Throughput Analysis of Multiple Input-Queueing in ATM Switches†

Christos Kolias  
Department of Computer Science  
University of California at Los Angeles, Los Angeles, CA 90024, U.S.A.  
email: kolias@cs.ucla.edu, tel.: +1 310 825-1563, fax: +1 310 825-2273

Leonard Kleinrock  
Department of Computer Science  
University of California at Los Angeles, Los Angeles, CA 90024, U.S.A.  
email: lk@cs.ucla.edu, tel.: +1 310 825-2543, fax: +1 310 825-2273

Abstract
In this paper we investigate various schemes of input-queueing ATM switching systems. We first give a brief description of the Odd-Even switch model which is followed by an approximate analysis for evaluating its throughput. We then introduce an extension of the Odd-Even model which employs a Multiple Input-Queueing strategy, where an input port is expanded into m queues. In fact, we consider two policies as far as arbitration among the input queues is concerned and we show how throughput can increase as m gets larger. We also comment on the special case where $m = N$, for an $N \times N$ switch, and show that the achieved throughput is actually 100%. We call this last scheme Virtual Output-Queueing. The models under examination assume a uniform output destination distribution and a Bernoulli process for the cell arrivals.

Keywords
Throughput, Odd-Even, Multiple Input-Queueing, ATM Switches.

1 INTRODUCTION

As we witness the proliferation of the ATM technology, it becomes quite evident the critical role switches are required to play to support the objectives of ATM networks. Switches, as an integral part of an ATM network not only support fast cell-relaying but also provide for network management, such as connection admission control, flow and congestion control, resource allocation, multicasting and other functions. This has allowed for various switch

†Supported by ARPA/CSTO under contract DABT63-93-C-0055, The Distributed Supercomputer Supernet - A Multi Service Optical Intelligent Network.
architectures to emerge and be proposed as standards. As a matter of fact, we cannot single out a particular switch design as a superior one, since each switching system is suitable, in terms of functionality, according to the specific network requirements. ATM switching systems have been covered quite extensively in the literature (Chen, 1991), (Onvural, 1994) and different schemes have been studied and characterized in terms of various aspects, such as location of buffers, switch fabric design, switching (i.e. time vs. space division), routing and congestion control (i.e. leaky bucket) and buffer management.

If cells are stored between the input ports and the switch fabric then the switching architecture is characterized as an input-queueing. On the other hand, if buffering of cells occurs between the switch fabric and the output trunks then the switch is classified as an output-queueing. A third option exists, where cell buffers are located inside the switch fabric (thus shared among the input and output ports), in which case this queueing discipline is described as internal or central queueing. Hybrid switch architectures exist adopting a combination or extension (i.e. recirculation) of these buffering strategies with the analogous cost in terms if complexity and management.

Input-buffered switches exhibit a blocking phenomenon known as head-of-line (HOL). HOL blocking occurs when more than one cell attempts to access the same output port. Also, since during the period of a time slot only one cell can be switched to the requested output, the remaining cells destined to the same output port can either remain buffered at their HOL positions (and try again in the next slot), thus blocking the cells queued behind them, or they can be simply discarded from the input queues, so no blocking occurs. As a result, HOL blocking can severely affect the throughput (which we define here as the utilization of the output ports averaged over the number of ports) of an input-queueing switch and normally large buffers are needed. For an ordinary input-buffered switch throughput is found to be $2 - \sqrt{2} \approx 0.586$ (Karol, 1987). In contrast, in output-buffering, each output is assigned a dedicated FIFO buffer where cells contending for the same output are all switched, within the same time slot, to the corresponding FIFO (but only one can be transmitted on the output link). Therefore, with output-queueing the ideal throughput of 100% can be actually achieved.

In this paper we demonstrate a few alternatives to ease the HOL blocking and consequently improve the switch throughput. The simulation results presented refer to the maximum attainable throughput (we consider a saturation point) for different values of $N$ (switch size). The analytical results show the asymptotic ($N \to \infty$) throughput behaviour of the switch under investigation. Switches are modeled as multi-queue, multi-server, time-slotted queueing systems. Queues at the inputs are FIFOs and assumed of infinite size. In clarifying some of the terminology we will be occasionally using throughout the paper, by new cells we mean any incoming cells that arrive at the input ports, while by fresh cells we describe these cells that move up to the HOL position (they were either queued or, if the FIFO was empty, they are new).

As for the organization of the paper: section 2 presents a rather concise description of the Odd-Even model and continues with an approximation analysis of the model. In section 3 we extend and generalize the Odd-Even model by introducing the Multiple Input-Queueing switch, which we study under two different arbitration policies. Finally, section 4 gives some concluding remarks.

*We would like to thank Larry Roberts of Connectware Inc. for bringing this model to our attention.*
2 THE ODD-EVEN MODEL

2.1 Switch Description

In this subsection we briefly present the Odd-Even switch (a more thorough treatment can be found in (Kolas, 1996)). We consider an $N \times N$ crossbar, nonblocking, input-buffered switch, where the innovation is that each input is split (expanded) into two FIFO queues which we call Odd and Even for this type of dual input queueing. By convention, we refer to outputs numbered as 2,4,6,... as even while to 1,3,5,... as odd ports. An incoming cell destined to an odd (even) numbered output port joins the odd (even) queue of the input port to which it was fed and waits for its turn to be switched to its output. Queued cells move up to the head of these queues that switched their HOL cells, so that they can get involved in the next output contention resolution cycles. Contention resolution for the HOL cells takes place in two consecutive contention resolution rounds. Arbitration during the first round involves the HOL cells at the even input queues. In the second round cells at the HOL of the odd input queues contend for the odd output addresses. However, those input ports whose even queues could not access an output port in the first round, because they either lost the contention (and therefore were blocked) or simply did not have any HOL cells present, are allowed to participate in a contention among their odd queues in the subsequent second round. In this fashion, an input port always gets a chance to route a cell either from an even or an odd queue, but not from both, within the same time slot. For fairness, if the even FIFOs were tried first in a time slot the odd ones are polled in the next one and vice versa.

Traffic intensity is equal for every input link (same traffic load is applied to each of the $N$ input ports). The output address of an incoming cell is assigned randomly among the $N$ output ports with equal probability. Both these assumptions characterize our system as a homogeneous one.

As far as the selection policy is concerned, the longest waiting in the HOL position cell is chosen among those contending for the same output, by the arbitration controller (when there is a tie the lowest numbered input queue is the winner). However, in the analysis that follows, we implicitly assume that the winning HOL cell is chosen randomly, since the selection policy has no bearing on the switch throughput (for this homogeneous system).

2.2 Approximate Analysis

Because of the various service dependencies introduced by the arbitration policy, where we assumed that even slots (during which even FIFOs can first transmit) are interchanged with odd slots (odd FIFOs can first transmit), an exact analysis of the switch's behaviour becomes intractable in terms of calculating the stochastic quantities involved. Therefore, we propose an approach where we analyze an approximate model, which is based on assuming a slightly modified arbitration policy. More specifically, we assume that all time slots are characterized as even, which means that the even FIFOs are given some priority in transmitting their HOL cells. The intuition behind this approach is that unclaimed output ports will still get the opportunity to receive (serve) HOL cells, given that output addresses are uniformly distributed. Due to this modified service policy we expect that the throughput achieved by the even output ports will be higher than the one for the odd outputs.
We make the following distinction among the $N$ input ports: we call an input port available (for the second round) if no HOL cell was switched from its even FIFO (either there was not one or it lost the contention). Conversely, it is unavailable if the corresponding even FIFO switched a cell through one of the even output ports. We denote the throughputs achieved by the even and the odd output ports, as $\gamma_E$ and $\gamma_O$ respectively.

We first derive an exact expression for $\gamma_E$ and then approximate $\gamma_O$, where, as we will see, the latter actually depends on $\gamma_E$. Our approach is similar to the one described in (Hui, 1987). We focus on an even numbered output $j$ (the “tagged” output) and we denote by $N_j^E$ the number of HOL cells at the even queues destined for output $j$. In steady state, $\gamma_E$ can be defined as:

$$\gamma_E = \lambda_E \frac{2}{N} \sum_{j \in E \text{ even}} E[\epsilon(N_j^E)] = E[\epsilon(N_j^E)],$$

where $\lambda_E$ is the arrival rate of new cells at the even queues, assuming a non-saturation situation and $\epsilon(x)$ is an indicator function ($\epsilon(x) = 0$ if $x \leq 0$, $\epsilon(x) = 1$ if $x > 0$). Note that the even HOL cells can be switched from any of the $N$ (even) input queues to any of $N/2$ even output ports.

Let also $N_b^E$ denote the total number of HOL cells blocked (i.e. lost the contention) at the even queues after the first contention round. Then:

$$N_b^E = \sum_{j \in E \text{ even}} N_j^E - \sum_{j \in E \text{ even}} \epsilon(N_j^E).$$

(2)

Taking now expectations in (2) and combining it with (1), we get, by symmetry:

$$\lambda_E = \gamma_E = E[N_j^E] - \frac{2E[N_b^E]}{N}. \tag{3}$$

Let $\rho_E$ be the steady-state probability that there is a fresh even cell at the HOL position of an even queue given that the HOL cell has departed and $M_E$ denote the number of released even HOL positions (those that had a cell switched) at the end of the first contention round, then clearly:

$$M_E = N - N_b^E. \tag{4}$$

The flow conservation relationship describing the system for the even queues states:

$$E[M_E] \rho_E = \frac{N}{2} \lambda_E, \tag{5}$$

where intuitively $\frac{E[M_E] \rho_E}{N/2} = \lambda_E$ is simply the expected fraction of the busy even servers (output ports), which defines the throughput of the even outputs. By taking expectations on both sides in (4) and using (5) we get:

$$\frac{E[N_j^E]}{N} = 1 - \frac{\lambda_E}{2\rho_E}. \tag{6}$$
THE ODD-EVEN MODEL

Let now \( K_j^E \) denote the number of HOL cells destined to output \( j \), in the next time slot, then:

\[
K_j^E = N_j^E - \epsilon(N_j^E) + A_j^E,
\]

where \( A_j^E \) is the number of fresh HOL cells destined for output \( j \). Forming the expectations on both sides of (7) and considering the steady-state case we have:

\[
E[K_j^E] = E[N_j^E].
\]

Taking into account (1) and (8), (7) yields\(^1\):

\[
E[A_j^E] = E[\epsilon(N_j^E)] = \lambda_E.
\]

In finding \( E[N_j^E] \) we first square (7) as follows:

\[
(K_j^E)^2 = (N_j^E)^2 + \epsilon(N_j^E) + (A_j^E)^2 - 2N_j^E + 2N_j^E A_j^E - 2\epsilon(N_j^E)A_j^E,
\]

where by taking expectations and then using (8) and (9) we further get:

\[
E[N_j^E] = E[A_j^E] + \frac{E[A_j^E(A_j^E - 1)]}{2(1 - E[A_j^E])}.
\]

In determining \( E[A_j^E(A_j^E - 1)] \), as in (Hui, 1987), we can argue that \( A_j^E \) becomes Poisson(\( \lambda_E \)) as \( N \to \infty \)\(^2\), thus we have:

\[
E[A_j^E(A_j^E - 1)] = \lambda_E^2,
\]

and \( E[N_j^E] \) can be directly expressed as:

\[
E[N_j^E] = \lambda_E + \frac{\lambda_E^2}{2(1 - \lambda_E)}.
\]

Applying (6) and (12) in (3), we further obtain:

\[
\lambda_E = \lambda_E + \frac{\lambda_E^2}{2(1 - \lambda_E)} - 2\left(1 - \frac{\lambda_E}{2\rho_E}\right),
\]

which leads to the following equation:

\[
(2 - \rho_E)\lambda_E^2 - 2(2\rho_E + 1)\lambda_E + 4\rho_E = 0.
\]

Since we are interested in maximizing throughput, then by letting \( \rho_E = 1 \) (i.e. all even HOL

\(^1\)Also derived from first principles about flow conservation.

\(^2\)Note that \( N \to \infty \), practically means a switch with sufficiently large \( N \), i.e. close or larger than 100.
positions are occupied), we get:

\[ \lambda_E = 3 - \sqrt{5} \approx 0.764, \]  

(15)

which, again, represents the throughput achieved by the even output ports.

It now remains to obtain the throughput of the odd output ports. Because not all the input ports are available for the second contention round it is clear that the throughput of the odd output ports can be potentially limited due to that unavailability factor. Let \( \delta \) be the fraction of the \( N \) input ports that are available for the second arbitration round, then \( \delta N \) is the expected number of odd queues that are allowed to switch a HOL cell to an odd destination port.

Setting up the same equations as we did for evaluating \( \lambda_E \), where \( i \) is now the “tagged” odd output, we have:

\[ \gamma_O = \lambda_O = \frac{2}{N} \sum_{i \text{ is odd}} E[\epsilon(N_i^O)] = E[\epsilon(N_i^O)], \]  

(16)

or

\[ \gamma_O = \lambda_O = E[N_i^O] - \frac{2E[N_i^O]}{N}, \]  

(17)

where \( \lambda_O \) is the input arrival rate at the odd queues and \( N_i^O \) is the number of available odd HOL cells destined to (odd) output \( i \) and \( N_i^O \) is the number of blocked odd HOL cells. Then, by letting \( M_O \) be the number of odd ports that became free at the end of the second contention round, we have:

\[ E[M_O] = \delta N - E[N_i^O]. \]  

(18)

The flow conservation for the odd outputs is expressed as follows:

\[ E[M_O] \rho_O = \frac{N}{2} \lambda_O. \]  

(19)

Then, by taking expectations in (18) and using (19) we have:

\[ \frac{E[N_i^O]}{N} = \delta - \frac{\lambda_O}{2 \rho_O}. \]  

(20)

If \( K_i^O \) is the r.v. for the number of available odd HOL cells destined for output \( i \) in the next time slot and \( A_i^O \) is the number of fresh odd HOL cells then, in general, we cannot claim that:

\[ K_i^O = N_i^O - \epsilon(N_i^O) + A_i^O, \]  

(21)

since queues that were available may become unavailable in the next slot (after the even queues arbitration round) while any of those that were unavailable may become available.
THE ODD-EVEN MODEL

However,

$$\sum_{i \text{ is odd}} K_i^O = \sum_{i \text{ is odd}} N_i^O - \sum_{i \text{ is odd}} \epsilon(N_i^O) + \sum_{i \text{ is Odd}} A_i^O + R - B$$

(22)

where $R$ denotes the number of input ports that become available in the next time slot while $B$ represents the number of input ports that become unavailable (and which may include any fresh HOL cells). At steady state the expected number of unavailable input ports should remain the same, therefore $E[R] = E[B]$. Now, taking expectations in (22) we have :

$$\sum_{i \text{ is Odd}} E[K_i^O] = \sum_{i \text{ is Odd}} E[N_i^O] - \sum_{i \text{ is Odd}} E[\epsilon(N_i^O)] + \sum_{i \text{ is Odd}} E[A_i^O],$$

(23)

and we can make the assumption that, by symmetry :

$$E[K_i^O] = E[N_i^O] - E[\epsilon(N_i^O)] + E[A_i^O],$$

(24)

and since $E[K_i^O] = E[N_i^O]$, (24) then implies :

$$E[A_i^O] = E[\epsilon(N_i^O)] = \lambda_O.$$

(25)

Proceeding as before, we can approximate the distribution of $A_i^O$ by a Poisson($\lambda_o$) distribution as $N \to \infty$ and conclude that :

$$E[N_i^O] = \lambda_O + \frac{\lambda_O^2}{2(1 - \lambda_O)}.$$

(26)

By substituting (20) and (26) in (17) we get :

$$\lambda_O = \lambda_O + \frac{\lambda_O^2}{2(1 - \lambda_O)} - 2(\delta - \frac{\lambda_O}{2\rho_O})$$

(27)

which, for $\rho_O = 1$, yields :

$$\lambda_O = 1 + 2\delta - \sqrt{4\delta^2 + 1}.$$

(28)

Now, since $\delta$ denotes the fraction of those input ports that are available at the end of the first arbitration round, then we can express it, using (6), as

$$\delta = 1 - \frac{\lambda_E}{2\rho_E}.$$ 

(29)

Then from (28) and (29) and for $\rho_E = 1$ we finally get $\lambda_O \approx 0.646$, which is the throughput for the odd output ports.

Note, that in (28) for $\delta = 1$ (i.e. all input ports are available at the beginning of the second round) $\lambda_O = 3 - \sqrt{5}$ and for $\delta = 0$, obviously, $\lambda_O = 0$. Also, for $\delta = \frac{1}{2}, (N/2$ input ports are available), we get $\lambda_O = 2 - \sqrt{2}$, which of course agrees with previous analysis (Karol, 1987) (i.e. consider an $\frac{N}{2} \times \frac{N}{2}$ switch).
We also observe that \( 1 + 2\delta - \sqrt{4\delta^2 + 1} \) is an increasing function of \( \delta \) and from (13) we actually have \( \delta = \frac{|\lambda E - \lambda O|}{2(\lambda E + \lambda O)} \) which is maximized when \( \lambda E \) is maximized, namely for \( \rho E = 1 \), which yields \( \lambda E = 3 - \sqrt{5} \). Thus, \( \rho E = \rho O = 1 \) yields the maximum achievable throughput for both the even and the odd outputs and therefore for the switch.

Since we now have the throughput for both the even and the odd output ports, we can finally find, as an approximation, the throughput for the whole system, namely the Odd-Even switch, by normalizing \( \lambda E \) and \( \lambda O \), as follows:

\[
\gamma_s = \frac{\lambda_O + \lambda_E}{2} \approx 0.705.
\]

We notice that \( \gamma_s \approx 0.705 \) is in a very good agreement (within 1% range) to the simulation result of 0.713 for the Poisson, uniform traffic case (Kolias, 1996).

3 \ MATERIAL-INPUT-QUEUEING SWITCHES

In this section we present an extension of the Odd-Even model. Assuming an \( N \times N \) single-stage, input-buffered switch, instead of having only two FIFOs (odd and even) per input port, we allow \( m \) queues per input port, where \( m \leq N \) (the switch size). It is essential to emphasize here the distinction between an input FIFO queue and an input port. As in the Odd-Even model, output ports are partitioned into \( m \) groups (it is not particularly important how we allocate the output ports to the groups, as long as we assume uniformity of the output addresses). That means each of the \( m \) queues is associated with one or more (depending on \( N \)) output ports, so that an incoming cell joins a particular queue according to its output port destination. We realize that the larger the \( m \) is, the more complex the system becomes, in terms of implementation of the arbitration scheme. There is no extra cost induced regarding the additional buffers, since each input port buffer (of size i.e. \( b \)) can be throught as being partitioned into \( m \) smaller buffers (of size \( b/m \)).

We study this Multiple Input-Queueing system in two different versions with respect to the arbitration policies.

3.1 \ Policy A

The arbitration process, under this policy, is a direct extension of the one mentioned in the Odd-Even scheme and takes place in \( m \) consecutive contention rounds (within the same time slot). Therefore, it can be considered as a generalization of the Odd-Even scheme. At the \( i \)-th round, only FIFOs that are available (whose input ports have not switched cells during the previous \( i-1 \) rounds) can compete for output ports. This kind of policy can limit the switch throughput, since even if there are HOL cells destined to a particular output, that output might stay idle. However, we notice from the simulation results (figure 1) that as \( m \) grows to infinity, the asymptotic throughput approaches 1 very quickly (always under the assumption of saturation at the inputs).

In fact, for \( m = N \), that is, the number of FIFOs per input port is equal to the total number of outputs, throughput is 1. Basically, in this form of input-queueing each input port has a dedicated buffer for each of the outputs. Therefore, the achieved throughput is
100% since, effectively, there is no HOL blocking. This type of buffering has exactly the same result as the pure output buffering scheme, where there is no blocking whatsoever and cells are switched immediately to their corresponding output buffers. We call this type of input queueing, where essentially output buffering is emulated by buffers at the inputs, Virtual Output-Queueing.

Figure 1 illustrates the simulation behaviour of this general model, for different values of \( m \), namely for \( m=1 \) (the ordinary input-buffered switch, (Karol, 1987) ), \( m=2 \) (the Odd-Even switch), and \( m=5, 10, 20, 50 \) and 100. We see that for \( m > 10 \) we obtain diminishing gains in throughput and most notably when \( m = 50 \) is doubled.

![Figure 1 Throughput of a Multiple Input-Queueing Switch under policy A (simulation).](image)

### 3.2 Policy B

Here we study a variation of policy A, where we allow all \( N \) input FIFOs (not only from those input ports that did not transmit a cell during the previous rounds) to participate in each of the arbitration rounds. Actually since we do not distinguish between available and unavailable input ports (all input ports are available), we can assume that the contention for the outputs is completed in just one round, where all \( mN \) input FIFOs can participate. In other words, we can view this system as an \( mN \times N \) switch where the \( N \) inputs are expanded into \( mN \) and a new cell joins a particular FIFO queue, depending on the cell’s output address. Under this policy the \( mN \) queues are considered to be independent.

Let us now concentrate on analytically obtaining the switch throughput for this type of multiple input queueing. As in subsection 2.2, let \( \gamma \) be the total throughput of the switch and \( N_j \) denote the number of HOL cells that are destined to a “tagged” output \( j \). Then, at
steady-state, we have:

\[ \gamma = \lambda = \frac{1}{N} \sum_{j=1}^{N} E[e(N_j)] = E[e(N_j)], \quad (31) \]

where \( \lambda \) is the probability that a new cell arrives at an input port at the beginning of a time slot.

If \( N_b \) is defined as the number of all HOL cells that became blocked at the end of a slot (after the arbitration phase is over), then clearly:

\[ N_b = \sum_{j=1}^{N} N_j - \sum_{j=1}^{N} e(N_j). \quad (32) \]

Taking expectations on both sides in (32) and using (31) we can then rewrite throughput as follows:

\[ \lambda = \gamma = E[N_j] - \frac{E[N_b]}{N}. \quad (33) \]

Recall that, if \( M \) represents the number of unblocked FIFOs then:

\[ E[M] = mN - E[N_b]. \quad (34) \]

The flow conservation rule implies:

\[ E[M] \rho = N \lambda, \quad (35) \]

where \( \rho \) is again the steady-state probability that a fresh cell occupies the HOL position immediately following the departure of the HOL cell. Taking into account (34) and (35), (33) can be modified as:

\[ \lambda = E[N_j] - (m - \frac{\lambda}{\rho}). \quad (36) \]

If \( K_j \) denotes the number of HOL cells destined to the \( j \)-th output port during the next time slot and by letting \( A_j \) be the number of the “fresh” HOL cells destined to \( j \), then:

\[ K_j = N_j - e(N_j) + A_j. \quad (37) \]

Taking expectations in Eq. (37) then, since at steady-state \( E[K_j] = E[N_j] \), we get:

\[ \lambda = E[e(N_j)] = E[A_j]. \quad (38) \]

By squaring Eq. (37) and taking expectations we finally obtain:

\[ E[N_j] = E[A_j] + \frac{E[A_j(A_j - 1)]}{2(1 - E[A_j])}. \quad (39) \]
MULTIPLE INPUT-QUEUEING SWITCHES

As we mentioned before, as \( N \to \infty \), \( A_j \) becomes Poisson\((\lambda)\), thus:

\[
E[A_j(A_j - 1)] = \lambda^2,
\]

then (39) becomes:

\[
E[N_j] = \lambda + \frac{\lambda^2}{2(1-\lambda)}.
\]  

(41)

By substituting (41) in (36) we get:

\[
\lambda = \lambda + \frac{\lambda^2}{2(1-\lambda)} - (m - \frac{\lambda}{\rho}),
\]

(42)

where for \( \rho = 1 \), solving (42) we find:

\[
\lambda = m + 1 - \sqrt{1 + m^2},
\]

(43)

which is the maximum throughput of the m-FIFO multiple input queueing switch.

Figure 2 Throughput of a Multiple Input-Queueing Switch under policy B (simulation).

Figure 2 shows simulation results for switches with \( m = 1, 2, 5, 10 \) and 100. Applying the result of (43) for these values of \( m \) the consistency between the theoretical and the simulation results becomes obvious.

Note that for \( m = 1 \), (43) yields \( \lambda = 2 - \sqrt{2} \), while for \( m = 2 \), \( \lambda = 3 - \sqrt{5} \) (see (15)). Also, in the limit, as \( m \to \infty \), (43) yields \( \lambda = 1 \). Let's give the physical explanation of this result: as \( m \to \infty \) then also \( N \to \infty \) which practically means that there are as many FIFOs per input port as the total number of output ports, namely \( m = N \). For \( m = N \), policies
A and B have ultimately the same effect, in terms of throughput, and under a saturation situation; so our remarks in the previous subsection are reinforced here.

4 CONCLUSION

It is apparent that buffering (location, size, contention resolution) in an ATM switching system is of great importance to the switch designer, in terms of complexity, efficiency and certainly cost (i.e. hardware implementation). In this paper we were concerned with the performance efficiency of input-buffered switches, acknowledging that the other factors are significant too. In that respect, we studied various schemes using multiple input-queueing as their buffering strategy by analyzing their performance in terms of throughput and demonstrated how throughput can be improved.

5 REFERENCES


6 BIOGRAPHIES

**Leonard Kleinrock** received his B.S. degree in electrical engineering from the City College of New York in 1957 and the M.S.E.E. and Ph.D.E.E. degrees from the Massachusetts Institute of Technology in 1959 and 1963, respectively. He is currently a Professor in the Computer Science Department at UCLA. His research interests focus on performance evaluation and design of many kinds of networks (e.g. packet switching networks, packet radio networks, local area networks, metropolitan area networks, broadband and gigabit networks) and of parallel and distributed systems. Dr. Kleinrock is a member of the National Academy of Engineering, a Guggenheim Fellow, recipient of the 1982 I.M. Ericsson Prize-Sweden, recipient of the 12th Marconi Award-1986 and an IEEE Fellow.

**Christos Kolias** received his B.S. degree in mathematics from University of Athens,
Greece in 1989. In 1991 he was awarded a Fulbright Scholarship. Since then he has been with the Department of Computer Science at University of California, Los Angeles where he is pursuing his Ph.D. in the area of Computer Networks and Communications. He received the M.S. and Engineers degrees both in Computer Science from UCLA in 1993 and 1995 respectively. His current research interests are in the areas of Asynchronous Transfer Mode (ATM) and high-speed networks and in particular of ATM switching. Mr. Kolas is a student member of the IEEE.