

# Performance of Output-Multibuffered Multistage Interconnection Networks Under General Traffic Patterns

Bin Zhou and M. Atiquzzaman

Dept. of Computer Science and Engineering  
La Trobe University  
Melbourne 3083, Australia  
atiq@LATCS1.lat.oz.au

## Abstract

*Multistage networks are strong candidates for implementation of ATM switching fabrics in Broadband ISDN networks. To prevent internal loss of data, buffers are often used inside the switching elements of the fabric. The objective of this paper is to develop models to evaluate the throughput and packet delay, in the presence of a general traffic pattern, of an ATM switching fabric using multiple buffers at the outputs of the switching elements. The models are based on Markov chains, and use several simplifying assumptions to make the model tractable. One of the models take into account of the blocked packets and the history of blocking. The models produce accurate results in the case of a general traffic pattern.*

## 1 Introduction

A multiprocessor system consists of a number of processors and memories connected together by an interconnection network. Multistage interconnection networks (MINs) have been found to be suitable for scalable large-scale systems. A MIN consists of several stages of small crossbar switching elements (SE), connected together by a permutation function. Because of their self-routing property, MINs have also been proposed as ATM switching fabrics in Broadband ISDN networks. The overall performance of a multiprocessor system or the Broadband ISDN depends significantly on the interconnection network or the switching fabric respectively. Hence, it is extremely important to evaluate the performance of the network and the fabric.

Performance analysis of unbuffered MINs have been reported in [1, 2, 3]. Buffered MINs prevent internal loss of data in the case of routing conflicts, and in most cases have a higher throughput than the unbuffered MINs. Dias [4] and Jenq [5] have analyzed MINs consisting of  $2 \times 2$  switching elements with single-packet input buffers and uniform traffic in the MIN. It has been shown that the maximum achievable throughput for large sized MINs having single input-buffered SEs is limited to approximately 0.45 under uniform input traffic pattern [5]. Even with infinite-sized input buffers, the maximum throughput of a large multibuffered MIN is limited to approximately 0.75 [4].

Kruskal and Snir [6] have discussed Banyans networks with output-buffered SEs for the case where the buffer capacity is infinite. Kim [7] reported a queueing analysis and a simulation study of output-buffered Banyans with an arbitrary (finite) buffer size. It was shown a maximum throughput of one can be achieved with infinite-sized output-buffered SEs. All the above performance analyses were made on the assumption of the MIN operating in the presence of a uniform traffic pattern.

Non-uniform traffic conditions can occur in a multiprocessor system, for example, in the case where shared variables are used for synchronization. Pfister [8] discussed the phenomenon of tree saturation arising as a result of a hot spot in a buffered MIN. Tree saturation results in degradation in the performance of the MIN. Nonuniform traffic conditions can also reflect the traffic patterns of ATM networks where a wide range of bandwidths need to be accommodated. Therefore, the performance of MINs in the presence of nonuniform traffic is an important issue to be studied.

Analysis of single input/output buffered MINs in a nonuniform traffic environment is described in [9, 10]. Atiquzzaman [11] proposed an efficient Markov chain model for the performance evaluation of a single input buffered Omega network in the presence of a hot spot. Performance studies of output-multibuffered MINs under nonuniform traffic patterns has been reported in [12].

Analytical models permit a fast and inexpensive method of performance evaluation, and provide insight into the factors that determine design trade-offs as well as quantitative estimates of their importance. Lin et.al [13] presented an analytical model for evaluating the performance of finite-buffered packet switching MIN under a general traffic pattern. In their model, there is no buffer space at the processing elements. Discarded packets are not re-submitted. Moreover, they have not considered the history of a blocked packet in an SE, and keep the values of the routing probabilities of blocked packets the same as the unblocked packets.

The aim of this paper is to develop an analytical model to evaluate the performance of an output-multibuffered MIN in the presence of a general traffic pattern. The objectives of the research work were as follows.

- To develop a generalized analytical model which can be applied to evaluate the performance of output-multibuffered MINs with arbitrary buffer sizes and input load under nonuniform traffic environment, particularly in the presence of a hot spot traffic.
- To allow resubmission of packets which are rejected by the network. The rejected packets are queued at the Input Buffer Controller for resubmission in the next cycle.
- To consider the history of a blocked packet in an SE.
- To check the validity of the proposed model by comparison with simulation results.

Two different models are presented in this paper. The basic model does not memorize the history of blocking, while the memorized model is capable of taking into account the fact that a blocked packet always hunts for the same output link of an SE during successive clock cycles.

In Section 2, the structure, operation, and the modeling assumptions of an output-buffered Omega network are described. A basic analytical model for the performance evaluation of an output-buffered Omega network under a general traffic environment is proposed in Section 3. A memorized model is introduced in Section 4 to account for the blocked packets and the history of blocking. Results obtained, in the presence of uniform and hot spot traffic, from the basic and memorized models are compared with those from stochastic simulations in Section 5, followed by concluding remarks in Section 6.

## 2 Modeling Assumptions

The Omega network, originally proposed by Patel [1], will be taken as an example of a MIN to be modeled. The model, in fact, applies equally well to all unique path networks.

We make the following assumptions regarding the operation and environment of the interconnection network as in [7, 11, 12].

1. There are  $N = 2^n$  inputs and  $N$  outputs of the network. Each input of the network has an Input Buffer Controller (IBC) of size  $m$ .
  2. The network is operated *synchronously*. This reflects the situation in an ATM environment where all packets have a fixed length, and fit exactly into one clock cycle. For modeling purposes, we split the clock cycle into two phases
    - In the first phase, the availability of buffer space at the succeeding stage along the destined path of a packet is determined; the packet is informed whether it may proceed to the succeeding stage or should stay in the current buffer during the current cycle.
    - Depending on the availability of buffer space in the succeeding stage, a packet may move forward one stage in the second phase.
  3. A *backpressure* mechanism ensures that no packets are lost within the network.
4. A buffer supports simultaneous *enqueueing* and *dequeueing* of packets during the same cycle.
  5. Packet arrival process at each input of the network is a simple *Bernoulli* process.
  6. There is no *blocking* at the output links of the network.
  7. The *conflict resolution* logic at each SE is unbiased.
  8. Since at least one time unit is spent in each buffer even when there is no waiting, the minimum possible delay of a packet is equal to  $n + 1$ , where  $n$  is the number of stages. It includes the delay at the IBC buffer.
  9. For a uniform traffic pattern, an incoming packet is equally likely to be directed to any output link. For a *hot spot traffic* pattern, the probability that an incoming packet is directed to a non-hot or a hot memory module are  $(1 - h)/N$  and  $h + (1 - h)/N$  respectively, where  $h$  is defined to be the hot spot probability.
  10. When two packets from different buffers in the same stage contend for the same output buffer in the next stage, a *contention* occurs. If there is more than one space available at the buffer in the next stage, the switch is assumed to be fast enough to accept both packets in one cycle. If only one space is available, a packet is randomly chosen to fill up this space; the other packet is then "blocked" and stays at the original buffer. However, if no space is available in the next stage then both the packets are blocked.
  11. We assume that there is an IBC at each input of the network. Packets buffered at the IBC are resubmitted in the next cycle.

## 3 Basic Model

In this section, we develop an analytical model to evaluate the performance of an Omega network in the presence of a general traffic pattern. Markov chain is used for the model, and is based on the methods given in Kim [9, 7] and Atiquzzaman [11].

### 3.1 Modeling a Single Queue

We first define the following variables in the same manner as in [9], and derive a set of state equations relating these variables. Figure 2 shows how the SEs are specified by parameters  $k$ ,  $l$ , input ports  $x$  and  $\underline{x}$ , and output ports  $y$  and  $\underline{y}$ . Figure 1 shows three consecutive SEs in successive stages of the network. We denote the input and output ports of the  $l$ -th SE at stage  $k$  by  $k l x$ ,  $k l \underline{x}$ ,  $k l y$  and  $k l \underline{y}$ . Let  $(k + 1) \hat{l} \hat{y}$  denote the input port of the  $\hat{l}$ -th switching element, at the  $(k+1)$ -th stage, which is fed by the output port  $k l y$ . Also, let  $(k - 1) \hat{l} \hat{x}$  denote the output port of SE, at the  $(k-1)$ -th stage, which feeds the input port  $k l x$ . We define the following variables to be used in our model.

$B_{klv}^i(t)$  = probability that there are  $i$  packets in the buffer at  $klv$ ,  $v \in \{y, \underline{y}\}$ , at the beginning of cycle  $t$ .

$\pi_{ij}$  = transition probability of a buffer from state  $i$  to state  $j$ , given that it is in state  $i$ .

$\beta_{ij}$  = probability that a packet arriving at the  $i$ -th input of the network is destined for the  $j$ -th output of the network. Note that  $\beta_{ij} = 1/N$  for a uniform traffic pattern.

$m$  = number of buffers at each output of an SE.

$n$  = number of switching stages in the network.

$P_{kluv}(t)$  = Probability that a packet, ready to come to  $klu$  during clock cycle  $t$  is destined to  $klv$ ,  $u \in \{x, \underline{x}\}$ ,  $v \in \{y, \underline{y}\}$ .

$Q_{klu}(t)$  = probability that a packet is ready to come to port  $klu$  during cycle  $t$ ,  $u \in \{x, \underline{x}\}$ .

$C_{klv}^i(t)$  = probability that  $i$ ,  $0 \leq i \leq 2$ , packets are ready to come to the buffer at  $klv$  during cycle  $t$ ,  $v \in \{y, \underline{y}\}$ .

$r_{klv}(t)$  = probability that a packet in buffer  $klv$  advances to the next stage during clock cycle  $t$ , given that there is a packet in the buffer at  $klv$ ,  $v \in \{y, \underline{y}\}$ .

$r_{kluv}(t)$  = probability that a packet from port  $klu$  advances to the buffer  $klv$  during clock cycle  $t$ , given that a packet, destined to  $klv$ , is ready to come to  $klu$  during cycle  $t$ ,  $u \in \{x, \underline{x}\}$ ,  $v \in \{y, \underline{y}\}$ .

$\rho$  = Probability of a packet coming to an IBC during a cycle.

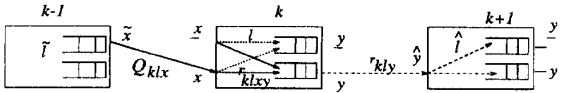


Figure 1: Illustration of  $r_{klxy}(t)$ , and  $r_{kly}(t)$

### 3.2 Routing Probabilities

The steady state value of any variable  $x(t)$  will be represented by  $x$ . Now, we obtain equations for  $r_{kluv}$ ,  $r_{klv}$  and  $Q_{klu}$ . Their relationship is shown in Figure 1. A packet, ready to come to the  $klu$  port and destined to  $klv$  port, can pass from  $klu$  to  $klv$  buffer when the following conditions are satisfied.

1. There are at least two packet spaces in the buffer of  $klv$  port, or
2. The buffer at  $klv$  has  $m - 1$  packets and a packet in the buffer advances to the next stage in the same clock cycle, or

3. The buffer at  $klv$  has  $m - 1$  packets and a packet can not be forwarded, and there is no conflict between the packet at  $klu$  and a possible packet at  $klv$ , or there is a packet at  $klu$  which is destined to  $klv$  and the conflict is resolved in favor of the packet at  $klu$ , or

4. The buffer at  $klv$  has  $m$  packets and a packet advances to the next stage in the same cycle, and there is no conflict between the packet at  $klu$  and a possible packet at  $klv$ , or there is a packet at  $klv$  which is destined to  $klv$ , and the conflict is resolved in favor of the packet at  $klu$ .

These conditions translate into the following equation for  $r_{kluv}$ .

$$r_{kluv} = 1 - \{B_{klv}^m + B_{klv}^{m-1}\} + B_{klv}^{m-1}r_{klv} + \{1 - Q_{kl\underline{u}} + Q_{kl\underline{u}}P_{kl\underline{uv}} + 0.5Q_{kl\underline{u}}P_{kl\underline{uv}}\} \{B_{klv}^{m-1} (1 - r_{klv}) + B_{klv}^m r_{klv}\}, 1 \leq k \leq n - 1 \quad (1)$$

Since the network output links are always ready to remove packets,  $r_{nlv} = 1$  and  $r_{nluv}$  is obtained from Equation (1) as

$$r_{nluv} = 1 - B_{nlv}^m Q_{kl\underline{u}} (1 - 0.5P_{nl\underline{uv}} - P_{nl\underline{uv}}) \quad (2)$$

$r_{klv}$  can then be obtained from  $r_{kluv}$  as follows

$$r_{klv} = P_{(k+1)\underline{iv}} r_{(k+1)\underline{iv}} + P_{(k+1)\underline{iv}} r_{(k+1)\underline{iv}} \quad (3)$$

$Q_{klu}$  and  $Q_{l\underline{lu}}$  can be expressed as

$$Q_{klu} = 1 - B_{(k-1)\underline{iu}}^0 \quad (4)$$

$$Q_{l\underline{lu}} = 1 - B_{0\underline{iu}}^0 \quad (5)$$

where  $B_{0\underline{iu}}^0$  is the probability of the IBC being empty. The probabilities of packets coming to a buffer at the  $klv$  port are given by

$$C_{klv}^0 = 1 - (C_{klv}^1 + C_{klv}^2) \quad (6)$$

$$C_{klv}^1 = Q_{kl\underline{u}} P_{kl\underline{uv}} (1 - Q_{kl\underline{u}}) + Q_{kl\underline{u}} P_{kl\underline{uv}} (1 - Q_{kl\underline{u}}) + Q_{kl\underline{u}} Q_{kl\underline{u}} (P_{kl\underline{uv}} P_{kl\underline{uv}} + P_{kl\underline{uv}} P_{kl\underline{uv}}) \quad (7)$$

$$C_{klv}^2 = Q_{kl\underline{u}} P_{kl\underline{uv}} Q_{kl\underline{u}} P_{kl\underline{uv}} \quad (8)$$

### 3.3 Destination Port Probabilities

$P_{kluv}$  is 0.5 for a uniform traffic pattern. On the contrary, a general traffic pattern implies that  $P_{kluv}$  may not be 0.5. We model this general traffic pattern by finding a mapping scheme that transforms a given memory referencing pattern into a set of  $P_{kluv}$ 's which reflects the given memory referencing pattern [13].

As an example, let us take an  $8 \times 8$  Omega network as shown in Figure 2. Since we assume that all processing elements generate the same general traffic pattern, we only discuss the mapping scheme for one processing element. We can represent the referencing pattern of processing element  $i$  in terms of the memory destination probability  $\beta_{ij}$ , the probability that a packet chooses memory module  $j$  as its destination.

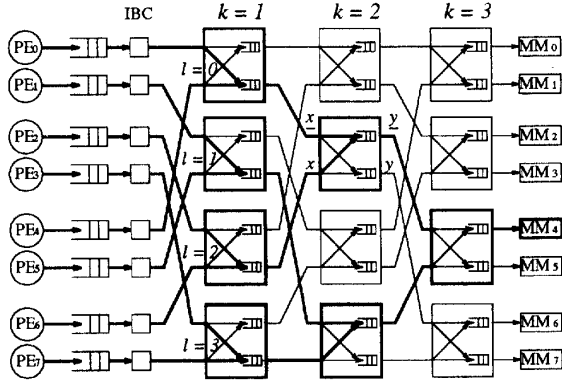


Figure 2: Omega network in hot spot traffic pattern

Consider a packet generated by processor 0 and observe the path it takes as it travels through the network to access a memory module. The packet chooses memory module 0 with probability  $\beta_{00}$  which equals  $P_{10xy}P_{20xy}P_{30xy}$ . Similarly, the packet from processing element 0 chooses memory module 1 with probability  $\beta_{01} = P_{10xy}P_{20xy}(1 - P_{30xy})$ . Using these two equations, we find  $P_{30xy}$  in terms of  $\beta_{00}$  and  $\beta_{01}$ .

$$P_{30xy} = \frac{\beta_{00}}{\beta_{00} + \beta_{01}} \quad (9)$$

The other destination port probabilities in the SEs can be found in a similar way, once the memory referencing patterns  $P_{kluv}$ 's are known.

### 3.4 Stage Buffers

A buffer at any SE at stage  $k$ ,  $1 \leq k < n$ , is modeled as a Markov chain. The state of a buffer is represented by the number of packets in the buffer. The probability of departure of a packet from a buffer in a stage is determined by the possibility of conflict with a packet at the same stage and the availability of buffer space at the destined SE of the next stage. The  $(m+1)$  state Markov chain of the buffer is shown in Figure 3, with  $\pi_{ij}$ ,  $i, j = 0, 1, \dots, m$  being the transition probability from state  $i$  to state  $j$ .

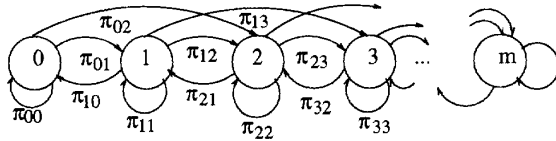


Figure 3: Markov state transition for stage buffers

Even a full buffer is able to accept a packet if a packet leaves the buffer in the same cycle. With

these considerations, the transition probabilities,  $\mathbf{\Pi} = [\pi_{ij}, 0 \leq i, j \leq m]$  at buffer  $klv$  are described by the following transition equations.

$$\mathbf{\Pi} = \begin{bmatrix} \pi_{00} & \pi_{01} & \dots & \pi_{0m} \\ \pi_{10} & \pi_{11} & \dots & \pi_{1m} \\ \vdots & \vdots & \dots & \vdots \\ \pi_{m0} & \pi_{m1} & \dots & \pi_{mm} \end{bmatrix}$$

where  $\pi_{00} = C_{klv}^0$ ,  $\pi_{01} = C_{klv}^1$ ,  $\pi_{02} = C_{klv}^2$ ,  $\pi_{03} = 0$ ,  $\pi_{10} = C_{klv}^0 r_{klv}$ ,  $\pi_{11} = C_{klv}^1 r_{klv} + C_{klv}^0 (1 - r_{klv})$ ,  $\pi_{12} = C_{klv}^1 (1 - r_{klv}) + C_{klv}^2 r_{klv}$ ,  $\pi_{13} = C_{klv}^2 (1 - r_{klv})$ ,  $\pi_{14} = 0$ . We obtain the steady state buffer state probabilities  $\mathbf{B} = [B_{klv}^0, B_{klv}^1, \dots, B_{klv}^m]$  for buffer  $klv$  by solving equations

$$\mathbf{B} = \mathbf{B}\mathbf{\Pi} \quad (10)$$

$$\text{and } \sum_{i=0}^m B_{klv}^i = 1 \quad (11)$$

The output links of the last stage buffers can always receive packets. The Markov chain remains the same as that of the intermediate stages, except that  $r_{nlv} = 1$ . We therefore obtain the steady buffer state probability vector  $\mathbf{B}$  for the  $k = n$  by substituting  $r_{nlv} = 1$  in Equation (10) and solving the Markov chain.

### 3.5 IBC Buffers

The IBC buffers are also modeled as a Markov chain. We assume the buffer as being driven by a Bernoulli process, with probability  $\rho$  of a packet arrivals during a cycle and with geometric departures. The probability of departure ( $r_{0lv}$ ) of a packet is determined by the availability of buffer space at the first stage. The  $m+1$  state Markov chain model of an  $m$ -buffer IBC buffer is shown in Figure 4. The transition

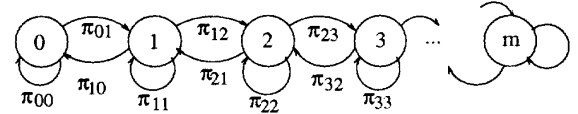


Figure 4: Markov state transition for IBC buffers

probabilities,  $\mathbf{\Pi} = [\pi_{ij}, 0 \leq i, j \leq m]$  are given by

$$\mathbf{\Pi} = \begin{bmatrix} 1 - \rho & \rho & \dots & \dots \\ (1 - \rho)r_{0lv} & (1 - \rho)(1 - r_{0lv}) + \rho r_{0lv} & \dots & \dots \\ \vdots & \vdots & \dots & \vdots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

The steady buffer state probabilities of the IBC are obtained by solving the Markov chain as in Section 3.4

### 3.6 Throughput and Delay

We use the normalized throughput ( $\mu$ ) and the delay ( $\delta$ ) as the performance criteria. Normalized throughput is defined as the number of packets leaving an output of the network during a cycle.

$$\mu = 1 - B_{nlv}^0 \quad (12)$$

The maximum normalized throughput is computed by increasing the arrival rate ( $\rho$ ) to the IBC until  $Q_{1lu}$  becomes one.  $Q_{1lu} = 1.0$  means that a packet from the IBC is ready to come to the first stage of the network at every clock cycle.

Delay ( $\delta$ ) is defined to be the number of clock cycles required by a packet to reach the destination port starting from the source port. Let  $R_k$  be the probability that a packet in the buffer of an SE in stage  $k$  is able to move forward. Therefore,

$$\delta = \sum_{k=1}^n \frac{1}{R_k} \quad (13)$$

$$\text{where, } R_k = r_{klv} \sum_{i=1}^m \left( \frac{B_{klv}^i}{1 - B_{klv}^0} \right) \frac{1}{i} \quad (14)$$

The single queue analyses are made consistent by forcing the single queue variables to yield certain known long term flows. The input traffic is described by the load matrix  $\beta = [\beta(i, j)]$ . The steady state flow at the  $kl$  switching element along  $uv$ ,  $P_{kluv}$  can easily be computed from Equation (9). The objective of the analysis is to determine the values of  $B_{klv}^i$ , the steady state probability of buffer occupancy at the  $k$ -th stage. Since the equations describing the dynamics of the network are described by recurrence relations, the solution is obtained by an iterative method.

## 4 The Memorized Model

The model described in Section 3 uses static routing probabilities for the packets at the SEs. It allows a packet at an SE to choose an output port of the SE according to the static routing probabilities. This permits a packet blocked at an SE to choose an output port, during the next cycle, which is independent of the port for which it was blocked. Consequently, the results obtained from the model in Section 3 are optimistic, since it allows a blocked packet to be routed around a congested queue. In practice, a blocked packet always hunts for the same output port during consecutive cycles, and hence does not obey the static routing probability. Moreover, a blocked packet at an SE has a higher chance of being blocked again than a new packet at the SE. To account for the above limitations of the previous model, we develop a *memorized* model which memorizes the history of a blocked packet. In this model, a blocked packet does not use the static routing probabilities for successive routing attempts. The memorized model is based on the work reported in [13].

In addition to the Markov chain representing the occupancy of a buffer (see Section 3), we use another Markov chain to define the blocking states of

the buffer. This chain memorizes whether the packet at the head of the buffer is a blocked one. The blocking status of a buffer is represented by the 3-state Markov chain shown in Figure 5. The states

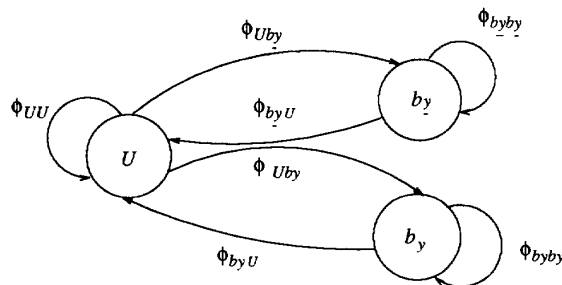


Figure 5: The states of a buffer during its busy period

are called “blocked” for buffer  $B_{klv}$  at the next stage ( $\theta_{klv}^b, v \in \{y, \underline{y}\}$ ) and “unblocked” ( $\theta_{klv}^U$ ) and represent the state of the packet at the head of the buffer at the beginning of a clock cycle. When a packet first comes into a buffer, the buffer remains in the unblocked state ( $\theta_{klv}^U$ ) and tries to route the packet. If it can be routed successfully, it remains in the unblocked state irrespective of whether the buffer becomes empty or a new packet moves to the head of the buffer. If the packet can not be routed, it enters one of the blocked state ( $\theta_{klv}^b$ ) and remains in the blocked state until the packet can be forwarded to the next stage, when it enters the unblocked state. While in the unblocked state, a packet obeys the static routing probabilities and chooses an output port according to  $P_{kluv}$ .

The transition probabilities between the three states are shown in Figure 5. For example,  $\phi_{byb_v}, v \in \{y, \underline{y}\}$  is the probability that a blocked packet is again blocked when it attempts to go to the same destination. The transition probabilities are given by

$$\Phi = \begin{bmatrix} \phi_{UU} & \phi_{Uby} & \phi_{Uby} \\ \phi_{byU} & \phi_{byby} & 0 \\ \phi_{byU} & 0 & \phi_{byby} \end{bmatrix}$$

where  $\phi_{UU} = r_{(k-1)lu}$ ,  $\phi_{Uby} = 1 - r_{kluy}$ ,  $\phi_{Uby} = 1 - r_{kluy}$ ,  $\phi_{byU} = r_{kluy}^{by}$ ,  $\phi_{byby} = 1 - r_{kluy}^{by}$ ,  $\phi_{byby} = 1 - r_{kluy}^{by}$ ,  $\phi_{byU} = r_{kluy}^{by}$ .  $r_{kluv}^{bv}$  is the probability that a blocked packet advances from  $klv$  to  $klv$ , given that a blocked packet which is destined to  $klv$  is ready to come to  $klv$ . The probability that a blocked packet in a stage can be forwarded to the next stage is given by  $r_{kluv}^{bv}, v \in \{y, \underline{y}\}$ .

Using Figure 5, the state probabilities at the end of a cycle are calculated from

$$\Theta = \Theta \Phi \quad (15)$$

where  $\Theta = [\theta_{klv}^U, \theta_{klv}^{by}, \theta_{klv}^{by}]$ . By solving the above

Markov chain, the steady state probabilities  $\theta_{klv}^{bv}$  and  $\theta_{klv}^U$  are obtained.

When a packet in stage  $(k-1)$  is in the blocked state at the beginning of cycle  $t$ , the length of the destination buffer in stage  $k$ , after the first phase of the same clock cycle will be either  $m$  (full) or  $m-1$  (since, at most one packet can leave during the first phase). The probability that a blocked packet in stage  $(k-1)$  will face the destination buffer with one place is  $\frac{B_{klv}^{m-1}}{B_{klv}^{m-1} + B_{klv}^m}$ . Moreover, the blocked packet may also be in routing conflict with the other buffer, in stage  $(k-1)$ , which feeds the same destination buffer in stage  $k$ . Considering the above two cases, we obtain  $r_{kluv}^{bv}$  as follows.

$$r_{kluv}^{bv} = (1 - Q_{kl\underline{u}} + Q_{k\underline{t}l\underline{u}}P_{k\underline{t}l\underline{u}} + 0.5Q_{k\underline{t}l\underline{u}}P_{k\underline{t}l\underline{u}}) \frac{B_{klv}^{m-1}}{B_{klv}^{m-1} + B_{klv}^m}, \quad 1 \leq k \leq n. \quad (16)$$

$r_{klv}^{bv}$  is the probability that a blocked packet in buffer  $klv$  can advance and is given by,

$$r_{klv}^{bv} = r_{(k+1)\underline{v}v}^{bv} \quad (17)$$

In the unblocked state, the queue length determines the possibility of a packet leaving the buffer. If there is a packet, the packet is routed according to the static routing probability. We now develop equations for the equivalent input rates  $Q'_{klu}$  and equivalent routing probabilities  $r'_{kluv}$ . The difference between the memorized model and the basic model is that buffer  $klv$  tries to transmit a packet to the next stage with probability  $(1 - B_{klv}^0)\theta_{klv}^U$  in the memorized model, whereas buffer  $klv$  in the basic model tries to transmit a packet with probability  $(1 - B_{klv}^0)$ . Therefore,

$$Q'_{klu} = (1 - B_{(k-1)\underline{u}}^0)\theta_{(k-1)\underline{u}}^U \quad (18)$$

We can obtain  $C_{klv}^{\prime 0}$ ,  $C_{klv}^{\prime 1}$  and  $C_{klv}^{\prime 2}$  by replacing  $C_{klv}^0$ ,  $C_{klv}^1$  and  $C_{klv}^2$  in Equations (6) - (8) by  $C_{klv}^{\prime 0}$ ,  $C_{klv}^{\prime 1}$  and  $C_{klv}^{\prime 2}$  respectively as follows.

$$C_{klv}^{\prime 0} = 1 - (C_{klv}^{\prime 1} + C_{klv}^{\prime 2}) \quad (19)$$

$$C_{klv}^{\prime 1} = Q'_{klu}P'_{kluv}(1 - Q'_{klu}) + Q'_{k\underline{t}l\underline{u}}P'_{k\underline{t}l\underline{u}v}(1 - Q'_{k\underline{t}l\underline{u}}) + Q'_{klu}Q'_{k\underline{t}l\underline{u}}(P'_{k\underline{t}l\underline{u}v}P'_{k\underline{t}l\underline{u}v} + P'_{k\underline{t}l\underline{u}v}P'_{k\underline{t}l\underline{u}v}) \quad (20)$$

$$C_{klv}^{\prime 2} = Q'_{klu}P'_{kluv}Q'_{k\underline{t}l\underline{u}}P'_{k\underline{t}l\underline{u}v} \quad (21)$$

The equivalent probability  $r'_{kluv}$  and  $r'_{klv}$  are obtained by substituting  $Q_{klx}(t)$  by  $Q'_{klu}$  in Equations (1)-(3) in Section 3 as follows.

$$r'_{kluv} = 1 - B_{klv}^m - B_{klv}^{m-1} + B_{klv}^{m-1}r_{klv} + (1 - Q'_{k\underline{t}l\underline{u}} + Q'_{k\underline{t}l\underline{u}}P'_{k\underline{t}l\underline{u}v} + 0.5Q'_{k\underline{t}l\underline{u}}P'_{k\underline{t}l\underline{u}v})(B_{klv}^{m-1}$$

$$(1 - r_{klv}) + B_{klv}^m r_{klv}), \quad 1 \leq k \leq n. \quad (22)$$

$$r'_{klv} = r'_{(k+1)\underline{u}v}P'_{(k+1)\underline{u}v} + r'_{(k+1)\underline{u}v}P'_{(k+1)\underline{u}v} \quad (23)$$

The equivalent routing probability is given by

$$P'_{kluv} = \frac{P_{kluv} + \theta_{kluv}^{bv}}{1 + \theta_{kluv}^{bv}P_{kluv}} \quad (24)$$

and can be explained as follows. When  $\theta_{kluv}^{bv} = 0$ , i.e., the server is in the "unblocked" state,  $P'_{kluv} = P_{kluv}$ . When  $\theta_{kluv}^{bv} = 1$ , i.e., the server is in the "blocked" state,  $P'_{kluv} = 1$ . The equivalent routing probability is changed from  $P_{kluv}$  to 1 according to the blocked probability. Thus, the model can store the history of blocking, i.e., the output link for which a packet was blocked. For calculating the time delay, we have

$$R'(k) = r_{klv}^b r_{klv} \sum_{i=1}^m \left( \frac{B_{klv}^i}{1 - B_{klv}^0} \right) \frac{1}{i} \quad (25)$$

Substituting Equations (16)-(24) in the basic model described in Section 3, the new set of equations for the memorized model can be obtained, which can be used to evaluate a multistage interconnection network with blocking behavior.

## 5 Results

The analytical models presented in Sections 3 and 4 can be applied to any general traffic pattern. In this section, we present results for the performance of the MIN, obtained from the two models, in the presence of uniform and hot-spot traffic patterns and compare them with simulation results. We compare the results

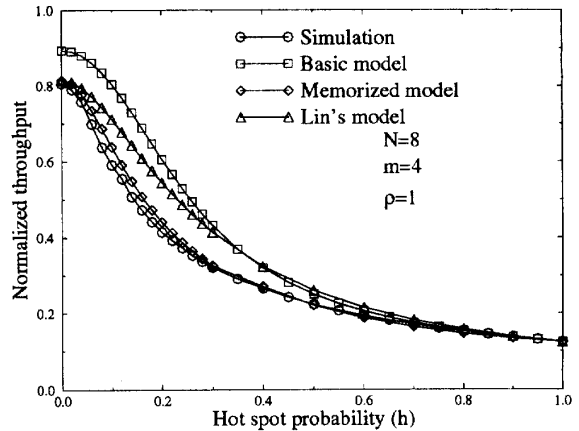


Figure 6: Comparison of throughput for  $\rho = 1$ ,  $m = 4$  and  $N = 8$ .

from the memorized and basic models with those from simulation in Figures 6-11. Figure 6 shows the normalized throughput versus the hot spot probability

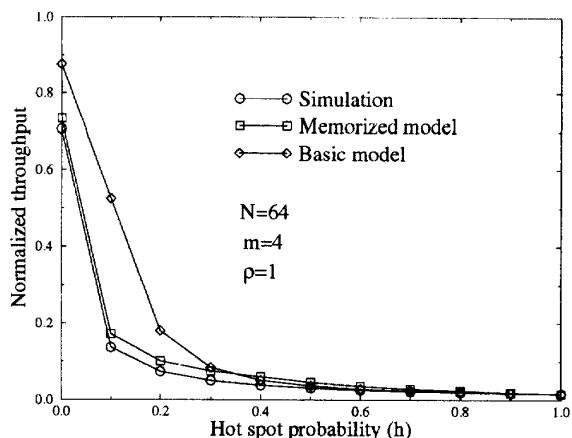


Figure 7: Comparison of throughput for  $\rho = 1.0$ ,  $m = 4$  and  $N = 64$ .

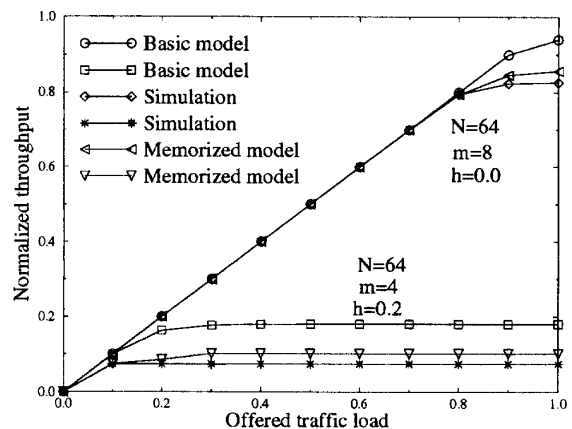


Figure 8: Throughput for different offered traffic load.

obtained from simulation, basic model, Lin's model [13], and the proposed memorized model for  $N = 8$ ,  $\rho = 1.0$ , and  $m = 4$ . When  $h = 0$ , the results from the basic model are much more optimistic than those from simulation, and the results from Lin's and memorized models show very good correspondence with simulation. As  $h$  increases from 0 to 0.5, there is a significant discrepancy between the simulation, basic model, and Lin's model. Lin's model does not store the history of blocking but partially takes into account of the blocking. Hence it produces better results than the basic model in low hot spot probabilities. The memorized model takes a rigorous account of the history of blocking and hence produces results which are close to simulation and are significantly better than those obtained from other models. It is important to store the history because a blocked packet always hunts for the same output link during successive cycles. When  $h$  increases from 0.5 to 1, the normalized throughput for all the models decreases to  $1/N$ . Figure 7 shows

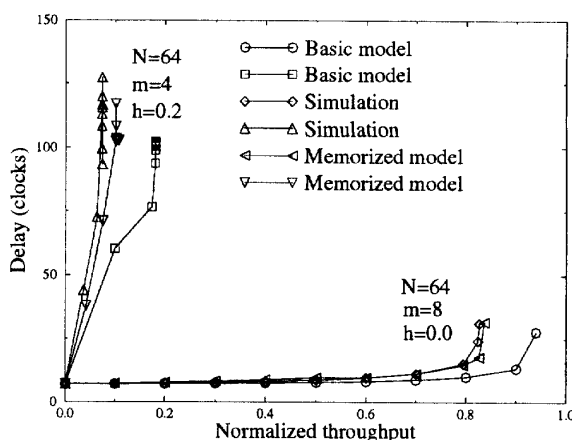


Figure 9: Mean delay vs. throughput  $\rho = 1.0$ ,  $m = 4, 8$  and  $N = 64$ .

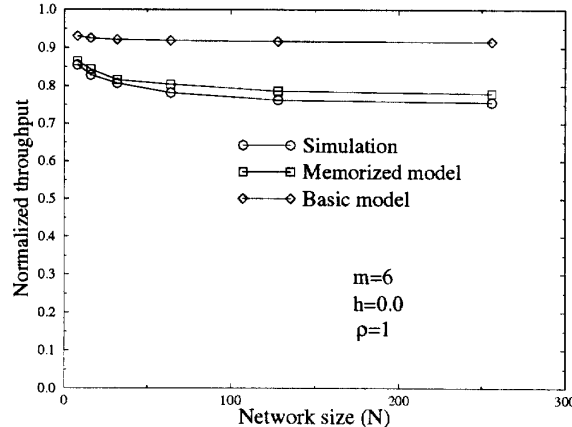


Figure 10: Variation of throughput for different network sizes with  $\rho = 1$ , and  $m = 6$ .

throughput comparison for  $N = 64$ . The results from the memorized model are found to be very accurate. Figure 8 shows the throughput for different offered traffic load with  $m = 8, h = 0$  and  $m = 4, h = 0.2$ . Throughput is plotted vs. mean delay in Figure 9. Figure 10 shows the throughput for different network sizes. The results from the memorized model are still optimistic for a large network. However the model is suitable for networks of various sizes. Figure 11 shows the throughput for various buffer sizes with  $\rho = 1$  and  $N = 64$ . It shows that the memorized model is suitable for studying the tradeoffs between the throughput and different buffer sizes.

## 6 Conclusions

Two Markov chain models have been proposed to evaluate the performance of output-multibuffered

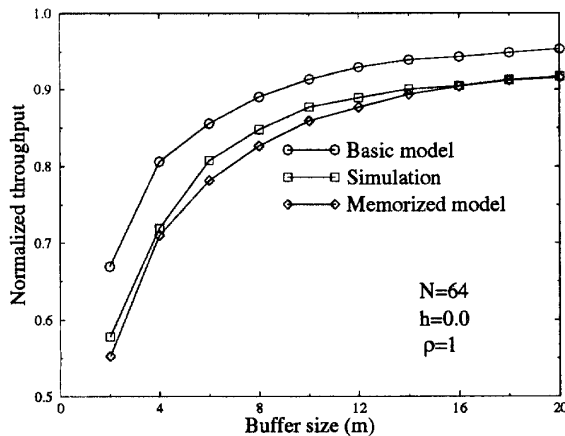


Figure 11: Variation of throughput for different buffer sizes for  $\rho = 1$  and  $N = 64$ .

MINs under a general traffic pattern. It can be applied to both uniform and nonuniform traffic patterns. Each buffer in an SE is modeled as a Markov chain, and the relationship between the SEs is described by average flow constraints. The analytic models are general enough to handle MINs with arbitrary buffer sizes and network sizes. It can be applied to other types of networks as well, such as the Banyan and Baseline networks. We have compared the analytic results to simulation results. The result have been found to be in close agreement. The basic model's low accuracy at high loads results from several independence assumptions that have been made in the model. The memorized model produces significantly better results than the basic model. The reason behind the memorized model's higher accuracy is its ability to take a rigorous account of blocking at the SEs by memorizing the output link of a SE for which a packet was blocked. Development of an analytical model for  $a_n \times a_n$  output multibuffered multistage interconnection networks by considering the correlations between consecutive clock cycles as well as the states of the buffers in the adjacent stages is currently underway.

## References

- [1] J.H. Patel, "Performance of processor-memory interconnections for multiprocessors," *IEEE Transactions on Computers*, vol. C-30, no. 10, pp. 771-780, October 1981.
- [2] M. Atiquzzaman and M.S. Akhtar, "Effect of hot spots on the performance of multistage interconnection networks," *FRONTIERS 92: The Fourth Symposium on the Frontiers of Massively Parallel Computation*, Virginia, pp. 504-505, October 19-21, 1992.
- [3] M. Atiquzzaman and M.S. Akhtar, "Effect of non-uniform traffic on the performance of unbuffered multistage interconnection networks," *IEE Proceedings Part-E*, To appear in 1994.
- [4] D.M. Dias and J.R. Jump, "Analysis and simulation of buffered Delta networks," *IEEE Transactions on Computers*, vol. C-30, no. 4, pp. 271-282, April 1981.
- [5] Y-C. Jenq, "Performance analysis of a packet switch based on single-buffered Banyan network," *IEEE Journal on Selected Areas in Communications*, vol. SAC-1, no. 6, pp. 1014-1021, December 1983.
- [6] C.P. Kruskal and M. Snir, "The performance of multistage interconnection networks for multiprocessors," *IEEE Transactions on Computers*, vol. C-32, no. 12, pp. 1091-1098, December 1983.
- [7] H.S. Kim, I. Widjaja, and A. Leon-Garcia, "Performance of output-buffered Banyan networks with arbitrary buffer sizes," *IEEE INFOCOM '91: Conference on Computer Communications*, Bal Harbour, Florida, pp. 701-710, April 9-11, 1991.
- [8] G.F. Pfister and V.A. Norton, "Hot spot contention and combining in multistage interconnection networks," *IEEE Transactions on Computers*, vol. C-34, no. 10, pp. 943-948, October 1985.
- [9] H.S. Kim and A. Leon-Garcia, "Performance of buffered Banyan networks under non-uniform traffic patterns," *IEEE Transactions on Communications*, vol. 38, no. 5, pp. 648-658, May 1990.
- [10] D.S. Meliksetian and C.Y.R. Chen, "A Markov modulated Bernoulli process approximation for the analysis of Banyan networks," *1993 ACM SIGMETRICS Conference on Measurement and Modeling of Computer Systems*, pp. 183-194, May 10-14, 1993.
- [11] M. Atiquzzaman and M.S. Akhtar, "Performance of buffered multistage interconnection networks in non uniform traffic environment," *7th International Parallel Processing Symposium*, California, pp. 762-767, April 13-16, 1993.
- [12] B. Zhou and M. Atiquzzaman, "Performance of output-multibuffered multistage interconnection networks under nonuniform traffic patterns," *International Workshop on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS'94)*, North Carolina, USA, pp. 405-406, January 31 - February 2, 1994.
- [13] T. Lin and L. Kleinrock, "Performance analysis of finite-buffered multistage interconnection networks with a general traffic pattern," *1991 ACM SIGMETRICS Conference on Measurement and Modeling of Computer Systems*, San Diego, CA, pp. 68-78, May 21-24, 1991.