Buffer dimensioning for rate adaption modules

B. Steyaert and H. Bruneel
SMACS Research Group, Lab. for Communications Engineering,
University of Ghent, Sint-Pietersnieuwstraat 41, B-9000 Gent,
Belgium. email : {bart.steyaert,herwig.bruneel}@lci.rug.ac.be

Abstract

In a telecommunication network, information (represented by fixed–length *packets*) sent by a source typically passes through a number of nodes before reaching its final destination. In a variety of networks, the communication links that interconnect multiple nodes do not necessarily transmit the same amount of information per time unit. In particular, when the speed of an incoming link(s) in a node exceeds the speed of the outgoing link(s), buffering of packets must be provided in order to avoid excessive packet loss. In this report, we examine the problem of dimensioning such a *rate–adaptation* buffer. The packet arrival stream is described by characterizing the length of consecutive *active* and *passive* periods (i.e., a series of consecutive slots during which packets, respectively no packets, are generated); the former quantities can have any distribution, while the latter are assumed to be geometrically distributed. Using a generating–functions approach, an expression for the steady–state probability generating function of the buffer occupancy is derived. From this result, a closed–form expression for the tail distribution of the buffer occupancy is derived, that is practical and easy to evaluate; this latter quantity is especially useful for buffer dimensioning purposes. In addition, an accurate approximations for this quantity, that reduces all numerical calculations to an absolute minimum, is established as well.

Keywords

Rate adapter, generating functions, tail approximation

1 SYSTEM DESCRIPTION

Let us consider a communication network, where information is segmented into fixed–length packets, and transmissions, as well as arrivals of packets are synchronized with respect to an infinite set of periodic (i.e., equidistant) time instants. The time period elapsed by two consecutive time instants is referred to as one *slot*, and one slot

The authors wish to acknowledge the support of Alcatel Bell Telephone Mfg. Co. (Antwerp, Belgium), and the National Fund for Scientific Research (NFWO).

L. Mason et al. (eds.), *Broadband Communications*

suffices for the transmission of exactly one packet, a situation which, for instance, occurs in an ATM–based B–ISDN network, where a packet then represents one ATM cell of 53 bytes (De Prycker, 1991).

The packets generated by a source typically pass through a number of nodes before reaching their final destination, and we consider the situation where at some point in the network, the amount of information per time unit that can arrive on the incoming link (*input link*) exceeds the amount of information that can be transmitted per time unit on the outgoing link (*output link*). This could be caused, for instance, by a difference in transmission speed between input and output link due to different clock rates, or because, due to some internal representation mode in a switching network, bits are added to the original packets before sending them through the network. In any case, if packet loss is to be avoided, incoming packets must be buffered in a so–called rate–adaptation buffer, and the difference in transmission rate can be incorporated in the corresponding discrete–time queueing model by choosing a different time scale, i.e., slot length, on the input and output link (meaning that we obtain longer slots on the output link compared to the input link), a situation which is depicted in Figure 1.a,b together with some additional quantities still to be defined in this and the following section. In the remainder of the paper, a slot corresponding to a time unit on the Input (Output) Link will be referred to as an IL (OL) slot, and, similarly, the associated Time Scale on the Input (Output) Link will be denoted by ILTS (OLTS).

The packet arrival process on the input link will be characterized by specifying the lengths of successive *passive* and *active periods*, being defined as a number of consecutive slots during which there is no packet arrival, respectively a number of consecutive slots each carrying exactly one packet (as we already mentioned, packet arrivals are synchronized with respect to the IL slot boundaries). In particular, we let the random variables b_n and i_n, $n \geq 1$, represent the lengths of successive active and passive periods, where at some initial time instant t=0, the first active period is initiated, each active period being followed by a passive period. It will now be assumed that $\{b_n \mid n \geq 1\}$ and $\{i_n \mid n \geq 1\}$ are two sets of i.i.d. random variables; in addition, the elements of these two sets of random variables are assumed to be mutually independent. This implies that the probability mass function of any random variable b_n (i_n) can be represented by one common probability generating function B(z) (I(z))

$$B(z) \triangleq E[z^{b_n}] \quad , \quad I(z) \triangleq E[z^{i_n}] \tag{1}$$

where E[.] denotes the expected value of the tagged quantity. In the analysis throughout the following sections, B(z) can take any form, whereas this is not the case as far as I(z) is concerned. For our purposes, it is sufficient to assume that I(z) has a geometric form with parameter α and mean $1/(1-\alpha)$, and thus can be written as

$$I(z) = \frac{(1-\alpha)z}{1-\alpha z} \quad , \tag{2}$$

although a more general, rational form for I(z) could also be taken into consideration.

The analysis presented in this paper, is a first step in the study of the rate–adaptation buffer–dimensioning problem. To that extent, we assume that the ratio of the transmission rate versus the arrival rate can be written as the fraction of two integers

$$\frac{\textit{Transmission Rate}}{\textit{Arrival Rate}} = \frac{k-1}{k} \quad , \quad k \geq 2 \quad , \tag{3}$$

meaning that during the time period required to transmit k–1 packets, exactly k packets could arrive. This assumption thus covers a considerable range of possible values for the above mentioned ratio, while keeping the analysis tractable. Now, assuming that the initial time instant t=0 coincides with both an IL and an OL slot boundary, then when the packet arrival process, originally generated on an ILTS–basis, is converted to the OLTS, the beginning (and ending) of a packet arrival will coincide with any of the k time instants within an OL slot that arise from devising the OL slot into k time periods of equal length; such a time instant will be referred to as a *microslot boundary*, while the time period elapsed between two successive microslot boundaries, is called a *microslot*. This is illustrated in Figure 1.a,b for k=6.

To the best of our knowledge, the buffer dimensioning of a rate adapter module has received only little attention in the literature. In Rothermel (1992), for a Bernoulli arrival process, (i.e., geometrically distributed active and passive periods, with parameters σ and $1-\sigma$ respectively, where σ is the load of the arrival stream on the ILTS), a simple approximate procedure was developed, which neglects the equidistant nature of slots during which multiple cell arrivals can occur on the OLTS. In Michiel (1990), also for a Bernoulli arrival stream, approximate results were derived that are sufficiently accurate, as long as the difference between input and output rates remains sufficiently small (less than 2%).

2 SYSTEM EQUATIONS

Since the transmission of packets is synchronized to the OL slot boundaries, and as such is essentially based on the OLTS, the packet arrival process, described by the random variables b_n and i_n, must be translated into corresponding quantities describing the arrival process on the OLTS. Let us therefore define $b_{n,o}$ and $i_{n,o}$ as the random variables describing the lengths (i.e., the numbers of OL slots) of successive active and passive periods on the OLTS, where an active (passive) period on the OLTS is defined as a number of consecutive OL slots during which at least one (no) packet arrives. Since it is quite possible that a packet arrival crosses an OL slot boundary (see Figure 1.b), it should be indicated that a packet is considered to be in the buffer only when its arrival is completed; therefore, the OL slot of arrival of a packet is the slot during which its arrival has ended. Also, note that during an OL slot, there are either 0, 1 or 2 packet arrivals. It is now possible to express $b_{n,o}$ and $i_{n,o}$ in terms of b_n and i_n respectively. For that purpose, let us also define R_n and P_n, $0 \leq R_n, P_n \leq k-1$, as the discrete random variables that represent the position of the microslot boundary within an OL slot that coincides with the beginning (on the ILTS) of the n–th active, respectively the n–th

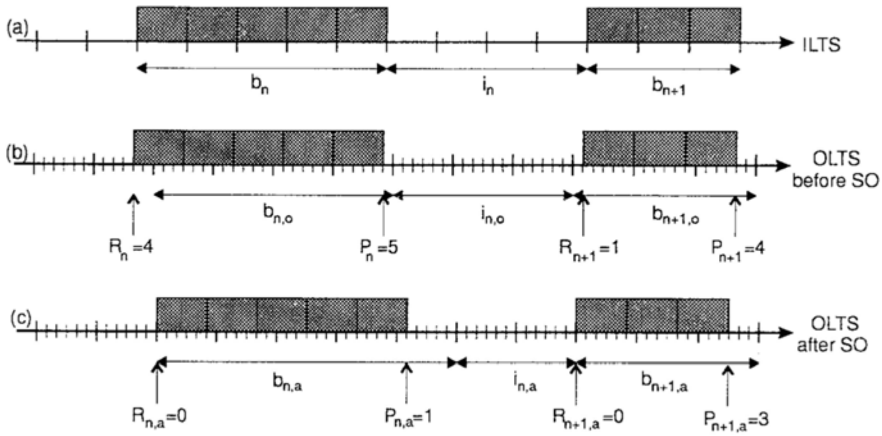

Figure 1 Active and passive periods on *Input Link Time Scale* (a), and on the *Output Link Time Scale* before (b) and after (c) the *Shift Operation*.

passive period. $R_n=0$ ($P_n=0$) would correspond with the beginning of an OL slot, and successive microslot boundaries within an OL slot are numbered sequentially, up to k–1, as shown in Figure 1.b. Obviously, the value of the steady–state joint probabilities

$$p(i,j) \triangleq \lim_{k \to \infty} \text{Prob}[R_n=i, P_n=j] \quad , \tag{4.a}$$

only depends on the number of IL slots enclosed by these two time instants (i.e., the length of b_n), and therefore only depends on the difference i–j. Consequently, defining $p_i = p(i,0)$, it can be verified that

$$\begin{aligned} p_i &= \frac{1}{k}\text{Prob}[(b_n \bmod k)=i] = \frac{1}{k}\sum_{j=0}^{\infty} b(jk+i) \quad , \quad 0 \le i \le k-1 \quad , \\ p(i,j) &= p_{i+k-j} \quad , \quad 0 \le i \le k-1 \quad \text{and} \quad i < j \le k-1 \\ p(i,j) &= p_{i-j} \quad , \quad 0 \le i \le k-1 \quad \text{and} \quad 0 \le j \le i \quad , \end{aligned} \tag{4.b}$$

where $b(j)$, $j \ge 1$, denotes the probability mass function of the random variable b_n, and (x mod y) represents the remainder of the fraction x/y of the integers x and y. From (4.b), it is also found that

$$\text{Prob}[R_n=i] = \text{Prob}[P_n=j] = 1/k \quad , \quad 0 \le i,j \le k-1 \quad , \tag{4.c}$$

as could be expected. After carefully examining the problem, the following relations between $\{b_{n,o}, i_{n,o}\}$ and $\{b_n, i_n\}$ were established :

$$b_{n,o} = b_n + \theta(R_n=0) + \theta(R_n=1) - ((b_n+k-R_n) \text{ div } k) \tag{5.a}$$

$$i_{n,o} = i_n + \theta(P_n=0) - ((i_n+k+1-P_n) \text{ div } k) \tag{5.b}$$

where (x div y) denotes the integer part of the fraction x/y, and where the indicator function $\theta(.)$ equals one if the Boolean argument is true, and zero otherwise.

Since we are interested in dimensioning the rate–adaptation buffer, let us define the random variable v_n as the buffer occupancy (i.e., the number of packets in the buffer, including the one that is currently being transmitted, if any) at the beginning of the first OL slot of the n–th active period (on the OLTS), and, similarly, u_n as the buffer occupancy at the beginning of the first OL slot of the n–th passive period. With the previous definitions and assumptions, a number of system equations that relate the sequence of random variables u_n and v_n can be established. First of all, since exactly one packet can be transmitted during each OL slot of a passive period, we find that

$$v_{n+1} = (u_n - i_{n,o})^+ \quad , \tag{6.a}$$

where $(.)^+ \triangleq \max\{0,.\}$. On the other hand, the buffer occupancy at the beginning of a passive period can be expressed in terms of the buffer occupancy at the beginning of the preceding active period as

$$u_n = (v_n - 1)^+ + a_n + 1 \quad , \tag{6.b}$$

where

$$a_n \triangleq b_n - b_{n,o} \quad , \tag{6.c}$$

is the number of packets that must be buffered during the active period due to the difference in arrival and transmission rate. Under the assumptions described in the previous section, it is not difficult to see from (5.a,b) and (6.a–c) that consecutive pairs of random variables $\{v_n, R_n\}$, $\{u_n, P_n\}$, for increasing values of n, form a two–

dimensional Markov chain. However, we will not try to solve the system equations for the exact problem. Instead, we first execute a so–called *Shift Operation* (SO), which means that we let the the first packet of a new sequence coincide with the beginning of the first OL slot of the corresponding active period, as shown in Figure 1.c. In the remainder, the random variables (b_n, i_n, R_n, P_n) will be tagged with the subscript $_a$ if necessary, to indicate that we consider these random variables *after* the SO has been executed. This SO implies that, since $R_{n,a}=0$, consecutive random variables v_n and u_n, for increasing values of n, form a Markov chain, thereby considerably reducing the complexity of the solution.

Note that, compared to the exact problem, this SO has almost no influence on the length of passive and active periods (i.e., a difference in length of at most one slot), and therefore, we expect to find a very accurate approximation while keeping the analysis tractable. First of all, when combining (6.c) and (5.a) while setting $R_n=0$, we obtain

$$a_n = b_n - b_{n,a} = (b_n \text{ div } k) \quad . \tag{7}$$

Using some standard techniques related to the calculation of z–transforms, it can be shown that the associated probability generating function A(z) is then given by

$$A(z^k) = \frac{1}{k}\sum_{s=0}^{k-1} \frac{1-z^{-k}}{1-(\mu^s z)^{-1}} B(\mu^s z) \quad , \quad \mu \triangleq \exp\{\tfrac{2\pi\iota}{k}\} \quad , \tag{8}$$

where ι is the imaginary unit

In addition, the SO can also alter the length of the passive periods, as becomes clear from Figure 1.b,c. The random variable $i_{n,a}$ that represents the length of the n–th passive period *after* the SO, is related to $i_{n,a}$ by

$$i_{n,a} = i_{n,o} + \theta(R_n=1,P_n=1) - \theta(R_n\geq 2,P_n=0) - \theta(R_n\geq 2,P_n>R_n) \quad , \tag{9}$$

and system equation (6.a) now becomes

$$v_{n+1} = (u_n - i_{n,a})^+ \quad . \tag{10}$$

The combination of (9) and (5.b) leads to an expression for $i_{n,a}$ in terms of i_n, the length of the original passive period on the ILTS. This relation can be transformed into a relationship between z–transforms, and it can be shown that the corresponding probability generating function $I_a(z)$ satisfies

$$I_a(z^k) = \frac{1}{k}\sum_{s=0}^{k-1} (\mu^s z^{k-1})^{-1} \left[\frac{1-z^k}{1-\mu^{-s}z}\right]^2 I(\mu^s z^{k-1})\, P(\mu^s z^{-1}) \quad , \quad P(z) \triangleq \sum_{i=0}^{k-1} z^i p_i \quad . \tag{11}$$

where the p_i's and I(z) were defined in (4.b) and (1) respectively. Expression (11) for $I_a(z)$ is convenient as far as the calculation of the mean value and higher order moments is concerned. However, as will become clear in the following analysis, we especially require the probability mass function that corresponds to $I_a(z)$. Assuming that I(z) indeed satisfies (2), $I_a(z)$ can be transformed into

$$I_a(z) = \frac{1-\alpha}{k-1}\sum_{t=0}^{k-2} \frac{s(t)}{1-\nu^{tk}\alpha^{k/(k-1)}z} - s_0 \quad , \quad \nu \triangleq \exp\{\tfrac{2\pi\iota}{k-1}\} \quad , \tag{12.a}$$

$$s_0 \triangleq \frac{1-\alpha}{\alpha}\left\{\sum_{j=k-1}^{2(k-1)} (2k-j-1)p_{j-k+1} + p_0/\alpha\right\}$$

$$s(t) \triangleq \left[\frac{1-\alpha^{-k/(k-1)}\nu^{-kt}}{1-\alpha^{-1/(k-1)}\nu^{-t}}\right]^2 P(\alpha^{1/(k-1)}\nu^t) \quad . \tag{12.b}$$

3 THE GENERATING FUNCTION OF THE BUFFER OCCUPANCY

Now we have everything at hand to be able to calculate expressions for $U_n(z)$ and $V_n(z)$, the generating functions corresponding to u_n and v_n respectively. First of all, system equation (6.b) can be transformed into a relation between z–transforms, which, in view of the statistical independence of a_n and v_n, leads to

$$U_n(z) = A(z)\,\{V_n(z) + (z-1)V_n(0)\} \quad . \tag{13}$$

On the other hand, from (10) and the statistical independence of u_n and $i_{n,a}$, we obtain

$$V_{n+1}(z) = U_n(z)\, I_a(\tfrac{1}{z}) + \sum_{i=1}^{\infty} \{I_i(1) - z^i I_i(\tfrac{1}{z})\}\ \mathrm{Prob}[u_n=i]$$

$$I_i(z) \triangleq E[z^{i_{n,a}} \mid i_{n,a} \geq i]\ \mathrm{Prob}[i_{n,a} \geq i] \quad , \quad 1 \leq i \ ,$$

which, due to expression (12.a) for $I_a(z)$ equals

$$I_i(z) = \frac{1-\alpha}{k-1} \sum_{t=0}^{k-2} s(t) \left\{ \frac{(xz)^i}{1 - xz} \right\}_{x=\nu^{tk}\alpha^{k/(k-1)}} \quad .$$

Combining the previous expressions for $I_i(z)$ and $V_{n+1}(z)$, we find

$$V_{n+1}(z) = U_n(z) I_a(\tfrac{1}{z}) + \frac{1-\alpha}{k-1} \sum_{t=0}^{k-2} s(t) \left\{ \frac{U_n(x)}{1-x} - \frac{U_n(x)}{1-xz^{-1}} \right\}_{x=\nu^{tk}\alpha^{k/(k-1)}} \quad . \tag{14}$$

The buffer occupancy will typically reach its highest values just after an active period, i.e., at the beginning of a passive period. It is appropriate to use the distribution of the buffer occupancy at these worst–case time instants for buffer–dimensioning purposes. We will therefore establish an expression for U(z), the steady–state probability generating function describing the buffer occupancy at the beginning of a passive period. The system will reach a steady–state only if the equilibrium condition is satisfied, meaning that ρ, the mean number of packet arrivals per slot (on the OLTS) in the rate adapter buffer must be less than 1, which is equivalent to requiring that

$$I_a'(1) > A'(1) \quad , \tag{15}$$

where primes denote derivatives with respect to the argument. In other words, the mean length of a passive period on the OLTS must exceed the mean number of packets that are accumulated during an active period. Equations (13) and (14) now lead to the following expression for U(z) :

$$U(z) = \frac{z(z-1)A(z) \sum_{j=0}^{k-2} r_j z^j}{z^{k-1} - (\alpha^k + A(z)\, I_a(\frac{1}{z}) Q(z))} \quad , \quad Q(z) \triangleq z^{k-1} - \alpha^k \quad . \tag{16}$$

While deriving (16), we have used the property that, just after an active period, there is always at least one packet in the buffer, implying that U(0)=0. The constants r_j, $0 \leq j \leq k-2$, that occur in (16) are linear combinations of the unknowns $U(\nu^{tk}\alpha^{k/(k-1)})$, $0 \leq t \leq k-2$, and V(0), where V(z) is the steady–state limit of $V_n(z)$. From Rouché's theorem, one can show that the denominator

$$D(z) \triangleq z^{k-1} - (\alpha^k + A(z) I_a(\tfrac{1}{z}) Q(z)) \tag{17}$$

of expression (16) for U(z) has k–1 zeros inside the unit disk (including z=1). Without giving a full prove, note that the zeros of $Q(z)$ cancel the poles of $I_a(1/z)$, which leads to the observation that $D(z)$ is analytic inside the complex unit disk, a necessary condition for applying Rouché's theorem. In the remainder, we we will denote by z_j, $1 \le j \le k-2$, the zeros of $D(z)$ inside the unit disk and different from 1. Since $U(z)$ is analytic inside the complex unit circle, these zeros must also make the numerator of (16) zero, and combined with the normalization condition, this completely determines the (k–2)–th polynomial in the numerator. We thus obtain

$$U(z) = D'(1)\frac{z(z-1)A(z)}{D(z)}\prod_{j=1}^{k-2}\frac{z-z_j}{1-z_j} \quad . \tag{18}$$

where the coefficient $D'(1) = Q(1)(I_a'(1)-A'(1))$ guarantees the normalization of $U(z)$. From this expression for the probability generating function, we readily obtain the moments of the buffer contents by taking the appropriate derivatives with respect to z for z=1 of $U(z)$. In this paper, we focus attention on the tail distribution of the buffer contents, a quantity which is very useful for buffer dimensioning purposes. Whatever performance characteristic we are interested in, the time consuming part of all numerical calculations remains finding the zeros z_j of $D(z)$ inside the complex unit disk. Therefore, we now show that these quantities can be accurately approximated by

$$z_j \cong \alpha^{k/(k-1)}\nu^j \quad , \quad 1 \le j \le k-2 \quad , \tag{19}$$

Indeed, from expressions (12.a) and (16) for $I_a(z)$ and $Q(z)$, it readily follows that

$$\left|D(\alpha^{k/(k-1)}\nu^t)\right| = (1-\alpha)\alpha^k\left|A(\alpha^{k/(k-1)}\nu^t)\right|\left|s(i)\right| \quad , \quad 1 \le t \le k-2 \quad , \tag{20}$$

where the integer i is such that $\nu^{ik} \equiv \nu^t$. Now, since

$$|s(i)| \le \sum_{i=0}^{k-1} p_i \sum_{j=0}^{2(k-1)} \min(j+1,2k-j-1)\alpha^{-j/(k-1)} = \frac{1}{k}\left[\frac{1-\alpha^{-k/(k-1)}}{1-\alpha^{-1/(k-1)}}\right]^2 \quad ,$$

and $|A(z)|<1$ if $|z|<1$, the above relations explain why (19) forms a very good approximation for the z_j's, which, due to the presence of the factor α^k/k, becomes better as k increases. The accuracy of this approximation is illustrated in Section 5.

4 TAIL DISTRIBUTION OF THE BUFFER OCCUPANCY

As has already been indicated in various papers (Woodside (1987), Desmet (1992), Sohraby (1992), Bruneel (1993)), if the probability generating function of the buffer occupancy has nothing but non–essential singularities of order 1 (i.e., simple poles), then the tail of the buffer–occupancy distribution can be approximated very accurately by a geometric form, implying in our specific case that the probability that the buffer occupancy just after an active period exceeds an integer threshold U, is given by

$$\text{Prob}[u>U] \cong \frac{-Cz_0^{-U-1}}{z_0-1} \quad , \tag{21}$$

where z_0 is the pole of $U(z)$ with the smallest modulus (i.e., the solution of $D(z)=0$ outside the unit disk with the smallest modulus), which is a real and positive quantity larger than 1. Due to the residue theorem, the constant C in the above expression is equal to

$$C \triangleq \lim_{z\to z_0}(z-z_0)U(z) = z_0(z_0-1)A(z_0)\frac{D'(1)}{D'(z_0)}\prod_{j=1}^{k-2}\frac{z_0-z_j}{1-z_j} \quad . \tag{22}$$

Using approximation (19) for the z_j's, a close approximation for C avoiding the calculation of these quantities can be derived. We obtain

$$C \cong z_0(z_0-1)A(z_0)\frac{D'(1)}{D'(z_0)}\frac{z_0^{k-1} - \alpha^k}{z_0 - \alpha^{k/(k-1)}}\frac{1 - \alpha^{k/(k-1)}}{1 - \alpha^k} \quad . \tag{23}$$

From equations (21) and (23), an approximation for the geometric–tail limit of the buffer occupancy distribution is obtained, which is easily evaluated, since it merely requires the calculation of z_0. As will be shown in the next section, this approximation is extremely close to the actual values of the geometric–tail approximation, calculated by combining (22) and (23).

5 SOME NUMERICAL EXAMPLES

From now on, we let σ denote the packet arrival rate on the input link. The rate adapter model previously described is completely specified, once the value of the parameter k (which determines the difference between input and output rate), and the active and passive period distributions (or, equivalently, B(z) and I(z)) are given. Up to now, B(z) could have any form, and in this section, we consider two cases, where either the input process is the output of a discrete–time M/D/1 queue with load σ (in the remainder referred to as *M/D/1–like arrivals*), or a model where passive and active periods are statistically independent and both geometrically distributed (in the remainder referred to as *geo–like arrivals*), with parameters α and β respectively, satisfying $\sigma = (1-\alpha)/(2-\alpha-\beta)$. Note that in the case of M/D/1–like arrivals, the lengths of active and passive periods indeed are statistically independent.

First of all, for M/D/1–like arrivals, the parameter α characterizing I(z) in (2) equals $\exp\{-\sigma\}$, the probability of having no packet arrivals during a slot in the M/D/1 queue, and on the other hand, it has been derived in Bruneel (1993) that in this case B(z) is given by

$$B(z) = \frac{R(z)/z - \alpha}{1 - \alpha} \quad , \tag{24.a}$$

where R(z) is implicitly defined by the equation

$$R(z) = z\exp\{\sigma(R(z)-1)\} \quad . \tag{24.b}$$

Furthermore, in Steyaert (1993) it was shown that for a generating function satisfying (24.b), the corresponding probability mass function, here denoted by $r(j)$, is given by

$$r(j) = \sigma\frac{(j\sigma)^{j-2}}{(j-1)!}\exp\{-j\sigma\} \quad , \quad j \geq 1 \quad . \tag{24.c}$$

From (24.a,c), the probability mass function corresponding to B(z) is readily obtained.

In the case of geo–like arrivals, I(z) is still given by (2), while B(z) now equals

$$B(z) = \frac{(1-\beta)z}{1 - \beta z} \quad . \tag{25}$$

An interesting quantity is the parameter L, which is defined as the ratio of the mean length of a passive (or, equivalently, active) period versus the mean length of a passive (active) period in the case of M/D/1–like arrivals and equal values of the load σ of the input process. Consequently, L=1 means that the geometrically distributed active and passive periods have the same average length as for the output process of an M/D/1 queue with equal load, and considering increasing values of L while keeping the ratio $\sigma = B'(1)/(I'(1)+B'(1))$ constant implies that the arrival rate in the rate adapter module remains constant, while the 'variability' in the arrival process increases. From the previous, we find that the parameter L can be calculated from

$$L = \frac{1 - \exp\{-\sigma\}}{1 - \alpha} \ . \tag{26}$$

The packet arrival process in the geo–like arrivals case can now be characterized by the pair (σ,L) instead of (α,β), for given values of k.

The results obtained throughout this paper have been illustrated in Figs. 2–8. In Figs. 2–3, for various values of k and $\rho=\sigma.k/(k-1)$, we have plotted the 'exact' geometric–tail approximation for Prob[u>U], obtained from (21) and (22) (full line) together with the simplified result obtained from (21) and (23) (marks), in the case where the arrivals process is either M/D/1–like (Fig. 2), or geo–like (Fig. 3) with L=1. First of all, it is found that, whatever the type of arrival process and whatever the values of the parameters characterizing the arrival process, no difference between the 'exact' and approximate results can be observed. These curves show that approximation (19) for the z_j's is extremely accurate, and can be used without any restriction. Furthermore comparing the curves of both figures, we may conclude that, although the active period distribution has the same mean value in both cases when considering equal values of ρ and k (due to L=1; note that this also implies the passive periods have identical geometric distributions), a substantial difference between both cases exists, i.e., second–order effects (such as higher–order moments of the active periods) still have a considerable impact on the buffer behavior of the rate–adapter, and, therefore, cannot be neglected in the buffer dimensioning process.

In Figs. 4–6, we have plotted the geometric–tail approximation for Prob[u>U] in the M/D/1–like arrivals case, for constant values of ρ and various values of k. It is observed that when, while large differences occur for low values of ρ, these diminish as ρ

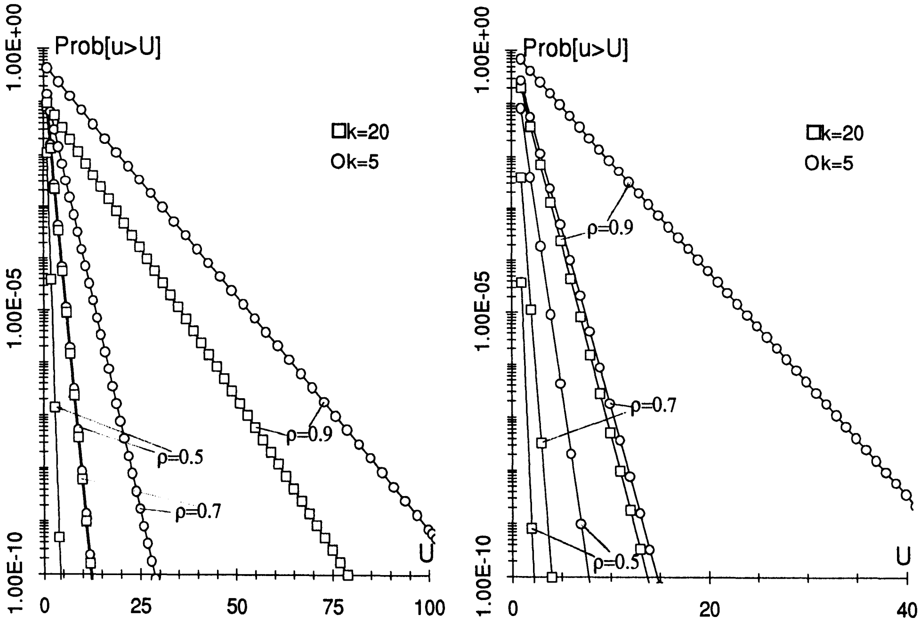

Figure 2 M/D/1–like arrivals, k=5,20, ρ=0.5,0.7,0.9.

Figure 3 Geo–like arrivals, k=5,20, ρ=0.5,0.7,0.9, L=1.

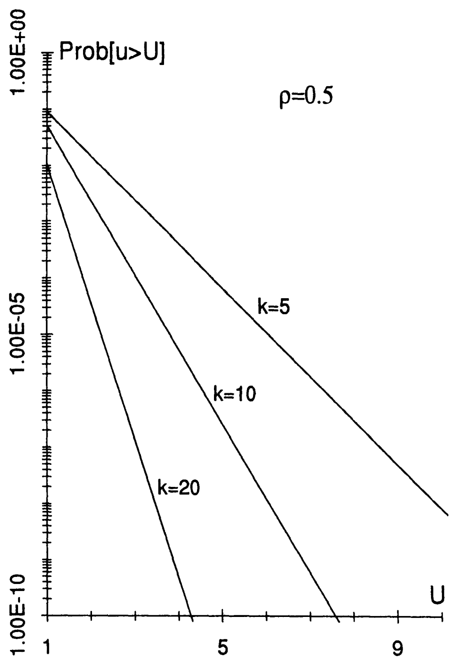

Figure 4 M/D/1–like arrivals, k=5,10,20, ρ=0.5.

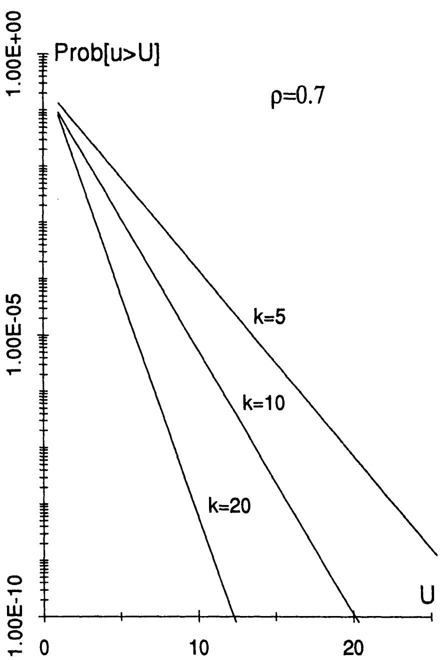

Figure 5 M/D/1–like arrivals, k=5,10,20, ρ=0.7.

increases, i.e., the results for the tail distribution become less sensitive to the exact value of the output rate/input rate ratio. Finally, in Figs. 7–8, we examine the impact of increasing values of L on the buffer requirements of the rate adapter, in the geo–like arrivals case. The case L=1 was plotted in Fig. 3; in Figs. 7–8, for various values of k and ρ, we considered values of L equal to 5 and 10. It is observed that, when comparing the respective curves for equal values of k and ρ, the required buffer space increases as L increases. Keeping in mind the conclusions in the discussion concerning Figs. 2–3, this is hardly surprising, since, again, increasing values of L implies increasing variability in the of the arrival process, and we observe that the impact of increasing values of L on the buffer beahvior is quite severe.

6 SUMMARY

In this paper, we have tackled the problem of dimensioning a rate adaption module, which arises in a node of a telecommunication networks if the arrival rate on the input link exceeds the transmission rate of the output link. Based on a generating functions approach, we obtained expressions for the mean and tail distribution of the buffer occupancy, which are easy to evaluate, since, due to the approximation for the z_j's, numerical calculations are reduced to a minimum. From the numerical results we may conclude that (1) even for small differences between input and output rate (less than 5%), the involved buffers can become quite large, and (2) second order effects, such as higher order moments of the lengths of active and passive periods are not negligible.

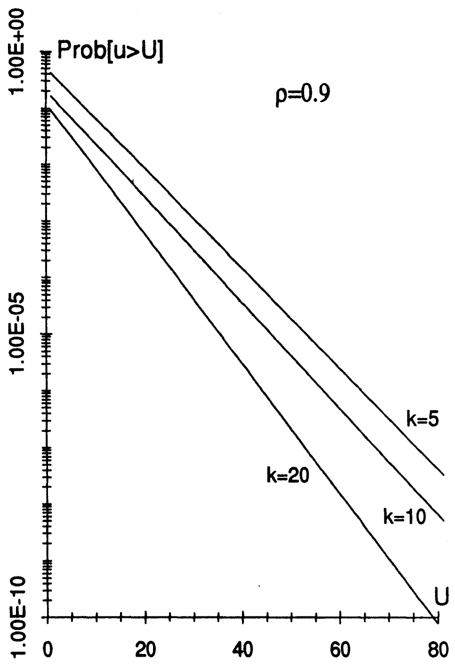

Figure 6 M/D/1–like arrivals, k=5,10,20, ρ=0.9.

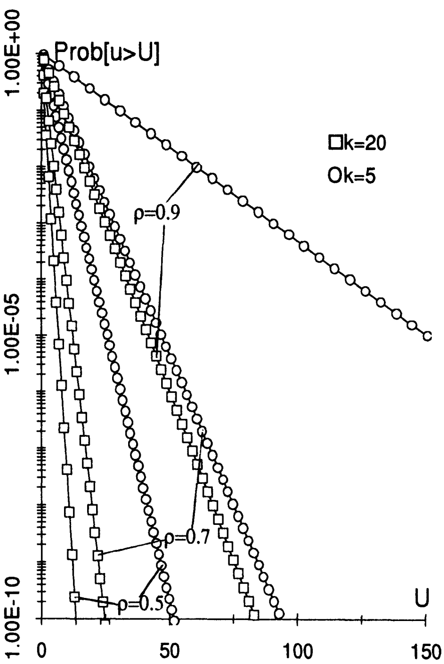

Figure 7 Geo–like arrivals, k=5.20, ρ=0.5,0.7,0.9, L=5.

7 REFERENCES

Bruneel, H., and Kim, B.G. "Discrete–time models for communication systems including ATM" (Kluwer Academic Publishers, Boston, 1993).

De Prycker, M. "Asynchronous Transfer Mode – Solution for Broadband ISDN", Ellis Horwood Ltd. (England, 1991).

Desmet, E.; Steyaert, B.; Bruneel, H. and Petit, G.H. "Tail distributions of queue length and delay in discrete–time multiserver queueing models, applicable ATM networks", *Proceedings of the 13th ITC* (Copenhagen, 1992), Vol. *Queueing, Performance and Control in ATM*, pp. 1–6.

Michiel, H. and Petit, G.H. "Dimensioning of a rate adapter", Alcatel–Bell Research Report ATG_029/HM_901019 (Antwerp, 1990).

Rothermel, K. "Traffic studies of transmission bit rate conversion in ATM networks", *Proceedings of the 13th ITC* (Copenhagen, 1992), Vol. *Queueing, Performance and Control in ATM*, pp. 65–69.

Sohraby, K. "On the asymptotic behavior of heterogeneous statistical multiplexer with applications, *Proceedings of Infocom'92*, pp. 839–847.

Steyaert, B. and Bruneel, H. "Delay characteristics of a CBR source with priority background traffic", LCI Research Report 18, June 1, 1993.

Woodside, C. and Ho, E. "Engineering calculation of overflow probabilities in buffers with Markov–interrupted service", *IEEE Trans. on Commun.*, Vol. COM–35 (1987), pp. 1272–1277.

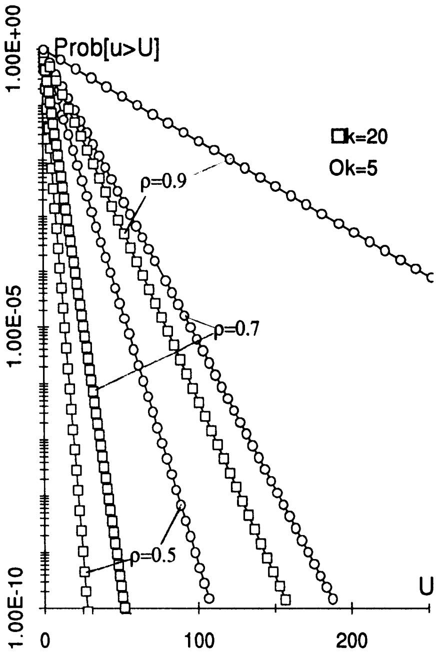

Figure 8 Geo–like arrivals, k=5,20, ρ=0.5,0.7,0.9, L=10.