THE UNIVERSITY OF MICHIGAN

COMPUTING RESEARCH LABORATORY[1]

A QUEUEING MODEL OF
DELTA NETWORKS

B. A. Makrucki
T. N. Mudge

CRL-TR-26-83

AUGUST 1983

Room 1079, East Engineering Building
Ann Arbor, Michigan 48109
USA
Tel: (313) 763-8000

A Queueing Model of Delta Networks

B. A. Makrucki

T. N. Mudge

January 1982

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING
SYSTEMS ENGINEERING LABORATORY
THE UNIVERSITY OF MICHIGAN, ANN ARBOR   48109

ABSTRACT

This report describes an analysis of multistage inter-
connection networks where queues are placed in the
$b \times b$ crossbar switches on which the networks are
based. A queueing analysis of the network is
presented, and results are obtained using approxima-
tions that are appropriate for network operation
parameters of primary interest. From the analysis
communication delay time and network throughput
are derived. Using the results obtained, queue
lengths may be chosen so that the network satisfies
certain performance requirements.

1. Introduction.

In high performance multiprocessor computer systems, a high capacity communication channel is typically required if processor communication/data movement is not to be a system bottleneck. Interconnection networks have been developed to provide such a high capacity channel [Clo53, Ben65, GoL73, Law75, Pat79]. Operation performance of these networks is a critical measurement of their usefulness (for example, one type of network may yield totally unacceptable performance under certain operating conditions). This report describes an analysis of a class of packet switched multistage interconnection networks appropriate for use in multiple instruction, multiple data stream (MIMD) multiprocessor systems.

Section 2 describes the multiprocessor model assumptions and the performance measures to be derived (several more are embedded in the analysis). Section 3 describes the queueing analysis (and approximations) of the model and derives expressions for the performance measures developed in section 2. Section 4 is the conclusion.

2. System Model, Operation, and Assumptions.

This section describes the multiprocessor system to be modeled. System operation in the context of the model is described and simplifying system operation assumptions are described.

The multiprocessor architecture is shown in Figure 1. The system consists of processing elements (PE's) connected by an interconnection network. The model described may also be used for analysis of processor-to-memory module connected systems if slight extensions to the model are made.

The system to be modeled is a packet switched system, in which n PE's ($n = b^z, b = 2^k$, k an integer, where z is the number of stages in the multistage interconnection network), or *packet sources*, emit communication packets of constant length. A packet consists of a destination address field and a data field For example, a packet may be a memory word or several memory
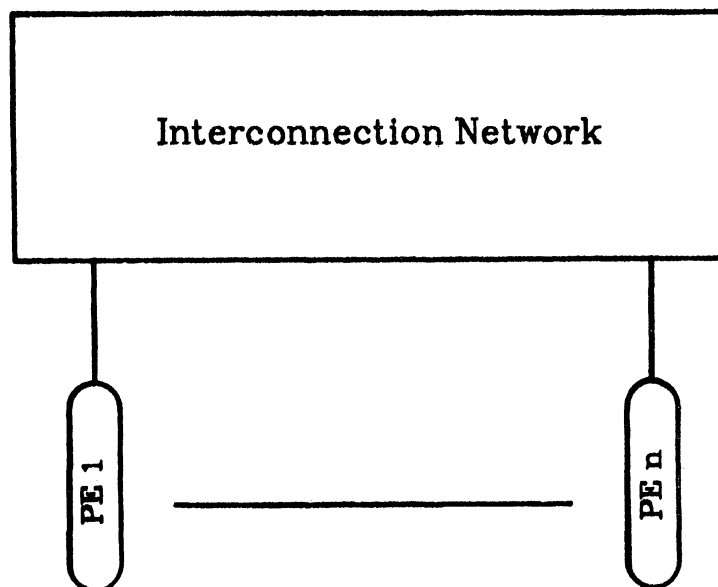

Figure 1. System Architecture.

words. When a source emits a packet, the packet is destined for one of the n *sink* devices (PE's here), i.e., packet broadcasting is not modeled here.

There are two network performance measures to be found, they are measures of communication delay time and network throughput. They are:

(1) *Packet delay time* (*PDT*) - this is the *average* time required for a packet to reach its destination. *PDT* is the average delay encountered by a packet from the time of emission from its source to the time of arrival at its destination.

(2) *Network throughput* (*NTP*) - this is the *average* rate of packet flow out of the network at the destination side.

The interconnection network for this analysis is assumed to have the unique transmission path property. That is, for any transmission path required there exists only a unique set of switch settings that will yield the transmission path; in other words, there is only one choice of a transmission path for every route (from PE to PE) required. Multistage networks which satisfy this property are those that are bit controlled, i.e., where each stage in the network determines its switch settings by using a unique bit field from the destination address tag. Furthermore, these fields are mutually disjoint. Multistage networks which exhibit this property include Omega, Delta and Generalized Cube Networks. The requirement of unique transmission paths will be seen later in the network analysis.

2.1. Assumptions.

(1) All processors behave independently. In particular this means interprocessor data dependencies are not modeled here.

(2) Each processor emits packets as a Poisson process with rate λ. Thus program behavior is modeled as follows: a program executing on processor i emits packets at random times, the average time between packet emissions is $\frac{1}{\lambda}$. This is simply a continuous time extension of discrete time processor models [SkA69, Str70, Bha75, BaS76, Hoo77, SeD79, Pat79, Rau79, MaM81a] where the packet emission process is a Bernoulli process. In discrete time processor models, it is assumed that processors emit packets during system cycles with probability p (which may be viewed as the fraction of instructions that are communication instructions, such fractions might be characterized by Gibson mix relative frequencies). The continuous time model is simply an extension of this idea, in fact it is the process that arises if the system cycle time goes to zero (with an appropriate adjustment in p) in the discrete time model.

(3) When a program emits a packet, the selection of a destination sink is assumed to be uniformly distributed over all sink devices. This is an approximation, since a processor will not emit packets destined for itself. Removing this approximation makes analysis unduly complex without gaining much.

From the bit controlled property, and the uniform distribution assumption, it may be seen that for every packet traversing the network (at stage i), its position at the next stage is distributed uniformly over all outputs of the next stage switch to which it is routed (each bit of the address field is either 0 or 1 with equal probability). Thus packet routing at any stage is *independent* of all other stage routings. This allows a stage-by-stage queueing analysis to be done. Figure 2 shows the queueing network configuration of a $b \times b$ crossbar switch on which the multistage networks considered here are based.
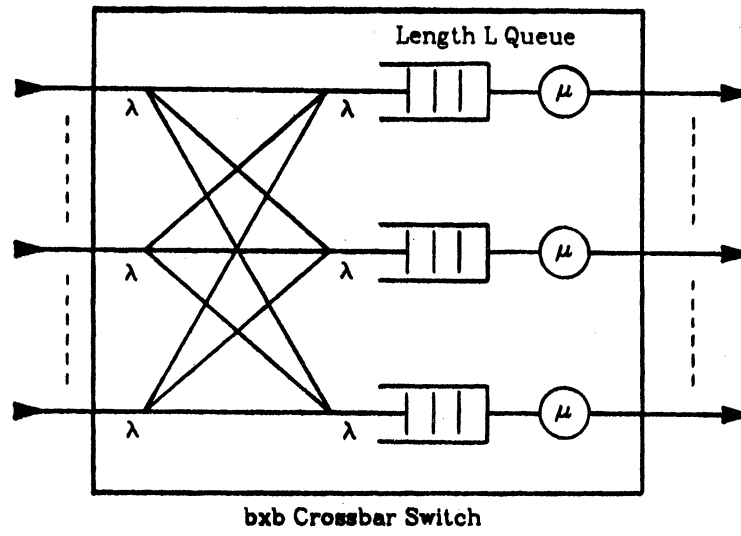
Figure 2. A *b* x*b* Switch with Queues.

3. Network Queueing Analysis.

This section analyzes interconnection network behavior to find *PDT* and *NTP*.

The analysis will be performed in a stage-wise fashion, starting with the source end of the network. The first stage will be analyzed (note that *all* queues in a given stage behave similarly due to symmetry in processor emission rates and uniform destination distributions) to find queue behavior. The results of the first stage may then be used to find second stage behavior. Likewise for all successive stages. Approximations will be made in order to make the analysis tractable.

Again, each packet entering a $b \times b$ switch is randomly destined (with uniform distribution) for one of the b output queues. All inputs to all $b \times b$ switches in the first stage are Poisson processes with rate λ. Thus, by decomposition and superposition of Poisson processes, each queue in the first stage sees a Poisson process with rate $\sum_{i=1}^{b} \frac{\lambda}{b} = \lambda$ at its input. Figure 2 shows the situation.

The exponential servers model randomness associated with the time required to move a packet from one stage to the next. Synchronous queue servers take a non-zero amount of time to move a packet to the next stage. Thus, multiple packets may try to enter a queue simultaneously (in a synchronous design), some of which are delayed. The use of an exponential server is an attempt to model this interstage data movement delay, without unduly complicating the analysis.

All queues in the first stage behave identically and independently; hence, it suffices to analyze a single queue of this stage.

3.1. First Stage, Single Queue Analysis.

For an $M/M/1/L$ queue, results from queueing theory are available [Coo72, GrH74, Kle76]. Let,

$$p_k = Pr\{ \text{k packets in the queueing station}\} \qquad 0 \leq k \leq L$$

be the steady-state, general-time, occupancy probabilities for a queueing station. Then,

$$p_k = \begin{cases} \left[\dfrac{1-\rho}{1-\rho^{L+1}}\right]\rho^k & 0 \leq k \leq L, \quad \rho \neq 1, \quad M/M/1/L \qquad (1.1) \\[2ex] \dfrac{1}{L+1} & 0 \leq k \leq L, \quad \rho = 1, \quad M/M/1/L \qquad (1.2) \end{cases}$$

Where $\rho = \dfrac{\lambda}{\mu}$.

Let N be the random variable representing the steady-state number of packets in a queueing station (N will be subscripted with a queue number later).

$$\begin{aligned} E[N] &= \sum_{k=1}^{L} k p_k \\ &= \frac{\rho[1 - (L+1)\rho^L + L\rho^{L+1}]}{(1-\rho)(1-\rho^{L+1})} \qquad \rho \neq 1, \quad M/M/1/L \qquad (2) \\ &= \frac{L}{2} \qquad \rho = 1, \quad M/M/1/L \end{aligned}$$

Thus the first stage queue is easily analyzed. For successive stages the situation is generally more complicated.

Consider the interdeparture process at the output of an $M/M/1/*$ (* denotes any queue length) queueing station with Poisson input rate λ and Poisson service rate μ, its probability density is found as follows [Kle76]:

Let E be the event that the queue is empty after a departure and let T be the interdeparture time random variable. Let $f_T(t)$ be the probability density of T, $f_{T|E}(t)$ be the probability density of T given that E occurred, and $f_{T|\sim E}(t)$

be the probability density of T given that E did not occur. Then,

$$f_T(t) = f_{T|E}(t)Pr\{E\} + f_{T|\sim E}(t)(1 - Pr\{E\})$$
$$= f_{T|E}(t)p_0 + f_{T|\sim E}(t)(1 - p_0)$$

And,

$f_{T|\sim E}(t)$ = probability density of the server

$f_{T|\sim E}(t) = \mu e^{-\mu t}$

$f_{T|E}(t)$ = probability density of service time plus arrival time

Taking Laplace transforms,

$$f_T{}^*(s) = p_0\left[\frac{\lambda}{s + \lambda}\right]\left[\frac{\mu}{s + \mu}\right] + (1 - p_0)\left[\frac{\mu}{s + \mu}\right]$$

So

$$f_T(t) = \begin{cases} \dfrac{1 - \rho - p_0}{1 - \rho}\, \mu e^{-\mu t} + \dfrac{p_0}{1 - \rho}\, \lambda e^{-\lambda t} & \rho \neq 1 \qquad (3.1) \\[2ex] \dfrac{1}{L+1}\, \mu^2 t e^{-\mu t} + \dfrac{L}{L+1}\, \mu e^{-\mu t} & \rho = 1 \qquad (3.2) \end{cases}$$

Note that here instead of using π_0, the departure point probability of an empty queue, the asymptotic limit (as $L \to \infty$), i.e., the general time probability, is used. This makes the density above approximate but allows bounds on the approximation regions (see later) to be derived in a simple manner. Note that the inter-departure process is a renewal process with a *non-memoryless* interdeparture time distribution. From this it may be seen that processes input to second stage queues are not renewal processes [Cin75] unless (to an approximation): $p_0 \approx 1 - \rho$, $p_0 \approx 0$, or $\rho = 1$ with $\dfrac{L}{L+1} \approx 1$. If processes input to queueing stations are not renewal processes, the analysis is very complex and simulation may be a better approach for obtaining accurate results.

The three cases $p_0 \approx 1 - \rho$, $p_0 \approx 0$, and $\lambda = \mu$ correspond to the following situations:

Case I

$p_0 \approx 1 - \rho$ corresponds to a *lightly loaded* $M/M/1/L$ queueing station with an output process that is close to a Poisson process with rate λ. Note that $p_0 \approx 1 - \rho$ is also an approximation to an $M/M/1/\infty$ queue, which has $p_0 = 1 - \rho$. Since $\frac{1}{\mu}$ is the average time required for a packet transfer between stages and $\frac{1}{\lambda}$ is the average time between packet emissions (i.e., message emissions) from processors, $\mu \gg \lambda$ in a typical multiprocessor system running in an independent processor, MIMD mode. μ models a high rate of operation while λ models a low rate of operation (relative to μ). Thus the lightly loaded situation is the one expected to be applicable in most systems.

Case II

$p_0 \approx 0$ corresponds to a *saturated* $M/M/1/L$ queueing station, here the output process is close to a Poisson process with rate μ because the queue is empty with low probability.

Case III

$\rho = 1$ ($\lambda = \mu$) corresponds to an $M/M/1/L$ queue where the input and output rates are about the same. In this case, the second expression (3.2) may be used to see that for L large (to be described later) the second term dominates the first so the output process is approximately Poisson with rate μ.

Notice that in these three cases, first stage output processes are approximately Poisson so the second stage has approximate Poisson inputs.

Obviously, there is some inaccuracy involved unless $p_0 = 1 - \rho$, $p_0 = 0$, or $\rho = 1$ with $L \to \infty$. The regions of validity for the approximations may be

characterized in terms of allowable values for p_0, ρ, and L.

First consider the region for which the lightly loaded $M/M/1/L$ approximation applies. Since an $M/M/1/\infty$ queue has a Poisson interdeparture process with rate λ, a simple bound on the difference between $p_{0_{M/M/1/\infty}}$ and $p_{0_{M/M/1/L}}$ will suffice (only p_0 affects $f_T(t)$ for both $M/M/1/L$ and $M/M/1/\infty$ queues when $\rho \neq 1$) for a bound on error. Define

$$D = \begin{cases} \text{the maximum allowable difference between} \\ p_{0_{M/M/1/\infty}} \text{ and } p_{0_{M/M/1/L}} \text{ in relative} \\ \text{error with respect to } p_{0_{M/M/1/\infty}}. \end{cases}$$

Then

$$D \geq \frac{p_{0_{M/M/1/L}} - p_{0_{M/M/1/\infty}}}{p_{0_{M/M/1/\infty}}} \qquad (4)$$

And

$$D \geq \frac{\dfrac{1-\rho}{1-\rho^{L+1}} - (1-\rho)}{1-\rho}$$

That is, $\rho \leq \left[\dfrac{D}{D+1}\right]^{\frac{1}{L+1}}$ will satisfy (4). Thus selecting D and L places a bound on the region of validity (with respect to D) for the light loading approximation.

Next consider the region of the saturation approximation validity. Define

$$p_{0_{max}} = \begin{cases} \text{maximum allowable value for} \\ p_0, \text{ i.e., it is the maximum} \\ \text{non-zero value deemed allowable.} \end{cases}$$

Then,

$$p_{0_{max}} \geq \frac{1-\rho}{1-\rho^{L+1}}$$

So for all $\rho \geq \rho_0$ the saturation model will hold with $p_0 \leq p_{0_{max}}$. ρ_0 is the solution to

$$p_{0_{max}}(1 - \rho_0^{L+1}) + \rho_0 = 1$$

This establishes a space of saturation approximation validity where L and $p_{0_{max}}$ are the parameters.

Finally consider the $\rho \approx 1$ case. Here it is desirable to require an approximate Poisson output process because a renewal process is needed at the next queue input. This approximation is satisfied as follows: Let

$$C = \text{minimum coefficient of the } \mu e^{-\mu t} \text{ term in (3.2)}.$$

Then $C \leq \dfrac{L}{L+1}$ for the $\rho \approx 1$ approximation to be accurate with respect to the chosen C.

Figure 3 shows the regions of analysis validity for $p_{0_{max}} = D = .05$, $C = .95$. As can be seen from the graph $L \geq 30 \Rightarrow$ the analysis is relatively accurate for almost all ρ.

3.2. Network Analysis.

Using the approximations from section 3.1, the network will now be analyzed to find PDT and NTP from section 2.

Since queues are finite in length, packets will be rejected when they attempt to enter full queues. Due to this effect, packets may be lost at any stage of the network. That is, the network does not exhibit the blocking property [DiJ80]. When a packet is lost, it is routed back to its source to be resubmitted (alternatively, buffers could be placed between stages for lost packets, but this amounts to lengthening queues). Figure 4 shows the return path configuration. The submission/resubmission process resembles a Bernoulli process where the probability of event occurrence is:
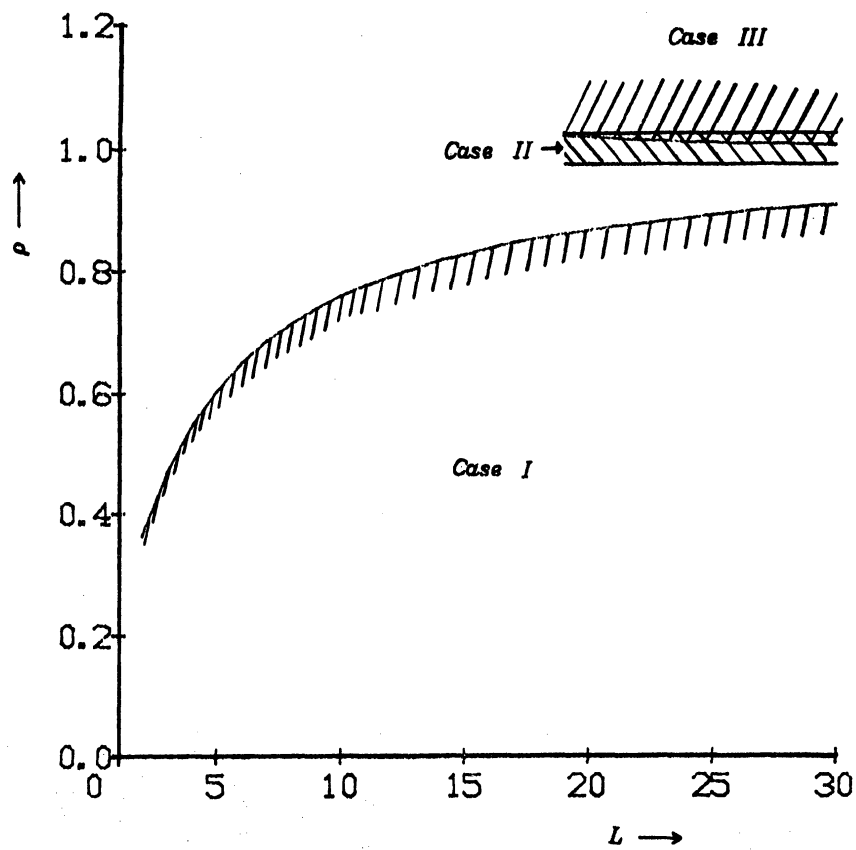
Figure 3. Approximation Validity Regions.

$p = Pr\{$ packet is not rejected at any queue $\}$.

This approximation is supported by discrete time analysis and simulation [MaM81a].

The average number of trials before the first event occurrence in a Bernoulli process is:

$$E[\text{number of trials before event occurrence}] = \frac{1 - p}{p}$$

Define the following:

$$p_{ki} = Pr\{\text{ k packets in queue } i \text{ at an arrival }\}$$

$$T_i = \begin{cases} \text{random variable representing the time spent} \\ \text{by a packet in queue i, service time} \\ \text{is included in this measurement.} \end{cases}$$
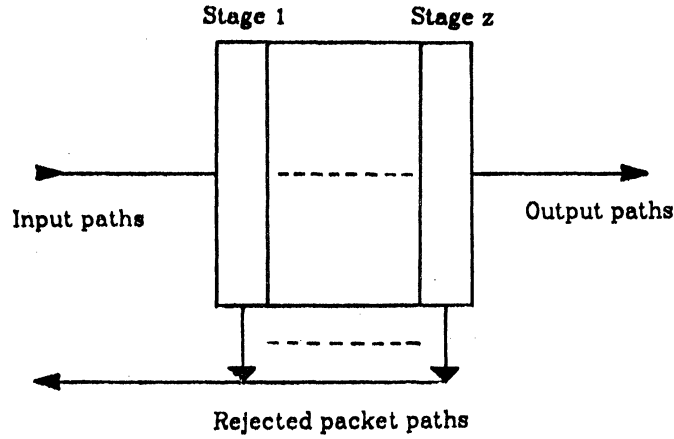
Figure 4. Resubmission Path.

$$E[T_{reject}] = \begin{cases} \text{the average time from packet emission} \\ \text{to return when the packet is lost (rejected)} \\ \text{somewhere in the network.} \end{cases}$$

Then,

$$p = \prod_{i=1}^{z} (1 - p_{Li})$$

With the Bernoulli submission/resubmission approximation

$$PDT \approx \sum_{i=1}^{z} E[T_i] + \frac{1-p}{p} E[T_{reject}] \tag{5}$$

And

$$
\begin{aligned}
E[T_{reject}] &= (1 - p_{L1})p_{L2}E[T_1] \\
&\quad + (1 - p_{L1})(1 - p_{L2})p_{L3}(E[T_1] + E[T_2]) \\
&\quad + \cdots \\
&\quad + (1 - p_{L1})(1 - p_{L2}) \cdots (1 - p_{L\,z-1})p_{Lz}(E[T_1] + \cdots + E[T_{z-1}]) \\
&= \sum_{i=2}^{z} \left[\sum_{j=1}^{i-1} E[T_j] \right] p_{Li} \prod_{j=1}^{i-1} (1 - p_{Lj}) \qquad z \geq 2
\end{aligned}
\tag{6}
$$

Here resubmission processing by sources is assumed to be instantaneous. See

[MaM81b] for a similar approximation. The Bernoulli resubmission process approximation is expected to hold when:

(1) p is small enough that p_{Li}'s are small (this is case I)

(2) and p_{Li}'s are the same for all packet submissions and resubmissions.

Case I Analysis.

For the light loading approximation Figure 5.a shows a series of z queues that represents a transmission path through the network. This equivalent series of queues relies on the equal rate, uniform destination distribution, and unique connection path assumptions. The analysis may be extended (by considering Poisson flow rates) to accommodate nonequal, general destination distributions but the analysis is then specific to the network topology, network rates and distributions considered.

From Little's formula

$$E[T_i] = \frac{E[N_i]}{\lambda(1 - p_{Li})} \qquad i = 1, 2, \cdots, z.$$

From (1.1)

$$p_L = p_{Li} = \frac{1 - \rho}{1 - \rho^{L+1}} \rho^L \qquad i = 1, 2, \cdots, z.$$

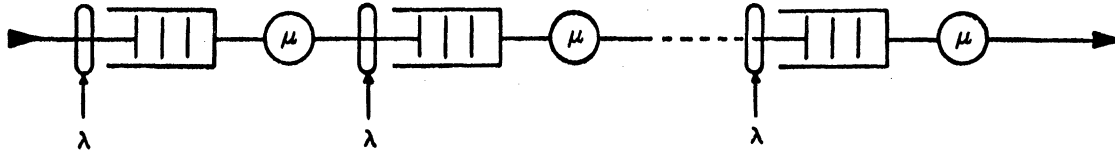$$1 - p_L = 1 - p_{Li} = \frac{1 - \rho^L}{1 - \rho^{L+1}} \qquad i = 1, 2, \cdots, z.$$

So from (2)

$$E[T] = E[T_i] = \frac{1 - (L+1)\rho^L + L\rho^{L+1}}{\mu(1 - \rho)(1 - \rho^L)} \qquad i = 1, 2, \cdots, z.$$

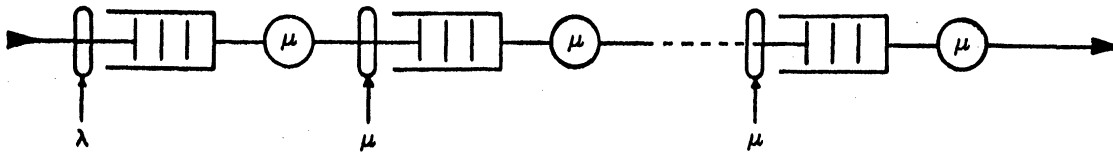And,

$$p = \prod_{i=1}^{z}(1 - p_{Li}) = \left[\frac{1 - \rho^L}{1 - \rho^{L+1}}\right]^z$$

From (6)

(5a)

(5b)

Figure 5. Equivalent Queue Series'.

$$E[T_{reject}] = p_L \sum_{i=2}^{z} (i - 1)E[T](1 - p_L)^{i-1}$$

$$= p_L E[T] \sum_{k=1}^{z-1} k(1 - p_L)^k$$

$$= E[T](1 - \beta)\beta \frac{1 - z\beta^{z-1} + (z-1)\beta^z}{(1 - \beta)^2}$$

$$= \frac{\beta}{1 - \beta}(1 - z\beta^{z-1} + (z-1)\beta^z)E[T] \qquad z \geq 2$$

(CI.1)

Where,

$$\beta = 1 - p_L = \frac{1 - \rho^L}{1 - \rho^{L+1}}$$

$$p = \beta^z$$

Therefore,

$$PDT = \left[z + \frac{1 - \beta^z}{(1 - \beta)\beta^{z-1}} \left(1 - z\beta^{z-1} + (z-1)\beta^z\right)\right] E[T] \qquad (CI.2)$$

To find NTP, simply subtract the rate of network packet loss from $n\lambda$, the packet input rate.

$$NTP = n\lambda - \sum_{i=1}^{z} n\lambda p_{Li}$$

$$= n\lambda \left[\frac{1 + (z-1)\rho^{L+1} - z\rho^L}{1 - \rho^{L+1}} \right]$$

PDT and NTP for $n = 64$, $b = 4$, verses ρ for several L are shown in Figure 6. Note that as $L \to \infty$, $NTP \to n\lambda$ as would be expected because no packets are lost when L is very large Note also that as $L \to \infty$ (or L gets large) the cost of the network becomes large for both delta networks and crossbars (i.e., the largest cost factor of the network becomes the queues) so it is actually less costly to use a crossbar because only n queues are required whereas for multistage networks $n\log_b n$ queues are needed. With a single stage, PDT is considerably better also. For a single stage, analysis is generally much more flexible [MaM81b].

Case II Analysis.

From Figure 5.b, the equivalent approximate series of queues.

$$E[T_1] = \frac{E[N_1]}{\lambda(1 - p_{L1})}$$

$$= \frac{1 - (L+1)\rho^L + L\rho^{L+1}}{\mu(1 - \rho)(1 - \rho^L)}$$

$$E[N_i] = \frac{L}{2} \qquad i = 2, 3, \cdots, z$$

$$E[T_i] = \frac{E[N_i]}{\mu(1 - p_{Li})}$$
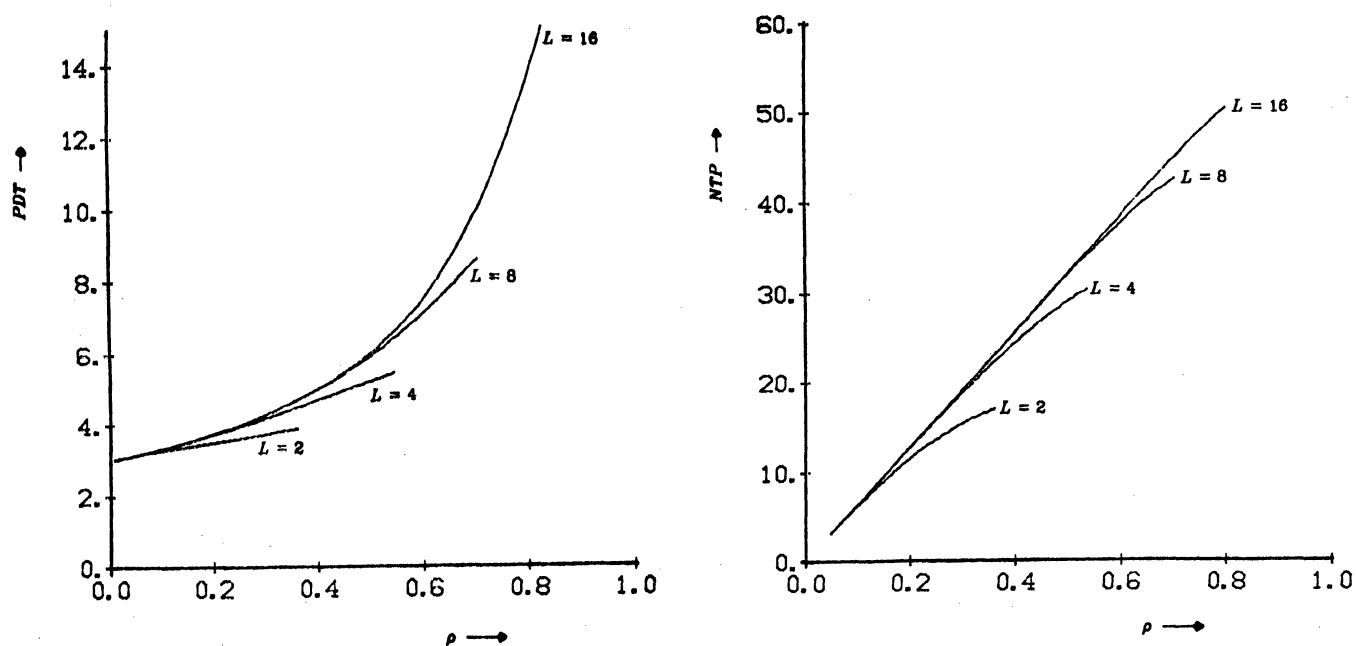
$$= \frac{L + 1}{2\mu} \qquad i = 2, 3, \cdots, z$$

Therefore,

Figure 6. *PDT* and *NTP* for $n = 64$, $b = 4$.

$$E[T_{reject}] = E[T_1](1 - p_{L1})(1 - (1 - p_{L2})^{z-1})$$

$$+ E[T_2](1 - p_{L1})(1 - p_{L2}) \frac{1 - (z-1)(1 - p_{L2})^{z-2} + (z-2)(1 - p_{L2})^{z-1}}{p_{L2}}$$

$$1 - p_{L1} = \frac{1 - \rho^L}{1 - \rho^{L+1}}$$

$$p = \frac{1 - \rho^L}{1 - \rho^{L+1}} \left[\frac{L}{L + 1} \right]^{z-1}$$

Finally, from (5)

$$PDT = E[T_1] + (z - 1)E[T_2] + \frac{1 - p}{p} E[T_{reject}]$$

Again, find NTP from

$$NTP = n\lambda - n\lambda p_{L1} - \sum_{i=2}^{z} n\mu p_{Li}$$

$$= n\lambda \left[\frac{1 - \rho^{L}}{1 - \rho^{L+1}} \right] - n\mu \left[\frac{z - 1}{L + 1} \right]$$

Provided $NTP \ge 0$ (incorrect use of the approximation can result in $NTP < 0$, in which case $NTP \approx 0$ in actuality). If a detailed analysis were available, this requirement would be satisfied because processes between stages would not be approximate.

Case III Analysis.

Here the equivalent series of queues is again shown by Figure 6.b with $\lambda = \mu$.

$$E[T] = E[T_i] = \frac{L + 1}{2\mu} \qquad i = 1, 2, \cdots, z.$$

$E[T_{reject}]$ may be found using (CI.1) with

$$\beta = \frac{L}{L + 1}$$

$$p = \beta^{z}$$

And PDT is given by (CI.2). Finally,

$$NTP = n\lambda \left[1 - \frac{z}{L + 1} \right]$$

Provided $0 \le 1 - \frac{z}{L + 1} \le 1$.

4. Conclusion.

An analysis of a class of packet switched, multistage interconnection networks that exhibit the bit controlled property was presented. The analysis (with approximations) allows network communication delay and network throughput to be evaluated for certain combinations of queue lengths, interstage transfer rates, and processor packet emissions rates. From the analysis, queue lengths that satisfy performance requirements may be chosen.

5. References.

[BaS76]

F. Baskett, and A. J. Smith, "Interference in Multiprocessor Computer Systems with Interleaved Memory," *CACM*, Vol. 19, No. 6, June 1976, pp. 327-334.

[Ben65]

V. E. Benes, *Mathematical Theory of Connecting Networks and Telephone Traffic*, Academic Press, New York, 1965.

[Bha75]

D. P. Bhandarkar, "Analysis of Memory Interference in Multiprocessors," *IEEE TC*, Vol. C-24, No. 9, Sept. 1975, pp. 897-908.

[Cin75]

E. Cinlar, *Introduction to Stochastic Processes*, Prentice-Hall Inc., Englewood Cliffs, N.J., 1975.

[Clo53]

C. Clos, "A Study of Non-blocking Switching Networks," *The Bell System Technical Journal*, Vol. 32, March 1953, pp. 406-424.

[Coo72]

R. B. Cooper, *Introduction to Queueing Theory*, The Macmillan Company, New York, 1972.

[DiJ80]

D. M. Dias, and J. R. Jump, "Analysis and Simulation of Buffered Delta Networks," *Proc. Workshop on Interconnection Networks*, Purdue University, April 21-22, 1980.

[GrH74]

D. Gross, and C. M. Harris, *Fundamentals of Queueing Theory*, John Wiley and Sons Inc., New York, 1974.

[Hoo77]

C. H. Hoogendorn, "A General Model for Memory Interference in Multiprocessors," *IEEE TC*, Vol. C-26, No. 10, Oct. 1977, pp. 998-1005.

[GoL73]

G. R. Goke, and G. J. Lipovski, "Banyan Networks for Partitioning Multiprocessor Systems," *Proc. First Annual Symp. on Computer Architecture*, IEEE, Dec. 1973, pp. 21-28.

[Kle75]

L. Kleinrock, *Queueing Systems Volume I: Theory*, John Wiley & Sons Inc., New York, 1975.

[Law75]

D. H. Lawrie, "Access and Alignment of Data in an Array Processor," *IEEE*

TC, Vol. C-24, No. 18, Dec. 1975, pp. 1145-1155.

[MaG81]
M. A. Marsan, and M. Gerla, *Markov Models for Multiple Bus Multiprocessor Systems*, Report No. CSD 810304, Computer Science Department, UCLA, Feb. 1981.

[MaM81a]
B. A. Makrucki, and T. N. Mudge, *Probabilistic Analysis of a Crossbar Switch*, SEL Report No. 150, Department of Electrical and Computer Engineering, University of Michigan, March 1981.

[MaM81b]
B. A. Makrucki, and T. N. Mudge, *A Multiple $M/D_z/1/L$ Queueing Model of Crossbar-based Multiprocessors*, SEL Report No. 157, Department of Electrical and Computer Engineering, University of Michigan, September 1981.

[Pat79]
J. H. Patel, "Processor-Memory Interconnections for Multiprocessors," *Proc. 6th Annual Symp. on Computer Architecture*, IEEE, April 1979, pp. 166-177.

[Rau79]
B. R. Rau, "Interleaved Memory Bandwidth in a Model of a Multiprocessor Computer System," *IEEE TC*, Vol. C-28, No. 9, Sept. 1979, pp. 678-681.

[SeD79]
A. S. Sethi, and N. Deo, "Interference in Multiprocessor Systems with Localized Memory Access Probabilities," *IEEE TC*, Vol. C-28, No. 2, Feb. 1979, pp. 157-163.

[Sie77]
H. J. Siegel, "Analysis Techniques for SIMD Machine Interconnection Networks and the Effects of Processor Address Masks," *IEEE TC*, Vol. C-26, No. 2, Feb. 1977, pp. 153-161.

[Sie80]
H. J. Siegel, (Ed.), *Proc. Workshop on Interconnection Networks*, Purdue University, April 21-22, 1980.

[SkA69]
C. E. Skinner, and J. R. Asher, "Effects of storage contention on system performance," *IBM Systems Journal*, No. 4, 1969, pp. 319-333.

[Str70]
W. D. Strecker, *Analysis of the Instruction Execution Rate in Certain Computer Structures*, Ph.D. dissertation, Carnegie-Mellon University, Pittsburgh, 1970.

Note on a Queueing Model of Delta Networks

B. A. Makrucki

T. N. Mudge

July 1982

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING
SYSTEMS ENGINEERING LABORATORY
THE UNIVERSITY OF MICHIGAN, ANN ARBOR   48109

This is intended as an addendum to *A Queueing Model of Delta Networks (SEL Report No. 159)*. where it was assumed that packets rejected from multistage interconnection networks (ICN's) are sent back to their source for resubmission. This approach "penalizes" packets lost in the latter stages of the network more severly and may be misleading in overestimating actual *PDT* (packet delay time) for ICN's which do exhibit blocking. This addendum describes a modification which may be more realistic in predicting *PDT* when blocking occurs.

As an approximation to the rejection/blocking phenomenon assume that a packet lost in state i ($2 \le i \le z$) is sent back to the input of its stage $i - 1$ queue. This should be a feasible approximation because a packet leaving stage $i - 1$ leaves an open position in its queue, and if it is rejected at the input of stage i, it may return to stage $i - 1$ where its position will still be available with high probability. If the rejection is practically instantaneous, then by the approximation of Poisson processes, another packet will arrive at the same time with probability 0 (or close to it). Hence the assumption of a single stage rejection delay seems reasonable.

Revising the Bernoulli resubmission approximation of equation (5) of SEL 159:

$$PDT \approx \sum_{i=1}^{z} E[T_i] + \sum_{i=2}^{z} \left(\frac{p_{Li}}{1 - p_{Li}} \right) E[T_{i-1}].$$

$\dfrac{p_{Li}}{1 - p_{Li}}$ is the average number of times that a packet must try to enter a queue in stage i before it is accepted. Each rejection takes $E[T_{i-1}]$ time units as its mean time for retry.

Or,

$$PDT \approx E[T_z] + \sum_{i=2}^{z}\left(1 + \frac{p_{Li}}{1 - p_{Li}}\right)E[T_{i-1}]$$

$$= E[T_z] + \sum_{i=2}^{z}\frac{E[T_{i-1}]}{1 - p_{Li}}.$$

Which leads to:

Case I

$$E[T_i] = E[T] \qquad i = 1,\dots, z$$

$$p_L = p_{Li} = \frac{1 - \rho}{1 - \rho^{L+1}}\rho^L \qquad i = 1,\dots, z$$

So

$$PDT = E[T] + \left[\frac{1 - \rho^{L+1}}{1 - \rho^L}\right](z - 1)E[T]$$

$$= \left[1 + (z - 1)\left[\frac{1 - \rho^{L+1}}{1 - \rho^L}\right]\right]E[T]$$

Case II

$$PDT = E[T_2] + \frac{E[T_1]}{1 - p_{L2}} + \sum_{i=3}^{z}\frac{E[T_2]}{1 - p_{L2}}$$

$$= \frac{E[T_1]}{1 - p_{L2}} + \left[\frac{z - 2}{1 - p_{L2}} + 1\right]E[T_2]$$

Case III is handled in a similar manner as discussed in SEL 159.

This technique for approximating network blocking seems to be a more reasonable technique than previously presented.

Note on SEL 159