Performance of Shared Buffer Switches Under Non-Uniform Traffic Pattern

Mahmoud Saleh
Dept. of Computer Science
La Trobe University
Melbourne 3083, Australia.
saleh@cs.latrobe.edu.au

Mohammed Atiquzzaman
Dept. of Electrical & Computer Systems Engg.
Monash University
Clayton, Melbourne 3168, Australia.
atiq@cs.latrobe.edu.au

ABSTRACT

ATM switches based on shared buffering are known to have better performance and buffer utilization than input or output queued switches. Although shared buffer switches do not suffer from head of line blocking which is a common problem in simple input buffering, their performance is adversly affected in the case of unbalanced traffic at the inputs of the switch. In this paper, we investigate the performance of a shared buffer ATM switch in the presence of a hot spot traffic pattern.

1 INTRODUCTION

In recent years, broadband ISDN (BISDN) has received increasing attention for its capability to provide a wide variety of services like video, audio, and data communications. One of the most promising approaches for BISDN is the Asynchronous Transfer Mode (ATM) which is based on fixed length packets called *cells*, and is characterized by simplified protocols, high speed links, and high capacity switching nodes. Multistage switches (MS) have attracted a great deal of attention as strong candidate for ATM switches, due to their self-routing capability and suitability for VLSI implementation. Performance of multistage ATM switches have been widely studied in the literature ([1, 2, 3, 4, 5] and references therein).

Turner [6] developed a model for shared buffer ATM switch. The model assumes that the input traffic is uniformly destined to all of the outputs, and uses backpressure to avoid cell loss inside the switch. In [7] we developed a similar model for *global flow control*. Gianatti [8] also developed an analytical model for shared buffer MS under unbalanced traffic pattern. However, for simplicity, the model assumes that a switching element contains either *hot* or *cold* traffic which results in a group of the output links being hot. In this paper we study the throughput, delay and buffer occupancy of shared buffer multistage ATM switches with single hot output. Single hot spot seems to happen often in telecommunications as well as computer communication when a particular output attracts more traffic than other outputs for a short time or continually. We show through simulation the performance degradation in a shared buffer switch due to *buffer hogging* and *tree saturation*.

In Section 2, we describe the construction of a shared buffer ATM switch and the operating assumptions under which we analyze the performance of the switch. The simulation methodology and the structure of the simulator is given in Section 3. Performance results, obtained from the simulator, and their analysis are presented in Section 4. Concluding remarks and suggestions for future work are outlined in Section 5.

2 DELTA SWITCH

A multistage switch consists of a number of stages of small switching elements (SEs) which are interconnected by a permutation function. A subgroup of multistage switches, known as Delta-d switch, consists of n stages of $d \times d$ SEs where $d^n = N$ is the number of inputs or outputs of the switch. In a shared buffer Delta switch, there are B buffer spaces in each SE which are shared among all of the inlets and outlets of the SE, i.e., a cell at any inlet may be placed in any buffer slot, and from there, it can be destined to any outlet of the SE.

In a Delta switch with hot spot traffic, one of the outputs known as the *hot* output receives cells with higher probability than the other *cold* outputs. This results in an unbalanced traffic inside the switch which different types of uniform or mixed traffic at different stages. An example of a Delta switch with hot spot traffic is given in Fig. 1.

2.1 Assumptions

The following assumptions are made regarding the switch and its operation:

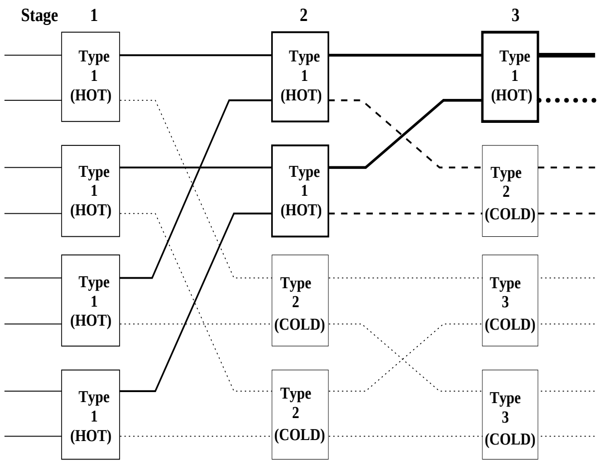

Figure 1: A Delta- 2 switch of size 8 with different types of traffics at different stages

- Each SE is of size $d \times d$ and contains B buffers which are shared by the d inlets and d outlets of the SE.

- The switch operates *synchronously*, i.e. the cells are submitted to the switch at the beginning of the time slots.

- *Destination tag* is used to route a cell. The conflict inside the switch is resolved *randomly*, i.e. if two or more cells are destined to the same output, one is chosen at random.

- The arrival of cells at each input of the switch is a *Bernoulli* process, i.e., the probability that a cell arrives during a time slot (ρ) is constant, and the arrivals are independent of each other.

- The probability of a cell arriving at a switch input and being destined to the *hot* output (p_h) or to the $N-1$ *cold* outputs (p_c) are given by:

$$p_h = \rho \left(h + \frac{1-h}{N} \right), \quad p_c = \rho \left(\frac{1-h}{N} \right) \quad (1)$$

  where h is the fraction of the traffic directed to the hot output of the switch.

- A *backpressure* mechanism with global flow control ensures that no cell is lost inside the switch. Thus, a cell leaves an SE if there is a space for it in the next stage's SE, or if a space becomes available during the same cycle. An acknowledgment policy is used to advise the receipt of a cell in the next stage's SE. Unacknowledged cells contend with other cells in subsequent cycles.

- There is no *blocking* at an output of the switch, i.e., an output can always accept a cell.

3 SIMULATION STUDY

We study delta switches with single hot spot output through a simulation program written in C language. The program can simulate delta switches with an arbitrary switch size, SE size, and buffer size depending on available memory of the computer on which it runs.

3.1 Structure of the Simulator

Our simulator is working based on the following procedures:

- At each cycle, a cell is generated with probability ρ (offered load to the switch input), and according to the given hot spot value. The generated cell is independent of the cells generated at previous cycles and other input ports. Each cell consists of the following information

  1. a source tag which represents the input link at which the cell arrived,
  2. a destination tag representing the output link to which the cell is destined, and
  3. the current clock, used for measurement of the cell delay through the switch.

- Simulation results from the first 500 cycles of the switch operation are ignored to allow the switch to reach a steady state condition. The simulation is then allowed to run until the change in the average throughput between the cycles becomes less than 10^{-6}.

- Conflict in the buffers for accessing a particular outlet as well as contention to seize a buffer space in the next stage is resolved using a random number generator with a different seed value from that of the cell generator.

The switch operates as follows:

1. The cells at the last stage's buffers are sent to the output links of the switch, and the instantaneous throughput and delay are measured for every link.

2. For each stage from stage $k-1$ to 1:

   - The SE's buffer is examined for cells destined to the different outlets of the SE, a copy of all cells destined to different outlets are placed in different lists, and the lists are sent to the corresponding inlets of the next stage.

- If the number of available buffer spaces in an SE is less than the number of cells in the different lists at the inlets to that SE, a number of cells equal to the number of available spaces are chosen at random from the available lists. Cells which are not accepted stay in the buffers at the previous stage until they can be forwarded in subsequent cycles.

3. A new set of cells are generated at the inputs of stage 1 with probability ρ, and cells are placed in the first stage's buffers if there is any room. Unaccepted cells are discarded.

It should be noted that there is no difference between a cell which is destined to the hot output and a cell destined to other outputs inside the SEs. Hot spot value contributes only at the time the cell is being generated.

4 RESULTS AND DISCUSSION

In this section we present the numerical results based on the assumptions in the previous section. We have mainly used the results from delta switches of 64 and 256 inputs and outputs. In the figures, N is the switch size, d is the number of input/output links of each switching element (SE), B is the buffer size, and h is the hotspot value. Fig. 2 illustrates the impact of hotspot traffic on the overall performance of a shared buffer ATM switch. As observed from the figure, for a particular switch size, the degradation in the presence of hotspot traffic is more severe when $\frac{B}{d}$ and/or the size of SE increases, which means that the proportion of throughput gain to hardware expenditure decreases more rapidly. Fig. 3 shows the average delay for the same values as in Fig. 2. The difference between the delay in the uniform input traffic and hot spot traffic is maximum at input load of around 0.7. When the input load approaches 1.0, the delay will not increase significantly. This is because the buffers at the SEs become saturated and the switch forwards the cells at some constant rate at which saturation took place. The effect of hot spot traffic on buffer occupancy in the hot SE of each stage is shown in Fig. 4 and Fig. 5 for $N = 256$ and different SE sizes and hot spot values. Under uniform traffic, the traffic destined to all of the outlets of an SE are the same, since no ultimate output is given any advantage. Moreover, total buffer occupancy decreases in the SEs as the traffic approaches the last stage. This can be explained by the assumption that each output of the Delta switch is able to accept a cell as it arrives. However, in the presence of hot spot traffic, the traffic for a designated hot outlet dominates in the SE. Contrary to the uniform traffic, buffers in the hot SEs start to saturate from the last stage, since the hot traffic becomes more concentrated as it passes from stage i to stage $i+1$. The traffic in cold SEs at stage i remains uniform, although the buffer occupancies in those SEs are affected by the hot traffic in their predecessor hot SE at stage $i-1$.

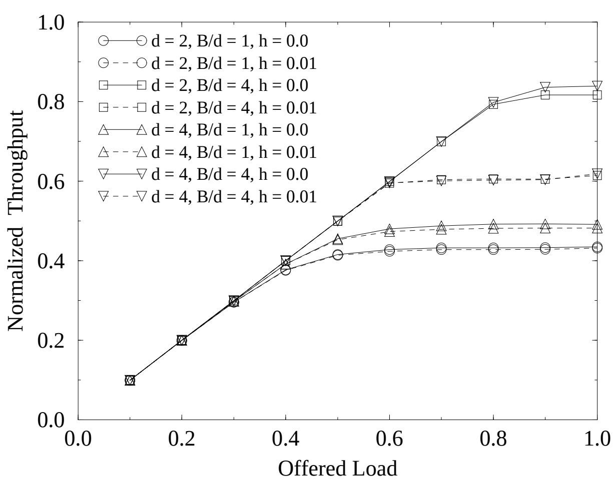

Figure 2: Normalized throughput versus offered load for $N = 64$

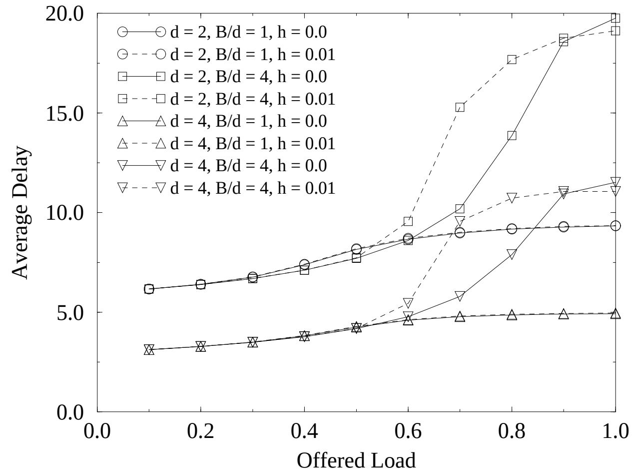

Figure 3: Average delay versus offered load for $N = 64$.

Fig. 6 illustrates the impact of hot fraction of the input traffic on the average throughput of the switch for $N = 256$. When the hot spot value is small, increasing the buffer size will result in an improvement of the throughput. As the hot spot value increases, increasing the buffer size can not compensate for the degradation caused by the hot traffic. For large enough hot spot values, increasing the buffer size has literally no effect on the improvement of the throghput.

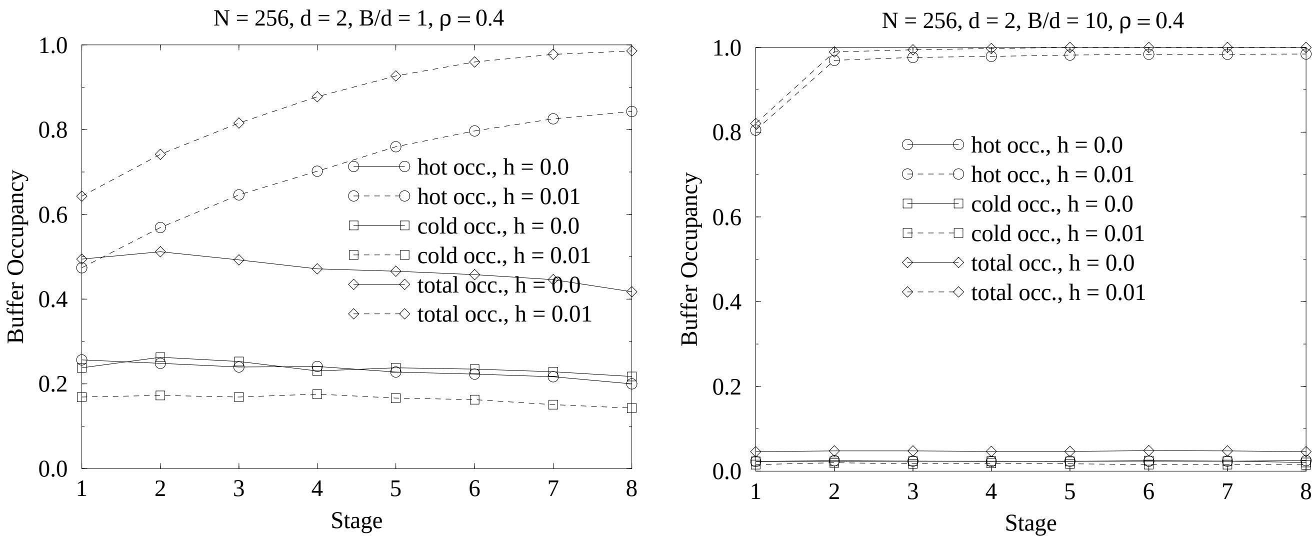

Figure 4: Buffer occupancy for hot SE at each stage for $N = 256$, $d = 2$, and offered load of $\rho = 0.4$.

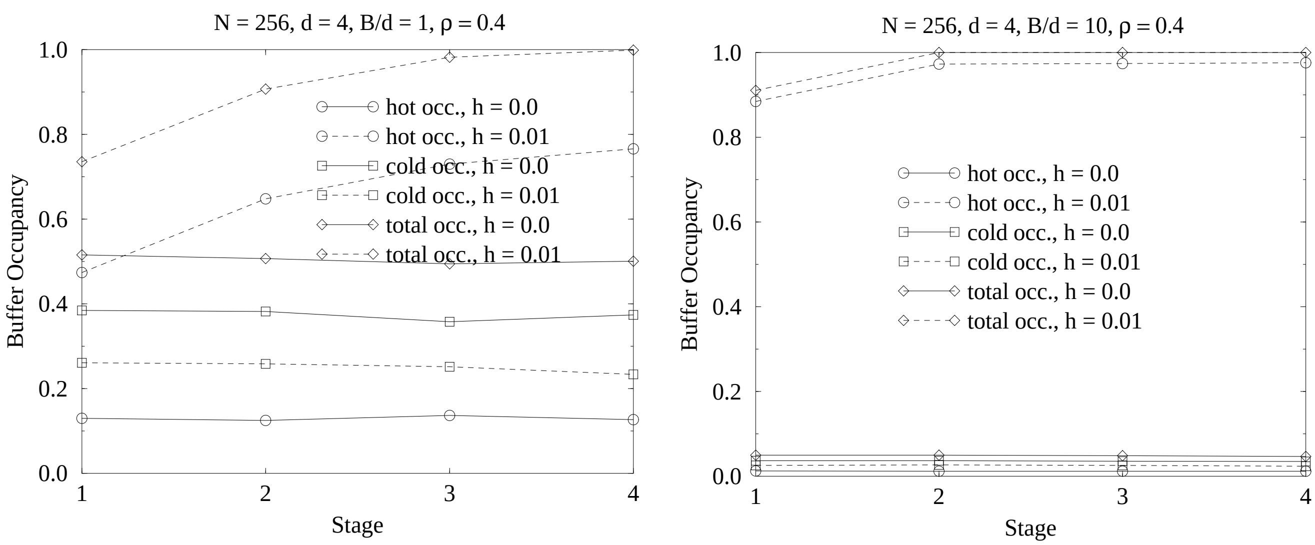

Figure 5: Buffer occupancy for hot SE at each stage for $N = 256$, $d = 4$, and offered load of $\rho = 0.4$

Increasing the buffer size has another effect, viz. sharp increase of the average delay per stage as illustrated in Fig 7. This increase is more significant when switch size and buffer size are bigger.

In Fig 8 the effect of some fixed value of hot spot on different switch sizes is illustrated. A specific hot spot value will degrade a bigger switch more significantly. This can be explained as if on the average one cell arrives at each output of a switch, that cell will be removed (read) from the output. If more than one cell arrive at an output, additional cells are piled up in the buffers and this build up propagates back in the switch. This phenomenon is similar to *tree saturation* addressed in [9]. If the total number of cells which are destined to the hot output satisfy the inequality

$$(h\rho + \frac{\rho}{N})N \leq 1 \qquad (2)$$

then there won't be any buffer saturation incurred by hot traffic in any SE. Solving this inequality for h gives:

$$h \leq \frac{1-\rho}{\rho N}. \qquad (3)$$

5 CONCLUSION

We have studied the impact of single hot spot traffic in shared buffer Delta switches through simulation. Shared buffer multistage switches have better performance and buffer utilization as compared to input or output buffer switches under uniform traffic pattern. However, since in a shared buffer switch, buffers are completely shared between all the inlets and outlets of an SE, a hot spot traffic may saturate buffer spaces in the SE and degrade the overall performance of the SE. The amount of degradation depends on

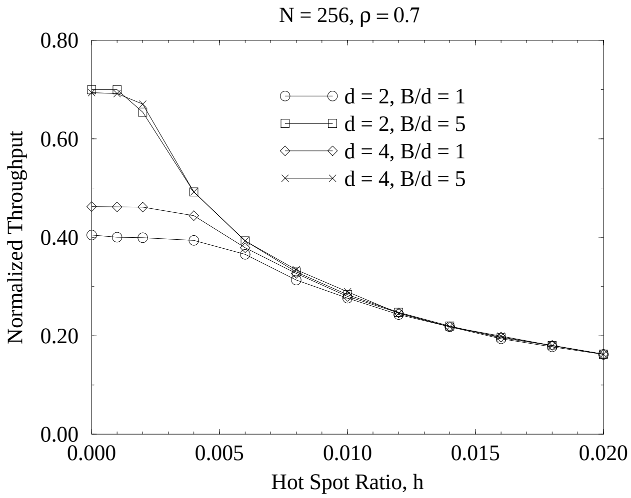

Figure 6: Normalized throughput versus hot spot for $N = 256$ and $\rho = 0.7$.

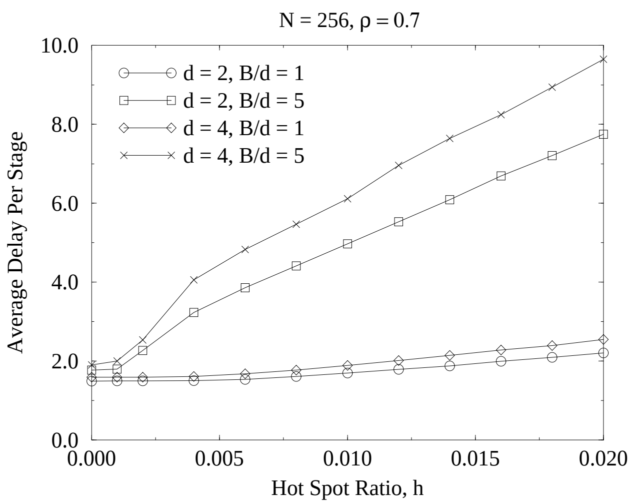

Figure 7: Average delay per stage versus hot spot for $N = 256$ and $\rho = 0.7$.

the fraction of hot spot traffic, switch size, and size of the SEs.

It is possible to alleviate the impact of buffer saturation by introducing some buffer management policy. A feasible solution is to limit the maximum number of buffers used by hot traffic to some fixed value. The maximum ratio of hot to cold traffic may vary in different stages and different SE types of a stage, since there are different traffics flowing in them. Therefore, it is possible to allow different limits of hot spot traffic for different SE types and different stages.

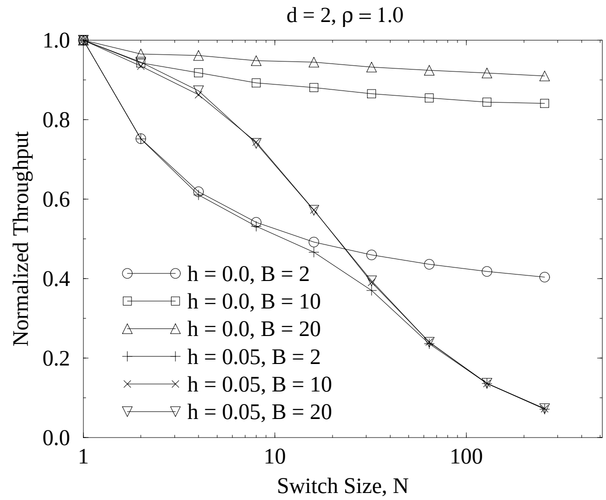

Figure 8: Normalized throughput versus switch size for $d = 2$, and $\rho = 1.0$.

References

[1] M.G. Hluchyj and M.J. Karol, "Queueing in high-performance packet switching," *IEEE Transactions on Communications*, vol. 6, no. 9, pp. 1587–1597, December 1988.

[2] I. Iliadis and W.E. Denzel, "Analysis of packet switches with input and output queueing," *IEEE Transactions on Communications*, vol. 41, no. 5, pp. 731–740, May 93.

[3] Y. Oie, T. Suda, M. Murata, and H. Miyahara, "Survey of switching techniques in high-speed networks and their performance," *International Journal of Satellite Communications*, vol. 9, no. 5, pp. 285–303, Sep-Oct 1991.

[4] M. Atiquzzaman and M.S. Akhtar, "Performance modeling of switching fabric for ATM switching node," *Australian Broadband Switching and Services Symposium*, Melbourne, pp. 581–588, July 15-17, 1992.

[5] B. Zhou and M. Atiquzzaman, "Improved performance model of multibuffered multistage interconnection network under general traffic patterns," *IEEE INFOCOM '94: Conference on Computer Communications*, Toronto, Canada, pp. 1448–1455, June 12-16, 1994.

[6] J.S. Turner, "Queueing analysis of buffered switching networks," *13th International Teletraffic Congress*, Copenhagen, Denmark, pp. 35–40, June 1991.

[7] 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.

[8] S. Gianatti and A. Pattavina, "Performance analysis of shared-buffered Banyan networks under arbitrary traffic patterns," *IEEE INFOCOM '93: Conference on Computer Commu-*

nications, San Francisco, California, pp. 943–952, March 30 - April 1, 1993.

[9] 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.