PERFORMANCE ANALYSIS OF CLOCK-REGULATED QUEUES WITH OUTPUT MULTIPLEXING IN THREE DIFFERENT 2 BY 2 CROSSBAR SWITCH ARCHITECTURES

Ora E. Percus and Susan R. Dickey

Courant Institute of Mathematical Sciences

New York University

251 Mercer Street

New York, NY 10012

Ultracomputer Note #169

April, 1990

Revised September, 1991

This work was supported in part by the Applied Mathematical Sciences Program of the U.S. Department of Energy under contract DE-FG02-88ER025052, and in part under contract DARPA F 49620-87-C-0065..

Performance Analysis of 2 by 2 Crossbar Switch Architectures

Susan R. Dickey
Courant Institute of Mathematical Sciences
New York University
251 Mercer Street
New York, NY 10012
(212) 998 -3354

Abstract

Switches in interconnection networks for highly parallel shared memory computer systems may be implemented with different internal buffer structures. For a 2×2 synchronous switch, previous studies have often assumed a switch composed of two queues, one at each output, each of which has unbounded size and may accept two inputs every clock cycle. Hardware implementations may actually use simpler queue designs and will have bounded size.

Two additional models for a 2×2 switch using queues that may accept only one input at a time are analyzed. The first uses four queues, one for each input/output pair. The second uses two queues, one at each input. In each case, a multiplexer blocks one queue if two queues desire the same output, making these models more difficult to analyze than the previous model. Maximum bandwidth, expected queue length, expected waiting time, and queue length distribution are presented for both models, with unbounded queue size and with queue size equal to 1.

Keywords: Interconnection Network, Clock-regulated Queues, Crossbar Switch

1. INTRODUCTION

Overall system performance for highly parallel shared memory computer systems critically depends on the message throughput that can be achieved by the interconnection network. The bandwidth of the network, the average number of messages per cycle that can be output by the network under maximum load, is an upper bound on the throughput. In practice, network latency, the time from the generation of a request message by a processor until a response message is received from memory, can affect message throughput, since processors will not generate requests indefinitely without receiving a response. The design of the interconnection network and its switching components thus requires careful consideration to achieve maximum bandwidth with minimum latency.

Multistage networks with the number of stages logarithmic in the number of processors provide a scalable interconnection scheme of reasonable cost and have been considered for many parallel computer architectures. The NYU Ultracomputer architecture, in particular, connects N independent processing elements to $M = N$ shared memory modules through a multi-stage interconnection network with the topology of an *omega* network [11] composed of 2×2 buffered crossbars (see Figure 1). Buffered switching components are used to avoid the bandwidth limitations of unbuffered networks [8]. Our analyses of switch design grew out of a consideration of the options for hardware implementation of the 2×2 buffered crossbar switches in the NYU Ultracomputer.

Different possible arrangements of the queues in a hardware implementation of the crossbar switch are illustrated in Figure 2. As discussed in [4], any 2×2 buffered crossbar switch used in a multi-stage interconnection network can be classified as one of these three types. The first design, Type A, contains two queues, each queue capable of accepting two messages simultaneously. This alternative has received the most analysis in previous work [9, 13] but has the disadvantage of being more difficult to implement than the designs we analyze in this paper. The capability of inserting two messages into a queue every cycle adds considerable complexity and may increase the cycle time of the component [15, 16].

The two alternative switch designs we present use one-input, one-output queues, which are simpler to design and need fewer hardware resources per queue item than the two-input queues required in the Type A design. Type B has four of these simpler queues, one for each input/output pair, and is the one that has been used to implement the NYU Ultracomputer switching node [3, 5]. Type C has only two queues, one at each input, and is very attractive to hardware designers because of its simplicity. However, the bandwidth of the Type C switch is limited due to blocking when both queues have a message for the same output. This bandwidth limitation was demonstrated in simulations by Kumar and Jump [10] comparing ‘‘buffers within the switches’’ to ‘‘buffers between the switches’’ and in network simulations by the NYU Ultracomputer group [2, 4].

The motivation for the research in this paper was to prove the bandwidth limitations for Type C switches rigorously and to quantify the performance differences between the three types of switches. Designers may use this information to weigh a higher cost implementation with higher performance (Type A) against lower cost implementations with lower performance (Types B and C), depending on the bandwidth and latency requirements of their system. We analyze queues of capacity one, as well as those with unbounded capacity, in order to characterize the performance of low cost hardware implementations.

Assumptions about arrival processes and service disciplines are reasonable for a multistage interconnection network like that of the NYU Ultracomputer (though the assumptions may certainly apply to other contexts as well). Since omega networks have a unique path from each network input link to each network output link, each message is assumed to include an address that controls which output from the switch will be taken [1]. The analysis of 2×2 switches in this paper thus assumes that the output by which a message must leave a switch is fixed before it enters the switch.

A systolic queue design is used in the NYU Ultracomputer's 2×2 buffered switching components [3]. This is a synchronous design, making the continuous time arrival assumptions of classical queueing theory inappropriate. Our analysis makes a direct but nontrivial extension of queueing theory to networks with quantized time.

Throughout the analysis we make use of generating functions. Readers without background in this area will find a good introduction in Wilf [17].

2. Type A switch with two two-input buffers

Sections 2.1 and 2.2 summarize some previous results for Type A switches, as well as giving the assumptions made in our analysis. Analyzing Type A switches with queues allowed to grow without bound (‘‘infinite buffers’’), Kruskal et. al [9] first found the generating function for the steady state distribution of waiting times in the first stage of the network. Using the same assumptions, Percus and Percus [13] determined the difference equations for the exact probability distribution of the number of waiting messages at the end of the n^{th} input cycle for both finite and infinite capacity output buffers, and solved the single queue or one dimensional distribution.

The latter paper also includes results on the output distribution, showing that it does not consist of independent identically distributed random variables. It also shows that, in steady state, the lengths of the intervals between cycles with no output are independent random variables and gives their probability distributions. Jaffe [7] has found the bivariate generating function of the joint distribution of queue lengths at equilibrium and solved for the joint queue length distribution; note that the two queue lengths are highly correlated.

2.1. Analysis with infinite buffers for the Type A switch

In the Type A switch, messages are queued at each output port. If two messages destined for the same output port arrive in the same cycle, they are queued in random order. The detailed assumptions with buffers of infinite capacity are as follows:

(1) At the (first stage) input, arriving messages are treated as independent identically distributed Bernoulli variables. At each cycle, a message arrives with probability p and enters one of the two output queues with probability $\frac{1}{2}$. Hence, the probability that j messages arrive in any cycle at a given queue is

$$f_j = \begin{bmatrix} 2 \\ j \end{bmatrix} \left(\frac{p}{2}\right)^j \left(1 - \frac{p}{2}\right)^{2-j} \quad j = 0,1,2.$$

(2) Every message requires one time unit (cycle) to be served.

(3) If S_n is the number of waiting messages at the end of the n^{th} cycle, and a_n is the number of messages arriving at the beginning of the n^{th} cycle, then

$$S_0 = 0 \text{ and}$$

$$S_n = \max[0, S_{n-1} + a_n - 1] \quad \text{for } n \geq 1,$$

i.e. a single message leaves during the n^{th} cycle if either at least one arrived or the queue was non-empty.

Let $p_n(j)$ represent $P(S_n = j)$, the probability of having j messages in the queue at the n^{th} cycle. Under the above assumptions

$$p_0(j) = \delta_{j,0,}$$

$$p_n(0) = f_0 p_{n-1}(1) + (f_0 + f_1) p_{n-1}(0) \quad \text{for } n > 0, \text{ and}$$

$$p_n(j) = f_0 p_{n-1}(j+1) + f_1 p_{n-1}(j) + f_2 p_{n-1}(j-1) \quad \text{for } n > 0,\ j \geq 1.$$

Percus and Percus [13] introduces the generating function

$$F(x,\lambda) = \sum_{n=0}^{\infty} \sum_{j=0}^{\infty} p_n(j) x^j \lambda^n$$

and evaluates it as

$$F(x,\lambda) = x - \frac{\dfrac{1-x}{1-x_0(\lambda)} x_0(\lambda)}{x - \lambda R(x)}$$

where

$$x_0(\lambda) = \frac{1 - \lambda p(1 - \frac{p}{2}) - [1 - 2\lambda p(1 - \frac{p}{2})]^{1/2}}{(\frac{p^2}{2}\lambda)}$$

and

$$R(x) = (1 - \frac{p}{2} + \frac{p}{2}x)^2.$$

The probability of the queue being empty is given by the expression (for $p < 1$)

$$p_n(0) = 1 - \frac{p/2}{1 - p/2} \sum_{k=1}^{n} \frac{\binom{2k}{k}}{k+1} \left(\frac{p}{2}(1 - \frac{p}{2})\right)^k \quad \text{for } n \geq 1.$$

The standard statistical parameters can readily be found. For example, the mean number of messages in the queue at the end of the n^{th} cycle is

$$E(S_n) = \sum_{k=0}^{n-1} (1 - \frac{p}{2})^2 \, p_k(0) - n(1 - p),$$

where $p_k(0)$ is given above.

The steady state limiting distribution $p(j) \equiv \lim_{n \to \infty} p_n(j)$ is then:

$$p(j) = \frac{(1 - p)}{(1 - p/2)^2} \left(\frac{p/2}{1 - p/2} \right)^{2j}. \qquad \textbf{(2-1)}$$

In particular,

$$p(0) = (1 - p)(1 - \frac{p}{2})^{-2}. \qquad (2\text{-}2)$$

The mean queue length in steady state is

$$\lim_{n \to \infty} E(S_n) = p^2 / 4(1 - p) \qquad (2\text{-}3)$$

and the variance is

$$\lim_{n \to \infty} Var(S_n) = \frac{p^4}{16(1 - p)^2} + \frac{p^2}{4(1 - p)}.$$

The latter two results were first obtained by Kruskal et. al [9].

2.2. Analysis with finite buffers for the Type A switch

For a queue with finite capacity N, the third assumption above must be changed. Percus and Percus [13] investigated two models:

$$\textit{Model I}: S_n = \min[N, \max(0, S_{n-1} + a_n - 1)]$$

$$\textit{Model II}: S_n = \begin{cases} \max[0, S_{n-1} + a_n - 1] & \text{if } S_{n-1} \le N - 1 \\ N - 1 & \text{if } S_{n-1} = N \end{cases}$$

The first model makes the assumption that a message is lost (and presumably retransmitted) if there is no room for it in the queue. The second model corresponds to a hardware implementation in which input to the queue is halted when the queue is full.

For model I, the difference equations are

$$p_0(0) = 1,$$

and, for $n \geq 1$,

$$p_n(0) = f_0 p_{n-1}(0) + f_1 p_{n-1}(0) + f_0 p_{n-1}(1), \tag{2-4}$$

$$p_n(j) = f_0 p_{n-1}(j+1) + f_1 p_{n-1}(j) + f_2 p_{n-1}(j-1) \quad \text{for } 1 \le j \le N-1, \text{ and}$$

$$p_n(N) = f_2 p_{n-1}(N) + f_1 p_{n-1}(N) + f_2 p_{n-1}(N-1).$$

This yields the steady state distribution

$$p(j) = \frac{1-\beta}{1-\beta^{N+1}}\beta^j \text{ for } j = 0,1,\ldots, N, \text{ where } \beta = \left[\frac{p/2}{1-p/2}\right]^2. \tag{2-5}$$

It follows that

$$E[\lim_{n\to\infty} S_n] = \sum_{j=0}^{N} jp(j) = \frac{\beta}{1-\beta}[1 - \frac{(N+1)(1-\beta)\beta^N}{1-\beta^{N+1}}], \tag{2-6}$$

with a more complicated expression for $Var[\lim_{n\to\infty} S_n]$. The steady state distribution and mean for Model II are close in value to these and are also found in [13].

2.3. The output for the Type A switch, Model I with $N=2$

For comparison with minimum size hardware implementations of the other two types of switches, we are interested in a Type A switch with N = 2. This is the smallest value for N which allows both messages arriving at a cycle to be accepted if the queue is not empty on the preceding cycle, and is comparable in hardware resources to a Type B switch with 4 queues each of size one. As in equations (2-4), it is assumed that a message is lost if two arrive at a full queue on a single cycle.

To find the expected output per cycle, let us introduce a tagging variable z in the difference equations (2-4) whenever we will have an output in that cycle. Then for $N=2$ equations (2-4) take the form:

$$p_0(0) = 1,$$

and, for $n \geq 1$,

$$p_n(0) = f_0 z p_{n-1}(1) + (f_0 + f_1 z) p_{n-1}(0),$$

$$p_n(1) = f_0 z p_{n-1}(2) + f_1 z p_{n-1}(1) + f_2 z p_{n-1}(0),$$

$$p_n(2) = f_2 z p_{n-1}(2) + f_1 z p_{n-1}(2) + f_2 z p_{n-1}(1).$$

Let

$$F_n(x,z) \equiv \sum_{k=0}^{2} p_n(k) x^k,$$

and

$$F(\lambda,x,z) \equiv \sum_{n=0}^{\infty} \lambda^n F_n(x,z).$$

Then the expected output per cycle, $E_p(O)$, in steady state will be

$$E_p(O) = \lim_{\lambda \to 1} (1-\lambda)^2 \frac{\partial}{\partial z} F(\lambda,x,z) \Big|_{\substack{z=1 \\ x=1}} \tag{2-7}$$

$$= p - \frac{p^6}{64 - 128p + 112p^2 - 48p^3 + 12p^4}$$

with $\max_p E_p(O) = \frac{11}{12}$.

Note that the expected output per cycle is equal to $1 - f_0 p(0)$, that is, the probability that either a message arrives or the queue was not empty at the start of the cycle.

3. Type B switch with four one-input buffers

The Type B switch has four queues, one for each input source/output destination pair. In the following analysis of this model, $q1$ and $q2$ will refer to two queues paired at an output port (see Figure 2). The same assumptions as in section 2 are made about arrivals at the input ports of the switch, but the service assumption is that a message at the front of $q1(q2)$ will be served in one cycle if $q2(q1)$ is empty and did not receive a message; otherwise only one message will be served, selected from the two queues with equal probability.

3.1. Analysis with infinite buffers for the Type B switch

Let $p_n(k,j)$ be the probability of k messages in $q1$ and j messages in $q2$ at the n^{th} cycle. It clearly satisfies

$$p_0(k,j) = \delta_{k,0}\delta_{j,0} .$$

For $n > 0$,

$$p_n(0,0) = f_0[p_{n-1}(1,0) + p_{n-1}(0,1)] + [f_0 + f_1]p_{n-1}(0,0),$$

for $k > 0$,

$$p_n(k,0) = f_0[p_{n-1}(k+1,0) + \frac{1}{2}p_{n-1}(k,1)] + \frac{3}{4}f_1 p_{n-1}(k,0)$$

$$+ \frac{1}{4}f_1 p_{n-1}(k-1,1) + \frac{1}{2}f_2 p_{n-1}(k-1,0),$$

for $j > 0$,

$$p_n(0,j) = f_0[p_{n-1}(0,j+1) + \frac{1}{2}p_{n-1}(1,j)] + \frac{3}{4}f_1 p_{n-1}(0,j)$$

$$+ \frac{1}{4}f_1 p_{n-1}(1,j-1) + \frac{1}{2}f_2 p_{n-1}(0,j-1)$$

for $kj > 0$,

$$p_n(k,j) = \frac{1}{2}f_0[p_{n-1}(k+1,j) + p_{n-1}(k,j+1)]$$

$$+ \frac{1}{2}f_1 p_{n-1}(k,j) + \frac{1}{4}f_1[p_{n-1}(k+1,j-1) + p_{n-1}(k-1,j+1)]$$

$$+ \frac{1}{2}f2[p_{n-1}(k,j-1) + p_{n-1}(k-1,j)],$$

where f_0, f_1, and f_2 have the same value as in section 2, but here express the probability of 0, 1 or 2 messages arriving at the inputs of the paired queues $q1$ and $q2$.

The above equations are solved by introducing an appropriate generating function:

$$F(x,y,\lambda) = \sum_{n=0}^{\infty} \lambda^n \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} p_n(k,j) x^k y^j ,$$

which on evaluation satisfies the functional equation:

$$F(x,y,\lambda) \left\{ xy - \frac{\lambda}{4}(x+y) \ [2(1-\frac{p}{2})^2 + p(1-\frac{p}{2})(x+y) + \frac{p^2}{2} xy] \right\}$$

$$= xy + \frac{\lambda}{4}(y-x) \ [2(1-\frac{p}{2})^2 + p(1-\frac{p}{2})x] \ F(x,0,\lambda)$$

$$+ \frac{\lambda}{4}(x-y) \ [2(1-\frac{p}{2})^2 + p(1-\frac{p}{2})y] \ F(0,y,\lambda)$$

$$+ \lambda(1-\frac{p}{2})^2 [xy - \frac{1}{2}(x+y)] \ F(0,0,\lambda).$$

The steady state limiting distribution is of particular interest; let $p(k,j) = \lim_{n \to \infty} p_n(k,j)$. Then the steady state generating function

$$F(x,y) = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} p(k,j) x^k y^j = \lim_{\lambda \to 1} (1-\lambda) F(x,y,\lambda)$$

satisfies

$$F(x,y)\left\{xy-\frac{1}{4}(x+y)[2(1-\frac{p}{2})^2 + p(1-\frac{p}{2})(x+y)+\frac{p^2}{2}xy]\right\}$$

$$= \frac{1}{4}(y-x)[2(1-\frac{p}{2})^2 + p(1-\frac{p}{2})x]\ F(x,0)$$

$$+ \frac{1}{4}(x-y)[2(1-\frac{p}{2})^2 + p(1-\frac{p}{2})y]F(0,y) \qquad \textbf{(3-1)}$$

$$+ (1-\frac{p}{2})^2[xy-\frac{1}{2}(x+y)]F(0,0)\ ;$$

whereas the 1-dimensional generating function will be

$$F(x,1)(\frac{p}{2}x-1+\frac{p}{2}) = [(1-\frac{p}{2})^2 + \frac{p}{2}(1-\frac{p}{2})x]F(x,0)$$

$$- [(1-\frac{p}{2})^2+\frac{p}{2}(1-\frac{p}{2})]F(1,0)-(1-\frac{p}{2})^2F(0,0).$$

But since $F(1,1)=1$, this implies $F(0,0)=(1-p)(1-\frac{p}{2})^{-2}$ and hence the probability that both queues are empty in the steady state $p(0,0)=(1-p)(1-\frac{p}{2})^{-2}$ which is the same as equation (2-2) in section 2.1.

Another quantity of interest is the probability that a total of l messages are waiting in the 2 queues ($q1$ and $q2$), i.e. $\sum_{k=0}^{l} p(k,l-k)$. This can easily be obtained from the generating function (3-1) when $y=x$ and we get

$$\sum_{k=0}^{l} p(k,l-k)=(1-p)(1-\frac{p}{2})^{-2}\left[\frac{p^2}{4}(1-\frac{p}{2})^{-2}\right]^l.$$

Let $E(S)$ be the expected queue length in the steady state. Then

$$2E(S) = (1-p)(1-\frac{p}{2})^{-2} \sum_{l=0}^{\infty} l(\left[\frac{p^2}{4}(1-\frac{p}{2})^{-2}\right]^l = \frac{p^2}{4(1-p)}$$

which is the same as (2-3) in section 2.1.

Notice that the expected queue length for a queue in the Type B switch is half that for a queue in a Type A switch, and that the probability that both queues paired at an output port in a Type B switch are empty is the same as that for one queue in a Type A switch. This corresponds to the intuitive perception that a Type B switch will output a message from a given port whenever a Type A switch does, though messages may be output in a different order. Simulations show that a Type B switch has a greater variance in the waiting time.

3.2. Analysis with queue size 1 for Type B switches

Consider the Type B switches with each queue of capacity $N = 1$. With this switch structure a message is lost if the queue it arrives at already has a message and does not receive service this cycle because the paired queue either had a message or receives a message and is given priority. Let $p_n(i,j)$ be the probability of *i* messages waiting at $q1$ and *j* messages waiting at $q2$ at the n^{th} cycle. Then $p_0(0,0) = 1$ and for $n \geq 1$

$$p_n(0,0) = f_0[zp_{n-1}(1,0)+zp_{n-1}(0,1)]+(f_0+f_1 z)p_{n-1}(0,0),$$

$$p_n(1,0) = \frac{1}{2}f_0 zp_{n-1}(1,1)+\frac{3}{4}f_1 zp_{n-1}(1,0)+\frac{1}{4}f_1 zp_{n-1}(0,1)$$

$$+\frac{1}{2}f_2 zp_{n-1}(0,0)+\frac{1}{2}f_2 zp_{n-1}(1,0)+\frac{1}{4}f_1 zp_{n-1}(1,1),$$

$$p_n(0,1) = \frac{1}{2}f_0 zp_{n-1}(1,1)+\frac{3}{4}f_1 zp_{n-1}(0,1)+\frac{1}{4}f_1 zp_{n-1}(1,0)$$

$$+\frac{1}{2}f_2 zp_{n-1}(0,0)+\frac{1}{2}f_2 zp_{n-1}(0,1)+\frac{1}{4}f_1 zp_{n-1}(1,1), \text{ and}$$

$$p_n(1,1) = \frac{1}{2} f_1 z p_{n-1}(1,1) + \frac{1}{2} f_2 z [p_{n-1}(1,0) + p_{n-1}(0,1)] + f_2 z p_{n-1}(1,1),$$

where z is a tagging variable for the output.

Let

$$F_0(x,y) = 1,$$

$$F_n(x,y) = \sum_{i=0}^{1} \sum_{j=0}^{1} p_n(i,j)\, x^i y^j \text{ for } n \geq 1,$$

and let

$$F(\lambda,x,y,z) = \sum_{n=0}^{\infty} \lambda^n F_n(x,y). \tag{3-2}$$

Then

$$F(\lambda,x,y,z) = (1,x,y,xy)(I - \lambda P)^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

where

$$P = \begin{bmatrix} f_0 + f_1 z & f_0 z & f_0 z & 0 \\ \frac{f_2 z}{2} & \frac{(3f_1 + 2f_2)z}{4} & \frac{f_1 z}{4} & \frac{(2f_0 + 1f_1)z}{4} \\ \frac{f_2 z}{2} & \frac{f_1 z}{4} & \frac{(3f_1 + 2f_2)z}{4} & \frac{(2f_0 + 1f_1)z}{4} \\ 0 & \frac{f_2 z}{2} & \frac{f_2 z}{2} & \frac{(f_1 + 2f_2)z}{2} \end{bmatrix}.$$

Hence,

$$F(\lambda,x,y,z) = \frac{A(\lambda,x,y,z)}{B(\lambda,z)}, \tag{3-3}$$

where

$$A(\lambda,x,y,z)=-z^2\ [\lambda^2 p^4 xy-2\lambda^2 p^3(x+y)-4\lambda^2 p^3+12\lambda^2 p^2]$$
$$-\ z[4\lambda p^2(x+y)+12\lambda p^2-48\lambda p]-32$$

$$B(\lambda,z)\ =\ z^3\ [\lambda^3 p^5-6\lambda^3 p^4+8p^3\lambda^3]\ +$$
$$z^2[-\lambda^3 p^5+7\lambda^3 p^4-4\lambda^2 p^4+32\lambda^2 p^3-16\lambda^3 p^3+12\lambda^3 p^2-52\lambda^2 p^2]\ +$$
$$z[3\lambda^2 p^4-24\lambda^2 p^3+60\lambda^2 p^2-28\lambda p^2+80\lambda p-48\lambda^2 p]\ +$$
$$8\lambda p^2-32\lambda p+32\lambda-32.$$

For $z=1$ we get

$$F(\lambda,x,y,1)=\frac{p^4\lambda^4 xy-(2\lambda^2 p^3-4\lambda p^2)(x+y)-4\lambda^2 p^3+12\lambda^2 p^2+12\lambda p^2-48\lambda p+32}{(1-\lambda)[\lambda^2 p^4-8\lambda^2 p^3+12\lambda^2 p^2+20\lambda p^2-48\lambda p+32]}$$

and therefore

$$\lim_{\lambda\to 1}(1-\lambda)F(\lambda,x,y,1)\ =\ \frac{p^4 xy-(2p^3-4p^2)(x+y)+32-48p+24p^2-4p^3}{32-48p+32p^2-8p^3+p^4}.$$

Let

$$p(i,j)=\lim_{n\to\infty}\ p_n(i,j).$$

Then

$$p(0,0)=\frac{32-48p+24p^2-4p^3}{32-48p+32p^2-8p^3+p^4},$$

$$p(0,1)=p(1,0)=\frac{4p^2-2p^3}{32-48p+32p^2-8p^3+p^4},$$

$$p(1,1)=\frac{p^4}{32-48p+32p^2-8p^3+p^4}.$$

Notice that for $p=1,\ p(1,1)=\frac{1}{9};\quad p(0,1)=p(1,0)=\frac{2}{9};\quad p(0,0)=\frac{4}{9}$ and for $p=0,\ p(0,0)=1$. Also, the one-dimensional probability

$$p(0)=p(0,0)+p(0,1)=\frac{32-48p+28p^2-6p^3}{32-48p+32p^2-8p^3+p^4}$$

is the probability that $q1(q2)$ is empty, and

$$p(1)=p(1,0)+p(1,1)=\frac{4p^2-2p^3+p^4}{32-48p+32p^2-8p^3+p^4} \tag{3-4}$$

is the probability that $q1(q2)$ is not empty, which, since the queue capacity is 1, is also the mean queue length $E(S)$.

Notice that for $p=1$, $p(1)=\frac{1}{3}$, from which we conclude that the average length of $q1$ is less than $\frac{1}{3}$ if $p<1$.

3.3. The output of the switch with queues of size 1

$F(\lambda,x,y,z)$ as defined in (3-2) can be written in the form

$$F(\lambda,x,y,z)=\sum_{n=0}^{\infty}\lambda^n\sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{k=0}^{\infty}p_n(i,j,k)z^kx^iy^j$$

where $p_n(i,j,k)$ is the probability of i messages waiting at $q1$ at the end of the n^{th} cycle, j messages waiting at $q2$ at the end of the n^{th} cycle and k messages leaving during the n^{th} cycle.

Then

$$\frac{\partial}{\partial z}F(\lambda,1,1,z)\big|_{z=1}=\sum_{n=0}^{\infty}\lambda^n\sum_{i=0}^{1}\sum_{j=0}^{1}\sum_{k=0}^{\infty}kp_n(i,j,k)$$

and

$$E_p(O) = \lim_{\lambda\to1}(1-\lambda)^2\frac{\partial}{\partial z}F(\lambda,1,1,z)\big|_{z=1}.$$

Using (3-3) to substitute for $F(\lambda, 1, 1, z)$ we get

$$E_p(O) = p\left[1 - \frac{p^3}{p^4 - 8p^3 + 32p^2 - 48p + 32}\right] \tag{3-5}$$

Note that

$$E_p(O|p=1) = \frac{8}{9}$$

gives the maximum bandwidth that can be obtained per output port for this switch structure with queue size 1. This is lower than the maximum bandwidth of $11/12$ which can be obtained from the Type A switch with queues of size 2.

4. Type C switch with two one-input buffers

This design (see Figure 2) is the simplest hardware implementation, but is the most difficult of the three to analyze, since the input queues may be in conflict for either of the two output servers. Simulations and approximations to the performance of this design are given in [1, 6, 10].

In the analysis of this design, $q1$ and $q2$ refer to the two queues in the switch, each with its own input port and each connected to both output ports. Since each message is predestined for a given port by its address, one may think of them as colored red and blue, with one output port serving red messages and the other blue messages. Intuitively, if the queues were never empty and all four possible patterns of the two colors were equally probable at each cycle, one would expect to output two messages half of the time, and only one message the other half, for a maximum expected output rate per port of 75 percent. In the following analysis we prove this formally and develop difference equations that do not depend on the pattern of colors in the queues.

4.1. Analysis with infinite queues for Type C switches

Assume that the messages arrive independently at the 2 queues with probability p at each cycle (i.e. $P(red)=P(blue)=p/2$). The service discipline is first come, first served, but when two identically colored items appear in front of the 2 queues at the same cycle one item is blocked (i.e, a message is selected for service from one of the two queues with equal probability).

The following nine possibilities for arrivals to the queue at the first cycle (or any cycle when both queues are empty) are

$q1$:	red	red	$blue$	$blue$	0	0	$blue$	0	red
$q2$:	red	$blue$	red	$blue$	0	$blue$	0	red	0
with probability	$\frac{p^2}{4}$	$\frac{p^2}{4}$	$\frac{p^2}{4}$	$\frac{p^2}{4}$	$(1-p)^2$	$\frac{p}{2}(1-p)$	$\frac{p}{2}(1-p)$	$\frac{p}{2}(1-p)$	$\frac{p}{2}(1-p)$

where 0 indicates no item arrives. The number of items served at the first cycle will be 2 for the cases $(red,blue)$ and $(blue,red)$, 0 for the case $(0,0)$ and 1 otherwise. The number of items left in the system after the first cycle will be 1 for the cases (red,red) and $(blue,blue)$ and 0 otherwise.

In a similar way, the effect of the nine different arrival possibilities on the state of the system can be analyzed for the case when only one queue is empty and the case when both queues are not empty. If difference equations that depend only on the number of items in each queue, and not on the colors of all items, can be constructed, the output probabilities for a given cycle must be independent of the exact color pattern in the preceding cycle. For the 2×2 switch, it is necessary to show that if items from both input streams appear at the output on the same cycle, they are equally likely to be of the same color or of different colors. This can be shown by analyzing the different cases.

Consider the case when both queues are non-empty and exactly one queue, $q1$, was non-empty on the preceding cycle. Suppose for definiteness that the item at the front of $q1$ was *red*. Then both queues will be non-empty only if a *red* item arriving at $q2$ was blocked by the item leaving from the front of $q1$ and another item either existed in or arrived at $q1$. The items remaining at the front of the queues will both be the same color if the next item in $q1$ was *red*, and they will be different colors if the next item in $q1$ was *blue*; the item is equally likely to be of either color because the input streams to the two queues are memoryless and independent. The other cases can be analyzed similarly.

Note that this analysis does not extend easily to switches of size greater than 2, where contention by more than two input queues for the same output on a particular cycle implies that some contention will continue to exist on the next cycle.

As before, let $p_n(k,j)$ be the probability that at the n^{th} cycle there are k messages in the first queue and j in the second. The difference equations constructed using the above case by case analysis are:

$$p_0(0,0) = 1.$$

For $n \geq 1$,

$$p_n(0,0) = (1-\frac{p^2}{2})p_{n-1}(0,0)+(1-p)(1-\frac{p}{2})[p_{n-1}(0,1)+p_{n-1}(1,0)] + \frac{1}{2}(1-p)^2\ p_{n-1}(1,1).$$

For $1 \leq j \leq n$,

$$p_n(0,j) = \frac{p^2}{4} p_{n-1}(0,j-1) + [\frac{p^2}{2} + \frac{5p(1-p)}{4}] p_{n-1}(0,j)$$

$$+ (1-p)(1-\frac{p}{2}) p_{n-1}(0,j+1) + \frac{p(1-p)}{4} p_{n-1}(1,j-1) \quad \textbf{(4-1)}$$

$$+ \frac{(1-p^2)}{4} p_{n-1}(1,j) + \frac{(1-p)^2}{2} p_{n-1}(1,j+1).$$

For $1 \le k \le n$,

$$p_n(k,0) = p_n(0,k).$$

For $k \ge 1, j \ge 1$,

$$p_n(k,j) = \frac{p^2}{4} p_{n-1}(k-1,j) + \frac{p(1-p)}{4} p_{n-1}(k-1,j+1) + \frac{p^2}{4} p_{n-1}(k,j-1)$$

$$+ \frac{p}{2} p_{n-1}(k,j) + \frac{(1-p^2)}{4} p_{n-1}(k,j+1) + \frac{p(1-p)}{4} p_{n-1}(k+1,j-1)$$

$$+ \frac{(1-p^2)}{4} p_{n-1}(k+1,j) + \frac{(1-p)^2}{2} p_{n-1}(k+1,j+1).$$

Note that $p_n(k,j) = 0$ for $k + j > n$.

Let

$$F(\lambda,x,y) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \lambda^n \, p_n(k,j) \, x^k y^j.$$

Then

$$F(\lambda,x,y)(xy - \lambda[\frac{p}{4}(1-p)(x^2+y^{2)}+\frac{p^2}{4}xy(x+y)+\frac{p}{2}xy+\frac{1}{4}(1-p^2)(x+y)+\frac{1}{2}(1-p)^2])$$

$$=xy + \lambda(1-p)^2F(\lambda,0,0)[xy+\frac{1}{2}-\frac{3}{4}(x+y)]$$

(4-2)

$$+ \lambda(1-p)F(\lambda,x,0)[\frac{3}{4}pxy-\frac{1}{2}(1-p)+\frac{3}{4}(1-p)y-\frac{p}{4}x^2-\frac{1+p}{4}x]$$

$$+ \lambda(1-p)F(\lambda,0,y)[\frac{3}{4}pxy-\frac{1}{2}(1-p)+\frac{3}{4}(1-p)x-\frac{p}{4}y^2-\frac{1+p}{4}y].$$

For $p=1$, (4-2) takes the simple form of

$$F(\lambda,x,y) = \frac{1}{1 - \frac{\lambda}{4}(x+y+2)}$$

with the one-dimensional generating function

$$F(\lambda,x,1) = \frac{1}{1 - \frac{\lambda}{4}(3+x)},$$

which implies, for $j+k \le n$,

$$p_n(k,j) = \frac{n!}{(n-k-j)!k!j!} 2^{-(n+k+j)}, \quad (4\text{-}3)$$

$$p_n(k,n-k) = \begin{bmatrix} n \\ k \end{bmatrix} 2^{-2n}, \quad (4\text{-}4)$$

$$\sum_{j=0}^{n} p_n(k,j) = \begin{bmatrix} n \\ k \end{bmatrix} 2^{-2n} 3^{n-k}, \quad (4\text{-}5)$$

$$\sum_{k=1}^{\infty} k \sum_{j=0}^{n} p_n(k,j) = \frac{n}{4}, \quad (4\text{-}6)$$

and

$$\sum_{k=0}^{l} p_n(k,l-k) \ = \ \begin{bmatrix} n \\ l \end{bmatrix} 2^{-n}, \qquad \textbf{(4-7)}$$

i.e., the probability that there are l messages waiting in the two queues at time n is $\begin{bmatrix} n \\ l \end{bmatrix} 2^{-n}$, and the average number of messages waiting in the two queues at time n, which we denote by E_1^n, is $n/2$.

Let us again define $F(x,y)$ to be $\lim\limits_{\lambda \to 1}(1-\lambda)F(\lambda,x,y)$ whenever the limit exists. Then from (4-2) we obtain the steady state functional equation for the generating function for the values of p that have a steady state distribution:

$$\begin{aligned} &F(x,y)(xy-[\frac{p}{4}(1-p)(x^2-y^2)+\frac{p^2}{4}xy(x+y)+\frac{p}{2}xy+\frac{1}{4}(1-p^2)(x+y)+\frac{1}{2}(1-p)^2]) \\ &=(1-p)^2F(0,0)[xy+\frac{1}{2}-\frac{3}{4}(x+y)] \\ &\quad +(1-p)F(x,0)[\frac{3}{4}pxy-\frac{1}{2}(1-p)-\frac{3}{4}(1-p)y-\frac{p}{4}x^2-(\frac{1+p}{4})x] \\ &\quad +(1-p)F(0,y)[\frac{3}{4}pxy-\frac{1}{2}(1-p)-\frac{3}{4}(1-p)x-\frac{p}{4}y^2-(\frac{1+p}{4})y] \end{aligned} \qquad (4\text{-}8)$$

Of special interest is the case $x=y$ or the generating function of $\sum_{l=0}^{k} p(k,l-k)$. For this case (4-8) is reduced to

$$\begin{aligned} &F(x,x)\left[p^2x^2 \ - \ 2x(1-p) \ - \ (1-p)^2\right] \\ &\quad = \ (1-p)^2(1-2x)F(0,0) \ - \ 2(1-p)(1-p+px)F(x,0). \end{aligned} \qquad (4\text{-}9)$$

Notice that if a steady state distribution exists F(1,1) = 1 hence we get

$$3 \ - \ 4p \ = \ (1-p)\left[(1-p)F(0,0) \ + \ 2F(1,0)\right]. \qquad (4\text{-}10)$$

Since the right hand side of (4-10) is non-negative, we have proved:

Theorem 1: For $p > \frac{3}{4}$ there is no steady state limiting distribution and the queues become infinitely large.

Let O_p^n be the average number of items exiting from each output port (server) at the n^{th} cycle given an input rate p and let O_p be $\lim_{n \to \infty} O_p^n$. From the definition of O_p^n:

Lemma 1: $O_p^n \le O_1^n$.

Lemma 2: $O_p \le O_1$.

Let $p_n(l) = \sum_{k=0}^{l} p_n(k, l-k)$, be the probability that at time n there are a total of l waiting messages in the two queues, and let

$$E_p^n(S) = \sum_{l=1}^{n} l p_n(l)$$

be the average length of the two queues at time n. Then

$$np - \sum_{i=1}^{n} O_p^i = \frac{1}{2} E_p^n(S). \tag{4-11}$$

Theorem 2: $O_p \le \frac{3}{4}$.

The proof of the theorem follows easily from (4-11). Let $p = 1$. Then (4-11) implies

$$\sum_{i=1}^{n} O_1^i = n - \frac{1}{2} E_1^n(S) = \frac{3}{4} n, \tag{4-12}$$

since E_1^n is $\frac{n}{2}$, see (4-6). Taking $n = 1$, this implies

$$O_1^1 = \frac{3}{4}.$$

For $n = 2$, using (4-12) $O_1^1 + O_1^2 = \frac{6}{4}$. Since $O_1^1 = \frac{3}{4}$, $O_1^2 = \frac{3}{4}$. Continuing by induction, we get $O_1^n = \frac{3}{4}$. Using Lemma 2, we get

$$O_p \le \frac{3}{4}. \qquad \text{QED}$$

Note that for $p < \frac{3}{4}$ there is a steady state limiting distribution and therefore $O_p = \lim_{n \to \infty} O_p^n = p$ for p less than $3/4$.

For the case that a limiting distribution exists, an upper bound for E(S) can easily be found. Since $2E(S) = \frac{d}{dx}F(x,x)|_{x=1}$, from (4-9), taking the derivative with respect to x at $x = 1$ we get

$$\begin{aligned} &E(S)(3 - 4p) + 1 - p - p^2 \\ &= (1 - p)\left[(1 - p)F(0,0) + pF(1,0) + \frac{d}{dx}F(x,0)|_{x=1}\right] \end{aligned} \qquad (4\text{-}13)$$

but

$$\frac{d}{dx}F(x,0)|_{x=1} = \sum_s sp(s,0) \le \sum_{s,s'} sp(s,s') = E(S).$$

Hence from (4-13) we get

$$E(S)\ (3 - 4p) + 1 - p - p^2 \le (1 - p)(1 + E(S))$$

or

$$E(S)\ (2 - 3p) \le p^2,$$

i.e., for $p \le 2/3$, $E(S) \le \frac{p^2}{2 - 3p}$. In particular if $p \le .56$, then $E(S) < 1$.

4.2. Analysis of Type C switch with queues of size 1

Now let us consider the Type C switch with each queue of capacity one. The conditions under which a message is lost now depend on the color of the items in the queue, as well as their number. For example, suppose $q1$ has a *red* item and $q2$ is empty. If a *red* item arrives at $q2$, it has a 50 percent chance of blocking the item in $q1$, causing the *red* item in $q1$ to remain for another cycle and any message of either color arriving at $q1$ to be lost. Considering each such case for all nine input possibilities and all nine initial queue configurations, and consolidating symmetric equations, we get the following difference equations:

$$p_0(0,0) = 1.$$

For $n \geq 1$,

$$p_n(0,0) = (1 - \frac{p^2}{2})p_{n-1}(0,0) + (1 - p)(1 - \frac{p}{2})[p_{n-1}(0,1) + p_{n-1}(1,0)]$$
$$+ \frac{(1 - p)^2}{2}p_{n-1}(1,1),$$

$$p_n(0,1) = \frac{p^2}{4}p_{n-1}(0,0) + (\frac{5p}{4} - \frac{p^2}{2})p_{n-1}(0,1)$$
$$+ \frac{p}{4}(1 - p)p_{n-1}(1,0) + \frac{(1 + p - 2p^2)}{4}p_{n-1}(1,1),$$

$$p_n(1,0) = \frac{p^2}{4}p_{n-1}(0,0) + \frac{p}{4}(1 - p)p_{n-1}(0,1)$$
$$+ (\frac{5p}{4} - \frac{p^2}{2})p_{n-1}(1,0) + \frac{(1 + p - 2p^2)}{4}p_{n-1}(1,1),$$

$$p_n(1,1) = \frac{p^2}{4}p_{n-1}(0,1) + \frac{p^2}{4}p_{n-1}(1,0) + \frac{p}{2}(1 + p)p_{n-1}(1,1).$$

Again we introduce the generating function

$$F_0(x,y) \equiv 1.$$

For $n \geq 1$,

$$F_n(x,y) = \sum_{k=0}^{1} \sum_{j=0}^{1} p_n(k,j)\, x^k y^j .$$

Then

$$F(\lambda,x,y) = \sum_{n=0}^{\infty} \lambda^n F_n(x,y)$$

$$= (1,\ x,\ y,\ xy)(I - \lambda Q)^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

where

$$Q = \begin{bmatrix} 1 - \frac{p^2}{2} & (1-p)(1-\frac{p}{2}) & (1-p)(1-\frac{p}{2}) & \frac{(1-p)^2}{2} \\ \frac{p^2}{4} & \frac{5p}{4} - \frac{p^2}{2} & \frac{p}{4}(1-p) & \frac{(1+p-2p^2)}{4} \\ \frac{p^2}{4} & \frac{p}{4}(1-p) & \frac{5p}{4} - \frac{p^2}{2} & \frac{(1+p-2p^2)}{4} \\ 0 & \frac{p^2}{4} & \frac{p^2}{4} & \frac{p(1+p)}{2} \end{bmatrix}.$$

As before, let F(x,y) be the steady state generating function of the limiting distribution. Then

$$F(x,y) = \lim_{\lambda \to 1} (1 - \lambda) F(\lambda,x,y) = \sum_{i=0}^{1} \sum_{j=0}^{1} p(i,j) x^i y^j$$

and we get

$$p(0,0) = \frac{8 - 16p + 7p^2 + 2p^3 - p^4}{8 - 16p + 11p^2 - 2p^4},$$

$$p(0,1) = p(1,0) = \frac{2p^2 - p^3 - p^4}{8 - 16p + 11p^2 - 2p^4},$$ and

$$p(1,1) \; = \; \frac{p^4}{8 \; - \; 16p \; + \; 11p^2 \; - \; 2p^4}.$$

From these we conclude

$$2E(S) \; = \; p(0,1) \; + \; p(1,0) \; + \; 2p(1,1)$$

hence the average queue length is

$$E(S) \; = \; \frac{2p^2 - p^3}{8 \; - \; 16p \; + \; 11p^2 \; - \; 2p^4} \qquad (4\text{-}14)$$

and the probability that one queue is empty is $p(0,0) + p(0,1) \; = \; 1 - E(S)$.

To find the expected output per cycle, $E_p(O)$, from both queues in steady state, we considered the three cases: both queues empty, one queue empty and neither queue empty. By considering the nine different input possibilities, we find that

$$E_p(O|(0,0)) = 2p - \frac{p^2}{2}$$

and

$$E_p(O|(0,1)) = E(O|(1,0)) = 1 + \frac{p}{2}.$$

When both queues are non-empty, the output will be 1 when the items at the front of each queue are the same color, and 2 when they are different colors. Since each case is equally likely, as discussed in the preceding section,

$$E_p(O|(1,1)) = \frac{3}{2}.$$

Then

$$E_p(O) \; = \; (2p - \frac{p^2}{2})p(0,0) \; + \; (1 \; + \; \frac{p}{2})(p(0,1) \; + \; p(1,0)) \; + \; \frac{3}{2}p(1,1) \qquad (4\text{-}15)$$

$$= \frac{3p}{2} + \frac{p}{2}(1-p)\frac{1-p+\frac{3}{8}p^2+\frac{1}{8}p^3-\frac{1}{8}p^4}{1-2p+\frac{11}{8}p^2-\frac{1}{4}p^4},$$

with

$$\max_p E_p(O) = \frac{3}{2},$$

and therefore the average output per queue (or output port) is $\leq 3/4$.

5. Comparison of the three implementations

Figures 3 through 9 display parameters of interest, using values obtained from the equations in sections 2, 3 and 4. All the results were checked with simulations. Data points from simulations are displayed along with the graphs from the equations in Figures 5 and 6, but are omitted elsewere for readability.

In the graphs, *offered load* (p) refers to the probability that a message will be generated at an input port to the switch on a given cycle, and *bandwidth* ($E_p(O)$) refers to the the expected output rate of the switch for that offered load; both these quantities are measured in messages per switch port per cycle. The distinction is important for switches with finite queue size, since the expected output rate will be less than the offered load due to blocking. *Latency* ($E(S)/E_p(O)$, by a variant of Little's Law) is the average waiting time of a message in the queue and was obtained by dividing the expected queue length by the bandwidth.

When examining the graphs, note that latency figures for finite queue sizes are somewhat misleading, since they do not take into account any waiting time which a message may experience when it is lost and must be retransmitted. Also the latency figures for a single switch in isolation are optimistic when projected to the performance of a multi-stage network of switches (see [9, 12, 14]).

Although much of the usefulness of this work is in specific results that can be used to predict behavior for particular loads, certain general conclusions can be drawn:

- Type A and Type B are very close in performance. For the parameters illustrated in Figures 5 and 6 their performance is identical, and thus only Type A is shown. Even for minimum size queues (Figures 7 and 8) the difference is slight, and the choice of one or another can reasonably be made on the basis of difficulty of implementation.
- For loads under 50 percent, the performance of Type C switches is reasonably close to that of Types A and B, and, because of its lower cost, Type C may be the implementation of choice. However, in addition to the bandwidth limitation at 75 percent, Type C switches show a significant increase in latency even at loads as low as 50 percent.
- Adding bigger buffers to a Type C switch will not bring about a substantial performance improvement, since the output rate is limited at 75 percent whether buffer size is 1 or unbounded. Type A and Type B switches with minimum size buffers have significantly better performance than Type C switches with unbounded buffer size. Thus if additional hardware resources are to be applied to improving network performance, they are better spent on implementing Type A or Type B designs than on making larger buffers for a Type C design.

We are currently trying to extend our analysis to larger crossbar sizes. Preliminary investigations indicate that the bandwidth per port limitation for Type C switches becomes even more severe as crossbar size grows. On the other hand, since even minimum sized Type A and Type B switches require buffer space that is quadratic in the size of the crossbar, it is expensive in area to implement a large crossbar on a chip with such designs. The best trade-off between cost and performance may require implementing a large crossbar switch as an on-chip omega network composed of 2×2 Type A or B switches.

Acknowledgements

We thank J.K. Percus whose many contributions and constant criticism were invaluable.

Ora Engelberg Percus received a M. Sc. in Mathematics, Physics and Statistics at Hebrew Univeristy, Jerusalem, and a Ph. D. in Mathematical Statistics from Columbia University in 1965. Currently she is a Research Associate Professor at Courant Institute of Mathematical Sciences, New York University. She has been active in several areas of mathematics and has contributed articles to scientific journals in probability, statistics, combinatorics, queueing theory and numerical simulation. Her previous article in this journal: Random Number Generators for MIMD Parallel Processors (with M. H. Kalos) 6, pp. 477 - 497 (1989).

Susan Dickey received a M. Sc. in Computer Science from New York University in 1984 and is currently working on her Ph. D. at NYU. She is an Assistant Research Scientist with the Ultracomputer Research Laboratory at the Courant Institute of Mathematical Sciences, NYU. Her research interests are VLSI design and parallel architectures.

[Legends for figures.]

Figure 1. Ultracomputer network, 8 processors, 8 memory models.

Figure 2. Three types of buffered 2×2 crossbar switches. (A) Two-input queues, one per output port (Type A). (B) One-input queues, two per output port (Type B). (C) One-input queues, one per output port (Type C).

Figure 3. Probability distribution $p(j)$ vs. Queue Length j. Unbounded queues, Type A, loads 0.2, 0.7 & 0.9. Data from equation 2-1. Sum of queue lengths probability distribution for Type B is identical.

Figure 4. Probability distribution $p(j)$ vs. Queue Length j. Unbounded queues, Type C, loads 0.2 & 0.7. Data computed from difference equations 4-1.

Figure 5. Mean Queue Length $E(S)$ vs. Offered Load p. Unbounded queues, Types A & C. For Type A, mean queue length was computed from equation 2-3. Type B mean sum of queue lengths is identical to Type A mean queue length. For Type C, the mean queue length was obtained from the probability distribution values computed from difference equations 4-1.

Figure 6. Latency $E(S)/p$ vs. Offered Load p. Unbounded queues, Types A & C. Type B values are identical to Type A. Latency values used in the line graphs were computed by taking the $E(S)$ values as in Figure 5 and dividing by p. (Simulation values for latency were simulated directly, not calculated from simulated values of queue length and bandwidth.)

Figure 7. Bandwidth $E_p(O)$ vs. Offered Load p. Minimum size queues, all 3 types. Data from equations 2-7, 3-5 and 4-15.

Figure 8. Latency $E(S)/E_p(O)$ vs. Bandwidth $E_p(O)$. Minimum size queues, all 3 types $E(S)$ from equations 2-6, 3-4 and 4-14; for Type B switches, $2E(S)$ is used. Bandwidth as in Figure 7.

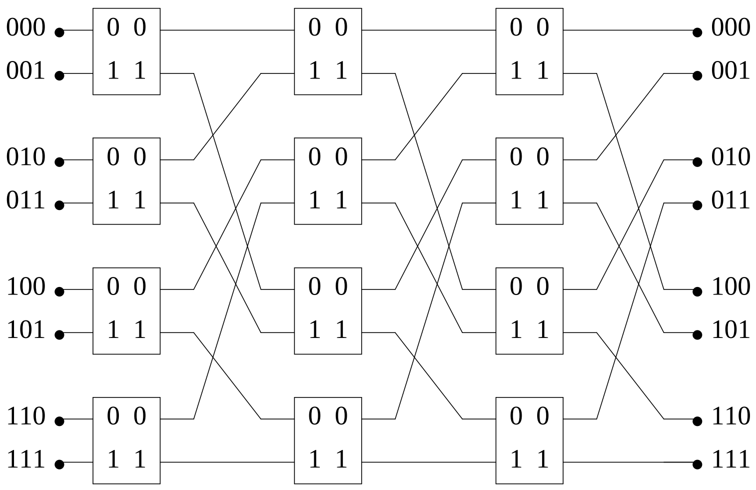

Figure 1.

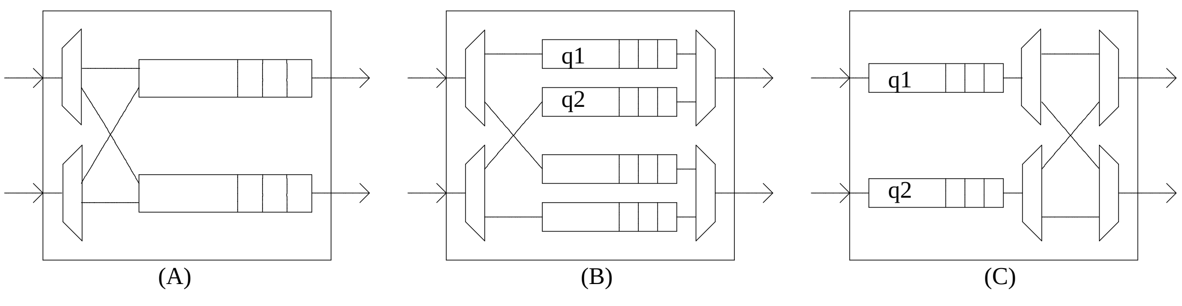

Figure 2.

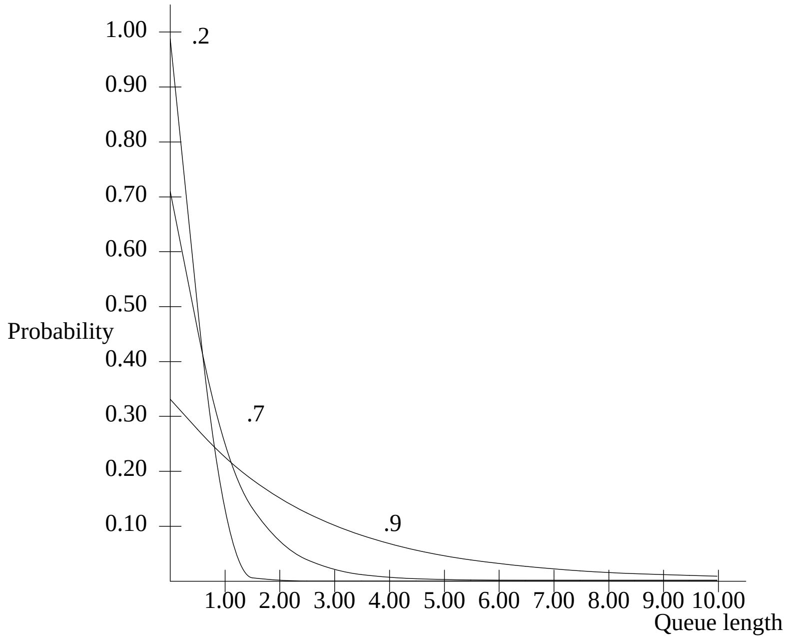

Figure 3.

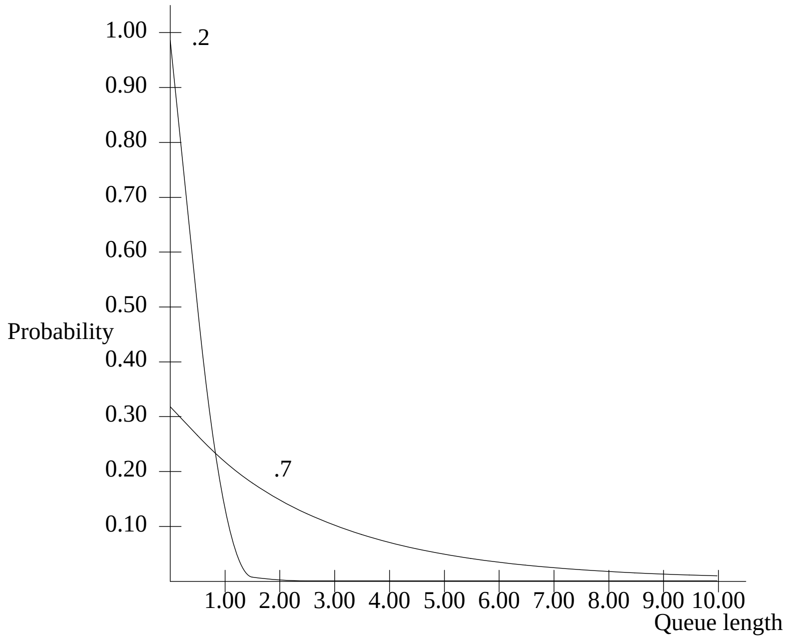

Figure 4.

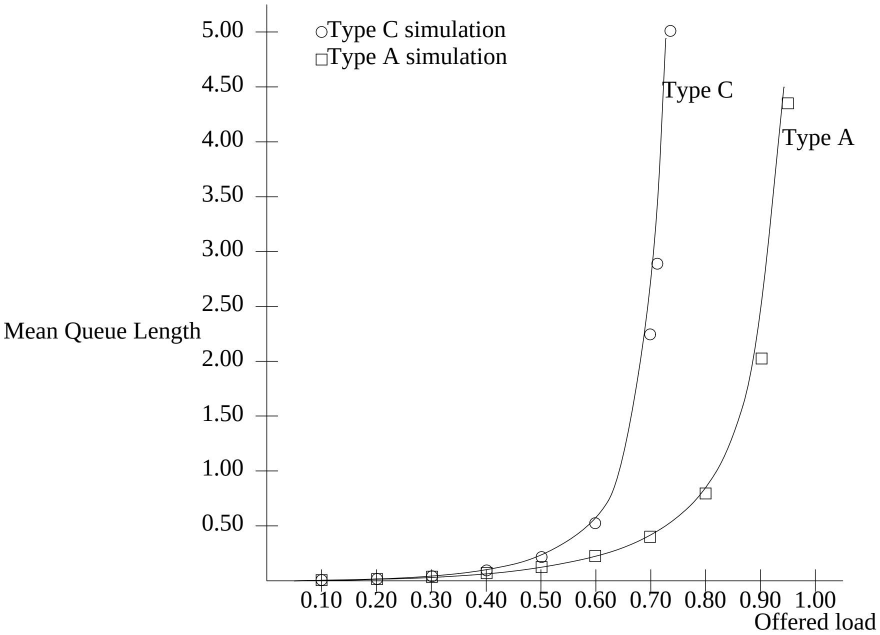

Figure 5.

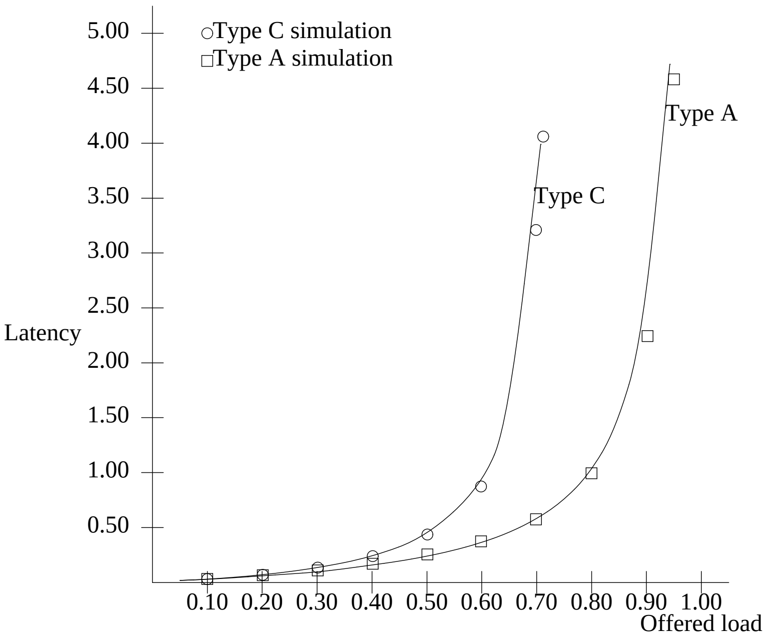

Figure 6.

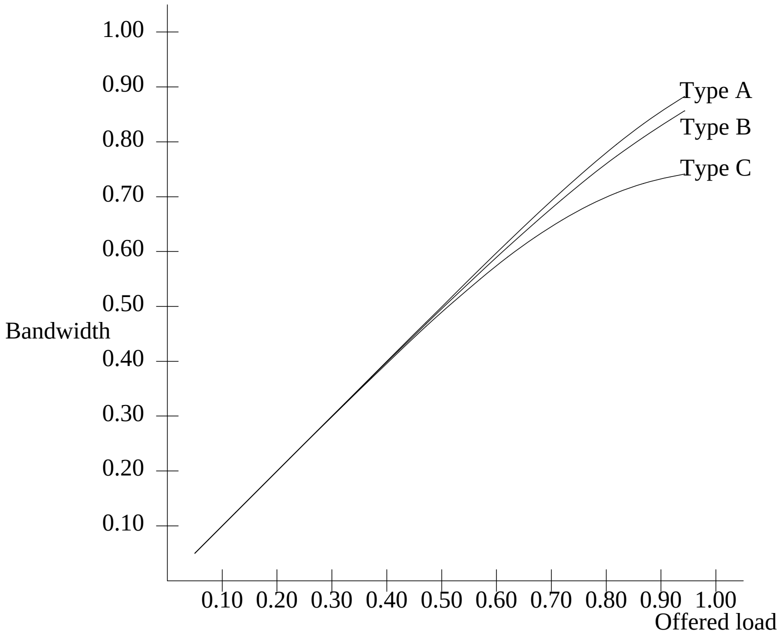

Figure 7.

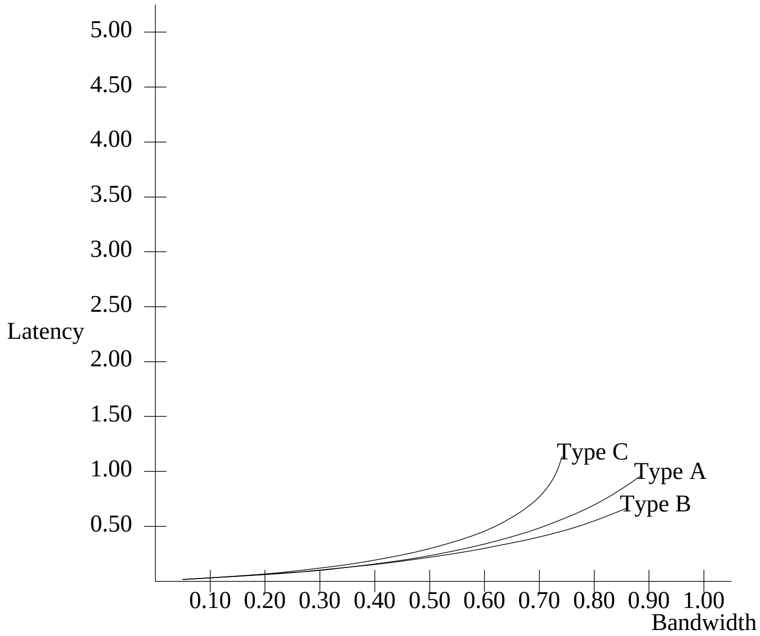

Figure 8.

References

1. Dias, D. and Jump, J. R. Analysis and Simulation of Buffered Delta Networks. *IEEE Transactions on Computers.* Vol C-30, No. 4, April 1981.

2. Dickey, S., Gottlieb, A., Kenner, R. and Liu, Y. S. Designing VLSI Network Nodes to Reduce Memory Traffic in a Shared Memory Parallel Computer. *Circuits, Systems and Signal Processing.* Vol. 6, No. 27, 1987, pp. 217-23.

3. Dickey, S., Kenner, R., Snir, M. and Solworth, J. A VLSI Combining Network for the NYU Ultracomputer. *Proceedings of the International Conference on Computer Design.* October 7, 1985, pp. 110 - 113.

4. Dickey, S. and Liu, Y. S. Simulation and Analysis of Different Switch Architectures for Interconnection Networks in MIMD Shared Memory Machines. Ultracomputer Note #141, Courant Institute, New York University, June 1988.

5. Gottlieb, A. An Overview of the Ultracomputer Project. In Dongarra, J. J., ed. *Experimental Parallel Computing Architectures.* Halstead Press, New York, 1987, pp. 25 - 95.

6. Harper, D. T. III and Jump, J. R. Performance Evaluation of Reduced Bandwidth Multistage Interconnection Networks. *The 14th Annual International Symposium on Computer Architecture.* May 1987, pp. 171-185.

7. Jaffe, S. Equilibrium Results for a Pair of Coupled Discrete Time Queues. Ultracomputer Note #160, Courant Institute, New York University, June 1989.

8. Kruskal, C.P. and Snir, M., The Performance of Multistage Interconnection Networks for Multiprocessors. *IEEE Trans. Comp.* **C-32**, 1983, pp. 1091-1098.

9. Kruskal, C. P., Snir, M. and Weiss, A. The Distribution of Waiting Times in Clocked Multistage Interconnection Networks. *Proc. 1986 Intern. Conf. on*

Parallel Processing.

10. Kumar, M. and Jump, J. R. Performance Enhancement in Buffered Delta Networks Using Crossbar Switches and Multiple Links. *Journal of Parallel and Distributed Computing.* Vol. 1, No. 1, August 1984, pp. 81-103.

11. Lawrie, D.H. Access and Alignment of Data in an Array Processor. *IEEE Trans. Comp.* **C-24** , Dec. 1975, pp. 1145-1155.

12. Liu, Y. S. Architecture and Performance of Processor-Memory Interconnection Networks for MIMD Shared Memory Parallel Processing Systems. Ph. D. Thesis, New York University, 1990.

13. Percus, O. E. and Percus, J. K. Elementary properties of Clock-regulated Queues. *SIAM Journal on Applied Mathematics.* 50, 1990, pp. 1166 - 1175.

14. O. E. Percus and J. K. Percus. Times Series Transformations in Clocked Queueing Networks. *Communications on Pure & Applied Math.* Vol 44, No 9, 1991.

15. Snir, M. and Solworth, J. The Ultraswitch -- A VLSI Network Node for Parallel Processing. Ultracomputer Note #39, Courant Institute, New York University, January 1984.

16. Snir, M., Switch Digest. Ultracomputer Hardware Note #27, Courant Institute, New York University, 1984.

17. Wilf, H. S. *Generatingfunctionology.* Academic Press, Inc., San Diego, 1990.