capacities, then

$$\bar{N}^* = M. \tag{2.5}$$

In fact the optimal choice of λ^* occurs when $\rho_i^* = 1/2$ (e.g. where $\lambda^* = 1/(2\bar{x}) = \mu C/2$). Also $\bar{N}_i^* = 1$ for $i = 1, \ldots, M$. This model of a message path in a computer network was further examined by Bharath-Kumar in [5]. He found that for such a path having *arbitrary* channel capacities

$$\bar{N}^* \leqslant M \tag{2.6}$$

with Kleinrock's "keep the pipe full" result of equation (2.5) holding when all capacities are equal.

In these results we already see invariant properties of the average number in system at maximum power. For example, equation (2.3) shows an invariance with respect to service time distribution, while equation (2.5) shows an invariance with respect to scaling of channel capacities. No other system variable shows such invariance. In the next two sections we introduce other simple networks and again emphasize such invariance properties in our analyses.

3. The M/D/1 Series Network

We choose to model a message path in a computer network as a series of M queues where, unlike the M/M/1 model discussed above, the length of a message remains *constant* as it traverses the path, this being a more realistic assumption (see Figure 1).

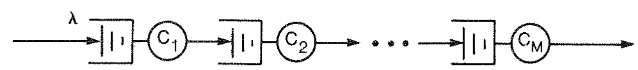

Figure 1 The M/D/1 Series Network

Note that the various channels need not have the same capacities. With Poisson arrivals the first node in the system is simply an M/D/1 queue. However, the subsequent nodes in the path are more complicated queueing systems (there is a dependence between the output process of one node and the input process of the next node). This model was extensively analyzed by Rubin [13]. He showed that the distribution of waiting time for the total system is identical to that for an M/D/1 queueing system with arrival rate λ and service time equal to the *maximum* of the individual service times at the nodes. Using Rubin's result we establish the following theorem in [14].

Theorem 1

For the M/D/1 series network, power is maximized when

$$\bar{N}^* = \sum_{i=1}^{M} \frac{C_{\min}}{C_i}. \tag{3.1}$$

Note that this expression does not involve λ^* and is invariant under scaling of channel capacities.

Let us express this important equation in another way which will lead to an interpretation which is intuitively pleasing. For each node $i = 1, \ldots, M$ set $C_i = \alpha_i C_{\min}$ where $\alpha_i \geqslant 1$. Then

$$\sum_{i=1}^{M} \frac{C_{\min}}{C_i} = \sum_{i=1}^{M} \frac{1}{\alpha_i}$$

and so we may write Theorem 1 in the form

$$\bar{N}^* = \sum_{i=1}^{M} \frac{1}{\alpha_i}. \tag{3.2}$$

Thus the slowest channel (capacity $C_{\min}$) contributes 1 to the average number in system while a channel which is $\alpha \geqslant 1$ times faster than the slowest channel contributes $1/\alpha \leqslant 1$ to this average number. This is similar to the "keep the pipe full" result of Kleinrock mentioned above. Also note that the average number does not depend on the *order* of the individual capacities C_i, but only on their values. For example, if the first channel is the slowest, no queueing takes place except at node 1. If the last channel is the slowest, queueing may take place at various nodes. Thus the *individual* values of $\bar{N}_i^*$ for $i = 1, \ldots, M$ *do* depend on the order of the capacities. But in all cases we have the invariance for the *total* number, i.e.,

$$\bar{N}^* = \sum_{i=1}^{M} \frac{1}{\alpha_i}.$$

Also clearly we have the same upper bound as for the M/M/1 series network, namely,

$$\bar{N}^* \leqslant M. \tag{3.3}$$

	M/M/1			M/D/1		
	Average Number In			Average Number In		
Node	Queue	Server	System	Queue	Server	System
1	$1-\rho^*$	ρ^*	1	$M(1-\rho^*)$	ρ^*	$M(1-\rho^*)+\rho^*$
2	$1-\rho^*$	ρ^*	1	0	ρ^*	ρ^*
3	$1-\rho^*$	ρ^*	1	0	ρ^*	ρ^*
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$M-1$	$1-\rho^*$	ρ^*	1	0	ρ^*	ρ^*
M	$1-\rho^*$	ρ^*	1	0	ρ^*	ρ^*
Total	$M(1-\rho^*)$	$M\rho^*$	M	$M(1-\rho^*)$	$M\rho^*$	M

Table 1

Other optimized variables of interest for the M/D/1 series net include λ^*, T^*, P^* and ρ_i^* for $i = 1, \ldots, M$. These are given in [14]. In particular we have

$$\rho^*_{\max} = \frac{\sqrt{2\bar{N}^*}}{1+\sqrt{2\bar{N}^*}}. \tag{3.4}$$

Equal Channel Capacities

Assuming $C_i = C$ for $i = 1, \ldots, M$ and thus $\rho_i \triangleq \rho$, Theorem 1 gives

$$\bar{N}^* = M. \tag{3.5}$$

This is the same result as Kleinrock obtained for the M/M/1 tandem [2]. Also equation (3.4) yields

$$\rho^* = \frac{\sqrt{2M}}{1+\sqrt{2M}} \tag{3.6}$$

which compares with the value of $\rho^* = 1/2$ mentioned above for the M/M/1 series net with equal channel capacities. Thus values of other variables (λ^*, ρ^*, T^*) differ between the M/M/1 and M/D/1 series models as do the individual node average numbers $\bar{N}_i^*$, $i = 1, \ldots, M$. It is all the more amazing that the expression $\bar{N}^* = M$ is invariant for both service time *distributions.*

Table 1 shows the differences in the values of various optimized variables for the two systems. This table graphically illustrates that, although $\bar{N}^* = M$ for both series networks, the behavior of the two systems at optimal power is quite different. Noting that $\rho^* = 1/2$ for the M/M/1 series network, and that $\rho^* = \frac{\sqrt{2M}}{1+\sqrt{2M}}$ for the series network with constant message length, we obtain the optimized values shown in Table 2. We see from Table 2 that in the M/M/1 tandem, each of the M channels contributes exactly 1 to the average number (for any value of M). In the M/D/1 series net with equal capacities there is no queueing at nodes $i = 2, \ldots, M$ (all queueing occurs at the first channel). In fact for the M/D/1 tandem, as $M \to \infty$, then $\rho^* \to 1$ (almost 1 in each server) and also the number in queue at the first node $M(1-\rho^*) \to \infty$. As the table indicates, other system variables (ρ^*, $\bar{N}_i^*$, etc.) differ greatly for the two models, but in both cases $\bar{N}^* = M$(!)

	M/M/1			M/D/1		
	Average Number In			Average Number In		
Node	Queue	Server	System	Queue	Server	System
1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{M}{1+\sqrt{2M}}$	$\frac{\sqrt{2M}}{1+\sqrt{2M}}$	$\frac{M+\sqrt{2M}}{1+\sqrt{2M}}$
2	$\frac{1}{2}$	$\frac{1}{2}$	1	0	$\frac{\sqrt{2M}}{1+\sqrt{2M}}$	$\frac{\sqrt{2M}}{1+\sqrt{2M}}$
3	$\frac{1}{2}$	$\frac{1}{2}$	1	0	$\frac{\sqrt{2M}}{1+\sqrt{2M}}$	$\frac{\sqrt{2M}}{1+\sqrt{2M}}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$M-1$	$\frac{1}{2}$	$\frac{1}{2}$	1	0	$\frac{\sqrt{2M}}{1+\sqrt{2M}}$	$\frac{\sqrt{2M}}{1+\sqrt{2M}}$
M	$\frac{1}{2}$	$\frac{1}{2}$	1	0	$\frac{\sqrt{2M}}{1+\sqrt{2M}}$	$\frac{\sqrt{2M}}{1+\sqrt{2M}}$
Total	$\frac{M}{2}$	$\frac{M}{2}$	M	$\frac{M}{1+\sqrt{2M}}$	$\frac{M\sqrt{2M}}{1+\sqrt{2M}}$	M

Table 2

4. The M/G/1 Parallel Network

This network model is for a parallel configuration of M message channels, where channel i has capacity C_i. The input to the system is again Poisson with parameter λ. The message length distribution is *arbitrary*. Probabilities $0 < p_i < 1$ for $i = 1, \ldots, M$ are given ($\sum_{i=1}^{M} p_i = 1$) which determine the channel a newly arrived message chooses. Thus each channel i acts as an M/G/1 queueing system with Poisson input rate $\lambda_i = p_i \lambda$, $i = 1, \ldots, M$ (see Figure 2). We wish to find that value of λ which will maximize the power of this parallel system. It can be shown [14] that the following equation relates various optimized variables, namely

$$\sum_{i=1}^{M} \frac{\rho_i^*}{1-\rho_i^*}\left(1-\bar{N}_i^*\right) = 0. \tag{4.1}$$

Let us exploit equation (4.1) to obtain various results about $\bar{N}^*$.

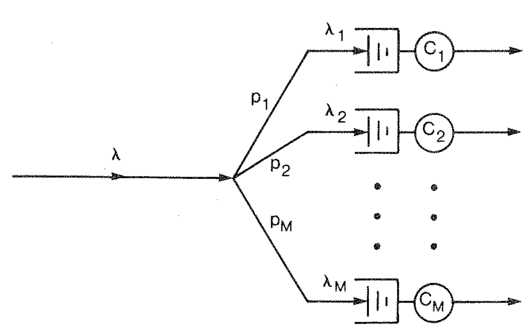

Figure 2 The M/G/1 Parallel Network

Case I: M/M/1 Parallel Network

Assume all M channels are modeled as M/M/1 queueing systems (exponential service). Then it is well-known that

$$\overline{N}_i^* = \frac{\rho_i^*}{1-\rho_i^*} \qquad i = 1, \ldots, M.$$

Thus equation (4.1) becomes

$$\sum_{i=1}^{M} \overline{N}_i^*\left[1-\overline{N}_i^*\right] = 0$$

or

$$\overline{N}^* = \sum_{i=1}^{M} \overline{N}_i^* = \sum_{i=1}^{M}\left[\overline{N}_i^*\right]^2. \tag{4.2}$$

(Note: this expression also appears in [5] as part of the M/M/1 *series* analysis.) From this equation we see

$$\overline{N}^* = \sum_{i=1}^{M} \overline{N}_i^* = \sum_{i=1}^{M}\left[\overline{N}_i^* + \left(\overline{N}_i^* - \left[\overline{N}_i^*\right]^2\right)\right] = \sum_{i=1}^{M}\left[1 - 1 + 2\overline{N}_i^* - \left[\overline{N}_i^*\right]^2\right]$$

or

$$\overline{N}^* = M - \sum_{i=1}^{M}\left[1-\overline{N}_i^*\right]^2. \tag{4.3}$$

This immediately yields

Theorem 2

For the M/M/1 parallel network, the average number in system at maximum power satisfies

$$\overline{N}^* \leqslant M. \tag{4.4}$$

This is the same bound mentioned above for the M/M/1 tandem.

Case II: Equal Loads (arbitrary service time distributions)

We now assume that all M channels have equal loads, that is, $\rho_i = \rho_j \triangleq \rho$ for $i, j \in \{1, \ldots, M\}$. This is similar to the series assumption of equal capacities, since in that case the assumption of equal capacities is equivalent to the assumption of equal loads (same input rate λ for all channels). Our point of departure is again equation (4.1). Since $\rho_i^* \triangleq \rho^*$ for $i = 1, \ldots, M$ we may write that equation in the form

$$\frac{\rho^*}{1-\rho^*}\sum_{i=1}^{M}\left[1-\overline{N}_i^*\right] = 0$$

and so

$$\sum_{i=1}^{M}\left[1-\overline{N}_i^*\right] = 0$$

or

$$\sum_{i=1}^{M}\overline{N}_i^* = M.$$

Hence, regardless of the service time distributions at the individual nodes, the equal load assumption gives

Theorem 3

For the equal load M/G/1 parallel network, power is maximized when

$$\overline{N}^* = M. \tag{4.5}$$

(Note: the special case $M = 1$ gives Kleinrock's result that $\overline{N}^* = 1$ for M/G/1.) Here again we see invariance with respect to service time distribution. There is also an invariance with respect to load scaling similar to the invariance with respect to scaling of channel capacities for the M/D/1 series net.

Other system parameters are derived in [14]. Of interest is

$$\rho^* = \frac{\sqrt{M}}{\sqrt{M} + \sqrt{\sum_{i=1}^{M}(1+\nu_i^2)/2}}. \tag{4.6}$$

Here ν_i is the coefficient of variation of the service time distribution for channel i. In general, the values of the average number at the *individual* nodes $\overline{N}_i^*$ will depend on the service time distribution. But amazingly, these quantities always add in such a way so as to force the total number $\overline{N}^* = M$.

In the case of equal coefficients of service time variation at the M nodes we can say more. We first express $\overline{N}_i^*$ by the Pollaczek–Khinchin mean-value formula as (recall that $\rho_i^* = \rho^*$)

$$\overline{N}_i^* = \frac{(\rho^*)^2}{2(1-\rho^*)}(1+\nu_i^2) + \rho^*. \tag{4.7}$$

Assume now that $\nu_i^2 = \nu^2$ for $i = 1, \ldots, M$ (which is certainly true if all service time distributions are of the same type). Under this additional assumption equation (4.6) gives

$$\rho^* = \frac{1}{1+\sqrt{(1+\nu^2)/2}}.$$

Also by equation (4.7) we clearly have

$$\overline{N}_i^* = \overline{N}_j^* \quad \text{for } i, j \in \{1, \ldots, M\}$$

and so

$$\overline{N}^* = \sum_{i=1}^{M}\overline{N}_i^* = M \cdot \overline{N}_i^*.$$

Thus by Theorem 3 we conclude

Corollary 4

For the M/G/1 parallel network with equal loads and equal coefficients of service time variation

$$\overline{N}_i^* = 1 \quad i = 1, \ldots, M. \tag{4.8}$$

5. Discussion

One of the difficult problems encountered in the design and analysis of computer networks is the issue of flow control. The comprehensive survey of Gerla and Kleinrock [8] describes the current "state-of-the-art" in this area. These authors state the main functions of flow control as:

1. prevention of throughput degradation and loss of efficiency due to overload
2. deadlock avoidance
3. fair allocation of resources among competing users
4. speed matching between the network and its attached users

From the extensive literature on flow control we also mention the book edited by Grange and Gien [15] which contains the proceedings of a recent symposium devoted solely to the subject.

As mentioned earlier, the M/D/1 series network of section 3 may be used to model a virtual circuit in a packet network. A window flow control scheme is often used with virtual circuits. Our result that $\overline{N}^* = \sum_{i=1}^{M}\frac{C_{\min}}{C_i}$ (equation (3.1)) may therefore be used as the window size setting for this path. This number is easily calculable, and uses only local path information in the sense of Jaffe [6].

The above M/D/1 expression for $\overline{N}^*$ may also be used as an *approximation* for the average number in system at maximum power of the M/M/1 series model with arbitrary capacities. As explained in [5] the calculation of $\overline{N}^*$ for the M/M/1 tandem involves the complexity of finding roots of polynomials. We have numerically calculated values for average number in system for the M/M/1 series model with up to 100 channels and various combinations of capacities. We have compared the results to values obtained by operating the M/M/1 tandem at a window size given from the M/D/1 equation (3.1). The error in value of power was found to be small, although other system parameters had larger error in some cases. The best cases were those with fewer channels (as one might expect) and those with one slow server and all the rest fast (the bottleneck case). This is pleasing because the bound $\bar{N}^* \leqslant M$ (as shown above for M/D/1 and given in [5] for M/M/1) can be quite bad. For example, an M/M/1 series network with one slow channel and all the other $M-1$ channels say β ($>> M$) times faster than this channel has a value for the average number in system which is close to 1 (nowhere near the bound of M). But the M/D/1 value of $\sum_{i=1}^{M} \frac{1}{\alpha_i} = 1 + \frac{M-1}{\beta} \cong 1$ as expected. Thus equation (3.1) may be used as an approximation to the M/M/1 tandem.

We have seen that the average number in system for certain simple networks operating at maximum power is a quantity that exhibits invariances in several ways and is numerically quite easy to evaluate. Unfortunately the models studied do not take into account interfering traffic. When more complicated networks are considered, it appears that maximizing global power may be difficult. Examples of networks have been given in which power cannot be maximized using only local information [6] or in which power is not a concave function [7]. But the beautiful results obtained for $\bar{N}^*$ in the simple networks above may perhaps be used as the basis for *approximations* to other nets (as in the use of the M/D/1 series net to approximate the M/M/1 series net).

References

[1] M. L. Molle, *On the Capacity of Infinite Population Multiple Access Protocols,* UCLA Computer Science Department, University of California, Los Angeles (April, 1980). to appear in *IEEE Transactions on Information Theory*

[2] L. Kleinrock, "On Flow Control in Computer Networks," *Conference Record, International Conference on Communications* **2**, pp.27.2.1-27.2.5 (June 1978).

[3] A. Giessler, J. Hanle, A. Konig, and E. Pade, "Free Buffer Allocation — An Investigation by Simulation," *Computer Networks* **1**(3), pp.191-204 (July 1978).

[4] L. Kleinrock, "Power and Deterministic Rules of Thumb for Probabilistic Problems in Computer Communications," *Conference Record, International Conference on Communications,* pp.43.1.1-43.1.10 (June 1979).

[5] K. Bharath-Kumar, "Optimum End-to-End Flow Control in Networks," *Conference Record, International Conference on Communications,* pp.23.3.1-23.3.6 (June 1980).

[6] J. M. Jaffe, "Flow Control Power is Non-Decentralizable," Research Report RC 8343, IBM Thomas J. Watson Research Center, Yorktown Heights, New York (July, 1980).

[7] K. Bharath-Kumar and J. M. Jaffe, "A New Approach to Performance Oriented Flow Control," Research Report RC 8307(#36055), IBM Thomas J. Watson Research Center, Yorktown Heights, New York (May, 1980).

[8] M. Gerla and L. Kleinrock, "Flow Control: A Comparative Survey," *IEEE Transactions on Communications* **COM-28**(4), pp.553-574 (April 1980).

[9] L. Kleinrock, *Queueing Systems, Vol I., Theory,* Wiley-Interscience, New York (1975).

[10] L. Kleinrock, *Queueing Systems, Vol II., Computer Applications,* Wiley-Interscience, New York (1976).

[11] J. Little, "A Proof of the Queueing Formula $L = \lambda W$," *Operations Research* **9**(2), pp.383-387 (March 1961).

[12] L. Kleinrock, *Communication Nets, Stochastic Message Flow and Delay,* McGraw-Hill, New York (1964). Out of print. Reprinted by Dover Publications, 1972.

[13] I. Rubin, "Communication Networks: Message Path Delays," *IEEE Transactions on Information Theory* **IT-20**(6), pp.738-745 (November 1974).

[14] L. Kleinrock and R. Gail, *An Analysis of Power for Simple Computer Network Configurations,* Computer Science Department, University of California, Los Angeles (March 1981).

[15] J. Grange and M. Gien, *Flow Control in Computer Networks,* IFIP, North-Holland Publishing Company, Amsterdam (1979).