Performance Modeling of Multistage ATM Switching Fabrics

Bin Zhou
Dept. of Computer Science
La Trobe University
Melbourne 3083, Australia.
binz@LATCS1.lat.oz.au

M. Atiquzzaman
Dept. of Electrical & Computer Systems Engg.
Monash University
Clayton, Melbourne 3168, Australia.
atiq@LATCS1.lat.oz.au

ABSTRACT

In this paper, a new and efficient Markov chain model for evaluating the performance of ATM switching fabrics based on multistage switches with 2×2 finite output-buffered SEs is proposed. By comparing the results obtained from the proposed model, previous models and simulations, it has been shown that the proposed model is much more accurate than previous models in the presence of a non-uniform traffic in the switch. The results from the previous models are unsatisfactory in the presence of an increased blocking in the switch arising due to a non-uniform traffic in the switch. On the contrary, the proposed model is very robust even under severe blocking in the switch.

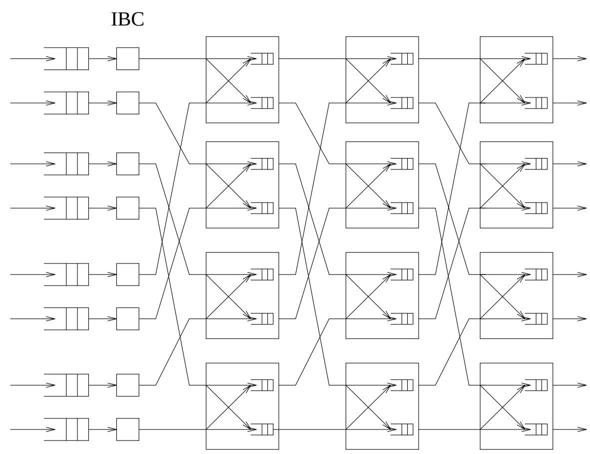

Figure 1: A three stages ATM switch fabric.

1 INTRODUCTION

Asynchronous Transfer Mode (ATM) has been proposed as a switching principle for the integrated broadband communication networks. To implement such an ATM network, high speed switching fabrics with high capacity are required. Self-routing multistage switches (MS) have been proposed as the switching fabrics in Broadband ISDN networks. An MS (Figure 1) consists of several stages of small crossbar switching elements (SE) connected by a permutation function. Buffer are often used in the SEs to increase the throughput of the switch and prevent internal loss of data. Performance of unbuffered and buffered multistage switches have been studied extensively in the literature [1, 2, 3].

To optimize the performance of an ATM network, it is vital to evaluate the performance of the switches, used in the network, under various traffic patterns and loads. It is also important to properly dimension the buffers in the switch to achieve an acceptable cell loss probability. Jenq [1] used a two-state Markov chain model consisting of the empty and new states to model the single input buffer of an SE. By iteratively evaluating a set of probabilistic formulas, he was able to obtain the latency and throughput of an MS constructed from single input buffered SEs. Kim and Garcia [2] extended Jenq's analysis to non-uniform traffic patterns.

Dedicated input buffers have a low buffer utilization in the case of a non-uniform traffic pattern. Moreover their throughput is limited due to the head-of-line blocking at the buffers. Shared buffer switches [4, 5] have a very high buffer utilization, but suffer from a phenomenon called *buffer hogging* in the case of a non-uniform traffic pattern. Additionally, buffer management in shared buffer switches is more complex and difficult to implement.

Due to the limited throughput caused by head-of-line blocking in input buffered SEs and complex buffer management requirements in shared buffer switches, a few analytical models have been proposed for MSs consisting of output-buffered SEs [6, 7, 8, 9]. Most of the models are based on the assumption of a traffic which is uniformly directed over the outputs of the switch. Moreover, they have not fully considered the blocking behavior of cells. Several non-uniform traffic patterns, viz. hot spot [10], favorite output [11] and single-source-to-single-destination [2] have been reported in the literature. Because of simplifying assumptions regarding blocked cells, the previous models are not accurate in the presence of severe blocking in the switch. In this work, we develop an analytical model which takes into account the blocking behavior of cells. The *objectives* of this work are as follows:

- To develop a highly accurate analytical model which can be applied to evaluate the performance of MSs with *multiple* buffers at SE *outputs* under a *general traffic* pattern, particularly in the presence of a hot spot traffic. Uniform traffic can also be analyzed by the model since the uniform traffic is a special case of the hot spot traffic.
- To compare the results obtained from the proposed model with those from the previous models for both uniform and hot spot traffic patterns.
- To check the validity of the proposed model by comparison with simulation results.

The main *contribution* of the proposed model lies in its ability to take a rigorous account of the behavior of blocked cells, resulting in more accurate results than previous models. The model can be used by switch designers to study the tradeoffs in varying the different switch parameters to arrive at a cost-effective design.

2 MODELING ASSUMPTIONS

We make the following assumptions regarding the operation and environment of the switch.

1. There are $N = 2^n$ inputs and N outputs of the switch. Each input of the switch has an Input Buffer Controller (IBC) of size m.
2. The switch operates *synchronously.*
3. A *backpressure* mechanism ensures that no cells are lost within the switch.
4. A buffer supports simultaneous *enqueueing* and *dequeuing* of cells during the same cycle.
5. Cell arrival process at each input of the switch is a simple *Bernoulli* process.
6. There is no *blocking* at the output links of the switch.
7. The *conflict resolution* logic at each SE is unbiased.
8. A cell spends at least one cycle in each buffer even when there is no waiting.
9. For a uniform traffic pattern, an incoming cell is equally likely to be directed to any output link. For a *hot spot traffic* pattern, the probability that an incoming cell 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.

3 THE MODEL

Fig. 1 shows a three stages ATM network fabrics. The state of a buffer is represented by a tuple (n, s) where n is the number of cells in hte buffer and s denotes the state of the cell at the head of the queue. A cell which has arrived at the head of the queue in the previous cycle is in a *new* state. The state of a cell which has been *blocked* in a previous routing attempt to the upper or lower output of the SE in the next stage will be represented by $\underline{y}$ or y respectively. The reader is referred to [12] for detail about the model. The model assumes a two phase operation of the buffer. Cells leave the buffer during the forward phase, and cells arrive during the receive phase. The state of a buffer at the end of the forward phase is called the *intermediate* state.

Figure 2 shows the possible state transitions from the initial states of a buffer to intermediate states, and from intermediate states to final states. To reduce the complexity of the figure, the state diagram is shown for a buffer of size two. If the size of a buffer is m, then there exist $3m$ possible states as follows. From initial state to intermediate state, a buffer at state (i, n) may go to state $(i-1, n)$ if the destined buffer at the next stage allows a cell at this stage to move forward, or to state $(i, \underline{y})$ if the destined upper buffer at the next stage does not allow a cell at this stage to move forward, or to state (i, y) if the destined lower buffer at the next stage does not allow a cell at this stage to move forward.

A buffer at state $(i, \underline{y})$ may go to state $(i-1, n)$ if the destined buffer at the next stage allows a cell at this stage to move forward, or may stay at state $(i, \underline{y})$ if the destined buffer at the next stage does not allow a cell at this stage to move forward. Similar reasoning applies for state (i, y).

From the intermediate state to the final state, a buffer at state 0 may go to state 0, $(1, n)$ and $(2, n)$ if none, one, or two cells respectively are offered to the buffer. A buffer at state $(1, n)$ may go to state $(2, n)$ if one or two cells are offered to the buffer and stay $(1, n)$ if no cell is offered.

A buffer at state $(1, \underline{y})$ and $(1, y)$ may go to state $(2, \underline{y})$ and $(2, y)$ respectively if one or two

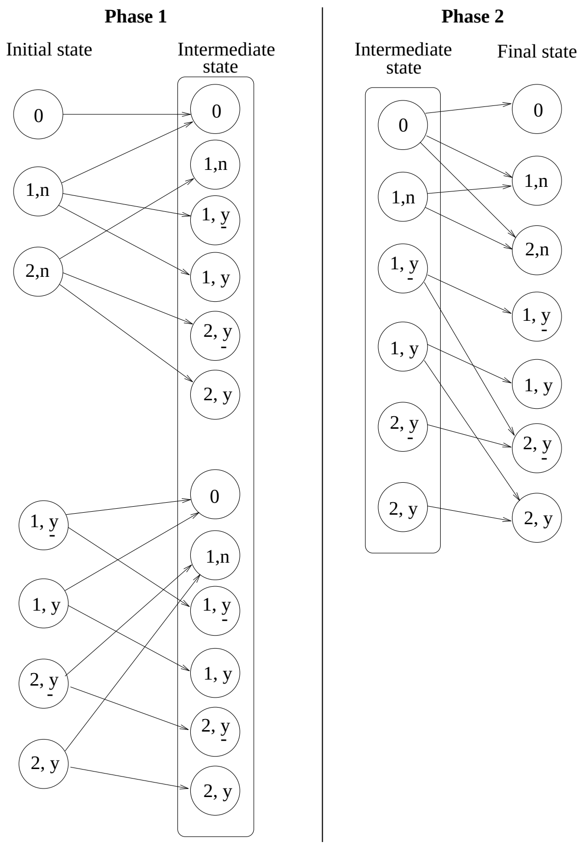

Figure 2: Markov state transition diagram for an SE having two buffers.

cells are offered to the the buffer, or may stay in the original state if no cell is offered. A buffer at state $(2, \underline{y})$ and $(2, y)$ will stay in the original state any way.

4 RESULTS

To validate the models presented in Section 3 we performed simulation of the switch using finite output-buffered SEs. The simulation methodology is described in [8, 13]. In this section, we investigate the accuracy of the proposed multi-buffer model by comparing the results with the basic model [8] (which, under a uniform traffic pattern, is similar to the model in [6]), the blocked model [14] (which uses three states per buffer and takes account of the blocked cells but does not store the history of blocking) and simulation results. The continuous lines in the figures represent analytical results whereas dashed lines are simulation data.

The effect of the buffer size on the throughput under uniform and hot spot traffic patterns are shown in Fig. 3. In the presence of a uniform traffic, an increase in the buffer size results in the normalized throughput increasing from a minimum value of 0.553 ($m = 2$) to an asymptotic value 0.92 ($m = 20$) for simulation and 0.95 ($m = 20$) for the blocked and proposed models, and from 0.74 up to 0.97 for the basic model. Due to a tree saturation in the presence of a hot spot traffic, the throughput reaches an asymptotic value of 0.61 from $m = 4$. Also, a very good match is found between the results from the proposed model and the simulation in the presence of a hot spot traffic. It is seen that buffer sizes greater than six do not significantly improve the throughput. Therefore, in the rest of this section, we investigate the performance of our proposed multi-buffer model for a buffer size of six.

Figure 4 plots the normalized throughput and delay obtained from the three different models and simulation for a 64×64 switch using six buffers per SE and operating under a uniform traffic. All the models are very accurate when the offered traffic load is small ($\lambda < 0.7$). However, if the load is high, the basic model is inaccurate since it does not consider the blocked cells. The results of the blocked model and the proposed model are identical, both of them producing very accurate results because of accounting for the blocked cells.

Figure 5 is similar to Figure 4 except that the traffic is due to a hot spot. It is observed that for the errors in the basic and blocked models rise drastically with an increase in the offered traffic load. Note that an increased offered traffic load causes blocking at all the SEs uniformly, while an increased hot spot traffic causes blocking at the hot SEs giving rise to tree saturation. Since the basic model does not account for blocked cells, the errors increase with an increase in blocking at the SEs as the load and the hot spot traffic increases. Blocked cells are partially taken into account in the blocked model, and hence the results are closer to simulation than those from the basic model. However, the blocked model does not store the history of blocking, i.e. the output link for which a cell was blocked. It is important to store this history because a blocked cell always hunts for the same output link during successive cycles.

In the case of a severe blocking in the switch (tree saturation) due to a hot spot traffic, the proposed model produces results which are very close to simulation and are significantly better than those obtained from the basic or blocked models. The reason for the improved results from our proposed model is the fact that the model memorizes the history of a blocked cell, whereas the blocked model does not memorize the history and the basic model does not even consider blocked cells.

A uniform traffic causes less blocking in the SEs than a hot spot traffic. Consequently, the accu-

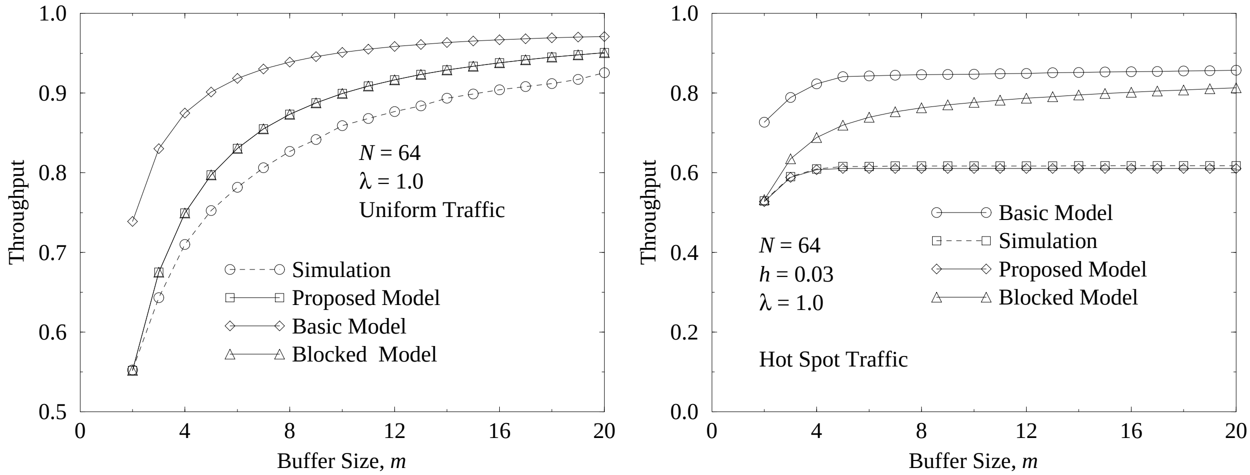

Figure 3: Throughput versus buffer size under uniform and hot spot traffic for $N = 64, \lambda = 1$.

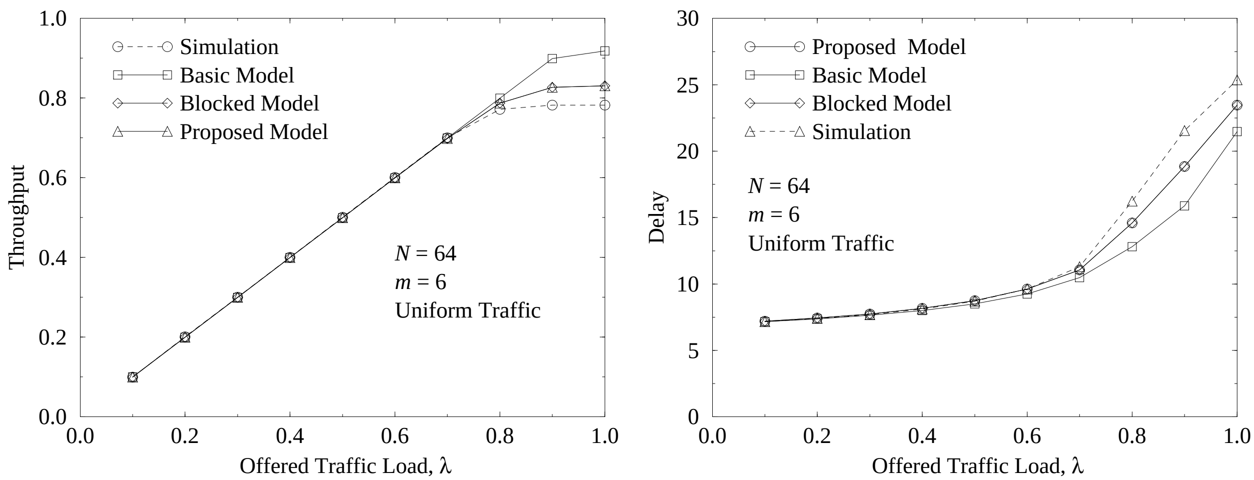

Figure 4: Throughput and delay versus offered load for $N = 64, m = 6$.

racy of the basic and blocked models are better in the presence of a uniform traffic than a hot spot traffic as is evident from Figs. 4 and 5. The blocked model and the proposed model produce similar results for uniform traffic. This is because the inaccuracy in the blocked model does not show up until there is a significant blocking in the switch. The effect of taking a rigorous account of blocking in the proposed model is reflected in an increased accuracy in the results in the presence of a significant blocking in the switch. In the presence of a hot spot traffic, the blocked model performs the same as the proposed model only when the blocking in the switch is insignificant. As an example, the blocked model produces good results for $0 \leq \lambda \leq 0.6$ and $h = 0.03$ (see Fig. 5), because of insignificant blocking for this range of offered traffic load. However, if the traffic load increases, the blocked model produces large errors even for the above hot spot traffic.

Fig. 6 shows the effect of buffer size on the cell loss probability for 64×64 and 1024×1024 switches operating under a uniform traffic. Due to excessive computing time required to simulate large switches, only results obtained from the proposed model are shown. When the offered load is below 0.4 for a 1024×1024 switch, the cell loss can be kept low (below 10^{-6}) if the buffer size is six. If such a low cell loss is desired at an offered load of 0.8, the buffer size has to be increased to 18. The effect of buffer size on the cell loss probability for various values of hot spot traffic is shown in Fig. 7. Note that in the case of a uniform traffic, the cell loss can be decreased by increasing the buffer size. However, in the presence of even 0.1% hot spot traffic, the increase in buffer size has no effect after $m = 10$. This is due to tree saturation in the case of a hot spot traffic, and further a increase in the buffer size only delays the onset of tree saturation.

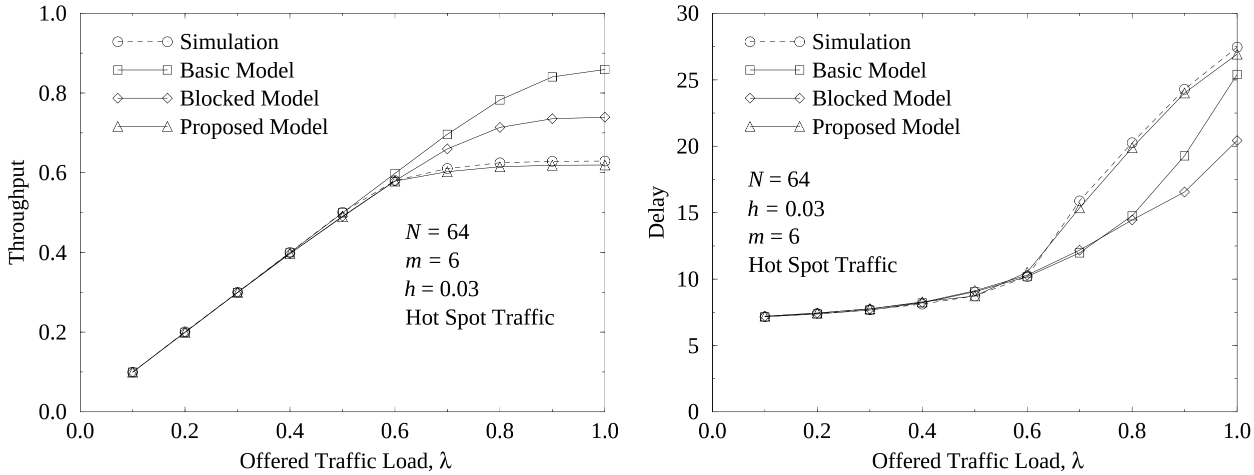

Figure 5: Throughput and delay versus offered load under hot spot traffic and $N = 64, m = 6$.

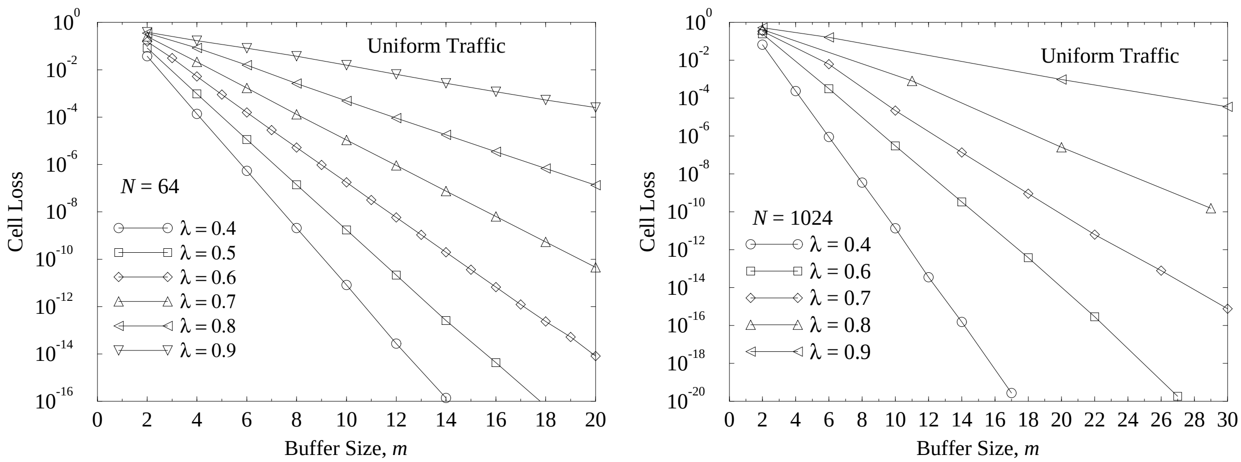

Figure 6: Cell loss versus buffer size under uniform traffic for $N = 64$ and 1024.

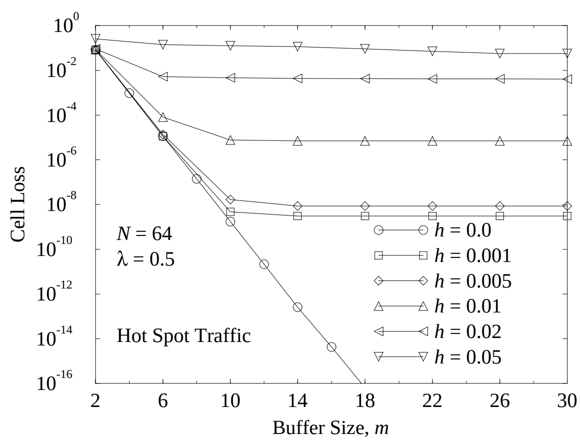

Figure 7: Cell loss versus buffer size under different hot spot traffic for $N = 64$, $\lambda = 0.5$.

5 CONCLUSIONS

Analytical models for the performance evaluation of multistage ATM switching fabrics with 2×2 finite output-buffered switching elements under uniform and hotspot traffic patterns have been developed. The models can be used to study the performance of fabrics in terms of throughput, delay and cell loss probabilities for arbitrary buffer size and switch sizes.

Earlier analytical models have, in most cases, assumed input buffered SEs. Some studies have been carried out for switches using output-buffered SEs. The models assume uniform traffic at the switch inputs or are inaccurate because of the simplicity in the assumptions. The main *contribution* of this work is the ability of the proposed model to take account of the blocked cells by storing the history of blocking. Moreover, it realistically models the transaction between buffers in adjacent stages and cycles. Results from the proposed model have been compared to simulation and previous models for different buffer and switch sizes under both uniform and hot spot traffic patterns. In the presence of a hot spot traffic, the proposed model produces significantly better results than those obtained from previous models. For a uniform traffic (which is a special case of hot spot traffic), the results obtained from the proposed model are the same as those from the best previous model. Uniform traffic does not cause as much blocking as hot spot traffic. Consequently, the results obtained from the proposed model in the presence of uniform traffic are similar to those obtained from the best previous model. It is only in the presence of considerable blocking in the switch that the inaccuracies of previous models become apparent. The reason behind the proposed model's higher accuracy in the presence of a hot spot traffic is its ability to take a rigorous account of blocking at the SEs by memorizing the output link of an SE for which a cell was blocked. The proposed model can be easily expanded to other switches having multistage topologies.

References

[1] Y. C. Jenq, "Performance analysis of a packet switch based on single-buffered banyan network," *IEEE J. Select. Areas Commun.*, vol. SAC-1, pp. 1014–1021, Dec. 1983.

[2] H. S. Kim and A. L. Garcia, "Performance of buffered banyan networks under nonuniform traffic patterns," *IEEE Trans. Commun.*, vol. 38, no. 5, pp. 648–658, May 1990.

[3] M. Atiquzzaman and M.S. Akhtar, "Effect of non-uniform traffic on the performance of unbuffered multistage interconnection networks," *IEE Proceedings - Computer and Digital Techniques*, vol. 141, no. 3, pp. 169–176, May 1994.

[4] M. Saleh and M. Atiquzzaman, "Queueing analysis of shared buffer switches for ATM networks," *GLOBECOM 94: IEEE Global Telecommunications Conference*, San Francisco, California, Nov 27 - Dec 1, 1994.

[5] J.S. Turner, "Queueing analysis of buffered switching networks," *IEEE Transactions on Communications*, vol. 41, no. 2, pp. 412–420, February 1993.

[6] 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.

[7] T. Lin and L. Kleinrock, "Performance analysis of finite-buffered multistage interconnection networks with a general traffic pattern," *ACM SIGMETRICS Conference on Measurement and Modeling of Computer Systems*, pp. 68–78, May 1991.

[8] 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.

[9] C. Kruskal and M. Snir, "The performance of multistage interconnection networks for multiprocessors," *IEEE Trans. Comput.*, vol. C-32, pp. 1091–1098, Dec. 1983.

[10] G.F. Pfister and V.A. Norton, "Hot spot contention and combining in multistage interconnection networks," *IEEE Trans. Comput.*, vol. C-34, no. 10, pp. 943–948, Oct. 1985.

[11] J. Ding, "Nonuniform traffic analysis of multistage interconnection networks with split buffers," *IEEE ICC'93*, Geneva, Switzerland, pp. 58–62, May 23-26, 1993.

[12] B. Zhou and M. Atiquzzaman, "Performance modeling of multistage ATM switching fabrics," Tech. Rep. 19/94, La Trobe University, Melbourne, Department of Computer Science, Sept. 1994.

[13] 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, Jan. 31-Feb. 2, 1994.

[14] 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.