Storage requirements in ATM switching elements with correlated arrivals and independent uniform routing *

Herwig Bruneel *, Bart Steyaert

Laboratory for Communications Engineering, SMACS Research Group, University of Ghent, Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium

Received 5 October 1992; revised 30 October 1994

Abstract

An ATM switching element with a finite number of independent and identical traffic sources (users), and an equally high number of destinations is considered. The sources are assumed to be bursty, and the burstiness of the sources is modeled by describing both their active periods (during which users generate one cell per slot) and their passive periods (during which users do not generate data) as geometric random variables. Arriving cells are assumed to be routed to their destinations in a uniform and independent way. As a result, a correlated-arrivals queueing model for the switch is obtained, which is analyzed in this paper by purely analytical means. Specifically, closed-form expressions for the mean and the variance of the buffer contents, and an accurate analytic approximation for the tail probabilities of the buffer contents are derived.

Keywords: ATM; Switching; Independent routing; Performance analysis; Queueing; Discrete time; Generating functions; Boundary probabilities

1. Introduction

In the past couple of years, discrete-time queueing models have become very popular in the performance analysis of ATM switching elements with output queueing. In many studies, the cells are assumed to be generated according to an uncorrelated (i.e., Bernoulli [1–5]) process by each of the sources. It has been observed however, that in practice, the arrival stream of the cells tends to be bursty, i.e., of a correlated nature. Therefore, several attempts have been made to incorporate bursty (nonindependent) cell arrival processes in the queueing analysis of buffers, both for single-server queues [6–10] and for multiple-server queues [11,12]. The discrete-time queueing models considered in these papers are analyzed by mainly analytical...
means (i.e., a generating-functions approach), and they have in common that they are applicable to ATM multiplexers, meaning that no routing mechanism of arriving cells to one of various output buffers (each corresponding with one possible destination) is taken into account. In this paper, we present the analysis of an ATM switching element with first-order Markovian arrivals on the inlets, and independent uniform routing. As will be shown in the following sections, the inclusion of this routing mechanism into the model increases the complexity of the analysis, since the calculation of a number of unknown probabilities using Rouché's theorem is now required. A related study, with a different type of correlated arrival process, is reported in [13].

As in most studies concerned with ATM-related discrete-time queueing models, the time axis is divided into slots of fixed length, where one slot suffices for the transmission of exactly one ATM cell. Cells enter the switch via one of its input links (sources), and are then routed to one of the output links, where they are temporarily buffered in a designated output queue to await transmission of the cells that arrived before them in this particular output queue and that are not yet sent. In this paper, we are concerned with single-server output queues, i.e., different output links correspond with different destinations. Due to the synchronous transmission mode of cells on the output links, a cell arriving in an empty queue will start its transmission only at the beginning of the slot following its slot of arrival, and will thus leave the buffer at the end of this slot.

In the remainder of this paper, \( N \) denotes the number of input and output links of a switch. Cell arrivals on different input links are assumed to be statistically independent, as they originate from different sources. It is assumed that the routing of cells from the input to the output links is performed in a uniform and independent way, i.e., each arriving cell is routed to a tagged output queue with probability \( 1/N \) (since there are \( N \) possible destinations), independently of the routing of other cells. This assumption is quite reasonable in a variety of self-routing switching networks, as was reported in [14]. In [15], a solution method (albeit approximate) was presented for the case of nonindependent routing, where all cells generated during a burst have the same destination, and therefore are routed to the same output link.

Since only one cell can be carried by each of the input links during one slot, the number of cell arrivals per slot on each of the input links is either 0 or 1. Nevertheless, the arrival process on an input link may be a time-correlated process, i.e., in general, the number of cell arrivals during one slot on an input link may depend on the number of cell arrivals during the previous slots on this input link. Throughout this paper, the arrival process on one particular input link will be modeled as a first-order Markov chain.

Let us now define an active period on an input link as a number of consecutive slots during which a cell arrival occurs. Similarly, a passive period is a number of consecutive slots during which there is no cell arrival on this input link. A slot belonging to an active (passive) period on one particular input link will be referred to as an active (passive) slot, and, similarly, the input link will be referred to as an active (passive) input link. In the following, we will assume that the lengths of both active and passive periods are geometrically distributed random variables with parameters \( \alpha \) and \( \beta \) respectively, i.e.,

\[
\begin{align*}
Pr[\text{active period} = n \text{ slots}] &= (1 - \alpha)\alpha^{n-1}, \quad n \geq 1; \\
Pr[\text{passive period} = n \text{ slots}] &= (1 - \beta)\beta^{n-1}, \quad n \geq 1.
\end{align*}
\]

Note that this assumption implies the above mentioned Markovian nature of the arrival processes on the individual inlets. The case of an uncorrelated Bernoulli arrival law on each input link of the switch is obtained for \( \alpha + \beta = 1 \).
2. System equations

Let us denote by \( a_k \) the number of active inlets during slot \( k \). Also, considering one particular output queue of the switch, let us denote by \( e_k \) the number of cell arrivals during slot \( k \) in the output queue, and by \( s_k \) the system contents of the tagged output queue at the beginning of slot \( k \); this is the number of cells buffered in the output queue at the beginning of slot \( k \), including the cell that is currently being transmitted (in the case that the output queue is nonempty). With the previous definitions, the following system equation holds for \( s_k \):

\[
s_k = (s_{k-1} - 1)^+ + e_{k-1},
\]

where \( (.)^+ \equiv \max(., 0) \). We have taken into account here that a cell cannot leave the output queue at the end of its slot of arrival, since the transmission of cells is synchronized and, consequently, always starts at a slot boundary.

We now define the joint probability generating function of the couple \((a_k, s_k)\) as

\[
P_k(x, z) \overset{\Delta}{=} E[x^{a_k} z^{s_k}],
\]

where \( E[.] \) denotes the expected value of the argument. In a similar way as was demonstrated in [9], a functional equation for this quantity can be established. After some standard manipulations, we then find the following expression for \( P_{k+1}(\cdot, \cdot) \) in terms of \( P_k(\cdot, \cdot) \):

\[
z P_{k+1}(x, z) = d(x)^N \left\{ \sum_{i=0}^{N} (u(x)f(z))^i \Pr[a_k = i, s_k = 0] \right\},
\]

where

\[
d(z) := \beta + (1 - \beta)z, \quad \text{(4a)}
\]

\[
u(z) := \frac{1 - \alpha + \alpha z}{\beta + (1 - \beta)z} \quad \text{(4b)}
\]

and

\[
f(z) := \frac{N - 1}{N} + \frac{z}{N}, \quad \text{(5)}
\]

describes the number of cell arrivals generated during an active slot on one input link destined for one particular output link. Although the analysis in this paper is carried out assuming that \( f(z) \) satisfies (5), it holds for general \( f(z) \) (that could be used, for instance, to describe a nonuniform routing of arriving packets to their destinations). When the system has reached its steady state, \( P_k(x, z) \) becomes independent of \( k \), and we obtain the following functional equation for \( P(x, z) \), the \( k \to \infty \) limit of \( P_k(x, z) \):

\[
z P(x, z) = d(x)^N \left\{ \sum_{i=0}^{N} (u(x)f(z))^i p_i \right\},
\]

where

\[
p_i \overset{\Delta}{=} \lim_{k \to \infty} \Pr[a_k = i, s_k = 0], \quad \text{(6b)}
\]
is the (unknown) steady-state probability of having an empty buffer at the beginning of an arbitrary slot and \( i \) active inlets during this slot. Eq. (6a) is a functional equation for the steady-state joint probability generating function \( P(x, z) \) of the couple \((a_k, s_k)\). Unfortunately, we are unable to solve \( P(x, z) \) from this equation. Furthermore, (6a) and (6b) still contain a number of unknowns that need to be determined. In Section 3, we will describe a technique that solves this problem. Nevertheless, putting \( z = 1 \) into (6a), we can derive an expression for the probability generating function of \( a \), the random variable describing the number of active inlets during an arbitrary slot in the steady state. One can easily verify that the equation

\[
P(x, 1) = d(x)^N P(u(x), 1)
\]

is satisfied by

\[
P(x, 1) = A(x) \frac{\Delta}{(1 - \sigma + \sigma x)^N},
\]

where

\[
\sigma \triangleq \frac{1 - \beta}{2 - \alpha - \beta}
\]

is the average load on one inlet, i.e., the steady-state probability that an inlet is active during an arbitrary slot, which makes (7a) intuitively clear.

3. Calculation of the unknowns \( p_l, 0 \leq l \leq N \)

Let us now consider those values of \( x \) and \( z \) for which the \( P \)-functions on both sides of the functional equation (6a) are equal, i.e., such that

\[
x = u(x) f(z).
\]

In view of the definition of \( u(.) \) and \( f(.) \), this relationship can be solved for \( z \) in terms of \( x \), giving

\[
z = T(x) \frac{\beta + (1 - \beta)x}{1 - \alpha + \alpha x} - (N - 1),
\]

or, vice versa, it can be solved for \( x \) in terms of \( z \), yielding two solutions

\[
x = r_{+,-}(z) = \frac{\alpha f(z) - \beta \pm \sqrt{(\alpha f(z) - \beta)^2 + 4(1 - \alpha)(1 - \beta)f(z)}}{2(1 - \beta)},
\]

with the agreement that the square root function is defined as \( \sqrt{\rho e^{i\theta}} = \rho^{1/2} e^{i\theta/2} \), for all \( \rho \geq 0, 0 \leq \theta < 2\pi \). One can easily verify that \( r_+(1) = T(1) = 1 \), which is not the case for \( r_-(1) \). Depending on the application, either one of the three possibilities in (8a) and (8b) may be used. With the expression for \( r_+(z) \) for instance, (6a) can be transformed into

\[
P(r_{+}(z), z) = \frac{d(r_{+}(z))^N (z - 1) \sum_{l=0}^{N} r_{+}(z)^l p_l}{z - d(r_{+}(z))^N}.
\]
Substituting $z = 1$ into this equation, we find
\[ \sum_{i=0}^{N} p_i = 1 - \sigma, \]  
(10)
due to the normalization condition $P(1, 1) = 1$. As is customary when some remaining unknown probabilities are to be determined, we will now invoke analyticity conditions of the generating functions to calculate the $p_i$'s in Eq. (9). From the definition of $P(x, z)$, it is clear that $P(r_+(z), z)$ must be analytic for all values of $z$ for which $|z| < 1$. It has been verified that the denominator of (9) has no zeros inside the unit disk; however, the proof of this property falls beyond the scope of this paper. Apparently, (9) does not contain enough information to allow us to calculate the values of the $p_i$'s. We have only found one relation between these unknowns, i.e., the normalization condition (10). Therefore, in order to obtain additional information, let us denote by $P(i)(x, z)$ the $i$th partial derivative of $P(x, z)$ with respect to $x$, and let us consider the $j$th derivative, $1 \leq j \leq N$, with respect to $x$ of the functional equation (6a). Using the property that $u'(x) = (\alpha + \beta - 1)/d(x)^2$ (where the prime denotes the first derivative with respect to $x$), we find the following equation:
\[ zP^{(j)}(x, z) = \sum_{i=0}^{j} c_j(i) f(z)^j d(x)^{N-i-j} \times \left[ P^{(i)}(u(x)f(z), z) + (z - 1) \sum_{l=i}^{N} \frac{l!}{(l-i)!} P^{(l)}(z)^{l-i} p_i \right], \]  
(11a)
for $1 \leq j \leq N$, where
\[ c_j(i) \triangleq \frac{j!}{i!(j-i)!} \frac{(N-i)!}{(N-j)!} (1-\beta)^j (\alpha + \beta - 1)^i, \quad 0 \leq i \leq j. \]  
(11b)
We note that with the above definitions, $P^{(0)}(x, z) \equiv P(x, z)$. Again, choosing $x = r_+(z)$, this time in (11b), we can derive an expression for $P^{(j)}(r_+(z), z)$, $1 \leq j \leq N$, in terms of $P^{(i)}(r_+(z), z)$, $0 \leq i \leq j - 1$, as follows:
\[ P^{(j)}(r_+(z), z) = \sum_{i=0}^{j-1} c_j(i) f(z)^i d(r_+(z))^{N-i-j} P^{(i)}(r_+(z), z) + \sum_{i=0}^{j} Q_{i,j}(z), \]  
(12a)
where
\[ Q_{i,j}(z) \triangleq c_j(i) f(z)^i d(r_+(z))^{N-i-j} (z - 1) \sum_{l=i}^{N} \frac{l!}{(l-i)!} r_+(z)^{l-i} p_i. \]  
(12b)
Together with (9), these equations allow the recursive calculation of $P^{(j)}(r_+(z), z)$, $1 \leq j \leq N$, in terms of the unknowns $p_i$, $0 \leq l \leq N$, for any value of $z$.

It is clear that $P^{(j)}(r_+(z), z)$ must be bounded within the complex unit disk, since this function is a polynomial in $r_+(z)$, and $r_+(z)$ is finite if $|z| < 1$. On the other hand, it has been verified that for each value of $j$, $1 \leq j \leq N$, the denominator of $P^{(j)}(r_+(z), z)$ in (12a) has exactly one zero inside the complex unit disk defined by $|z| < 1$. Let us denote by $z_j$, $1 \leq j \leq N$, the zero of the denominator of the expression (12a) for $P^{(j)}(r_+(z), z)$, with $|z_j| < 1$. For each value of $j$, $z_j$ must also be a zero of the numerator of the
expression (12a) for $P^{(j)}(r_+(z), z)$. Thus, for each value of $j$, $1 \leq j \leq N$, we obtain a linear equation for the $p_l$'s, $0 \leq l \leq N$. Together with the normalization condition (10), they constitute a set of $(N + 1)$ linear equations that allow us to calculate the $(N + 1)$ unknowns $p_l$, $0 \leq l \leq N$.

4. Mean and variance of the system contents

Let us again consider those values for $x$ and $z$ for which $z = T(x)$. From (6a) we find

$$P(x, T(x)) = \frac{d(x)^N(T(x) - 1)W(x)}{T(x) - d(x)^N} = \frac{N(x)}{D(x)},$$

(13a)

where

$$W(x) \triangleq \sum_{l=0}^{N} x^l p_l$$

(13b)

is considered to be a known polynomial, as an algorithm for calculating the $p_l$'s has been given in Section 3.

Let us denote by $S(z) \triangleq P(1, z)$ the steady-state probability generating function of the system contents in an output queue of the switch. Furthermore, let $P_z(x, z)$ denote the first-order partial derivative of $P(x, z)$ with respect to $z$; similar definitions hold for higher order derivatives and/or partial derivatives with respect to $x$. Taking the first-order derivative with respect to $x$ of (13a), and using (7a), we find an expression for the mean system contents

$$E[s] = S'(1) = \left[ \frac{N''(1) - D''(1)}{2D'(1)} - N\sigma \right] / T'(1).$$

Using the definitions for $N(z)$ and $D(z)$, and expression (8a) for $T(z)$, we obtain after some calculations

$$S'(1) = \frac{N - 1}{N} \left[ \sigma + \frac{\sigma^2}{2(1 - \sigma)} - \frac{\sigma^2}{1 - \beta} \right] + \frac{W'(1)}{N(1 - \sigma)}.$$  

(14)

Similarly, taking the second-order derivative with respect to $x$ of (13a), we obtain

$$S''(1) = \left[ \frac{2(D'(1)N'''(1) - D'''(1)N'(1)) - 3D''(1)(N''(1) - D''(1))}{6D'(1)^2} 
- N(N - 1)\sigma - 2T'(1)P_{xz}(1, 1) - T''(1)S'(1) \right] / T'(1)^2,$$

(15)

where

$$N'(1) = (1 - \sigma)T'(1),$$

(16a)

$$N''(1) = 2T'(1)[N(1 - \beta)(1 - \sigma) + W'(1)] + (1 - \sigma)T''(1),$$

(16b)

$$N'''(1) = 3T'(1)[N(N - 1)(1 - \beta)^2(1 - \sigma) + 2N(1 - \beta)W'(1) + W''(1)]
+ 3T''(1)[N(1 - \beta)(1 - \sigma) + W'(1)] + (1 - \sigma)T'''(1),$$

(16c)

and

$$D'(1) = T'(1) - N(1 - \beta),$$

(16d)
\[ D''(1) = T''(1) - N(N - 1)(1 - \beta)^2, \quad (16e) \]
\[ D''''(1) = T''''(1) - N(N - 1)(N - 2)(1 - \beta)^3. \quad (16f) \]

Furthermore, from (8a), we obtain
\[ T'(1) = N(2 - \alpha - \beta), \quad (17a) \]
\[ T''(1) = -2N(1 - \alpha)(\alpha + \beta - 1), \quad (17b) \]
\[ T''''(1) = 6N\alpha(1 - \alpha)(\alpha + \beta - 1). \quad (17c) \]

The variance of the system contents can then be calculated from
\[ \text{var}[s] = S''(1) + S'(1) - S'(1)^2. \quad (18) \]

What still remains is the calculation of the value of \( P_{xz}(1, 1) \) in expression (15). Taking the second-order partial derivative with respect to \( z \) of (6a), we obtain an expression for \( P_{xz}(1, 1) \), given by
\[ P_{xz}(1, 1) = NS'(1) - \frac{N - 1}{2}\sigma^2 - W'(1). \quad (19) \]

Eqs. (15)–(19) allow the calculation of the variance of the system contents. Although these equations seem rather involved, the calculations are quite straightforward, once the \( N \)th degree polynomial \( W(x) \) has been determined.

5. The tail distribution of the system contents

In this section, we will derive an approximate expression for the tail distribution of the buffer contents. In many cases (see, e.g. [3,10,13,16]), it has been shown that approximating the tail distribution of the buffer contents by a geometric form is quite accurate. Comparison with simulation results has revealed that this is also the case for the system under study, as will be shown in Section 6 for some values of the parameters.

Approximating the distribution of the system contents by a geometric form corresponds to approximating the probability function \( S(z) \) by
\[ S(z) \approx \frac{C}{z - z_0} = -z_0^{-1}C \sum_{n=0}^{\infty} \left[ \frac{z}{z_0} \right]^n, \quad (20) \]
where we are particularly interested in sufficiently large values of \( n \). The parameter \( z_0 \) in the above expression is the pole of \( S(z) \) with the smallest modulus. The terms in the above expression for \( S(z) \) that have been neglected correspond to the poles of \( S(z) \) "larger" than \( z_0 \); it is obvious that these terms converge much faster to zero than the one in Eq. (20). By a similar argumentation as in [10,13], it can be shown (or at least expected) that \( z_0 \), the pole of \( S(z) \) with the smallest modulus also is the pole of \( P(r_+(z), z) \) with the smallest modulus. Furthermore, in a similar way as in [3], it can be shown that this pole is a real and positive quantity (note that \( z_0 > 1 \) since \( S(z) \) is analytic inside the unit disk). Summarizing, we may conclude that \( z_0 \) can be found by evaluating the real zeros of the denominator of (9) outside the unit circle; this can be easily achieved using a Newton–Raphson scheme.
What remains now is to find an algorithm for the calculation of the C-factor in (20). It can be shown, using the residue theorem, that

\[ C = \lim_{z \to z_0} (z - z_0)S(z). \]

Let us therefore define \( R_j(z_0) \) as

\[ R_j(z_0) \triangleq \lim_{z \to z_0} (z - z_0)P^{(j)}(r_+(z_0), z_0). \]

Using (9) and (12a), we respectively obtain

\[ R_0(z_0) = \frac{z_0(z_0 - 1) \sum_{i=0}^{N} r_+(z_0)^i p_i}{1 - N r'_+(z_0)(1 - \beta)d(r_+(z_0))^{N-1}}, \]

and

\[ R_j(z_0) = \frac{\sum_{i=0}^{j-1} c_j(i) f(z_0)^i d(r_+(z_0))^{N-i-j} R_i(z_0)}{z_0 - ((\alpha + \beta - 1)f(z_0))^j d(r_+(z_0))^{N-2j}}, \quad 1 \leq j \leq N, \]

where \( r'_+(z_0) \) is the first derivative of \( r_+(z) \) in \( z = z_0 \), and can be easily calculated from (8b). In order to derive the above equations, we have used the property that \( z_0 \) is a zero of \( z - d(r_+(z))^N \), but not of the denominators of the expressions (12a) for \( P^{(j)}(r_+(z), z) \), \( 1 \leq j \leq N \). Furthermore, defining \( W_i(z) \) as

\[ W_i(z) \triangleq \lim_{k \to \infty} \sum_{j=0}^\infty z^j \Pr[a_k = i, s_k = j], \quad 0 \leq i \leq N, \]

and

\[ Y_i(z_0) \triangleq \lim_{z \to z_0} (z - z_0)W_i(z), \]

we have

\[ P(x, z) = \sum_{i=0}^{N} x^i W_i(z), \]

\[ S(z) = \sum_{i=0}^{N} W_i(z), \]

and

\[ R_j(z_0) = \sum_{i=j}^{N} \frac{i!}{(i - j)!} r_+(z_0)^{i-j} Y_i(z_0), \quad 0 \leq j \leq N. \]

This can then be written as

\[ Y_j(z_0) = \left[ R_j(z_0) - \sum_{i=j+1}^{N} \frac{i!}{(i - j)!} r_+(z_0)^{i-j} Y_i(z_0) \right]/j!. \]
Starting from $j = N$, the above equation forms a recursive algorithm for the calculation of all $Y_j(z_0)$'s, for decreasing values of $j$. Finally, using (23), we obtain

$$C = \sum_{j=0}^{N} Y_j(z_0).$$

Once $z_0$ and $C$ have been derived, we find the following approximation for the tail distribution of the system contents (which corresponds to (20)):

$$\Pr[s = n] \approx -Cz_0^{-n-1}.$$ 

This, in turn, yields a formula for the cumulative tail probabilities:

$$\Pr[s > n] \approx -\frac{Cz_0^{-n-1}}{z_0 - 1}.$$  \hspace{1cm} (26)

Fig. 1. Mean buffer contents versus $\sigma$, $N = 8$, $K = 1, 10, 20, 30$. 
6. Results

Let us define the parameter $K$ as the ratio of the mean length of an active period (or a passive period) in the correlated-arrivals case versus the uncorrelated-arrivals case, i.e.,

$$K \triangleq (1 - \sigma)/(1 - \alpha) \quad \text{or equivalently} \quad K \triangleq \sigma/(1 - \beta).$$

It is clear that the couple $(K, \sigma)$, just as the couple $(\alpha, \beta)$, fully describes the first-order Markov process on each of the input links generating the cell arrivals: $\sigma$, as defined in (7b), denotes the fraction of active slots on an input link and thus is a measure for the relative lengths of the active and passive periods, while $K$ characterizes the absolute lengths of these periods and thus is a measure for the degree of correlation in the arrival process ($K = 1$ characterizes uncorrelated arrivals).

First of all, in Fig. 1 we have plotted the mean buffer contents versus the load $\sigma$, for $N = 8$, and $K = 1, 10, 20, 30$. We observe that still there is a considerable influence of the value of $K$ on the mean buffer

![Graph showing mean buffer contents versus load for different values of K and N.](image)

Fig. 2. Mean buffer contents versus $\sigma$, $N = 2, 4, 8, 16$, $K = 10$. 
contents, especially for relatively large values of \( \sigma \), in spite of the fact that the uniform and independent routing mechanism of the arriving cells to their destinations "diminishes" the correlation in the cell arrival process in the output queues. In Fig. 2, the mean buffer contents is plotted versus \( \sigma \), for \( K = 10, N = 2, 4, 8, 16 \). It is worth noting, for this correlated-arrivals case (as opposed to the uncorrelated-arrivals case), that in the low load region (i.e., \( \sigma < 0.8 \)), higher values of \( N \) imply lower values of the mean buffer contents (this behavior is caused by the uniform and independent routing process: the higher the value of \( N \), the number of destinations, the weaker the correlation in the cell arrival process in the output queues will be); this relation is inverted in the high load region (i.e., \( \sigma > 0.85 \): in this case, the behavior of the mean buffer contents is dominated by the high number of cell arrivals per slot, rather than by the correlation in the arrival process; the higher the value of \( N \), the number of sources, the larger the number of possible cell arrivals per slot). Similar remarks can be made in the discussion for Figs. 3 and 4, where the variance of the buffer contents is plotted versus the load, for the same parameter values.

In Figs. 5 and 6, we compare the geometric approximation (26) for \( \Pr[s > n] \), the cumulative tail distribution of the system contents (solid line) with the exact values, obtained via simulation (dashed line),
Fig. 4. Variance of the buffer contents versus $\sigma$, $N = 2, 4, 8, 16, K = 10$.

for $N = 16, K = 1, 10, 20$, and $\sigma = 0.5$ (Fig. 5) and 0.8 (Fig. 6), respectively. These results confirm that for the model under study, the asymptotic behavior of the distribution of the system contents can indeed be described by a geometric form, with values for $z_0$ and $C$ as derived in Section 5. Considering high values of $\sigma$ or low values of $K$, it is observed that for increasing values of $n$, $\Pr[s > n]$ converges quite rapidly to the asymptotic values given by (26); for low values of $\sigma$ combined with high values of $K$, the convergence is somewhat slower. Notice that for probability values lower than $10^{-4}$, the simulation results become unstable, due to the limited length of the simulation runs.

In the analysis throughout this paper, we have assumed an infinite storage capacity for a tagged output queue in a switching element. In reality, buffers are always of finite size and a fraction of the arriving cells will be lost. This fraction is usually referred to as the Cell Loss Ratio (CLR), and its value, for a given set of system parameters, is one of the main performance measures of the switch. One way of calculating the CLR is by solving the balance equations for the system with storage capacity $S$ by numerical means, from which the exact CLR value can be obtained. However, this becomes extremely memory and/or CPU-time consuming (and error prone) for high values of $N$ and $S$. Therefore, we prefer an alternative approach,
Based on results that were derived for the case of an uncorrelated arrival process (i.e., $K = 1$) in [17,18]. In [17], a relationship (exact, in the case of single-server output queues) between the system contents in an infinite capacity buffer and the CLR in a buffer with storage capacity $S$ was established, from which it is found that

$$\text{CLR} = \frac{1 - \sigma}{\sigma} \frac{\Pr[s > S]}{1 - \Pr[s > S]},$$

(27)
a formula that will be used here to approximate the CLR in the case of correlated arrivals. We are mainly interested in the very low CLR-region, in which case the denominator $1 - \Pr[s > S]$ in (27) can be set equal to 1. In Figs. 7 and 8, we compare the exact values of the CLR (dashed line), obtained through simulation with the approximate values $(1 - \sigma) \Pr[s > S]/\sigma$ (solid line), where $\Pr[s > S]$ is calculated from (26), for $N = 8$, $K = 1$, 10, 20, and $\sigma = 0.5$ (Fig. 7) and 0.8 (Fig. 8), respectively. We notice that in the case of uncorrelated arrivals (i.e., $K = 1$), the CLR nearly coincides with the latter approximation, which could be expected, since the relationship (27) then holds. Furthermore, for all values of the system parameters
Fig. 6. Exact (dashed) and approximate (solid) cumulative tail probabilities of the buffer contents, \( N = 16, \sigma = 0.8, K = 1, 10, 20 \).

considered here with \( K > 1 \), if the buffer size \( S \) is sufficiently high, the approximation

\[
\text{CLR} \equiv -C \frac{1 - \sigma}{\sigma} \frac{z_0^{-S-1}}{z_0 - 1},
\]

forms an upper bound for the CLR which is quite close to the actual value of the CLR, and this observation is especially valid for high values of \( \sigma \). Therefore, from a practical point of view, (28) forms an efficient tool for dimensioning purposes.

Also, Figs. 5–8 reveal that the buffer requirements can severely increase for increasing values of \( K \), in spite of the fact that the routing of cells from the inlets to their destination is uniform and independent. Furthermore, comparing Fig. 5 with Fig. 7 (similarly Fig. 6 with Fig. 8), we observe that, for these particular values of \( \sigma \), as soon as the cell arrival process becomes correlated, the buffer space required to guarantee that the CLR will not exceed a given threshold (when using (28) as an upper bound for the CLR in a buffer with storage capacity \( S \)) decreases for increasing values of \( N \); however, it was verified that this relation
is inverted for high values of $\sigma$, thus leading to a similar behavior as discussed with respect to the mean buffer contents.

7. Conclusions

In this paper, we have analyzed a discrete-time queueing model for an ATM switching element with correlated arrivals and independent uniform routing. The model has been analyzed according to a generating-functions approach, and this has resulted in explicit expressions for the mean and the variance of the buffer contents, and a geometric tail approximation of its probability mass function. It has been observed that, although the routing of the cells to their destinations is uniform and independent, the correlation in the cell arrival process has a far from negligible influence on the buffer requirements of the switching element. In addition, an accurate practical approximation for the CLR in a switching element with finite-capacity output queues has been established.
Fig. 8. Exact (dashed) and approximate (solid) CLR, \( N = 8, \sigma = 0.8, K = 1, 10, 20 \).

Acknowledgements

Herwig Bruneel wishes to acknowledge the support of the Belgian National Fund for Scientific Research (NFWO). The support of Alcatel Bell Mfg. Co., Antwerp, Belgium, is also gratefully acknowledged.

References


**Herwig Bruneel** was born in Zottegem, Belgium, in 1954. He received the M.S. degree in Electrical Engineering, the degree of Licentiate in Computer Science, and the Ph.D. degree in Computer Science in 1978, 1979 and 1984, respectively, all from the University of Ghent, Belgium. Since 1979, he has been working as a researcher for the Belgian National Fund for Scientific Research (NFWO) at the University of Ghent, where he currently leads the SMACS Research Group of the Laboratory for Communications Engineering. He is also a Professor in the Faculty of Applied Sciences at the same university. His main research interests include stochastic modeling of digital communication systems, discrete-time queueing theory, and the study of ARQ protocols. He has published more than 80 papers on these subjects and is coauthor of the book *H. Bruneel and B.G. Kim, Discrete-Time Models for Communication Systems Including ATM* (Kluwer Academic Publishers, Boston, 1993).

**Bart Steyaert** was born in Roeselare, Belgium, in 1964. He received the degrees of Licentiate in Physics and Licentiate in Computer Science in 1987 and 1989, respectively, both from the University of Ghent, Gent, Belgium. Since 1990, he has been working as a Ph.D. student at the Laboratory for Communications Engineering as a member of the SMACS Research Group. His main research interests include performance evaluation of discrete-time queueing models, applicable in B-ISDN networks.