

Dimensioning of an ATM switch with shared buffer and threshold priority *

J.F. Meyer ^a, S. Montagna ^b, R. Paglino ^b

^a Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI, USA

^b Central Research Laboratories, Italtel, Milano, Italy

Abstract

A number of recent studies have addressed the use of priority mechanisms in Asynchronous Transfer Mode (ATM) switches. This investigation concerns the performance evaluation and dimensioning of a shared-buffer switching element with a threshold priority mechanism (partial buffer sharing). It assumes that incoming ATM cells are distinguished by a space priority assignment, i.e., loss of a high priority cell should be less likely than loss of a low priority cell. The evaluation method is analytic, based on an approximate discrete-time, finite-state Markov model of a switch and its incoming traffic. The development focuses on the formulation of steady-state loss probabilities for each priority class. Evaluation of delay measures for each class is also supported by the model; results concerning the latter are illustrated without development. The analysis of loss probabilities is then used to dimension the buffer capacity and threshold level such that required maximum cell loss probabilities are just satisfied for each cell type. Moreover, when so dimensioned with respect to relatively stringent loss requirements, i.e., probabilities of 10^{-10} and 10^{-5} for high and low priority cells, respectively, we find that both loss performance and resource utilization are appreciably improved over a comparable switch without such a mechanism.

Keywords: Asynchronous transfer mode; Switching element; Partial buffer sharing; Loss performance

1. Introduction

Broadband Integrated Services Digital Networks (B-ISDNs) are intended to provide a variety of different teleservices, all on a single “universal” network. Moreover, such services can have widely differing Quality of Service (QoS) requirements which, at the cell (packet) level, translate to differences in permissible cell losses and cell transfer delays. Accordingly, B-ISDN architectures, notably those employing Asynchronous Transfer Mode (ATM) techniques, should be able to accommodate such differences without appre-

ciable reductions in resource utilization or significant increases in network complexity.

Generally, two types of priorities are considered in a teletraffic context: *time priority* and *space priority*. In a time priority system, cells are distinguished according to transfer delay requirements, where the higher the priority, the lower the average delay. Space priority (also referred as “loss” or “semantic” priority; see [2], for example) distinguishes cells according to loss requirements, where the higher the priority, the lower the intended loss probability. Historically, time priority schemes have received greater attention but, with the advent of strict loss requirements associated with certain vital services, concern with space priorities has begun to increase.

Specifically, with regard to various ATM switching and multiplexing elements, a number of recent studies [1,4–10] have evaluated space priority mechanisms (in certain cases, time priorities

* This work was partially carried out within the framework of the Telecommunication Project of the Italian National Research Council (CNR).

Correspondence to: J.F. Meyer, Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI, USA.

are also considered) via analysis and/or simulation. The investigation that follows likewise concerns a space priority mechanism and presumes that incoming cells are in one of two priority classes: high (Class 1) and low (Class 2). The switching element in question has N input ports, R output ports, a shared internal buffer of finite capacity K , and a threshold priority mechanism. Use of such a mechanism in conjunction with a shared buffer is sometimes referred to as *partial buffer sharing*. The principal objective of this study is to accurately assess the performance of such a switch, as compared to one without a priority mechanism. This includes discussion of how the shared buffer (capacity and threshold level) are dimensioned so as to optimize its effectiveness with respect to specified Class 1 and Class 2 cell loss requirements. In addition, we seek to esti-

mate the robustness of the system, as indicated by the effects of variations in switch and traffic parameter values.

The modeling approach we take is similar to that reported in [10]. However, unlike the latter, which assumes Poissonian cell arrivals, we account more precisely for the slotted nature of the ATM protocol format through use of a binomial traffic model. Our analysis also relates to that presented in [6]. There, the same kind of switch is considered with respect to a more general bursty traffic model. The switch model, on the other hand, is less refined than that described below, consisting of a single state variable that represents the total number of cells stored in the shared buffer. To obtain more exact results, we choose to represent the status of the buffer by a pair of state variables that convey (i) the number



John F. Meyer received the B.S. degree from the University of Michigan, Ann Arbor, the M.S. degree from Stanford University, Stanford, CA, and the Ph.D. degree in communication sciences from the University of Michigan in 1957, 1958, and 1967, respectively. He is currently a Professor in the Department of Electrical Engineering and Computer Science at the University of Michigan. He has been active in computer and systems research for over 30 years and has published widely in the areas of fault-tolerant computing and model-based evaluation of system performance, dependability, and performability. He joined the Michigan faculty in 1967 and, in addition to his university appointment, has held visiting research positions at laboratories in the USA, England, France, Italy, and Japan. Prior to 1967, he was a Research Engineer at the California Institute of Technology Jet Propulsion Laboratory where his contributions included the first patent issued to the National Aeronautics and Space Administration.



Sergio Montagna was born in Redavalle, Italy in 1955. He received a degree in physics from the University of Pavia (Italy) in 1978. He joined the Central Research Laboratories of Italtel in January 1983. His current research is concerned with performance/performability evaluation of Asynchronous Transfer Mode systems and networks.



Roberto Paglino was born in Italy in 1962. He graduated from the Università Statale di Milano (Milan, Italy), where he obtained a degree in computer science. Since 1987 he has been with Italtel as a researcher in the System Engineering Department of the Central R&D Laboratories. His interests are in the performance and architecture of ATM systems.

of cells tagged for some arbitrary but fixed output port and (ii) the number of cells tagged for the remaining $R - 1$ output ports.

Section 2 describes the above assumptions in greater detail. The system model (switch plus traffic) is then developed in Section 3, along with the solution method and the resulting formulations of state-occupancy probabilities and steady-state loss probabilities. Applying this analysis technique, Section 4 presents evaluation results aimed at satisfying the above-stated objectives and, in particular, describes how the buffer capacity and threshold level can be dimensioned via application of the evaluation data. Conclusions drawn from the study are then summarized in the closing section of the paper.

2. System assumptions

As indicated in the introduction, the system in question is an ATM switching element (the object system) as it operates in the presence of space prioritized ATM cell traffic (the environment). Assumptions regarding the nature of each are as follows, where discussion is limited, for the most part, to details that bear on the subsequent evaluation study.

2.1. Object system

The object of the evaluation is an $N \times R$ ATM switch, where N and R are positive integers denoting the number of input ports and output ports, respectively. Although our analysis applies, in theory, to any such values of N and R , all the evaluation results (Section 4) pertain to a 16×8 switch. The bit-rate capacities of all input and output ports are assumed to be identical. Switching of a cell from an input port to an output port is tag-based, with temporary internal storage provided by a buffer having a finite capacity K . Logically, the buffer is organized as R first-come-first-served queues, one for each output port. Physically, however, the capacity is shared, i.e., the number cells of stored in any one logical queue can vary, subject only to the constraint that, when summed over all R queues, the total number of cells does not exceed K .

Fixed-length ATM cells enter and depart the switch during successive operational cycles; each

such cycle, called a *time slot*, is likewise fixed in duration. A *threshold priority mechanism* controls access to the shared buffer according to the priority status of an incoming cell, indicating by a priority bit in the cell's header. A low priority cell is accepted during a time slot if and only if the number of cells (both high and low priority) in the buffer is less than some designated *threshold level* S , where S is an integer in the range $0 \leq S \leq K$. The ratio S/K of threshold level to buffer capacity is referred to as the *threshold ratio*. In the degenerate case where $S/K = 0$, all Class 2 traffic is rejected. At the other extreme where $S/K = 1$, Class 1 and Class 2 cells receive equal treatment, i.e., the switch behaves as one without a priority mechanism. During each time slot, operation of the switch can be viewed, effectively, as a sequence of three primitive operations.

(1) **Send:** The cell at the head of each non-empty logical queue is transmitted and the corresponding storage space in the buffer is freed.

(2) **Check:** The total number of cells remaining in the buffer (after the send operation) is checked.

(3) **Receive:** Each input port is scanned and incoming cells are accepted or rejected (lost) according to the following criteria. A cell of high priority is accepted as long as space remains in the buffer; otherwise it is rejected. A cell of low priority is accepted if the value recorded during the check operation is less than S ; otherwise it is rejected, regardless of any free space that may remain in the buffer.

2.2. Environment

Due to the discrete-time nature of cell processing in an ATM environment, the number of incoming cells during a time slot is random, ranging from 0 to N (the number of input ports). More precisely, looking ahead to model construction (Section 3), we assume that the probability of a cell arriving at a given input port during a time slot is a fixed value p ($0 < p \leq 1$). Moreover, such an arrival event is assumed to be statistically independent of past arrivals, i.e., cells arrive at a port as a Bernoulli process with parameter p . Assuming further that arrival streams among different ports are independent, the number of cell arrivals during a time slot is a binomially distributed random variable, the parameters of the

distribution being N and p . Each arriving cell is tagged for transmission (departure) from a particular output port, where this destination is assumed to be uniformly distributed over the R output ports.

Given these assumptions, it follows that the average number of cell arrivals per time slot is pN . Assuming no cell losses, this translates into a average cell departure rate, per output port, of

$$\rho = \frac{pN}{R} \quad \text{cells per time slot.}$$

With a slight abuse of terminology, ρ is referred to as the switch's *offered load* (or simply *load*) since it expresses service demands inherited by an output port under ideal, no-loss conditions. Equivalently, ρ may be viewed as output port utilization, i.e., the steady-state fraction of time that an output port is busy (again assuming no cell losses).

As for priorities, we presume that incoming cells have been sufficiently merged via statistical multiplexing so that the priority class of an arriving cell has a fixed distribution and, moreover, is independent of the history of prior arrivals. Letting p_h and p_l denote the probabilities of an arrival being high priority (Class 1) and low priority (Class 2), respectively ($p_h + p_l = 1$), it follows that, during each time slot, a high priority cell arrives at an input port with probability $p \cdot p_h$. Similarly, the probability of a low priority arrival is $p \cdot p_l$.

Relative to the switch's steady-state operation, the probability p_h can also be interpreted as the ratio between rates of high priority traffic and total traffic, i.e., the *load ratio*

$$p_h = \alpha_h / \alpha,$$

where α_h is the arrival rate (average number of arrivals per time slot) of high priority cells and α is the arrival rate of all cells. This follows immediately from the fact that $\alpha_h = p \cdot p_h N$ and $\alpha = pN$. This parameter, along with the switch's threshold ratio S/K and offered load ρ (a derived parameter involving both switch and traffic; see above), are key considerations in both the construction and solution of an appropriate system model.

3. System model

The model we use to evaluate loss and delay performance is a discrete-time, finite-state Markov process. This is a natural choice since we are modeling a slotted system with a finite buffer capacity. Its more precise nature is described at the outset of this section. This is followed a discussion of how the model is solved.

3.1. Model construction

To begin, we let

$$T = \{0, 1, 2, \dots\}$$

denote the time domain of the model where, following initial time 0, time n ($n \geq 1$) marks the end of the n th time slot. Hence, the difference between any two successive elements of T has the interpretation of a fixed time interval, its length being the duration of a time slot. To support exact evaluations of cell loss probabilities and cell delay distributions, a relatively refined notion of state is required for the switch's shared buffer. Specifically, given that output ports are named by integers from the set $\{1, 2, \dots, R\}$, for output port i and time $n \in T$, we let $X_{i,n}$ denote the random variable

$X_{i,n}$ = the number of cells in the buffer,
at time n , tagged for output port i .

Note that since the buffer is shared and has capacity K , each such variable takes values in the set $\{0, 1, \dots, K\}$. If, further, we define the state of the model at time n to be the vector-valued random variable

$$X_n = (X_{1,n}, X_{2,n}, \dots, X_{R,n}),$$

then the stochastic process

$$\{X_n | n \in T\}$$

is sufficiently detailed to permit exact solutions. However, for realistic values of R and K , the state space of this process can become intractably large.

Alternatively, it is possible to focus on the behavior of a single output port, thereby simplifying the model considerably. But to support evaluation of the desired measures, one must then assume that the (logical) queues associated with distinct output ports behave independently, i.e.,

at any time n , the collection of variables $\{X_{1,n}, X_{2,n}, \dots, X_{R,n}\}$ is statistically independent. Unfortunately, this is not the case. In addition to some dependence that results from a finite buffer capacity, i.e., the condition

$$\sum_{i=1}^R X_{i,n} \leq K, \quad \text{for all } n \in T,$$

even with unbounded capacity, there is dependence inherited from the incoming traffic. As first observed in [3] for traffic of the type assumed here (although not prioritized), this is due to the fact that at most N cells arrive during a given time slot. Accordingly, knowledge that certain arrivals are tagged for some designated output port i reduces the probability of receiving a cell (during the same time slot) that's destined for another port j .

To illustrate this for a specific (albeit extreme) case, suppose that the system is empty at time n and, during the next time slot, there are N arrivals, all of which are tagged for output port i . Then, for any other port j , the conditional probability that $X_{j,n+1} > 0$ given that $X_{i,n+1} = N$ is 0. As shown in [3], such dependence results in a negative correlation between the variables $X_{i,n}$, $X_{j,n}$ associated with any pair of output ports i and j ($i \neq j$). Moreover, if this correlation is ignored via an assumption of independence, the buffer capacity required for a specified steady-state loss probability can be appreciably overestimated, especially under heavy load conditions. For example, a comparison made in [3] for a 16×16 switch and a specified loss probability of 10^{-6} reveals a 30% discrepancy in buffer capacity estimates (uncorrelated as compared to negatively correlated) when $\rho = 0.9$.

As a compromise between these two extremes, we employ the following approximation. The queue status for one of the output ports, say port 1, is represented exactly. Those of the remaining ports are then approximated via a second variable that is the sum of individual state variables for ports 2 through R . More precisely, in terms of the variables $X_{i,n}$ defined above, we represent the state of the system at time n by the ordered pair of random variables

$$Z_n = (Z_{1,n}, Z_{2,n}),$$

where

$$Z_{1,n} = X_{1,n} \text{ and } Z_{2,n} = \sum_{i=2}^R X_{i,n}.$$

In other words,

$Z_{1,n}$ = the number of cells in the shared buffer, at time n , tagged for output port 1,

and

$Z_{2,n}$ = the number of cells in the shared buffer, as time n , tagged for the remaining $(R - 1)$ output ports.

Accordingly, the system model we consider is the discrete-time, finite-state process

$$Z = \{Z_n | n \in T\},$$

where, due to the aggregation of states induced by $Z_{2,n}$, Z is no longer a Markov process. Further, as with the exact model and for the same reasons, the state variables $Z_{1,n}$ and $Z_{2,n}$ are dependent. Nevertheless, by employing an approximate formulation of the cell departure probabilities (see Section 3.2.2), we find that the probabilistic nature of Z can be approximated by a process which is Markovian (although not time-homogeneous). Moreover, as borne out by comparisons with simulation results in regions where the latter are feasible (see Section 4, Figs. 5 and 6), this approximation appears to capture much of the influence of interqueue correlation, at least to the extent that it affects the measures in question.

To obtain an explicit approximation of Z as a (non-homogeneous) Markov process, one approach would be to formulate the latter's time-dependent, 1-step transition matrix P_n in terms of n (current time, as expressed by the number of elapsed time slots) and the underlying model parameters. However, since the random variables Z_n take values in the set

$$\{(i, j) | 0 \leq i + j \leq K\},$$

there are $(K + 2)(K + 1)/2$ distinct states, thus involving a matrix with approximately $K^2/4$ entries. Although this could be reduced by lumping states which are (probabilistically) equivalent vis-à-vis the measures in question, direct formulation of P_n can be avoided via an iterative formula that expresses the state-occupancy probability distri-

bution at time $n + 1$ in terms of that at time n . In doing so, we thus defer, to the solution phase, much of the work that's typically associated with model construction.

3.2. Model solution

Let (i, j) be a state of the model Z just described, i.e., $0 \leq i + j \leq K$, and for each $n \in T$ let

$$p_n(i, j) = P[Z_{1,n} = i, Z_{2,n} = j]$$

be the probability of occupying state (i, j) at time n . For the distribution of the initial state Z_0 , we assume that the buffer is empty with probability 1, i.e.,

$$p_0(i, j) = \begin{cases} 1 & \text{if } (i, j) = (0, 0), \\ 0 & \text{otherwise.} \end{cases}$$

Once these state-occupancy probabilities are determined for some time n , their values at time $n + 1$ (the end of the next time slot) are obtained via the iterative formula developed below. The limiting distribution we seek is then approximated by letting n become sufficiently large.

To formulate $p_{n+1}(i, j)$ in terms of probabilities of the form $p_n(r, s)$, we first recall the basic operations of the switch (see Section 2) and examine them more closely in terms of their effects on state transitions. In doing so, it is helpful to introduce the following notation for the larger and smaller of two integers x and y :

$$(x, y)_+ = \max\{x, y\}$$

$$(x, y)_- = \min\{x, y\}$$

Then, given that the state of the system is (r, s) at time n , the consequence of the send, check, and receive operations during time slot $n + 1$ are the following.

(1) **Send:** The cells at the head of each logical queue are transmitted; the system is then in the temporary state $(r - u_r, s - v)$, where $u_r = (1, r)_-$ and v is the number of nonempty logical queues (prior to this operation) associated with output ports 2 through R ($0 \leq v \leq R - 1$).

(2) **Check:** The number of cells now in the buffer, namely $r - u_r + s - v$, is compared with the threshold level S to determine whether low priority cells should be accepted.

(3) **Receive:** The N input lines are scanned and incoming cells are stored or lost according to

the state of the buffer and a cell's priority; the maximum number of accepted cells is $(K - r + u_r - s + v, N)_-$.

Suppose now that (i, j) is the resulting state at time $n + 1$. Then, at time n , the states (r, s) that need be considered are those having a nonzero probability of making a transition to (i, j) . Letting $S_{i,j}$ denote the set of all such states, from the above observations it follows that $(r, s) \in S_{i,j}$ if and only if

$$(0, i - N)_+ \leq r \leq (K, i + 1)_-$$

and

$$(0, i + j + u_r - N - r)_+$$

$$\leq s \leq (K - r, j + R - 1)_-$$

Further, if we let

$d_n(v|s)$ = the probability, during the next time slot, of v departures from output ports 2 through R , given that $Z_{2,n} = s$,

and

$b(l, m|k)$ = the probability, during a time slot, of accepting l cells tagged for output port 1 and m cells tagged for the remaining output ports, given that there are k cells in the buffer at time of the check operation,

then, for all states (i, j) , $p_{n+1}(i, j)$ can be formulated as

$$p_{n+1}(i, j) = \sum_{(r, s) \in S_{i,j}} \sum_{v=(0, s-j)_+}^{v_m} p_n(r, s) \cdot d_n(v|s) \cdot b(i - (r - u_r), j - (s - v) | r + s - (u_r + v)),$$

where u_r is as defined earlier (indicating whether a cell departs from output port 1) and v_m is the maximum number of cells that can depart from output ports 2 \cdots R , given that $Z_n = (r, s)$ and $Z_{n+1} = (i, j)$. v_m is thus the smallest integer among $s, R - 1$, and $c_m + s - j$, where c_m is the maximum number of subsequently accepted cells destined for ports 2 through R . Since at most N cells can be accepted during a time slot and since $i - (r - u_r)$ of these are tagged for output port 1, it follows that $c_m = N - i + r - u_r$. Hence, $v_m = \min\{s, R - 1, N + r + s - i - j - u_r\}$.

Given the above, it remains only to formulate the conditional “send” probabilities $d_n(v|s)$ and “receive” probabilities $b(l, m|k)$. The latter follow directly from assumptions regarding the input traffic and threshold mechanism, and hence we choose to consider them first.

3.2.1. Receive probabilities

Due to the facts that arrivals, per time slot, are binomially distributed and output port destinations are equally likely, it follows that

$$b(l, m|k) = \binom{l+m}{l} (1/R)^l (1-1/R)^m c(l+m|k).$$

where

$c(x|k)$ = the probability that x cells are stored in the buffer during the receive operation, given there are k cells in the buffer at the time of the check operation.

To formulate the probabilities $c(x|k)$, we recall that during each time slot and at each input port, a cell arrives with probability p and, given it arrives, it has probability p_h of being high priority. If the number of cells k in the buffer, when checked, is less than the threshold level S , then all arrivals are stored until the buffer becomes full. On the other hand, if $k \geq S$, then only high priority cells are stored, again until there is no free space in the buffer. Accordingly, $c(x|k)$ can be expressed as follows in terms of model parameters N, K, S, p , and p_h .

$$c(x|k) = \begin{cases} \binom{N}{x} p^x (1-p)^{N-x} & \text{if } k < S \text{ and } k < K-x, \\ \sum_{y=K-k}^N \binom{N}{y} p^y (1-p)^{N-y} & \text{if } k < S \text{ and } k = K-x, \\ \binom{N}{x} (p \cdot p_h)^x (1-p \cdot p_h)^{N-x} & \text{if } k \geq S \text{ and } k < K-x, \\ \sum_{y=K-k}^N \binom{N}{y} (p \cdot p_h)^y (1-p \cdot p_h)^{N-y} & \text{if } k \geq S \text{ and } k = K-x, \\ 0 & \text{otherwise.} \end{cases}$$

3.2.2. Send probabilities

By its definition, $d_n(v|s)$ is the probability that v cells depart (are sent) from output ports 2 through R during the send operation of time slot $n+1$, given that a total of s cells in the buffer are tagged for these ports at the end of time slot n , i.e., $Z_{2,n} = s$.

Formulation of these probabilities is therefore more complex and must be approximated since, for an exact solution, we would have to know the joint distribution of the queue-size variables associated with output ports 2 through R at each time n . However, due to the symmetry of the model, we do know the marginal distributions of each such variable, since each of these queues behaves exactly as the one associated with output port 1. We then approximate the joint distribution by assuming that these $R-1$ queues behave independently.

To express this more precisely, let $q_n(i)$ denote the probability that, at time n , i cells in the shared buffer are tagged for a given output port h . Since this probability is the same for all ports, its value is given by our analysis of port 1, i.e.,

$$q_n(i) = P[Z_{1,n} = i] = \sum_{j=0}^{K-i} p_n(i, j).$$

Suppose further that

$$\bar{i} = (i_2, i_3, \dots, i_R)$$

denotes the joint status of the queues associated with output ports 2, 3, ..., R , where i_h is the number of cells tagged for output port h ($2 \leq h \leq R$). Finally, if we let ν and σ be the function

$$\nu(\bar{i}) = |\{h | i_h > 0\}| = \text{the total number of nonempty queues in state } \bar{i},$$

and

$$\sigma(\bar{i}) = \sum_{h=2}^R i_h$$

= total number of cells stored in state \bar{i}

then, via the independence approximation just mentioned, it follows that

$$d_n(v|s) = \frac{\sum_{\bar{i} \ni \nu(\bar{i})=v, \sigma(\bar{i})=s} q_n(i_2) q_n(i_3) \cdots q_n(i_R)}{\sum_{\bar{i} \ni \sigma(\bar{i})=s} q_n(i_2) q_n(i_3) \cdots q_n(i_R)}$$

3.3. Steady-state distributions and measures

For specific values of the model parameters, the formulation of $p_n(i, j) = P[Z_n = (i, j)]$ developed in the previous subsection permits iterative numerical calculation of the distribution of Z_n , beginning with its specified initial distribution p_0 . Iteration is continued until steady-state conditions are closely approximated. More specifically, we increase n until a value is reached where the maximum absolute % difference between two successive state occupancy probabilities, i.e., the quantity

$$\max \left\{ \frac{|p_{n+1}(i, j) - p_n(i, j)|}{p_{n+1}(i, j)} \mid 0 \leq i + j \leq K \right\}$$

is less than some very small number (typically 10^{-7}). This limiting distribution can be used, in turn, to calculate steady-state loss and delay measures for both high and low priority cells.

The development that follows considers loss probabilities only, since these are the principal concern in the context of a space priority mechanism. (Although omitted here, the mean and variance of transfer delay have also been formulated; results of applying these formulae are illustrated in Figs. 7 and 8 of the following section.) Specifically, our goal is the determination of the quantities

B_h = the steady-state probability that an arriving high priority cell is lost,

and

B_l = the steady-state probability that an arriving low priority cell is lost.

Due to the discrete-time, binomially-distributed nature of the incoming traffic, the formulation of these probabilities is far from immediate, unlike continuous-time models (see [7], for example) where arrivals and departures occur singly.

To begin, let $p(r, s)$ and $d(v|s)$ denote the limiting values, as $n \rightarrow \infty$, of the distributions $p_n(r, s)$ and $d_n(v|s)$ determined earlier (see Section 3.2) and let

$e(k)$ = the steady-state probability of there being k cells in the buffer at the time of the check operation,

where $0 \leq k \leq K - 1$. (Note that the condition $k = K$ need not be considered since at least one cell must depart during the send operation whenever the buffer is nonempty.) Further, for a given value of k , let S_k denote those combinations of current state (r, s) and cells v departing output ports 2-R which leave k cells in the buffer, i.e., $(r, s, v) \in S_k$ if and only if

$$0 \leq r + s \leq K, (1, s)_- \leq v \leq (R - 1, s)_-,$$

and

$$r + s - u_r - v = k,$$

where, as earlier, $u_r = (1, r)_-$. Then, by the definitions of $p(r, s)$ and $d(v|s)$, it follows that

$$e(k) = \sum_{(r, s, v) \in S_k} p(r, s) \cdot d(v|s).$$

In terms of this distribution, let us now consider the probabilities of cell losses that occur during the subsequent receive operation. (Hence, in what follows, "lost" has the implicit meaning of "lost during a time slot"). We formulate these for each priority class and for each of two buffer conditions, beginning with the condition where, at check time, $k < S$. In this case, high and low priority cells are not distinguished and losses occur if and only if an arriving cell encounters a full buffer (due to earlier arrivals during the same time slot). More precisely, let $l_b(y)$ (b for "below threshold") denote the joint probability

$l_b(y)$ = the probability that $k < S$ and exactly y arriving cells are lost,

and let $a(w)$ denote the probability of exactly w cell arrivals during the receive operation, i.e.,

$$a(w) = \begin{cases} \binom{N}{w} p^w (1-p)^{N-w} & \text{if } 0 \leq w \leq N, \\ 0 & \text{otherwise.} \end{cases}$$

Then, for $y \geq 0$, it is easily verified that

$$l_b(y) = \sum_{k=0}^{S-1} e(k) \cdot a(y + K - k).$$

Note, however, that the values $l_b(y)$ of interest lie in the range where $1 \leq y \leq N - 1$ since (i) we seek probabilities of losing a nonzero number of cells and (ii) in the below-threshold case, at most $N - 1$ cells can be lost.

In turn, the probability of losing exactly x cells of a given type when $k < S$ can be formulated in

terms of the above probabilities. Specifically, for high priority cells, the joint probability

$l_{b,h}(x)$ = the probability that $k < S$ and exactly x high priority cells are lost.

is given by the formula

$$l_{b,h}(x) = \sum_{y=x}^{N-1} l_b(y) \binom{y}{x} p_h^x p_l^{y-x}.$$

Similarly, for low priority cells

$$l_{b,l}(x) = \sum_{y=x}^{N-1} l_b(y) \binom{y}{x} p_l^x p_h^{y-x}.$$

In the second case, where the number of cells k in the buffer at check time is at or above the threshold level S , the situation is easier to analyze. Here, high and low priority cells no longer compete for space in the buffer since all low priority cells are rejected. Accordingly, if $l_{a,h}(x)$ and $l_{a,l}(x)$ denote the analogous "at or above threshold" distributions for high and low priority cell losses, and we let $a_h(w)$ and $a_l(w)$ be the probabilities of w high and low priority cell arrivals, respectively (formulated in a manner similar to that of $a(w)$ above) then, for $x \geq 0$,

$$l_{a,h}(x) = \sum_{k=S}^{K-1} e(k) \cdot a_h(x + K - k)$$

and

$$l_{a,l}(x) = \sum_{k=S}^{K-1} e(k) \cdot a_l(x).$$

Combining the above results, the average number L_h of high priority cell losses during a time slot is

$$L_h = \sum_{x=1}^{N-1} x \cdot (l_{b,h}(x) + l_{a,h}(x)).$$

Similarly, for low priority cells, the average number L_l of losses per time slot is

$$L_l = \sum_{x=1}^N x \cdot (l_{b,l}(x) + l_{a,l}(x)).$$

Note that an upper limit of N is needed for low priority losses since $l_{a,l}(N) > 0$, i.e., there's the possibility of N low priority arrivals and, when threshold rejection is in effect, all of these are lost. It is also worth noting that, when losses are so averaged, one can obtain simpler expressions of L_h and L_l with respect to the below-threshold probabilities, thus eliminating the need for separate calculations of $l_{b,h}$ and $l_{b,l}$. Specifically, in the case of high priority losses, it can be shown that

$$L_h = \sum_{x=1}^{N-1} x \cdot (l_b(x) \cdot p_h + l_{a,h}(x)).$$

An analogous formula holds for L_l .

The loss probabilities we seek (see the outset of this development) are then obtained by dividing each of these averages by the corresponding average number of cell arrivals per time slot, i.e.,

$$B_h = L_h / p \cdot p_h N \text{ and } B_l = L_l / p \cdot p_l N.$$

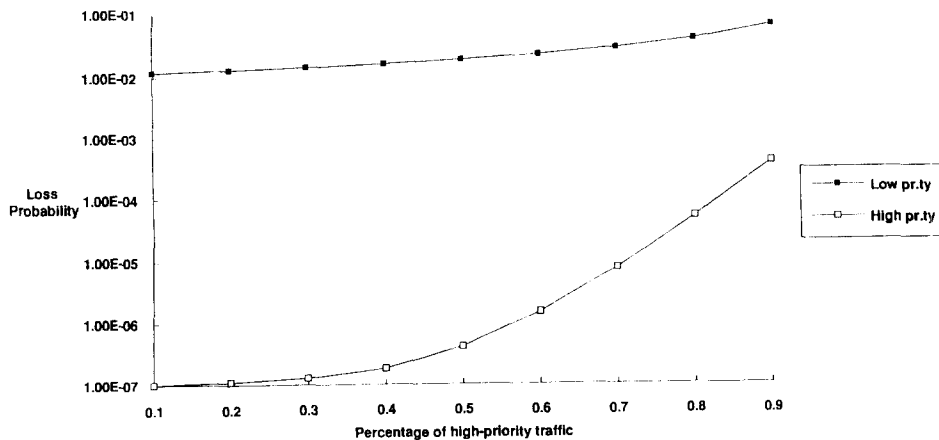


Fig. 1. Loss probability versus load ratio ($N = 16, R = 8, K = 50, S = 37, \rho = 0.9$).

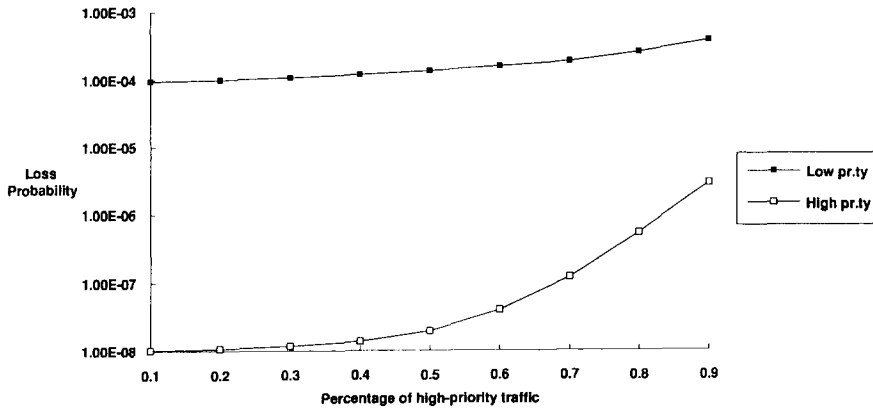


Fig. 2. Loss probability versus load ratio ($N = 16, R = 8, K = 50, S = 68, \rho = 0.9$).

Applying this analysis to various instances of traffic and switch parameter values, we obtain the results described in the section that follows.

4. Evaluation results

As noted in the introduction, the principal objective of this study is to accurately assess the performance of an ATM switch with a shared buffer and a threshold priority mechanism (partial buffer sharing). In particular, we wish to assess the system's robustness by varying the values of selected model parameters. We also want to evaluate its relative effectiveness via comparisons with a "no-priority" switch. As a consequence of properties revealed by such data, we are able to show that, in spite of the number of parameters involved, the buffer capacity and threshold level can be optimally dimensioned in a

relatively straightforward manner. Throughout this process, we choose to limit our attention to a 16×8 switch. Obviously, switches of either smaller or larger size (provided the latter are within bounds of computational feasibility) could be assessed in a similar manner.

To begin, we examine effects of varying the traffic mix, where Fig. 1 displays high and low priority cell loss probabilities as a function of the load ratio p_h (fraction of high priority traffic) for buffer capacity $K = 50$ and threshold level $S = 37$ (thus the threshold ratio S/K is 0.74). Fig. 2 displays the same information for $K = 80$ and $S = 68$ ($S/K = 0.85$). In both cases, the offered load is $\rho = 0.9$. As these curves illustrate, the switch is quite robust with respect to variations in the Class 1/Class 2 mix when traffic is heavy and the load ratio is less than 0.4. This observation is quite important, since we are considering a fixed (non-adaptive) threshold ratio that will hopefully

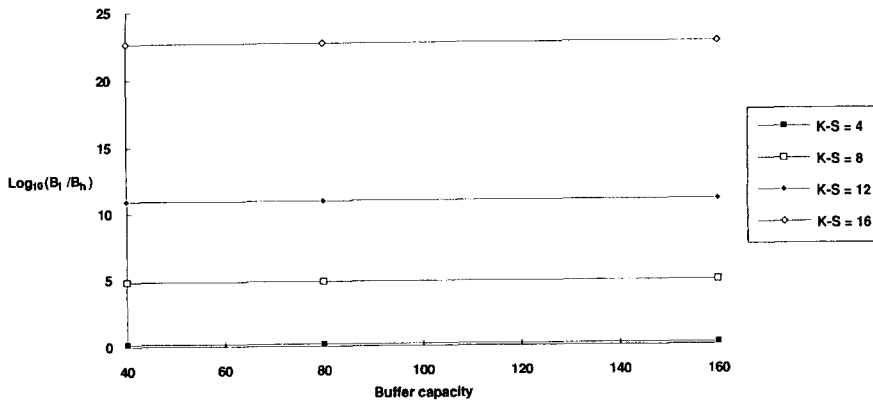


Fig. 3. Difference between the loss probability for high and low priority cells $\rho = 0.2, p_h = 0.15$).

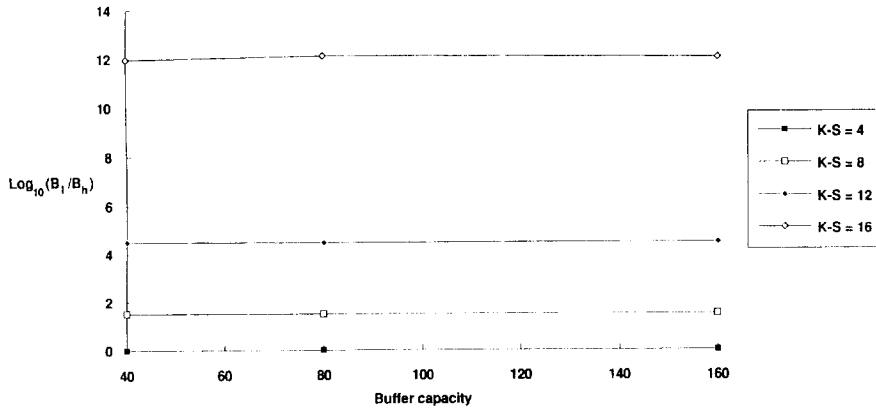


Fig. 4. Difference between the loss probability for high and low priority cells ($\rho = 0.8, p_h = 0.15$).

accommodate different traffic mixes resulting from the introduction of new services. (The same observation, based on a Poisson arrival model, is noted in [7]). Moreover, since typical ATM traffic mixes are anticipated to lie in a range where

15–20% of the traffic is vital, it is reasonably safe to assume that p_h will lie in this favorable region. As a consequence, the remainder of our parametric studies assume that the load ratio is fixed at the value $p_h = 0.15$.

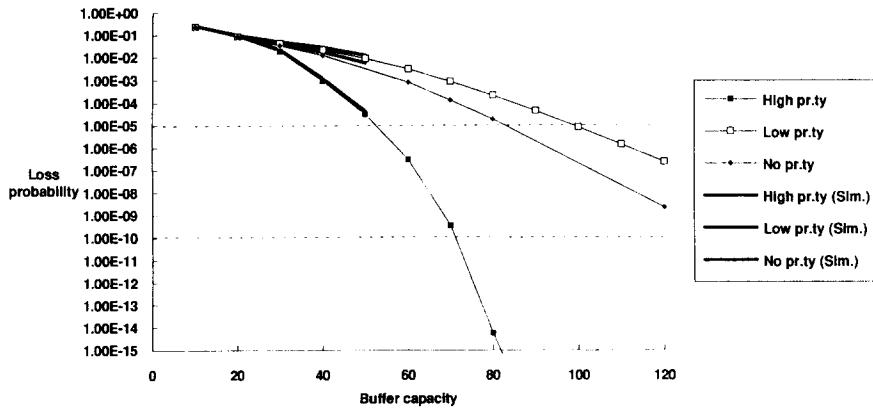


Fig. 5. Loss probability versus buffer capacity ($N = 16, R = 8, p_h = 0.15, S/K = 0.8, \rho = 0.9$).

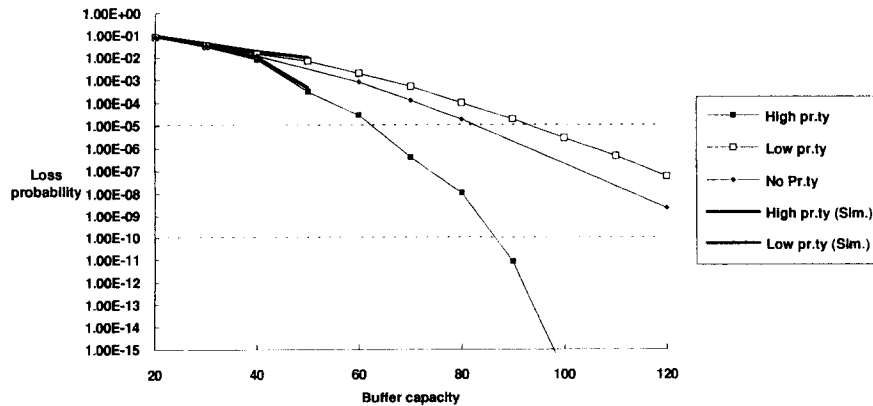


Fig. 6. Loss probability versus buffer capacity ($N = 16, R = 8, p_h = 0.15, S/K = 0.85, \rho = 0.9$).

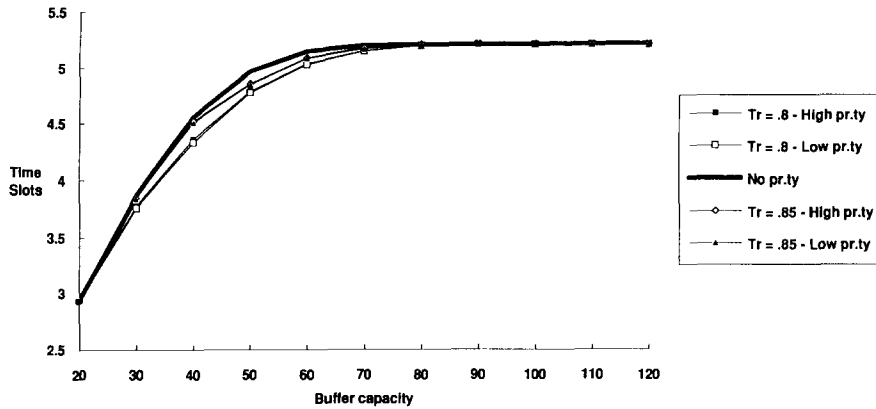


Fig. 7. Mean time in the shared buffer ($N = 16, R = 8, p_h = 0.15, \rho = 0.9$).

Another important relationship, which has likewise been observed for continuous-time models (again see [7]) is illustrated in Figs. 3 and 4. Specifically, they show that, for a given offered load ρ , the difference Δ in the order of magnitude of the loss probabilities for low and high priority cells ($\Delta = \log_{10}(B_l) - \log_{10}(B_h)$) remains constant as buffer capacity varies, provided the threshold level S is such that $K - S$ remains constant. Fig. 3 displays this interesting property for a light load ($\rho = 0.2$); Fig. 4 demonstrates that the same is true for a heavier load ($\rho = 0.8$). Supposing further that the target loss probabilities are 10^{-10} for Class 1 cells and 10^{-5} for Class 2 cells, it is therefore reasonable to assume that an optimal choice of the threshold level S , relative to a given K and ρ , will be given by the difference $K - S$ that corresponds to $\Delta = 5$. Accordingly, by examining similar data for other values of ρ , we obtain Table 1 which, for a given

load ρ , indicates the value of $K - S$ for which $\Delta = 5$. This table thus provides an empirical means of optimizing the choice of S in concert with K , when dimensioning the buffer for a specified admissible load (as further discussed below).

Let us now examine how Class 1 and Class 2 cell loss probabilities vary as a function of the buffer capacity K . Results in this regard are given by Figs. 5 and 6, where Fig. 5 compares the loss performance of a switch with threshold ratio $S/K = 0.8$ to that of a no-priority switch ($S/K = 1$). Fig. 6 displays similar information, where in this case, the threshold ratio of the priority switch is 0.85. Both figures presume that $p_h = 0.15$ and $\rho = 0.9$. Comparing, within each figure, the loss probabilities incurred with vs. without priority discrimination, we find that the buffer capacity required to guarantee acceptable loss probabilities (10^{-10} for Class 1; 10^{-5} for Class 2) is considerably less when selective discarding is ex-

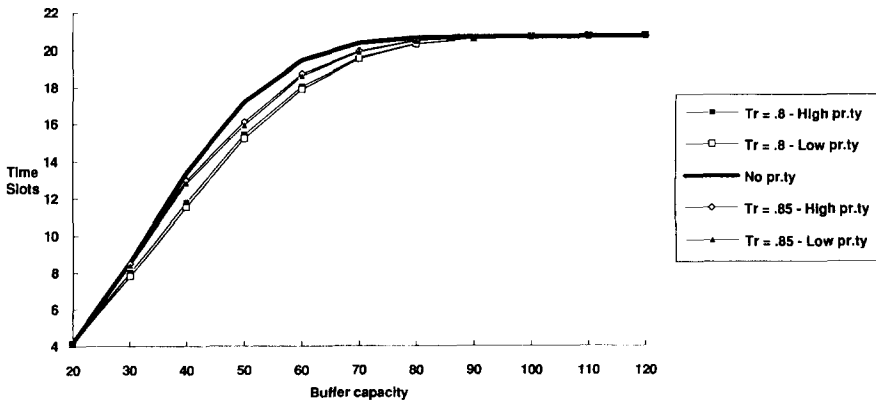


Fig. 8. Variance of the time in the shared buffer ($N = 16, R = 8, p_h = 0.15, \rho = 0.9$).

Table 1
 $K - S$ as a function of ρ for $\Delta = 5$.

ρ	0.15-0.25	0.25-0.35	0.35-0.45	0.45-0.60	0.60-0.80	0.80-1.00
$K - S$	8	9	10	11	12	13

exercised. Specifically, for the comparison made in Fig. 6, we see that a buffer capacity of 90 suffices if the priority mechanism is used. Without such a mechanism, the required capacity lies well beyond the maximum value plotted (Fig. 9, discussed below, pinpoints this more exactly). This presumes, of course, that the stricter 10^{-10} loss probability must be held to when cell classes are not distinguished. Figs. 5 and 6 also indicate comparisons with simulation data for capacity values that permit reasonably accurate simulation results. Here we note, at least for the capacities considered, very close agreement with results obtained from the (approximate) analytic model.

Although formulation of delay measures (mean time in the shared buffer, variance of time in the shared buffer) was not included in the paper, Figs. 7 and 8 illustrate that delay performance is not severely altered by the threshold mechanism. Specifically, Fig. 7 displays mean delay as a function of buffer capacity for threshold ratios that yield good loss performance ($S/K = 0.8, 0.85$; see Figs. 5 and 6) and for a switch with no priority ($S/K = 1$). When compared to the latter, differences in mean delays are relatively small, the most appreciable reductions occurring with moderate values of capacity. Moreover, in a region where the priority mechanism is effective (with

respect to loss performance), mean delay appears to be quite insensitive to changes in the threshold ratio. Similar comments apply to the variance of delay, as evidenced by the curves of Fig. 8.

Finally, let us return to the important question of buffer dimensioning, i.e., selection of values of K and S which optimize the switch's effectiveness with regard to specified cell loss requirements. More precisely, we seek values of K and S such that that the loss probabilities experienced by Class 1 and Class 2 cells coincide exactly with the specified maximums of 10^{-10} and 10^{-5} , respectively. Although such a choice would normally involve a considerable amount of trial and error, by employing Table 1, the dimensioning problem reduces to a consideration of K only. In other words, for a given load ρ , the appropriate difference $d_\rho = K - S$ is selected from the table; K is then varied, with its corresponding threshold level $K - d_\rho$ following in concert, to determine the smallest K which just admits ρ . Fig. 9 displays results of this dimensioning procedure (the "optimal" curve) where maximum admissible load is plotted as a function of buffer capacity. A second curve, in close agreement with the first, displays the results of a suboptimal determination, where the capacity-threshold difference is fixed at $K - S = 13$ throughout. These are then compared

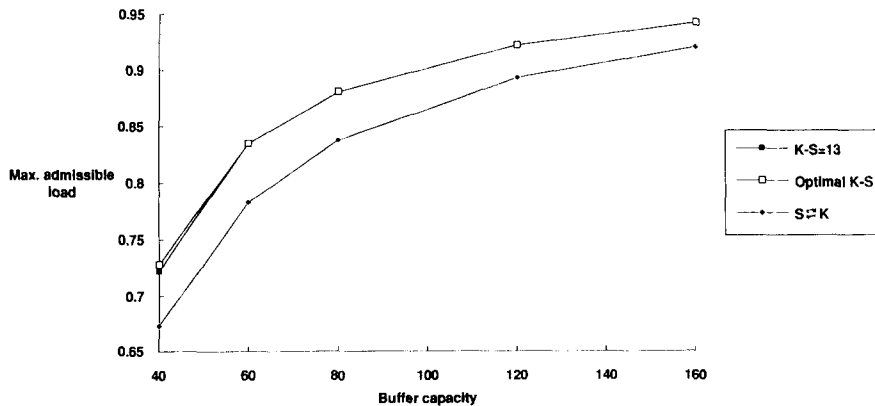


Fig. 9. Shared-buffer dimensioning with threshold priority mechanism ($N = 16, R = 8, \rho_{11} = 0.15$)

with the load admitted by a no-priority switch, where the more stringent 10^{-10} requirement applies to both cell classes.

5. Conclusions

As a consequence of the results presented in Fig. 9, the advantages afforded by the combined use of a threshold priority mechanism and a shared buffer become quantifiably obvious. From a performance perspective, there's an appreciable increase in admissible load for a specified buffer capacity, especially for smaller capacities. From a resource utilization point of view, i.e., the buffer capacity required for a given admissible load, the improvement is even more striking, particularly in the case of heavy loading. For example, if $\rho = 0.9$, the resulting reduction in buffer size is approximately 25%. Also, as testified to by other evaluation data, this combination is robust with respect to load ratio variations in regions of practical interest. Further, for threshold ratios that yield good loss performance, the priority mechanism is relatively benign with respect to its effect on transfer delays. Finally, due to a fortunate invariance of the loss difference Δ with changes in buffer capacity, the switch can be dimensioned in a straightforward, practical manner.

Beyond these conclusions regarding application results, we have also learned something about modeling the types of object systems and environments that are encountered in an ATM context. To accurately represent the discrete-time nature of ATM cell arrivals, unless the number of input ports is sufficiently large to justify a Poisson approximation, there is growing evidence of the need to employ discrete-time models. The latter, however, call for effective means of accommodating subtle consequences of slotted traffic (e.g. negatively correlated output queues) and simplifying large state spaces, either by lumping (exact) or aggregation (approximate) techniques. Choosing the second of these alternatives for the model developed herein, we have shown that implications of discrete-time, prioritized traffic can be indeed by accounted for by a state space of reasonable size.

Although this approximate model appears to be satisfactory in regions where high-confidence simulation data is obtainable (loss probabilities $\leq 10^{-4}$), there remains the question of whether its accuracy is sustained when loss probabilities are very small (e.g., 10^{-10}). We are currently addressing this question via construction of an exact model (see the outset of Section 3.1) which can be used as a reference. The problem here, as noted earlier, is that the number of states becomes excessively large for realistic values R and K . To overcome this difficulty, we seek exact reductions of the state space, via lumping, that (i) support the loss and delay measures of interest and (ii) admit feasible solutions of the type afforded by contemporary model-based evaluation tools.

References

- [1] F. Bonomi, L. Fratta, S. Montagna and R. Paglino, Priority on cell service and cell loss in ATM-based networks, *Proc. 7th ITC Seminar*, Morristown, NJ, October 1990.
- [2] M. de Prycker, *Asynchronous Transfer Mode: Solution for Broadband ISDN* (Ellis Horwood, Chichester, UK, 1991).
- [3] A.E. Eckberg and T.-C. Hou, Effects of output buffer sharing on buffer requirements in an ATDM packet switch, *Proc. IEEE INFOCOM'88*, New Orleans, LA, March 1988.
- [4] A. Gravey and G. Hébuterne, Analysis of a priority queue with delay and/or loss sensitive customers, *Proc. 7th ITC Seminar*, Morristown, NJ, October 1990.
- [5] G. Hébuterne and A. Gravey, A space priority queueing mechanism for multiplexing ATM channels, *Comput. Networks and ISDN Systems* **20** (1990) 37–43.
- [6] H. Kröner, Comparative performance study of space priority mechanisms for ATM networks, *Proc. IEEE INFOCOM'90*, San Francisco, CA, June 1990, pp. 1136–1143.
- [7] H. Kröner, G. Hébuterne, P. Boyer and A. Gravey, Priority management in ATM switching nodes, *IEEE J. Selected Areas Comm.* **9** (3) (1991) 418–427.
- [8] J.-Y. Le Boudec, An efficient solution method for Markov models of ATM links with loss priorities, *IEEE J. Selected Areas Comm.* **9** (3) (1991) 408–417.
- [9] D.W. Petr and V.S. Frost, Optimal threshold-based discarding for queue overload control, *Proc. 7th ITC Seminar*, Morristown, NJ, October 1990.
- [10] K. Rothermel, Priority mechanisms in ATM networks, *Proc. IEEE GLOBECOM'90*, San Diego, CA, December 1990.