A fluid model of an ATM traffic shaper

EINDHOVEN UNIVERSITY OF TECHNOLOGY
Department of Mathematics and Computing Science

Memorandum COSOR 95-08

A fluid model of an ATM traffic shaper

J.L. van den Berg
E.B. Diks
J.A.C. Resing
J. van der Wal

Eindhoven, March 1995
The Netherlands

Eindhoven University of Technology
Department of Mathematics and Computing Science
Probability theory, statistics, operations research and systems theory
P.O. Box 513
5600 MB Eindhoven - The Netherlands

Secretariat: Main building 9.10
Telephone: 040-47 3130

ISSN 0926 4493

A fluid flow model of an ATM traffic shaper

J.L. van den Berg*, E.B. Diks†, J.A.C. Resing† and J. van der Wal †

Abstract

In this paper we study a fluid flow model of an ATM traffic shaper. In particular, we are interested in the relation between the buffer capacity of the shaper, the shaper rate, the cell arrival process and the cell loss probability. The main conclusion of the paper is that cell loss probabilities are very sensitive to several traffic characteristics of the arrival process, like e.g. the squared coefficient of variation of the durations of on-periods in an on-off arrival process.

Keywords : Traffic shaper, cell loss, fluid flow model.

1 Introduction

Public ATM backbone networks will be operational within a few years. In the first years of the introduction mainly semipermanent virtual path connections (VPCs) will be used, e.g. for the interconnection of Local Area Networks (LANs). In order to reduce the required VPC bandwidth it is necessary to reduce the rate with which LAN data bursts are offered to the ATM network. This can be achieved by buffering data bursts at the entrance to the network.

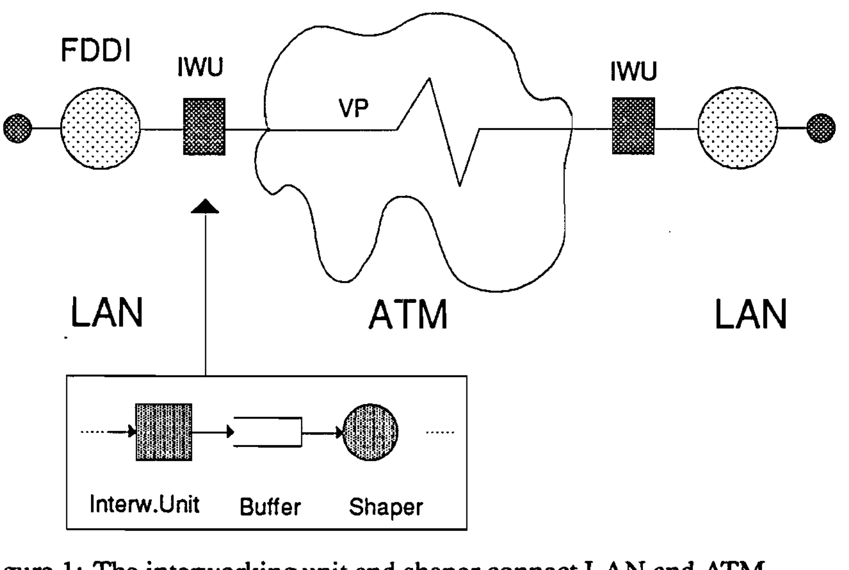

Figure 1: The interworking unit and shaper connect LAN and ATM.

In this paper, we consider the situation in which two LANs are connected via an ATM VPC, see figure 1. Protocol conversion is performed in the interworking units (IWUs) between the LANs and the ATM network. In particular, in the IWUs the relatively large packets coming from the LANs are segmented

*PTT Research, P.O. Box 421, 2260 AK, Leidschendam, The Netherlands.
†Eindhoven University of Technology, P.O.Box 513, 5600 MB, Eindhoven, The Netherlands.

into short, fixed length, ATM cells. The cells are stored in a buffer (traffic shaper) before they are offered to the network. The traffic shaper forwards the cells to the network at a fixed (cell) rate equal to the ATM VPC rate. Cells that arrive at a full buffer are discarded. In fact the traffic shaper takes care of adapting the speed of the LANs to the bandwidth provided by the ATM VPC. For instance, in the case of an FDDI LAN, packets arrive at the IWU with a speed of 100 Mbit/s while the mean arrival rate may be much smaller, e.g. 2 Mbit/s. Hence traffic shaping is needed in order to reduce the required VPC bandwidth.

The present study is concerned with the following important question: what is the relation between the buffer capacity, the shaper rate (VPC bandwidth), the cell arrival process and the cell loss probability? Insight into this problem is important for making a proper choice for the bandwidth of the ATM VPC. Obviously, to minimize costs, the VPC bandwidth should be chosen as low as possible under the condition that the resulting cell loss probability is smaller than a certain predefined value and the cell delay is acceptable.

Different approaches for modelling ATM systems are used in the literature (see Roberts [6]), in particular discrete-time queueing models (see e.g. Gravey, Louvion and Boyer [2]) and fluid flow models (see e.g. Anick, Mitra and Sondhi [1]). In this paper we have chosen to use a fluid flow approach. This approach enables us to derive expressions for the cell loss probability in the case that the arrival process is given by an N-state Markov process. We come back to related literature on fluid flow models in section 2.3.

We will now give an overview of the contents of the rest of the paper. In section 2 the traffic shaper is modelled as a (continuous time) fluid flow model and expressions for the cell loss probability are derived. In section 3 we present and discuss numerical results. We restrict our attention mainly to on-off arrival streams. First, we study the influence of the variability of the length of the on-periods of the arrival process (i.e. the packet size) on the cell loss probability. Next, we consider the relation between the shaper rate and the buffer size for given cell loss probability and arrival process. Finally, we compare the results of the fluid flow model with the results of a discrete-time MMBP queueing model. The conclusions of our study are given in section 4.

2 The fluid flow approach

2.1 Model description

In the fluid flow model data enters the buffer with a rate depending on the state of the arrival process. The buffer is emptied, say at rate r (here r equals the shaper rate). The buffer is assumed to have finite capacity K. If the buffer is empty and the arrival rate is less than r then the output rate of the buffer reduces to the arrival rate. If the buffer is full and the arrival rate exceeds r then the arrival stream is split into two parts, one with rate r which enters the buffer, and a remainder which overflows and is lost.

The state of the arrival process is described by a Markov process with N states. The transition rate from state i to state j, $j \neq i$, of this chain equals λ_{ij}. The Markov chain regulates the arrival rates, i.e. when the chain is in state i the arrival rate is equal to μ_i. We will assume $\mu_i \neq r$ for all i. (If one or more arrival rates are equal to r the system can be reduced by eliminating these components.)

We write L and H for the sets of low and high arrival rates respectively:

$$L := \{i \mid \mu_i < r\}\ ,\ H := \{i \mid \mu_i > r\}. \tag{1}$$

The problem is only interesting if both L and H are nonempty. If H is empty then the buffer will never be filled. If L is empty then the average overflow rate is equal to the average arrival rate minus r. For convenience, we assume $\mu_N > r$.

2.2 Analysis

For the analysis of our system, we follow the renewal approach for the analysis of a manufacturing system as presented in Wijngaard [8]. The state of the system is a pair (i, x) with i the state of the arrival process and x the buffer content. In order to obtain the overflow (loss) fraction we look at the overflow in a renewal cycle. As renewal points we take the entrance into the state (N, K), i.e., the arrival process is in state N and the buffer is full. (Note that also other choices for the renewal points are possible.)

Wijngaard introduces general "cost" functions to obtain performance measures for his manufacturing system. We use these cost functions to obtain the loss fraction for the traffic shaper. In order to find the loss fraction two cost functions are considered. One will give the expected duration of a cycle, the other the average loss in a cycle. Obviously the loss fraction is equal to the quotient of the average loss in a cycle and the product of the mean cycle duration and the average arrival rate. The average arrival rate is equal to $\sum \pi_i \mu_i$, with π_i the equilibrium probability for the arrival process to be in state i. In any realistic situation we will have $\sum \pi_i \mu_i < r$.

For both cost functions we have to define a cost rate: when the arrival state is equal to i and the buffer content is equal to x the cost rate is denoted by $c_i(x)$. For the cost function "time" we will have $c_i(x) = 1$ for all i and $0 \le x \le K$. For the cost function "loss" we define $c_i(K) := \mu_i - r$ for $i \in H$ and 0 everywhere else.

Define $v_i(x)$ as the residual cost until the next renewal if the arrival process is in state i and the buffer content is equal to x. For $v_i(x)$, $0 < x < K$, a differential equation can be obtained by looking ahead an infinitesimal amount of time Δ:

$$v_i(x) = \Delta c_i(x) + \sum_{j \neq i} \lambda_{ij} \Delta v_j(x + \Delta\mu_i - \Delta r) + (1 - \sum_{j \neq i} \lambda_{ij} \Delta) v_i(x + \Delta\mu_i - \Delta r) + o(\Delta). \quad (2)$$

Taking the limit for Δ towards 0 we obtain

$$(r - \mu_i) v_i'(x) = c_i(x) + \sum_{j \neq i} \lambda_{ij} v_j(x) - \sum_{j \neq i} \lambda_{ij} v_i(x). \quad (3)$$

Defining

$$\begin{aligned} d_i(x) &:= c_i(x)/(r - \mu_i), \\ a_{ij} &:= \lambda_{ij}/(r - \mu_i),\ j \neq i \\ a_{ii} &:= -\sum_{j \neq i} \lambda_{ij}/(r - \mu_i), \end{aligned}$$

we can rewrite this as

$$v_i'(x) = d_i(x) + \sum_{j \neq i} a_{ij} v_j(x) + a_{ii} v_i(x). \quad (4)$$

Or, with A the matrix with elements a_{ij}, in vector-matrix notation

$$v'(x) = d(x) + Av(x). \quad (5)$$

The equations hold for all i and all $0 < x < K$ but also on a part of the boundary. For the components $i \in L$ the boundary $x = K$ can be included, and for $i \in H$ the boundary $x = 0$. This whole region will be called the interior. For the components $i \in L$ in $x = 0$ and for $i \in H$ in $x = K$ boundary conditions will be formulated later on.

Note that the cost rate "time" does not depend on i and x at all. For given i the cost rate "loss" is equal to 0, thus independent of x except for $i \in H$ in the boundary points $x = K$. So the functions $d_i(x)$ are all independent of x on the interior and the set of differential equations simplifies to

$$v'(x) = d + Av(x). \tag{6}$$

Since the matrix A is a scaled Markov generator it has row sums equal to 0. Hence e , the column vector $(1, 1, ..., 1)$, is an eigenvector of A for eigenvalue 0. We assume that A has N independent eigenvectors u_ℓ with corresponding eigenvalues ρ_ℓ with (for convenience) $\rho_1 = 0$ and $u_1 = e$. If we write d as

$$d = \sum_\ell \gamma_\ell u_\ell, \tag{7}$$

then the set of solutions of (6) can be written as

$$v(x) = \sum_{\ell=1}^{N} D_\ell u_\ell e^{\rho_\ell x} + f(x), \tag{8}$$

with

$$f(x) = u_1 \gamma_1 x - \sum_{\ell=2}^{N} u_\ell \gamma_\ell / \rho_\ell. \tag{9}$$

The N unknown constants D_l have to follow from the boundary conditions. We will define $v_N(K) := 0$, since we use this expression at the end of the cycle and not at the beginning of it. The boundary conditions are

$$v_i(0) = \frac{1}{\sum_{k \neq i} \lambda_{ik}} \left(c_i(0) + \sum_{j \neq i} \lambda_{ij} v_j(0) \right), \; i \in L, \tag{10}$$

$$v_i(K) = \frac{1}{\sum_{k \neq i} \lambda_{ik}} \left(c_i(K) + \sum_{j \neq i} \lambda_{ij} v_j(K) \right), \; i \in H, \; i \neq N \tag{11}$$

and

$$v_N(K) = 0. \tag{12}$$

So in order to obtain the loss fraction we have to solve a number of subproblems:

- Obtain the eigenvalues and eigenvectors of the matrix A.

For both cost functions "time" and "loss":

- Express the vector d in the eigenvectors (i.e. find the N coefficients γ_i , see (7)).
- Obtain the N constants D_l from the boundary conditions (10), (11) and (12).

Finally the total expected costs during one cycle, denoted by C , can be obtained from

$$C = \frac{1}{\sum_{k \neq N} \lambda_{Nk}} \left(c_N(K) + \sum_{j \neq N} \lambda_{Nj} v_j(K) \right). \tag{13}$$

Writing C_T and C_L for the total expected cycle time and total expected loss in a cycle, the loss fraction α is obtained from

$$\alpha = \frac{C_L}{C_T \sum \pi_i \mu_i}. \tag{14}$$

2.3 Related literature

In the context of communication systems, the fluid flow model was introduced by Anick, Mitra and Sondhi [1]. They considered a system with an infinite buffer in which the arrival process consists of a superposition of identical on-off sources with exponential on-times and off-times. Later on, Kosten [4] extended the arrival process to a general "Markov-driven" source. In the work of Mitra [5], the model of Anick, Mitra and Sondhi is extended in the way that also the service process can be in different states (and hence you can have different service rates). Furthermore, both the finite and the infinite buffer case are considered.

We have chosen to use the renewal approach of Wijngaard in contrary to the above mentioned references, which use a more direct approach to calculate the distribution of the buffer contents and, in the finite buffer case, the cell loss.

3 Numerical results

In this section we will show some numerical results for the cell loss of a traffic shaper in the case of an on-off arrival process using the fluid flow model of the previous section.

3.1 Variability of the arrival process

First, we are interested in the influence of the variability of the arrival process on the cell loss probability in the shaper. The arrival process consists of successive off-periods and on-periods. During an off-period the arrival rate equals 0 and during an on-period the arrival rate equals 1, i.e. every time unit one cell arrives. The length of the off-periods is exponentially distributed, with mean τ_{off}. For the on-periods, we keep the mean, τ_{on}, fixed. However, we vary the squared coefficient of variation, c^2. In our numerical example, we choose $\tau_{on} = 30$ and $\tau_{off} = 170$ and the shaper rate equals 0.5, i.e. every two time units a cell leaves the buffer (if the buffer is non-empty). Hence the total load of the traffic offered to the shaper is equal to 0.30. For the duration of the on-periods, we choose different distributions depending on the value of c^2.

In the case that $c^2 = 1$ we choose an exponential distribution. The arrival process consists of two states. The arrival rates satisfy $\mu_1 = 0$ and $\mu_2 = 1$. The transition rates are chosen as $\lambda_{21} = (\tau_{on})^{-1}$ and $\lambda_{12} = (\tau_{off})^{-1}$.

When $c^2 < 1$ we choose a Coxian-2 distribution for the duration of the on-periods. The arrival process consists of three states. The arrival rates satisfy $\mu_1 = 0$ and $\mu_2 = \mu_3 = 1$. The transition rates λ_{13}, λ_{21} and λ_{32} are 0. The other transition rates are chosen as:

$$\lambda_{12} = \frac{1}{\tau_{off}}, \ \lambda_{23} = \frac{\alpha + 1}{\alpha \tau_{on}}, \ \lambda_{31} = \frac{\alpha + 1}{\tau_{on}},$$

where

$$\alpha = \frac{c^2 + \sqrt{2c^2 - 1}}{1 - c^2}.$$

Remark that this only works in the case $c^2 \geq 1/2$.

Finally, when $c^2 > 1$ we choose a hyper-exponential distribution with balanced means for the duration of the on-periods (see e.g. page 399 of Tijms [7]). Once again, the arrival process consists of three states. The arrival rates satisfy $\mu_1 = 0$ and $\mu_2 = \mu_3 = 1$. The transition rates λ_{23} and λ_{32} are equal to 0. The other transition rates are chosen as:

$$\lambda_{12} = \frac{q}{\tau_{off}}, \ \lambda_{13} = \frac{(1-q)}{\tau_{off}}, \lambda_{21} = \frac{2q}{\tau_{on}}, \ \lambda_{31} = \frac{2(1-q)}{\tau_{on}},$$

where

$$q = \frac{1}{2}(1 + \sqrt{\frac{c^2 - 1}{c^2 + 1}}).$$

Comparison of the cell loss probabilities for these distributions of the length of the on-periods gives us insight into the impact of c^2 on the cell loss.

Figure 2 depicts the behaviour of the cell loss for a range of buffer values K ($30 \leq K \leq 270$).

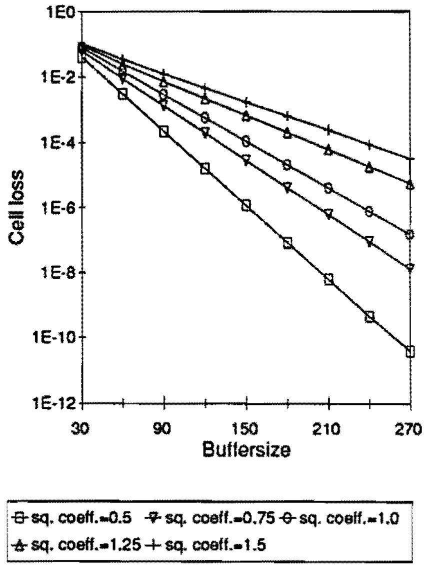

Figure 2: Cell loss as function of the buffer size for various values of c^2 for a system with load=0.30, shaper rate=0.5, $\tau_{on} = 30$ and $\tau_{off} = 170$.

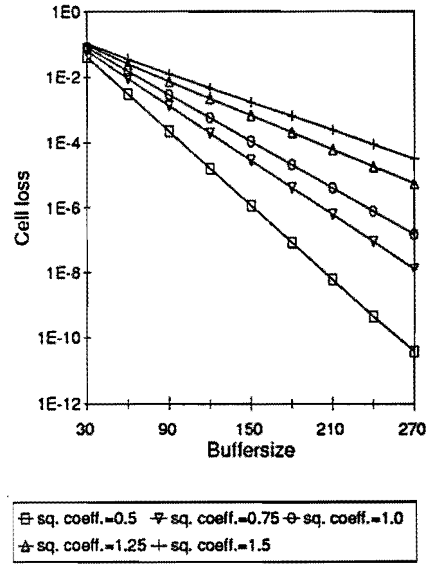

Figure 3: Cell loss as function of buffer size for various values of q for a system with load=0.30, shaper rate=0.5, $\tau_{on} = 30$, $\tau_{off} = 170$ and $c^2 = 1.5$. (BM = Balanced Means)

Note that the buffer size in Figure 2 is given in multiples of 30, the number of arriving cells during the on-period. In practical applications the length of the on-period corresponds to the number of cells in a packet. From Figure 2 we conclude that the cell loss decreases exponentially when the buffer size increases (cf. Mitra [5]). Furthermore we conclude that the squared coefficient of variation of the on-periods has a large impact on the cell loss. For example, when the required (maximum) cell loss equals 10^{-5} the buffer sizes needed for the different values of c^2 are 130 for $c^2 = 0.5$, 200 for $c^2 = 1.0$ and 330 for $c^2 = 1.5$.

So far, we used the balanced means fit for the hyper-exponential distributed on-periods. However, note that in general we have one degree of freedom in fitting a hyper-exponential distribution on the two quantities τ_{on} and c^2. In Figure 3 we show the influence of this degree of freedom on the loss probabilities by varying the parameter q. Note that also λ_{21} and λ_{31} need to be adapted in order to satisfy the fit on τ_{on} and c^2.

We conclude that the effect of q on the cell loss is considerable. Remark that in our numerical example we could choose q arbitrarily between 0 and 1 (of course, by symmetry this reduces essentially to q between $\frac{1}{2}$ and 1). However, in general, there can be some restrictions on the possible values of q.

So far, we considered on-periods with exponentially distributed tails. Next, we want to investigate the influence of the tail of the on-periods on the cell loss probabilities. This is important because in practice data packets have a maximum length and hence the length of on-periods will be bounded. Therefore, we compare two distributions for the on-periods: an exponential distribution with mean τ_{on} and a Bernoulli (two-point) distribution, i.e. with probability p the on-period equals 1 and with probability $1-p$ the on-period equals a. The parameters p and a are chosen such that the mean and squared coefficient of variation of the on-periods are τ_{on} and 1, respectively, i.e.

$$p = \frac{\tau_{on}^2}{2\tau_{on}(\tau_{on}-1)+1}, \quad a = \frac{2p \pm \sqrt{\frac{p}{1-p}}}{2p-1}.$$

The cell loss of the system with the above described distribution for the on-periods is determined by simulation. Figure 4 depicts the results of the simulation in comparison with the analytical results for the system with exponentially distributed on-periods. As expected we see that in the case of exponentially distributed on-periods the cell loss is considerably higher.

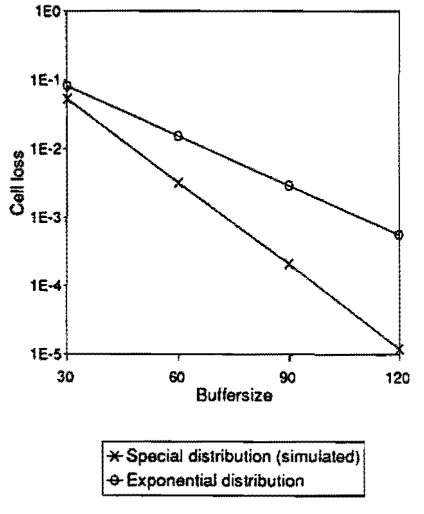

Figure 4: Influence of tail of on-periods on cell loss for a system with load=0.30, shaper rate=0.5, $\tau_{on}=30$, $\tau_{off}=170$ and $c^2=1$.

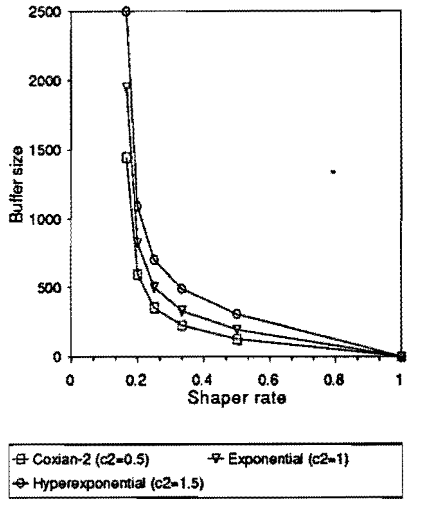

Figure 5: Buffersize versus shaper rate for a system with $\tau_{on}=30$, $\tau_{off}=170$ and cell loss probability 10^{-5}.

3.2 Buffer size versus shaper rate

In this subsection we come back to the question raised in the introduction how the VPC rate should be chosen such that, on one hand the cell loss probability does not exceed a predefined value and, on the other hand the cell delay remains acceptable. For several shaper rates we have determined the buffer size needed to obtain a given cell loss probability. Clearly, the size of a buffer guarantees a maximal cell delay. A buffer should not be too large, otherwise this maximal cell delay becomes unacceptable.

Figure 5 depicts the relation between the buffer size and the shaper rate for a cell loss probability of 10^{-5} and three different arrival processes. These arrival processes are on-off processes with $\tau_{on}=30$, $\tau_{off}=170$ and $c^2=0.5$, 1 and 1.5, respectively.

Remark that a shaper rate smaller than 0.15 is not allowed because otherwise the load of the shaper exceeds one. The required buffer size has been calculated for shaper rates equal to 1, $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{5}$ and $\frac{1}{6}$ (this is because the shaper operates in discrete time). We conclude that the required buffer size grows very rapidly when the shaper rate decreases below a certain value. This effect is enhanced in the case of high variability of the length of the on-periods.

3.3 Discrete-time MMBP versus continuous-time fluid flow

The main difference between the shaper and the fluid flow model is that the shaper operates in discrete time, i.e. time is divided into slots, while the fluid flow model operates in continuous time. In this section, we compare the cell loss of our fluid flow model with the cell loss in a discrete time model with arrivals according to a two-state Markov modulated Bernoulli process (MMBP). In an MMBP the arrival probability in a time slot depends on the state of an underlying Markov chain, i.e. when the chain is in state i the arrival probability equals p_i, $i = 1, 2$. For a more precise description of an MMBP and the analysis of an MMBP/D/1/K queue, we refer to Guillemin, Boyer and Dupuis [3].

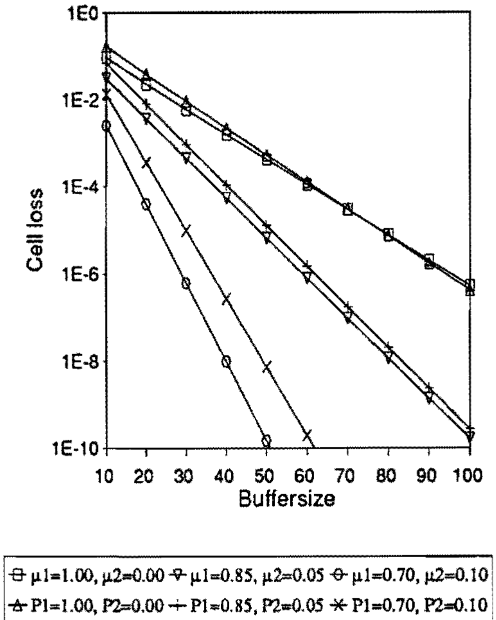

Figure 6: Comparison of fluid flow model with MMBP model for a system with load=0.5, shaper rate=0.5, mean length phase 1 = 10, mean length phase 2 = 30.

For three different cases the results of the comparison are shown in Figure 6. Of course, the system parameters have been chosen such that mean time in state i, shaper rate and load are the same in the fluid flow model and the MMBP model. We see that in the case of on-off arrival processes (i.e. $p_1 = 0$, $p_2 = 1$, resp. $\mu_1 = 0$, $\mu_2 = 1$) the behaviour of the cell loss is identical in both models. However, when the arrival rates differ from 1 resp. 0, the cell loss probabilities do not coincide in the two models. This is caused by the extra cell level randomness of the arrival process in the MMBP model. In the MMBP model there is, except from randomness in the length of the periods in the different states, also randomness in the number of arrivals during a period (this in contrary to the fluid flow model). Furthermore, in the MMBP model there is also randomness in the positions where cells arrive within a period in a certain state.

It might be expected that during burst periods traffic streams resemble CBR streams. Hence the cell level randomness is limited, even for a superposition of several of these streams. Therefore we consider the fluid flow model as a practically useful model.

4 Conclusions

In this paper we use a fluid flow queueing model to study the behaviour of an ATM traffic shaper. Aim of the paper is to get insight into the relation between buffer capacity, shaper rate, cell loss and cell arrival process.

Our main conclusion is that in order to make sensible predictions about the cell loss at a shaper, detailed information is needed about the traffic characteristics of the arrival process. For example, in the case of on-off arrival processes with exponentially distributed off periods, we have seen that not only the mean and the coefficient of variation of the length of the on-periods play an important role, but also other quantities like e.g. higher moments and the tail of the distribution.

Concerning the relation between the shaper rate and the buffer capacity we have seen that the buffer capacity needed to obtain a predefined cell loss probability grows very rapidly when the shaper rate decreases below a certain point. Hence, when we choose to hire a small part of the bandwidth we have to take into account that the cell delay becomes very large.

Finally, in order to study the influence of cell level randomness (i.e. randomness both in the number and in the position of cells within a period) on the cell loss probability, we have compared results obtained by a fluid flow model and a discrete time two-state MMBP/D/1/K model. We have seen that the cell loss probabilities for these models may differ considerably, depending on the cell rates during the two states of the arrival process. Due to the, probably, limited cell level randomness of the streams offered to the traffic shaper we expect that the behaviour of the cell loss probabilities is well predicted by the fluid flow model.

References

1. Anick, D., Mitra, D. and Sondhi, M.M. (1982). Stochastic theory of a data-handling system with multiple sources. Bell System Tech. J. 61, pp 1871-1894.
2. Gravey, A., Louvion, J.-R. and Boyer, B. (1990). On the Geo/D/1 and Geo/D/1/n queues. Performance Evaluation 11, 117-125.
3. Guillemin, F., Boyer, J. and Dupuis, A. (1992). Burstiness in broadband integrated networks. Performance Evaluation 15, pp 163-176.
4. Kosten, L. (1986). Liquid models for a type of information storage problems. Delft Progress Report 11, pp 71-86.
5. Mitra, D. (1988). Stochastic theory of a fluid model of producers and consumers coupled by a buffer. Adv. Appl. Prob. 20, pp 646-676.
6. Roberts, J.W. (ed.) (1992). COST 224 : Performance evaluation and design of multiservice networks, EUR 14152 EN.
7. Tijms, H.C. (1986). Stochastic Modelling and Analysis: A Computational Approach. John Wiley & Sons, Chichester.
8. Wijngaard, J. (1979). The effect of interstage buffer storage on the output of two unreliable production units in series, with different production rates. AIIE Trans. 11, pp 42-47.