Accurate Analysis of Multistage Interconnection Networks Using Finite Output-Buffered Switching Elements

Bin Zhou
Dept. of Comp. Science and Comp. Engg.
LaTrobe University
Melbourne 3083, Australia
Email: binz@latcs1.lat.oz.au

M. Atiquzzaman
Dept. of Elec. & Computer Systems Engg.
Monash University, Clayton,
Melbourne 3168, Australia
Email: atiq@eng.monash.edu.au

Abstract

Many of the existing analytical models for output buffered switching elements (SE) assume uniform traffic and infinite buffers at each output port of an SE. Moreover, because of simplifying assumptions, the results are not accurate. It is important to develop an accurate analytical model to tailor the design of the network parameters and optimize the network performance by proper dimensioning of the buffers. The objective of this paper is to develop an accurate model for the performance of MINs using finite output buffered SEs, and operating in the presence of nonuniform traffic patterns. It is shown that the proposed analytical model is much accurate than existing models.

1 Introduction

Multistage interconnection networks (MIN's) are used to connect processors and memories in large-scale scalable multiprocessor systems. They have also been proposed as ATM switching nodes in Broadband ISDN networks. A MIN (Figure 1) consists of several stages of small crossbar switching elements (SE). The overall performance of a multiprocessor system or the B-ISDN network depends heavily on the performance of the interconnection network or the switching node respectively. Several models have been proposed for the performance evaluation of output buffered MINs [1, 2]. The models are either not very accurate or assume a traffic which is uniformly distributed over the outputs of the MIN. The inaccuracies are mainly due to simplifying assumptions regarding the behavior of blocked packets in the MIN.

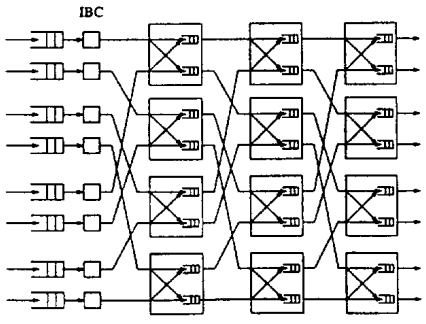

Figure 1: A three stage MIN.

In this paper, an accurate model for evaluating the performance of MINs using finite output-buffered SEs is introduced. It *differs* from existing models in the sense that it accounts for the fact that a packet which is blocked in a buffer of an SE always hunts for the same output link of the SE in subsequent clock cycles. The model takes a rigorous account of this behavior of blocked packets and also considers the dependency between packets in consecutive clock cycles and states of the buffers in adjacent stages. Due to the above considerations, it *produces results which are more accurate* than existing models. Moreover, it is *applicable to both uniform and non-uniform traffic* patterns in the network. The proposed model presented in this paper will serve as a *tool* for the network designer, and will help in obtaining a *better insight* into the performance of the network under different non-uniform traffic patterns.

In Section 2, the modeling assumptions of an output-buffered Omega network are described. The basic concepts of the model are illustrated using a two-buffered model in Sections 3. Results obtained from the proposed model and a previous model are compared with those from simulations in Section 4, followed by concluding remarks in Section 5.

2 Model Platform

We make the following assumptions regarding the operation and environment of the MIN. There are $S = 2^s$ inputs and S outputs of the MIN. Each input of the MIN has an Input Buffer Controller (IBC) of finite size. The network operates *synchronously*. 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. Depending on the availability of buffer space in the succeeding stage, a packet may move forward one stage in the second phase. A *back-*

0-7803-2018-2/95/$4.00

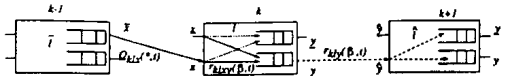

Figure 2: Illustration of $Q_{klx}(*,t)$ and $r_{kly}(\beta,t)$.

pressure mechanism ensures that no packets are lost within the MIN. Packet arrival at the inputs of the MIN are *Bernoulli* processes and are independent of each other. Each input link of the MIN is offered the same traffic *load*. There is no *blocking* at the output links of the MIN. The *conflict resolution* logic at each SE is unbiased. A packet must spend at least one clock cycle in each buffer including the IBC. If there are more than two spaces available at the destination buffer in the next stage, the buffer is assumed to be fast enough to accept both the packets in one clock cycle.

3 The Model

For the sake of simplicity, we illustrate the underlying concepts of our model using a MIN having two buffers at each output of an SE. Figure 2 shows three SEs in consecutive stages of the network. In defining the state variables, a *new* state means that the packet at the head of the queue has moved to the head of the queue in the previous clock cycle. A *blocked* state implies that the packet at the head of the queue has made at least one unsuccessful routing attempt. State 0 means the buffer is empty. State (i,n) means the buffer contains i, $1 \leq i \leq 2$, packets and the packet at the head of the queue is in a new state. State (i,b) means the buffer contains i, $1 \leq i \leq 2$, packets and the packet at the head of the queue is in a blocked state. From the definitions, it is clear that when a buffer is not empty, it will be in either state (i,n) or (i,b). When a buffer is in state (i,b), the packet at the head of queue either lost the contention among the packets competing for the same output link or the receiving buffer at the succeeding stage was full. When a buffer is in state (i,n), the packet can be destined to any one of the outputs in the next stage.

When there is a packet in a buffer, the probability that this packet can move forward during this cycle will strongly depend on the routing history during the previous cycle. Models which consider this dependency between consecutive clock cycles will have higher accuracy.

The state of a buffer at the beginning of phase 1, at the end of phase 1, and at the end of phase 2 of a stage cycle will be called the *initial* state, the *intermediate* state, and the *final* state respectively.

3.1 Buffer State Transition

Figure 3 shows the possible state transitions from the initial states of a buffer to intermediate states and from intermediate states to final states for the buffers in the intermediate stages, i.e. the stages other than the first and the last. Some of the state transitions in Figure 3 are explained as follows.

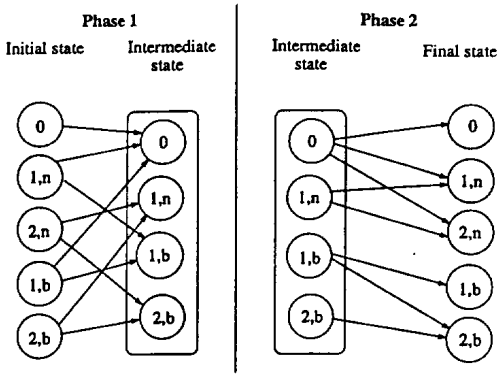

Figure 3: Markov state transition diagram.

A buffer at initial state (i,n) may go to intermediate state $(i-1,n)$ if the destined buffer at the next stage allows a packet at this stage to move forward, and the packet can win any possible conflict with another packet in the same stage or may go to state (i,b) if the destined buffer at the next stage does not accept a packet or the packet loses any possible conflict. A buffer at initial state (i,b) may go to intermediate state $(i-1,n)$ if the destined buffer at the next stage allows a packet at this stage to move forward and the packet can win any possible conflict with another packet in the same stage, or may stay at state (i,b) if the destined buffer at the next stage does not allow a packet at this stage to move forward or the packet loses any possible conflict. A buffer may go from intermediate state 0 to final state 0, $(1,n)$ or $(2,n)$ if none, one or two packets are offered to the buffer respectively. A buffer from intermediate state $(1,n)$ may go to final state $(2,n)$ if one or two packets are offered to the buffer, and stay in the final state $(1,n)$ if no packet is offered. A buffer may go from intermediate state $(1,b)$ to final state $(2,b)$ if one or two packets are offered to the buffer, or stay in state $(1,b)$ if no packet is offered. A buffer at state $(2,b)$ will stay in the original state in all cases. The balance equations from the state transition diagram are given in [3].

3.2 Throughput and Delay

The normalized throughput of the MIN at the output is given by the number of packets leaving an output port during a clock cycle. Packet delay is defined to be the number of clock cycles taken by a packet to reach the destination port from the source port. The model is described in detail in [3].

4 Results

In this section, we investigate the accuracy of the model described above. The results from the proposed model are compared with the basic model [2] (which is similar to the analysis in [1]

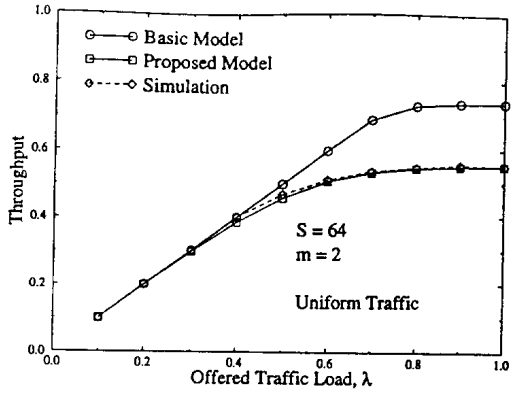

Figure 4: Throughput versus offered load.

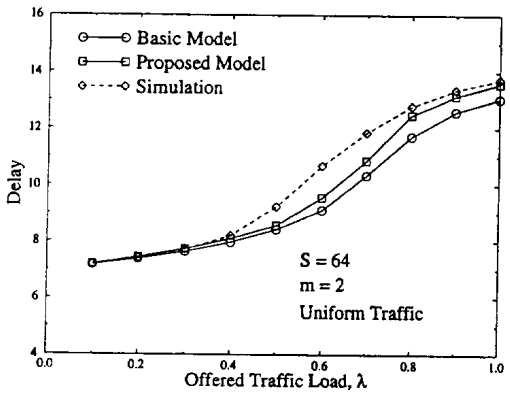

Figure 5: Delay versus offered load.

when the network is operated under uniform traffic) and simulations. In the figures, λ is the offered traffic load, λ_{sssd} is the offered load of the single source to single destination (SSSD) and m is the buffer size [2]. Due to space limitations, we only present the results of the MIN with buffer two and operating under uniform and SSSD traffic patterns. The simulation methodology is described in [4, 2] and is omitted here for the sake of brevity. Figures 4 and 5 show the normalized throughput and delay of a two buffered Omega network of size 64×64 under uniform traffic. It is found that the accuracy of all the models are very good when the offered traffic load is small. However, the basic model is inaccurate under heavy traffic loads because it does not take a rigorous account of blocked packets. The proposed model produces very accurate results because it accurately models the behavior of blocked packets. The influence of the size of the network on the performance is shown in Figure 6. Since every additional stage of the network introduces further collisions, the maximum throughput decreases when the size of the network increases. It is seen that the basic model is much less accurate than the proposed model.

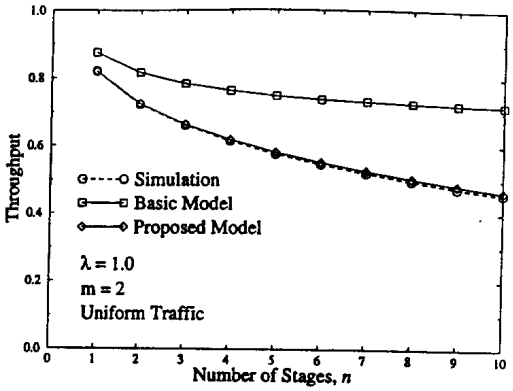

Figure 6: Throughput versus network size.

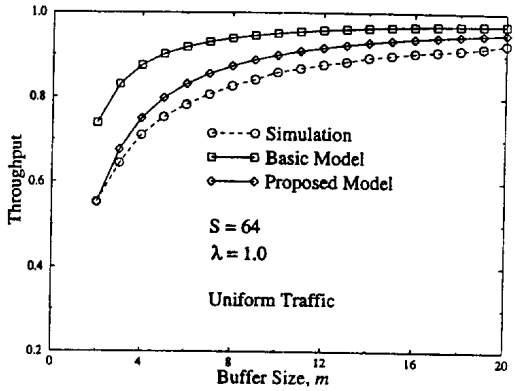

Figure 7: Throughput versus buffer size.

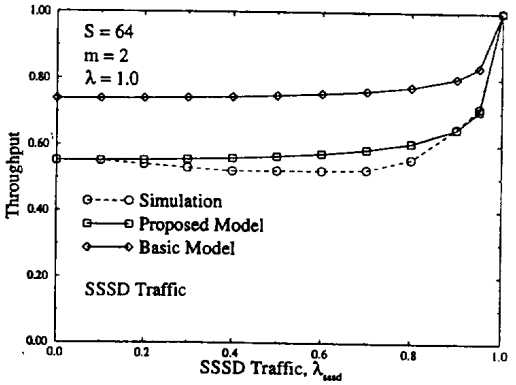

Figure 8: Throughput versus SSSD traffic for $N = 64$ and $m = 2$.

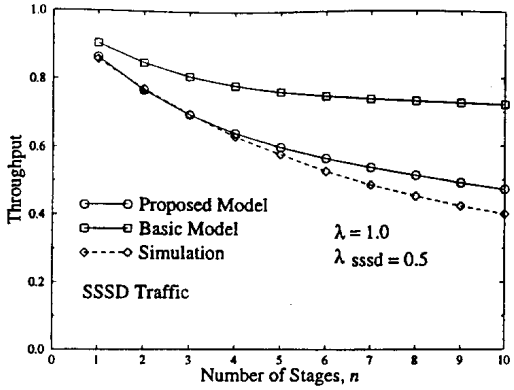

Figure 9: Throughput versus network size for SSSD traffic.

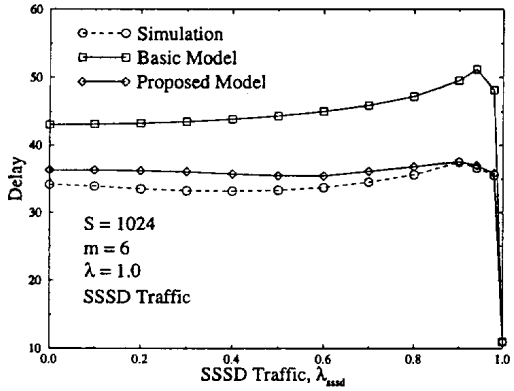

Figure 10: Delay versus SSSD traffic $N = 1024$, $m = 6$.

Figure 7 shows the effect of buffer size on the throughput under uniform traffic patterns. It is seen that the multi-buffer extension of the proposed model is much more accurate than the basic model, and buffer sizes greater than six do not result in a significant increase in the throughput. Figure 8 shows the normalized throughput versus the SSSD traffic pattern. Since blocked packets are taken into account, the results from the proposed model are closer to simulation than those from the basic model. The influence of the size of the network on the performance under SSSD traffic is shown in Figure 9. Since every additional stage of the network introduces further collisions, the maximum throughput decreases when the size of the network increases. It is seen that the basic model is much less accurate than the proposed model. Figure 10 shows the normalized delay for a large network size ($S = 1024$) under SSSD traffic load. Note that $\lambda_{sssd} = 0$, corresponds to the uniform traffic load. The delay decrease with an increase in λ_{sssd} until $\lambda_{sssd} = 0.6$ when the delay increases. When $\lambda_{sssd} = 1$, the uniform background traffic becomes zero, the delay drops to minimal. The reason is that increased offered λ_{sssd} traffic causes more packet blocking in the SEs under lower λ_{sssd} load. But when λ_{sssd} load increases to high, the conflicts between different input in a SE is reduced, and no conflict arises at $\lambda_{sssd} = 1$. Therefore, there are two factors which are influencing the performance of the switch. First, the tree saturation due to the background traffic, which has a negative effect on the delay. The second factor is that there is no internal routing conflict inside the switch in the case of purely SSSD traffic, which has a positive effect on the delay. The net result of the above two factors determines the normalized delay.

5 Conclusions

A new analytic model for MINs using finite output-buffered SEs is introduced in this paper. The higher accuracy of the model has been illustrated using SEs with two buffers per output. The model can be easily extended to multiple buffers per output of an SE [3]. The proposed model produces accurate results for both uniform and nonuniform traffic patterns. The proposed model demonstrates that the blocking behavior of packets, the correlation of packet movement between subsequent clock cycles, and the states of the buffers of two adjacent stages need to be carefully taken into consideration in order to obtain accurate results from Markov chain-based analytical models. The proposed model is general enough to analyze MINs with arbitrary buffer sizes and traffic patterns.

References

[1] 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 1991.

[2] B. Zhou and M. Atiquzzaman, "Performance of output-multibuffered multistage interconnection networks under general traffic patterns," *IEEE INFOCOM '94: Conference on Computer Communications*, Toronto, Canada, pp. 1448–1455, June 14-16, 1994.

[3] B. Zhou and M. Atiquzzaman, "Efficient analysis of multistage interconnection networks using finite output-buffered switching elements," Tech. Rep. 14/94, La Trobe University, Melbourne, Department of Computer Science, July 1994.

[4] B. Zhou and M. Atiquzzaman, "Impact of switch architectures on the performance of multistage interconnection networks," *IEEE TENCON: Region 10's Ninth Annual International Conference*, Singapore, pp. 365–369, August 22-26, 1994.