Realistic modelling of blocked packets for accurate performance evaluation of ATM switches

M.Atiquzzaman and C.-K.Chen

Abstract: Multistage switches have been used as ATM switching fabrics in broadband ISDN networks, and also to connect processors to memories in massively parallel multiprocessor systems. Previous performance models for multistage switches have been neither accurate enough nor based on realistic assumptions regarding modelling the correlation of the blocked packets in successive stages of the switch. Two new analytical models for accurate analysis of multistage switches are proposed. The new models reflect the realistic behaviour of blocked packets, and take into account the fact that a blocked packet always hunts for the same output link in successive clock cycles. The results obtained from the models are more accurate than those available in the literature.

1 Introduction

Multistage switches (MS) have been used to connect processors to memories in massively parallel multiprocessor systems and also in ATM switching fabrics in broadband ISDN networks. An MS consists of several stages of small crossbar switching elements (SE). The adjacent stages are connected by a permutation function. Decentralised routing using destination tag routing is used to route packets from the input to the output of the network. Because of their self-routing property and the existence of a unique path between an input–output pair, Banyan-type switches have been found to be suitable for high speed ATM switches [1].

Routing conflicts inside the MS prevent simultaneous arbitrary input–output connection requests. Buffers are used to store the packets which lose the routing conflicts. The primary purpose of buffers in an SE is to prevent loss of packets which lose routing conflicts. A packet is forwarded only if the link to which it is to be routed is able to accept the packet. The performance evaluation of MS in terms of traffic characteristics and switch parameters is critical for comparing alternative switch designs, planning appropriate traffic measurements, and for developing network congestion control strategies.

Previously, the approximate performance of packet switched banyans based on analytical modelling under uniform, [2] nonuniform, and hot-spot traffic references [3–7] have been presented. Because of simplifying assumptions in the models, the results obtained were not accurate. Previous models have not seriously accounted for the fact that a blocked packet always hunts for the same output link in subsequent clock cycles. In this paper, we propose an accurate model, which takes into account the correlation between the states of the buffer of adjacent stages and account for the packets blocked during routing conflicts. The novelty of the model lies in removing the assumptions of previous models regarding blocked packets. The results obtained from the proposed model are compared with those from the most recently [2] published model. The proposed model can be used to compare different switch architectures and the effects of various parameters (e.g. buffer size, SE size) and various traffic characteristics on the switch performance. This will also permit the development of network congestion control strategies.

IEE Proceedings online no. 19990594

DOI: 10.1049/ip-com:19990594

Paper first received 7th January 1998 and in revised form 4th May 1999

M. Atiquzzaman is with the Department of Electrical and Computer Engineering, University of Dayton, Dayton, Ohio 45469-0226, USA

C.-K. Chen is with the Department of Comnputer Science and Computer Engineering, La Trobe University, Melbourne 3083, Australia

2 Previous models for multistage switches

Jenq [8] proposed two models for analysing the performance of a banyan switch constructed of single input-buffered 2 × 2 SEs. Jenq concluded that the assumption of independence between buffers of an SE is reasonable. Jenq's analytical models were not verified by simulation, and later studies revealed that the model exhibits significant errors for high input traffic.

Jenq's work has been extended by Yoon, Lee and Liu [9]. The same independence assumptions, and the states as used by Jenq, were used, and hence the model did not show any significant improvement over Jenq's model in terms of accuracy. The model did not take into consideration the blocking state of the buffers, and hence was inaccurate.

Theimer, Rathgeb and Huber [10] proposed a model (we refer to it as the TRH model) to eliminate the assumption of independence between the buffers in the same SE and to partially take into account the effect of blocking. The TRH model showed significant improvement in terms of accuracy over Jenq's model. However, because of its complexity, expansion of the model for switches of arbitrary size SEs or buffers is very difficult.

Hsiao and Chen [11] pointed out the factors which caused low accuracy in Jenq's single-buffered and uniform traffic model, and the impossibility of generalising the TRH model. They proposed (we call it the HC model) an extension to the YLL model by taking into account the dependencies between the states of the buffers of an SE.

The authors partially considered the effects of blocking and buffer dependence within the same SE in consecutive clock cycles.

Even though the HC model is relatively simple, it is based on several unrealistic assumptions. From the state transition diagram of the HC model [12], two abnormal phenomena are observed. First, a blocked packet cannot move to the succeeding stage, even though there is buffer space available. It has to spend at least one clock cycle in first changing itself to the normal state. Secondly, it assumes that a blocked packet will first compete with other blocked packets in the same SE. Blocked packets are discriminated such that they compete first. Thirdly, the blocked packets are not resubmitted to the same output. Moreover, the analysis is for single-buffered SEs, which is not realistic for practical switches. The model cannot accurately describe the real operation, and thus the model had to employ some empirical parameters to get reasonable results.

Mun and Youn [2] proposed an improved model (we call the MY model) over the HC model by introducing a correlation between the blocked state and the empty state. The MY model uses three states to model a buffer. It uses the empty, normal and blocked states. The model uses the blocked state information to justify that, if a stage contains a blocked packet, the destined buffer must be in the normal or blocked states ([2], p. 156). However, the model with only one blocked state is not capable of remembering the output for which a packet was blocked. We argue that, without storing the blocking history of a packet, it is not possible to resubmit a packet to its original destination in a subsequent clock cycle. For example, the probability that a blocked packet in a buffer can get to the desired output port when the other buffer has a blocked packet (r_{bb} of eqn. 7 in [2]) is shown as

$$r_{bb}(k,t) = 0.75P_b(k,t)$$

which is actually obtained from

$$r_{bb}(k,t) = \underbrace{[P_b(k,t)]}_{\text{part A}}\underbrace{[(0.5+0.5\times 0.5)]}_{\text{part B}}$$
$$= [\Pr(\text{other buffer is blocked})][\Pr(\text{no contention}) + \Pr(\text{contention}) \times \Pr(\text{packet wins conflict})] \quad (1)$$

Since the probability of contention is taken to be 0.5 in part B of the above equation, it implies that a blocked packet is not resubmitted to the original destination in the subsequent cycle. It is simply treated as a normal packet. The fact that a blocked packet is treated as a normal packet in eqn. 7 of [2] is even more clear from the $0.75P_n(k, t)$ term in eqn. 6 of [2], which accounts for the routing probability of a normal packet. The same 0.75 term is incorrectly used in eqn. 7 of [2] to account for a blocked packet. In the analysis, the MY model implicitly assumes that a blocked packet has equal probabilities of requesting the output links of the SE during subsequent cycles as shown above. The impact of this implicit assumption is reflected in the increased errors with increased number of stages and an increased blocking in the switch.

In this paper, we propose a Markov chain model which takes into account the fact that a blocked packet in an SE will request the same output of the SE at successive cycles. The proposed model explicitly stores the information regarding the output for which a packet was blocked. The proposed model captures resubmission of a blocked packet to the original destination. The results obtained from our model have higher accuracy than those obtained from the MY model. We have chosen the MY model for comparison with our model, since it is the most recently published model and produces better results than previous models. Therefore, by comparing our model with the MY model, we show that our proposed model produces more accurate results than all previous models.

3 Network and modelling assumptions

A banyan switch can be used to connect N inputs to N outputs using a multistage interconnection network. It consists of $n = \log_a N$ stages of N/a SEs per stage, each SE having a input and a output lines. A perfect shuffle interconnection is used to connect the different stages of a switch as shown in Fig. 1 for $N = 8$ using 2×2 SEs.

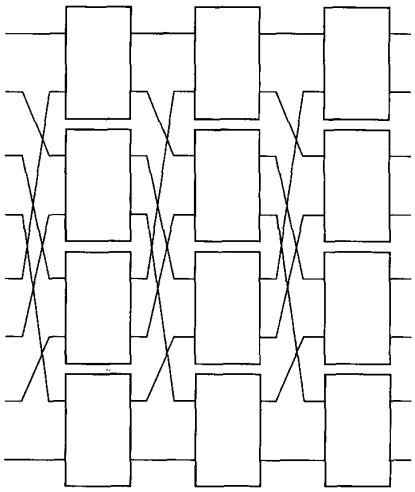

Fig. 1 *Multistage ATM switch*

A clock cycle is assumed to be composed of two phases [5]. In phase 1, packets are forwarded and packets are received during phase 2. The following assumptions are used for developing the proposed model.

(i) Each input of an SE in the first stage receives a packet during a cycle with probability r. The packets are uniformly distributed over all the outputs for uniform output reference traffic, or biased towards one particular output for a hot spot traffic.

(ii) The SEs are finite input buffered (i.e. there are fixed size buffers at the inputs of every SE).

(iii) There is a backpressure mechanism between successive stages of the switch. A packet can move to the next stage only if there is enough space in the next stage.

(iv) The queueing and dequeueing of packets in a buffer take place during the same cycle.

(v) Every packet has equal probability of winning the contention and the blocked packets are resubmitted to the original destination.

(vi) All input requests are synchronised at the beginning of a cycle.

(vii) The arrival of a packet during a cycle is independent of whether the packet received at the previous cycle was accepted or rejected.

(viii) The packet received at an input is independent of the packets received by other inputs.

We represent the states of a buffer in an SE by four states as follows.

State 0: Buffer is empty.

State n: Buffer has a new packet which has just arrived in the previous cycle.

State n: Buffer has a packet which is blocked for the upper output link.

State l: Buffer has a packet which is blocked for the lower output link.

Fig. 2 shows the state transition diagram of the proposed model. We first develop a model for a switch consisting of single buffered SEs and then extend it to multiple buffered SEs in Section 5.

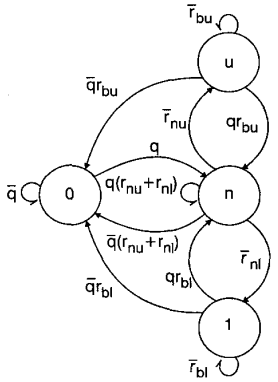

Fig. 2 *Markov chain for proposed single buffer model*

4 Single buffer model

The following notations will be used in developing the model. Without loss of generality, we assume a 2 × 2 SE to illustrate our model. The two outputs of an SE will be called the upper and lower outputs. To make it easier for the reader to compare the proposed model with the MY model, wherever possible, we have used the same notations as in [2].

SE(k) = a switching element at stage k. The inputs of each SE have dedicated buffers.

$q(k, t)$ = probability that a packet is ready to come to SE(k) during cycle t. $q(1, t)$ is therefore the offered load at the input of the switch.

$\rho(k, t)$ = probability that a buffer at SE(k) receives a packet during cycle t.

$\pi_0(k, t)$ = probability that a buffer of SE(k) is empty at the beginning of cycle t.

$\pi_n(k, t)$ = probability that a buffer of SE(k) has a new packet at the beginning of cycle t.

$\pi_u(k, t)$ = probability that a buffer of SE(k) has a packet which is blocked for the upper output at the beginning of cycle t.

$\pi_l(k, t)$ = probability that a buffer of SE(k) has a packet which is blocked for the lower output at the beginning of cycle t.

$r_{nu}(k, t)$ = probability that a new packet in SE(k) moves to the upper output during cycle t.

$r_{nu}^{0nl}(k, t)$ = probability that a new packet in SE(k) is able to get to the upper output during cycle t and the state of the other buffer in the same SE is either 0, n or l.

$r_{nu}^{u}(k, t)$ = probability that a new packet in SE(k) is able to get to the upper output during cycle t and the state of the other buffer in the same SE is u.

$r_{bu}(k, t)$ = probability that a packet which is blocked for the upper output in SE(k) moves during cycle t.

$r_{bu}^{0n}(k, t)$ = probability that a packet which is blocked for the upper output in SE(k) is able to get to the upper output during cycle t and the state of the other buffer in the same SE is 0 or n.

$r_{bu}^{ul}(k, t)$ = probability that a packet which is blocked for the upper output in SE(k) is able to get to the upper output during cycle t and the state of the other buffer in the same SE is u or l.

$r_{nl}(k, t)$ = probability that a new packet in SE(k) moves to the lower output during cycle t.

$r_{nl}^{0nu}(k, t)$ = probability that a new packet in SE(k) is able to get to the lower output during cycle t and the state of the other buffer in the same SE is either 0, n or u.

$r_{nl}^{l}(k, t)$ = probability that a new packet in SE(k) is able to get to the lower output during cycle t and the state of the other buffer in the same SE is l.

$r_{bl}(k, t)$ = probability that a packet which is blocked for the lower output in SE(k) moves during cycle t.

$r_{bl}^{0n}(k, t)$ = probability that a packet which is blocked for the lower output in SE(k) is able to get to the lower output during cycle t and the state of the other buffer in the same SE is 0 or n.

$r_{bl}^{ul}(k, t)$ = probability that a packet which is blocked for the lower output in SE(k) is able to get to the lower output during cycle t and the state of the other buffer in the same SE is u or l.

$\alpha(k, t)$ = probability that at least one buffer space in SE(k) is available during cycle t, given that there is no blocked packet in SE($k-1$) which is destined to this buffer.

$\alpha^{b}(k, t)$ = probability that at least one buffer space in SE(k) is available during cycle t, given that only one blocked packet in SE($k-1$) is destined to this buffer.

$\alpha^{bb}(k, t)$ = probability that at least one buffer space in SE(k) is available during cycle t, given that SE($k-1$) contains two blocked packets which are destined to this buffer.

Next we derive the different state and routing probabilities based on the state diagram of Fig. 2.

4.1 State probabilities

$$\pi_0(k,t) = 1 - \pi_n(k,t) - \pi_u(k,t) - \pi_l(k,t) \quad (2)$$

$$\begin{aligned}\pi_n(k,t) = q(k,t-1)\big[&\pi_0(k,t-1) + \pi_n(k,t-1)\\ &\times [r_{nu}(k,t-1) + r_{nl}(k,t-1)]\\ &+ \pi_u(k,t-1)r_{bu}(k,t-1)\\ &+ \pi_l(k,t-1)r_{bl}(k,t-1)\big]\end{aligned} \quad (3)$$

$$\begin{aligned}\pi_u(k,t) = \pi_n(k,t-1)&\overline{r_{nu}(k,t-1)}\\ &+ \pi_u(k,t-1)\overline{r_{bu}(k,t-1)}\end{aligned} \quad (4)$$

$$\begin{aligned}\pi_l(k,t) = \pi_n(k,t-1)&\overline{r_{nl}(k,t-1)}\\ &+ \pi_l(k,t-1)\overline{r_{bl}(k,t-1)}\end{aligned} \quad (5)$$

where, unless otherwise specified, $\overline{x(k,t)} = 1 - x(k,t)$ for any variable $x(k, t)$.

4.2 Routing and acceptance probabilities

The relationship between the routing and blocking probabilities for new and blocked packets is expressed as follows:

$$r_{nu}(k,t) + r_{nl}(k,t) + \overline{r_{nu}(k,t)} + \overline{r_{nl}(k,t)} = 1 \quad (6)$$

$$r_{bu}(k,t) + \overline{r_{bu}(k,t)} = 1 \quad (7)$$

$$r_{bl}(k,t) + \overline{r_{bl}(k,t)} = 1 \quad (8)$$

where r and $\bar{r}$ are the routing and blocking probabilities respectively. For uniform incoming traffic, which is uniformly distributed over the outputs, $r_{nu}(k, t) = r_{nl}(k, t)$ and $\overline{r_{nu}(k, t)} = \overline{r_{nl}(k, t)}$. Therefore,

$$\overline{r_{nu}(k,t)} = 0.5 - r_{nu}(k,t) \quad (9)$$

$$\overline{r_{nl}(k,t)} = 0.5 - r_{nl}(k,t) \quad (10)$$

The probability of successfully routing a packet in an SE to the next stage depends on the probability of the packet winning any possible contention in the SE and the probability of the next stage buffer being able to accept it.

When considering $r_{nu}^{0nl}(k, t)$, the new packet under consideration is able to get to its desired output when the other buffer in the SE is empty or if there exists a new packet in the other buffer which is destined to the lower output port or the packet in the other buffer is blocked for the lower output. However, if the new packet in the other buffer is destined to the same upper output, internal conflict between packets occurs and each packet has the same probability to win the contention. Therefore,

$$\begin{aligned} r_{nu}^{0nl}(k,t) &= 0.5\pi_0(k,t) + 0.5 \times 0.5 \times 0.5\pi_n(k,t) \\ &\quad + 0.5 \times 0.5\pi_n(k,t) + 0.5\pi_l(k,t) \\ &= 0.5\pi_0(k,t) + 0.375\pi_n(k,t) + 0.5\pi_l(k,t) \end{aligned} \quad (11)$$

The probability that the new packet under consideration and a blocked packet in the other buffer are both destined to the upper output and the new packet can get to the desired upper output by winning any possible contention is

$$r_{nu}^{u}(k,t) = 0.5 \times 0.5\pi_u(k,t) = 0.25\pi_u(k,t) \quad (12)$$

Similarly, the routing probability of a new packet to the lower output port at SE(k), $1 \le k \le n-1$, is

$$r_{nl}^{0nu}(k,t) = 0.5\pi_0(k,t) + 0.375\pi_n(k,t) + 0.5\pi_u(k,t) \quad (13)$$

$$r_{nl}^{l}(k,t) = 0.25\pi_l(k,t) \quad (14)$$

When the two buffers of SE($k-1$) contain a new and a blocked packet during cycle t and both of them are destined to the same output link of the SE, the probability that the buffer at SE(k) is able to accept one of these packets is

$$\begin{aligned} \alpha^b(k,t) &= 0.5\rho(k,t-1)[r_{nu}(k,t) + r_{nl}(k,t)] \\ &\quad + \left[\frac{1 - 0.5\rho(k,t-1)}{2}\right][r_{bu}(k,t) + r_{bl}(k,t)] \end{aligned} \quad (15)$$

where $0.5\rho(k, t-1)$ is the probability that the buffer at SE(k) has received a packet during cycle $t-1$ from the buffer in SE($k-1$). Since the buffer at SE(k) cannot be in an empty state, the probability that the buffer is in a blocked state (either u or l) is therefore $1 - 0.5\rho(k, t-1)$. For uniform output-destined traffic at the input of the switch, the probability that the buffer is in state u or 1 is equal and is given by $[1 - 0.5\rho(k, t-1)]/2$.

The probability that a buffer in SE(k) is able to accept a packet during phase 2 of cycle t is given by

$$\begin{aligned} \alpha(k,t) &= \pi_0(k,t) + \pi_n(k,t)[r_{nu}(k,t) + r_{nl}(k,t)] \\ &\quad + \pi_u(k,t)r_{bu}(k,t) + \pi_l(k,t)r_{bl}(k,t) \end{aligned} \quad (16)$$

Finally, for a packet to move to the succeeding stage, it should be able to get to the desired output port and the destined buffer should be available. Thus for $1 \le k \le n-1$

$$r_{nu}(k,t) = r_{nu}^{0nl}(k,t)\alpha(k+1,t) + r_{nu}^{u}(k,t)\alpha^b(k+1,t) \quad (17)$$

$$r_{nl}(k,t) = r_{nl}^{0nu}(k,t)\alpha(k+1,t) + r_{nl}^{l}(k,t)\alpha^b(k+1,t) \quad (18)$$

When a packet is in state u and the other buffer in the same SE is in state 0 or n, the probability that the blocked packet is able to get to the desired upper output port depends on whether the other buffer in the same SE is empty or non-competing n-state or competing n-state and the blocked packet wins the conflict. Therefore, for $1 \le k \le n-1$:

$$\begin{aligned} r_{bu}^{0n}(k,t) &= \pi_0(k,t) + 0.5\pi_n(k,t) + 0.5 \times 0.5\pi_n(k,t) \\ &= \pi_0(k,t) + 0.75\pi_n(k,t) \end{aligned} \quad (19)$$

If there exist two blocked packets in a buffer at SE(k), the blocked packet under consideration is able to get to its destined upper output when there is no contention or there is contention and it wins the contention. For $1 \le k \le n-1$:

$$r_{bu}^{ul}(k,t) = \pi_l(k,t) + 0.5\pi_u(k,t) \quad (20)$$

If both buffers at SE($k-1$) are blocked, the acceptance probability of the corresponding destination buffer at stage k, $2 \le k \le n$, which must be in a blocked state, is given by

$$\alpha^{bb}(k,t) = \frac{1 - 0.5\rho(k,t-1)}{2}(r_{bu}(k,t) + r_{bl}(k,t)) \quad (21)$$

Therefore, for $1 \le k \le n-1$:

$$r_{bu}(k,t) = r_{bu}^{0n}(k,t)\alpha^b(k+1,t) + r_{bu}^{ul}(k,t)\alpha^{bb}(k+1,t) \quad (22)$$

Similar logic applies for the routing of a packet blocked for the lower output port. Hence, for $1 \le k \le n-1$:

$$r_{bl}^{0n}(k,t) = \pi_0(k,t) + 0.75\pi_n(k,t) \quad (23)$$

$$r_{bl}^{ul}(k,t) = \pi_u(k,t) + 0.5\pi_l(k,t) \quad (24)$$

$$r_{bl}(k,t) = r_{bl}^{0n}(k,t)\alpha^b(k+1,t) + r_{bl}^{ul}(k,t)\alpha^{bb}(k+1,t) \quad (25)$$

All the necessary equations for determining the throughput and delay of the MS have been derived in this Section. In the next Section, we use the different state, routing and blocking probabilities derived in this Section to derive the performance criteria of the switch. The equations derived in this Section are solved iteratively as in [8] until all the variables converge to their steady-state values.

4.3 Throughput and delay

The bandwidth of the switch is obtained by determining the traffic rates at the output links of the switch. For uniform traffic, the rates will be the same for all output links. The traffic rate at any link connecting the SEs between SE($k-1$) and SE(k) can be obtained by finding the probability that a packet is received at SE(k) from SE($k-1$). For a buffer at the first stage (i.e. for $k = 1$):

$$\begin{aligned} \rho(k,t) &= q(k,t)[\pi_0(k,t) \\ &\quad + \pi_n(k,t)[r_{nu}(k,t) + r_{nl}(k,t)] \\ &\quad + \pi_u(k,t)r_{bu}(k,t) + \pi_l(k,t)r_{bl}(k,t)] \end{aligned} \quad (26)$$

For SE(k) $2 \le k \le n$,

$$\rho(k,t) = \pi_n(k-1,t)[r_{nu}(k-1,t) + r_{nl}(k-1,t)] + \pi_u(k-1,t)r_{bu}(k-1,t) + \pi_l(k-1,t)r_{bl}(k-1,t) \quad (27)$$

The probability that a packet is offered to a buffer at SE(k) is the ratio of the probability that a packet is received by the buffer to the probability that the buffer can accept the packet. For SE(k), $2 \le k \le n$,

$$q(k,t) = \frac{\rho(k,t)}{\pi_0(k,t)+\pi_n(k,t)[r_{nu}(k,t)+r_{nl}(k,t)] + \pi_u(k,t)r_{bu}(k,t)+\pi_l(k,t)r_{bl}(k,t)} \quad (28)$$

where $\rho(k, t)$ is obtained from eqn. 27.

The boundary conditions required for the first and last stage are as follows.

- Stage 1: Since there is no preceding stage to it, $q(1, t)$ is specified as an input to the model.
- Stage n: Since a packet at the last stage can always proceed and both buffers in an SE can not be in the blocked state, all the acceptance probabilities at the last stage are always equal to one, i.e. $\alpha(n, t) = \alpha^b(n, t) = \alpha^{bb}(n, t) = 1$.

The routing probabilities at the last stage are therefore, obtained by setting the acceptance probabilities of the next stage equal to unity in eqns. 17, 18, 22 and 25.

$$r_{nu}(n,t) = r_{nu}^{0nl}(n,t) + r_{nu}^{u}(n,t) \quad (29)$$

$$r_{nl}(n,t) = r_{nl}^{0nu}(n,t) + r_{nl}^{l}(n,t) \quad (30)$$

$$r_{bu}(n,t) = r_{bu}^{0n}(n,t) + r_{bu}^{ul}(n,t) \quad (31)$$

$$r_{bl}(n,t) = r_{bl}^{0n}(n,t) + r_{bl}^{ul}(n,t) \quad (32)$$

Normalised throughput (λ) of a switch is defined to be the probability of a packet leaving an output port of the switch during a cycle. Thus the normalised throughput at time t is given by:

$$\lambda(n,t) = \pi_n(n,t)[r_{nu}(n,t) + r_{nl}(n,t)] + \pi_u(n,t)r_{bu}(n,t) + \pi_l(n,t)r_{bl}(n,t) \quad (33)$$

The mean delay (δ) is defined to be the number of cycles taken by a packet to traverse the switch.

$$\delta = \sum_{k=1}^{n} \delta(k) \quad (34)$$

where $\delta(k)$ is the delay for a packet at stage k in the steady state. Thus, for $1 \le k \le n$:

$$\delta(k) = \frac{\pi_n(k) + \pi_u(k) + \pi_l(k)}{\rho(k)} \quad (35)$$

where $x(k, t)$ represents the time independent steady state value of the variable $x(k, t)$.

5 Multiple buffer model

The model for the analysis of finite-buffered switches will be developed, based on the concepts in the previous Section. The assumptions and switch operation employed in the single-buffer model are still valid for the finite-buffered model. However, the number of possible states of a buffer depends on the size of the buffer. We define the states of a buffer as follows.

State 0: The buffer is empty.

State *ni*: The buffer has i, $1 \le i \le m$, packets and the packet at the head of the buffer is in a new state.

State *ui*: The buffer has i, $1 \le i \le m$, packets and the packet at the head of the queue is blocked for the upper output of the switch, either because it lost a contention or because the next stage buffer could not accept it in a previous cycle.

State *li*: The buffer has i, $1 \le i \le m$, packets and the packet at the head of the queue is blocked for the lower output of the switch.

For a buffer of size m, there are $3m + 1$ possible states. The buffers are served on a first-come-first-served basis. Fig. 3 shows the state transition diagram of the proposed multiple input buffer model. Although the assumptions made in the single buffer model are applicable in the multi-buffer model, some of the definitions need to be modified to reflect the multiple buffers.

- The size of a buffer at an SE input is m.
- $\pi_n(k, t)$, $\pi_u(k, t)$ or $\pi_l(k, t)$ will represent the probabilities that the packet at the head of a buffer in SE(k) is in state n, u or l, respectively.
- $r_{nu}(k, t)$, $r_{nu}^{0nl}(k, t)$, $r_{nu}^{u}(k, t)$, $r_{bu}(k, t)$, $r_{bu}^{0n}(k, t)$, $r_{bu}^{ul}(k, t)$, $r_{nl}(k, t)$, $r_{nl}^{0nu}(k, t)$, $r_{nl}^{l}(k, t)$, $r_{bl}(k, t)$, $r_{bl}^{0n}(k, t)$, and $r_{bl}^{ul}(k, t)$ will represent the routing probabilities of the packet at the head of the queue.
- $\alpha(k, t)$, $\alpha^b(k, t)$, and $\alpha^{bb}(k, t)$ are used for an entire buffer.

As seen in Fig. 3, three sets of state equations are required depending on the number of packets in the buffer at any SE.

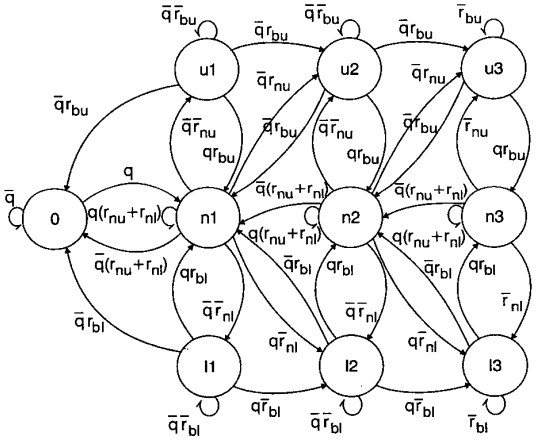

Fig. 3 *Markov chain for proposed multiple buffer model*

5.1 State probabilities

5.1.1 State probabilities of zero or one packet: For a buffer which contains exactly one packet, the probabilities that this packet is blocked for either the upper or lower output are

$$\pi_{u1}(k,t) = [1 - q(k,t-1)] \times [\pi_{n1}(k,t-1)[0.5 - r_{nu}(k,t-1)] + \pi_{u1}(k,t-1)[1 - r_{bu}(k,t-1)]] \quad (36)$$

$$\pi_{l1}(k,t) = [1 - q(k,t-1)] \times [\pi_{n1}(k,t-1)[0.5 - r_{nl}(k,t-1)] + \pi_{l1}(k,t-1)[1 - r_{bl}(k,t-1)]] \quad (37)$$

Alternatively, the packet could be in the new state or the buffer is empty. The probability of a buffer being empty or having a new packet is given by

$$\pi_0(k,t) = [1-q(k,t-1)]\Big[\pi_0(k,t-1) + \pi_{n1}(k,t-1)[r_{nu}(k,t-1)+r_{nl}(k,t-1)] + \pi_{u1}(k,t-1)r_{bu}(k,t-1) + \pi_{l1}(k,t-1)r_{bl}(k,t-1)\Big] \quad (38)$$

$$\pi_{n1}(k,t) = q(k,t-1)\Big[\pi_0(k,t-1) + \pi_{n1}(k,t-1)[r_{nu}(k,t-1)+r_{nl}(k,t-1)] + \pi_{u1}(k,t-1)r_{bu}(k,t-1) + \pi_{l1}(k,t-1)r_{bl}(k,t-1)\Big] + [1-q(k,t-1)] \times \Big[\pi_{n2}(k,t-1)[r_{nu}(k,t-1)+r_{nl}(k,t-1)] + \pi_{u2}(k,t-1)r_{bu}(k,t-1) + \pi_{l2}(k,t-1)r_{bl}(k,t-1)\Big] \quad (39)$$

5.1.2 State probabilities of i, 2 ≤ i ≤ m − 1, packets: When a buffer contains more than one packet but is not full, the correlation between the transition of the buffer states are reduced since the buffer cannot be emptied in a single transition. Hence, the state probabilities are given by

$$\pi_{ni}(k,t) = q(k,t-1)\Big[\pi_{ni}(k,t-1) \times [r_{nu}(k,t-1)+r_{nl}(k,t-1)] + \pi_{ui}(k,t-1)r_{bu}(k,t-1) + \pi_{li}(k,t-1)r_{bl}(k,t-1)\Big] + [1-q(k,t-1)]\Big[\pi_{n(i+1)}(k,t-1) \times [r_{nu}(k,t-1)+r_{nl}(k,t-1)] + \pi_{u(i+1)}(k,t-1)r_{bu}(k,t-1) + \pi_{l(i+1)}(k,t-1)r_{bl}(k,t-1)\Big] \quad (40)$$

$$\pi_{ui}(k,t) = [1-q(k,t-1)]\Big[\pi_{ni}(k,t-1) \times [0.5-r_{nu}(k,t-1)] + \pi_{ui}(k,t-1)[1-r_{bu}(k,t-1)]\Big] + q(k,t-1)\Big[\pi_{n(i-1)}(k,t-1) \times [0.5-r_{nu}(k,t-1)] + \pi_{u(i-1)}(k,t-1)[1-r_{bu}(k,t-1)]\Big] \quad (41)$$

$$\pi_{li}(k,t) = [1-q(k,t-1)]\Big[\pi_{ni}(k,t-1) \times [0.5-r_{nl}(k,t-1)] + \pi_{li}(k,t-1)[1-r_{bl}(k,t-1)]\Big] + q(k,t-1)\Big[\pi_{n(i-1)}(k,t-1) \times [0.5-r_{nl}(k,t-1)] + \pi_{l(i-1)}(k,t-1)[1-r_{bl}(k,t-1)]\Big] \quad (42)$$

5.1.3 State probabilities of m packets: The probability of the buffer being full and in a new state is given by

$$\pi_{nm}(k,t) = q(k,t-1)[\pi_{nm}(k,t-1) \times [r_{nu}(k,t-1)+r_{nl}(k,t-1)] + \pi_{um}(k,t-1)r_{bu}(k,t-1) + \pi_{lm}(k,t-1)r_{bl}(k,t-1)] \quad (43)$$

and the state probabilities of the buffer being blocked for the upper or lower output are

$$\pi_{um}(k,t) = \pi_{nm}(k,t-1)[0.5-r_{nu}(k,t-1)] + \pi_{um}(k,t-1)[1-r_{bu}(k,t-1)] + q(k,t-1)\Big[\pi_{n(m-1)}(k,t-1) \times [0.5-r_{nu}(k,t-1)] + \pi_{u(m-1)}(k,t-1)[1-r_{bu}(k,t-1)]\Big] \quad (44)$$

$$\pi_{lm}(k,t) = \pi_{nm}(k,t-1)[0.5-r_{nl}(k,t-1)] + \pi_{lm}(k,t-1)[1-r_{bl}(k,t-1)] + q(k,t-1)\Big[\pi_{n(m-1)}(k,t-1) \times [0.5-r_{nl}(k,t-1)] + \pi_{l(m-1)}(k,t-1)[1-r_{bl}(k,t-1)]\Big] \quad (45)$$

Therefore, the probability that the buffer is not full is simply

$$\overline{\pi_{*m}(k,t)} = 1 - \pi_{nm}(k,t) - \pi_{um}(k,t) - \pi_{lm}(k,t) \quad (46)$$

where $\pi_{*m}(k, t)$ and $\overline{\pi_{*m}(k, t)}$ are the probabilities that the buffer is full and not full, respectively.

The probability that the packet at the head of a buffer is in state n, u or l is represented by $\pi_n(k, t)$, $\pi_u(k, t)$, and $\pi_l(k, t)$ and are given by

$$\pi_n(k,t) = \sum_{i=1}^{m} \pi_{ni}(k,t) \quad (47)$$

$$\pi_u(k,t) = \sum_{i=1}^{m} \pi_{ui}(k,t) \quad (48)$$

$$\pi_l(k,t) = \sum_{i=1}^{m} \pi_{li}(k,t) \quad (49)$$

5.2 Acceptance and routing probabilities

The acceptance probabilities $\alpha(k, t)$ and $\alpha^b(k, t)$ for the multiple buffer model are obtained in the same way as the single buffer model. If there exists a blocked packet competing with a new packet in SE(k − 1) for the same output port, the destined buffer in SE(k) must be in either state $\pi_{nl}(k, t)$, $\pi_{ul}(k, t)$ or $\pi_{ll}(k, t)$. If the destined buffer received a packet in the previous cycle, it can be in state $\pi_{ni}(k, t)$, $1 \le i \le m$, or $\pi_{ui}(k, t)$, $2 \le i \le m$, or $\pi_{li}(k, t)$, $2 \le i \le m$. If it has not

received a packet, it must be in state $\pi_{um}(k, t)$ or $\pi_{lm}(k, t)$. Therefore,

$$\alpha^b(k,t) = 0.5\rho(k,t-1)A + \frac{1-0.5\rho(k,t-1)}{2}\left(r_{bu}(k,t) + r_{bl}(k,t)\right) \quad (50)$$

where $0.5\rho(k, t-1)$ is the probability that the buffer at SE(k) received a packet during cycle $t-1$ from the buffer in SE($k-1$), which contains the new packet at cycle t, and the buffer in SE($k+1$) is in state $\pi_{ni}(k, t)$ and A is given by

$$A = \frac{\sum_{i=1}^{m-1}\pi_{ni}(k,t) + \sum_{i=2}^{m-1}(\pi_{ui}(k,t)+\pi_{li}(k,t)) + \phi_1 + \phi_2}{1-\pi_0(k,t)-\pi_{u1}(k,t)-\pi_{l1}(k,t)} \quad (51)$$

where

$$\phi_1 = \pi_{nm}(k,t)\left(r_{nu}(k,t)+r_{nl}(k,t)\right)$$
$$\phi_2 = \pi_{um}(k,t)r_{bu}(k,t) + \pi_{lm}(k,t)r_{bl}(k,t) \quad (52)$$

Similar logic applies for $\alpha(k, t)$. Thus

$$\alpha(k,t) = \pi_0(k,t) + \sum_{i=1}^{m-1}\pi_{ni}(k,t) + \sum_{i=1}^{m-1}[\pi_{ui}(k,t)+\pi_{li}(k,t)] + \pi_{nm}(k,t)[r_{nu}(k,t)+r_{nl}(k,t)] + \pi_{um}(k,t)r_{bu}(k,t) + \pi_{lm}(k,t)r_{bl}(k,t) \quad (53)$$

The routing probabilities $r_{nu}^{0nl}(k, t)$, $r_{nu}^{u}(k, t)$, $r_{nu}(k, t)$, $r_{nl}^{0nu}(k, t)$, $r_{nl}^{l}(k, t)$, and $r_{nl}(k,t)$ for a new packet are the same as eqns. 11, 12, 17, 13, 14 and 18, respectively for the single buffer model.

Before obtaining the routing probabilities of a blocked packet, the acceptance probability of its destined output buffer for the case where both of its originating buffers are in the blocked state, need to be determined.

Based on the derivation of $\alpha^{bb}(k, t)$ for the single buffer model, the required equations for the multiple buffer model are obtained as follows:

$$r_{bu}^{0n}(k,t) = \pi_0(k,t) + 0.75\pi_n(k,t) \quad (54)$$

$$r_{bu}^{ul}(k,t) = \pi_l(k,t) + 0.5\pi_u(k,t) \quad (55)$$

$$\alpha^{bb}(k,t) = \frac{\pi_u(k,t)r_{bu}(k,t) + \pi_l(k,t)r_{bl}(k,t)}{\pi_u(k,t)+\pi_l(k,t)} \quad (56)$$

Therefore, $r_{bu}(k, t)$, $r_{bl}^{0n}(k, t)$, $r_{bl}^{ul}(k, t)$ and $r_{bl}(k, t)$ are given by eqns. 22–25 for the single buffer case.

5.3 Throughput and delay

The traffic rate and the probability of offered load for multiple buffer model are obtained as follows. For the first stage of the switch (i.e. $k = 1$):

$$\rho(k,t) = q(k,t)\left[\overline{\pi_{*m}(k,t)} + \pi_{nm}(k,t)[r_{nu}(k,t)+r_{nl}(k,t)] + \pi_{um}(k,t)r_{bu}(k,t) + \pi_{lm}(k,t)r_{bl}(k,t)\right] \quad (57)$$

For stage k, $2 \le k \le n$:

$$\rho(k,t) = \pi_n(k-1,t)[r_{nu}(k-1,t)+r_{nl}(k-1,t)] + \pi_u(k-1,t)r_{bu}(k-1,t) + \pi_l(k-1,t)r_{bl}(k-1,t) \quad (58)$$

Therefore, $q(k, t)$ for stage k, $2 \le k \le n$, is given by

$$q(k,t) = \frac{\rho(k,t)}{\overline{\pi_{*m}(k,t)}+\pi_{nm}(k,t)[r_{nu}(k,t)+r_{nl}(k,t)] + \pi_{um}(k,t)r_{bu}(k,t)+\pi_{lm}(k,t)r_{bl}(k,t)} \quad (59)$$

where $\rho(k, t)$ is obtained from eqn. 58. The normalised throughput and the mean delay of a packet are given by eqns. 33 and 34 where the delay at stage k for the multiple buffer model is given by

$$\delta(k) = \frac{\sum_{i=1}^{m} i[\pi_{ni}(k)+\pi_{ui}(k)+\pi_{li}(k)]}{\rho(k)} \quad (60)$$

6 Results

The equations for the throughput and delay of a MS with finite input buffered SEs have been derived in Section 5.3. We use the above equations to determine the throughput and delay of the MS. The results are compared with those obtained from simulation and the model described in [2].

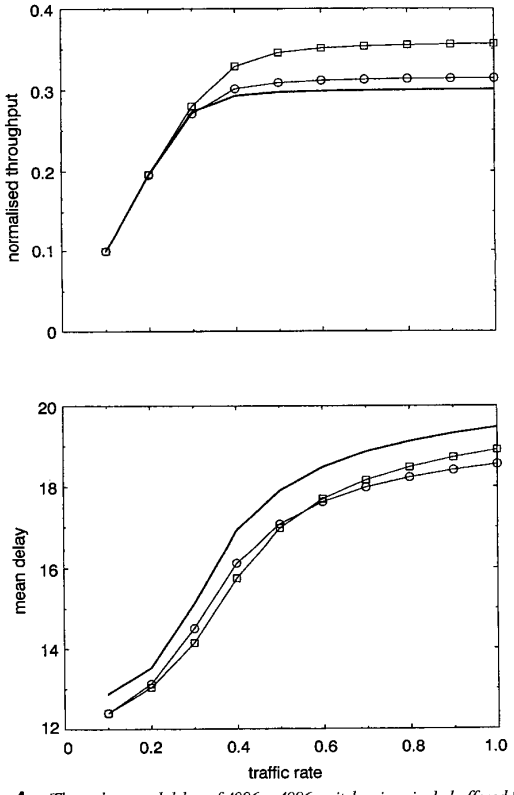

Fig. 4 *Throughput and delay of 4096 × 4096 switch using single-buffered SEs*
——— simulation
○——○ proposed
□——□ MY model

Fig. 4 shows the normalised throughput and the mean delay (obtained from the proposed model and the MY model) as a function of the traffic rates for a 4096 × 4096 switch using single buffered SEs. The proposed model takes

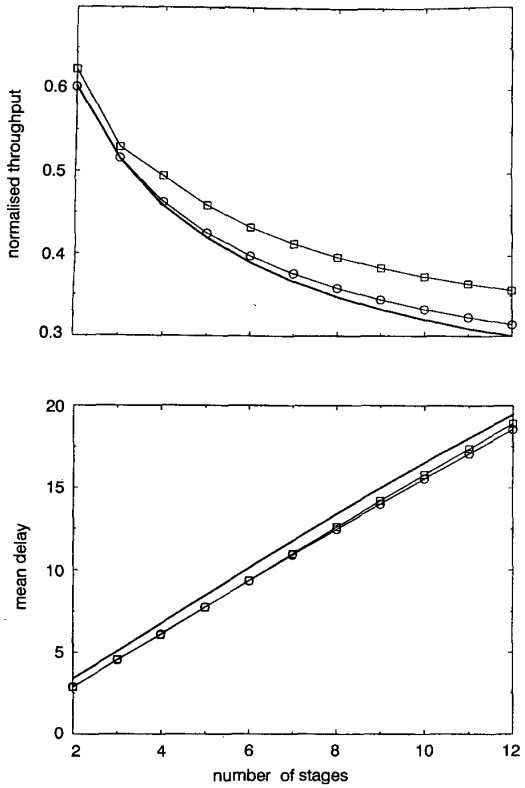

Fig. 5 *Throughput and delay of switches of different sizes using single buffered SEs*
——— simulation
○—○ proposed
□—□ MY model

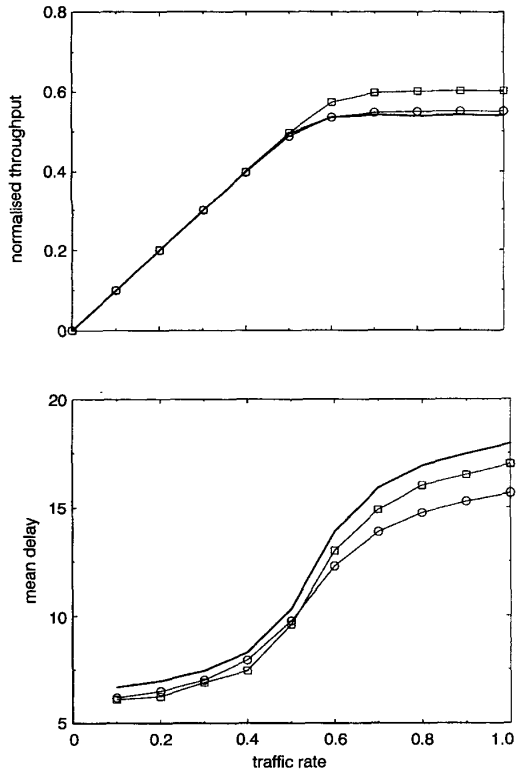

Fig. 6 *Throughput and delay for 64 × 64 switch using SEs with three buffers*
——— simulation
○—○ proposed
□—□ MY model

account of the transaction between buffers in adjacent stages and cycles and is more accurate than the MY model. The reason for the improved results from the proposed model is the fact that the model memorises the output for which a packet was blocked in an SE.

Fig. 5 shows the normalised throughput and the mean delay as a function of the number of stages obtained from the proposed model, the MY model and simulation. Fig. 5 reveals that the proposed model is much more accurate than MY model when the switch size is large. This is because, as the switch size increases, the number of stages also increase and results in an increased blocking in the switch. By taking a rigorous account of blocking in the proposed model, increased accuracy is obtained.

Fig. 6 shows the normalised throughput and the mean delay obtained from the proposed finite-buffer model against the MY finite buffer model for switches having three buffers per SE. The throughput from the proposed model is very close to the simulation results and better than the MY model. In the MY model, the analysis of the acceptance probabilities and the transition rate revealed that the destination of a blocked packet is randomly selected in succeeding cycles rather than being to the same output. This exemplifies the dependencies between buffer operations in consecutive cycles.

7 Conclusions

An analytical model permits a fast and inexpensive method of performance evaluation of switches. Previous analytical models, in most cases, have not taken a rigorous account of the blocked state of the packets, and hence the results have not been accurate. Models which produce good results are based either on exhaustively enumerating the possible states of a buffer in an SE, which makes the models too complex for large sized switches (say 1024 × 1024), or on unrealistic assumptions and thus require the use of empirical parameters [12].

New analytical models for Banyan-type switches constructed from single and finite input-buffered switching elements have been presented in this paper. The proposed models produce better results than the most recently published model. It has been shown that, in order to accurately and effectively model the behaviour of switches, the blocked packets need to be properly accounted for.

Performance measurements obtained from the proposed single buffer model were compared against the MY [2] single buffer model. The proposed model has been found to have higher accuracy. Comparison with simulation results revealed that the proposed model is very accurate when the switch size is large. The reason for the improved results from the proposed model is the fact that the model memorises the history of a blocked packet, whereas other models do not.

The second proposed model is for switches with finite-buffered SEs. It is a generalisation of the proposed single buffer model. Results demonstrate that the proposed finite-buffered model is more accurate than the MY model. The throughput obtained from the proposed model is very close to simulation results.

Through extensive simulations and comparison with the best available model in the literature, it has been shown that the model proposed in this paper is better than previously published models in terms of accuracy. The proposed models can be used to evaluate the influence of buffer size and the number of stages on the switch performance. Such evaluations can be used to optimise the switch for the best cost/performance ratio. The models can be extended to

other types of nonuniform traffic patterns [13], structures, and operating conditions of switches.

8 References

1 TOBAGI, F.A.: 'Fast packet switch architectures for broadband integrated services digital networks', *Proc. IEEE,* 1990, **78**, (1), pp. 133–167

2 MUN, Y., and YOUN, H.Y.: 'Performance analysis of finite buffered multistage interconnection networks', *IEEE Trans. Comput.,* 1994, **43**, (2), pp. 153–162

3 KIM, H.S., and LEON-GARCIA, A.: 'Performance of buffered Banyan networks under non-uniform traffic patterns', *IEEE Trans. Commun.,* 1990, **38**, (5), pp. 648–658

4 ZHOU, B., and ATIQUZZAMAN, M.: 'Performance of output-multibuffered multistage interconnection networks under general traffic patterns'. Proceedings of IEEE INFOCOM '94: conference on *Computer communications*, Toronto, Canada, 1994, pp. 1448–1455

5 ATIQUZZAMAN, M., and AKHTAR, M.S.: 'Performance of buffered multistage interconnection networks in non-uniform traffic environment'. Proceedings of Seventh international *Parallel processing* symposium, California, 1993, pp. 762–767

6 ATIQUZZAMAN, M., and AKHTAR, M.S.: 'Effect of non-uniform traffic on the performance of unbuffered multistage interconnection networks', *IEE Proc., Comput. Digit. Tech.,* 1994, **141**, (3), pp. 169–176

7 ATIQUZZAMAN, M., and AKHTAR, M.S.: 'Effect of hot spots on the performance of multistage interconnection networks'. Proceedings of FRONTIERS 92: the fourth symposium on the *Frontiers of massively parallel computation*, Virginia, 1992, pp. 504–505

8 JENQ, Y-C.: 'Performance analysis of a packet switch based on single-buffered Banyan network', *IEEE J. Sel. Areas Commun.,* 1983, **SAC–1**, (6), pp. 1014–1021

9 YOON, H., LEE, K.Y., and LIU, M.T.: 'Performance analysis of multibuffered packet-switching networks in multiprocessor systems', *IEEE Trans. Comput.,* 1990, **39**, (3), pp. 319–327

10 THEIMER, T.H., RATHGEB, E.P., and HUBER, M.N.: 'Performance analysis of buffered Banyan networks', *IEEE Trans. Commun.,* 1991, **39**, (2), pp. 269–277

11 HSIAO, S.H., and CHEN, R.Y.: 'Performance analysis of single-buffered multistage interconnection networks', *IEEE Trans. Commun.,* 1994, **42**, (9), pp. 2722–2729

12 HSIAO, S.H., and CHEN, R.Y.: 'Performance analysis of single-buffered multistage interconnection networks'. Proceedings of Third IEEE symposium on *Parallel and distributed processing*, 1991, pp. 864–867

13 WU, L.T.: 'Mixing traffic in a buffered Banyan network'. Proceedings of 9th *Data communications* symposium, 1985, pp. 134–139