



Load control in atm networks

Jacobsen, S.B.; Moth, Klaus; Dittmann, Lars; Sallberg, K.

Published in:
Proceedings of the 8th international Switching Symposium

Publication date:
1990

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Jacobsen, S. B., Moth, K., Dittmann, L., & Sallberg, K. (1990). Load control in atm networks. In *Proceedings of the 8th international Switching Symposium* (Vol. Volume 5, pp. 131-138). IEEE.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LOAD CONTROL IN ATM NETWORKS

S. B. Jacobsen, K. Moth, L. Dittmann
Technical University of Denmark

K. Sällberg
ELLEMTTEL, S-223 70 Lund, Sweden

ABSTRACT

This paper considers preventive load control in ATM networks. A one link performance evaluation based on on/off traffic sources is carried out. Approximative models are derived, and a specific admission control algorithm based on one of the models is presented and the obtained transfer efficiency is derived. A Flow Enforcement function to supervise each connection is suggested and the correspondence between modelling parameters, flow enforcement parameters and parameters used by the admission control algorithm is considered.

1. INTRODUCTION

In a future B-ISDN based on the Asynchronous Transfer Mode (ATM), a large variety of services with different characteristics, such as bandwidth demand, burstiness and quality of service (QOS) requirements will be integrated. This implies that new challenges will face the network designers. This contribution will address the problems related to load control in ATM networks and approximative models concerning performance will be elaborated in detail.

ATM is considered as a packet-oriented switching and multiplexing technique, which potentially offers full bandwidth flexibility at a high transmission rate and a higher transfer efficiency than a Synchronous Transfer Mode (STM) through statistical multiplexing. The gain in bandwidth utilisation and - more important - achieved bandwidth flexibility will cost in a more complicated Load Control mechanism.

Load control may be partitioned into two categories:

- 1) reactive load control by the use of backwards signaling during connection phase to reduce or even terminate individual connections during intervals of high load on the actual part of the network. As explained in [1] the use of reactive load control in high speed ATM network introduces major problems.
- 2) preventive load control (PLC) guarantees a certain Quality of Service (QOS) during the duration of a connection by admission or blocking of new connection requests. A key function of the PLC is the Flow Enforcement (FE) function which sometimes is denoted the police function.

In general it is possible to employ both categories in the same network. However, this paper considers preventive load control, only.

At connection set-up phase, the requested connection is specified by a set of parameters (denoted Flow Enforcement Parameter, FEP), and on each link on the route through the ATM network a decision is taken whether or not to accept the requested connection. This local decision is taken by the algorithm called Admission Control (AC). The criterion of admission is

based on sets of parameters describing the aggregated traffic on the link and a set of parameters describing the requested new connection. It should be noticed that no steps are taken to use the momentary load state on the links. The estimated load state only changes when a new connection is established or disconnected. Moreover, the concept fully relies on the FEPs declared during connection set-up - there will be no changes in AC criteria whether or not the users utilise the allocated bandwidth provided through the declared FEPs. This idea makes the AC independent of actual connection behaviour such that it can be based solely on the declared FEPs.

In order to maintain a certain QOS within the network, the FE function is needed to supervise that the characteristics declared during set-up are not violated. The function bounds the characteristics of a connection and, thereby, prevents that the AC underestimates the load state of the link.

The most well-known FE algorithm is the Leaky Bucket [2], but lately a number of alternative algorithms such as the Running Window and Exponentially Weighted Moving Average have been introduced.

A second type of function is proposed in [3] to be used by the user of a connection to adapt to the characteristics declared at connection set-up. The function is called Flow Throttling and further details can be found in [3].

In the following section a background is given and the assumptions for the following analysis of the link performance are motivated. The objective of the analysis is to get a better understanding of the QOS and transfer efficiency in an ATM network, but also to show the principle of AC and evaluate the resulting performance for some traffic mixes. In section 3, a number of approximative models are presented and evaluated. Section 4 explains and evaluates the AC and FE concept in more detail by applying the results obtained in section 3.

2. NETWORK ASPECTS

The approach to be taken in the rest of the paper is illustrated in fig. 1. A terminal generates a certain traffic, described by an equivalent model with a parameter set (m,p) . At the entrance of the network the cell stream is supervised by the FE with parameter set (D_m, D_p) . There is a one-to-one correspondence between the (m,p) and (D_m, D_p) , which will be illustrated in section 4.

2.1 Basic Assumptions on Traffic

Three different levels of traffic analysis has been identified when dealing with performance analysis in an integrated broadband communication environment. That is the cell level, the burst level and the connection (call) level, see [4]. The time scale differs significantly among the three levels - see fig. 2.

At the cell level the individual cells and the structures of the switches are taken into account and queueing theory is used to obtain performance data. The analysis is based on the assumption

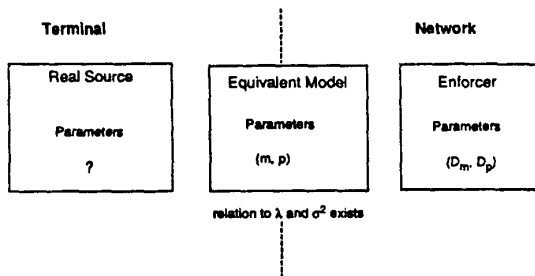


Figure 1 A terminal generates a certain traffic, described by the equivalent model with parameter set (m, p) . The traffic is enforced based on the FEP (D_m, D_p)

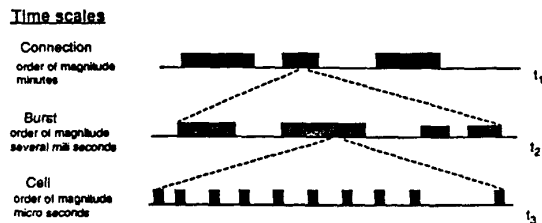


Figure 2 Illustration of the different traffic levels and the associated time scales

that the number of bursts is fixed. [4], [5] and [6] discusses performance at the cell level.

At the burst level the number of established connections is assumed fixed, but the cell rate is varying due to internal variations in each source. In [4] the burst scale is investigated using the theory of large deviations while [7] uses an approximation based on a multi-server loss system.

At the connection level the number of established connections is varying. [8],[9] and [10] calculates the blocking probability in the connection scale in different contexts.

In this paper only the burst level and the cell level is considered.

2.1.1 Source models

Various kinds of traffic ranging from CBO services to bursty or variable bitrate services is expected to be offered to the ATM-network.

An important requirement for a source model is therefore versatility. The model should include both smooth and bursty traffic. However, it is desirable if an analytical solution of the corresponding queueing model exists such that the quality of service can be derived.

A source model which is versatile and for which an analytical solution exists, is the on/off model. In this model a source can be in two states, an on-state in which the source is transmitting cells with a fixed rate p , and an off-state in which the source is silent. The duration of the on-state (off-state) is exponentially distributed with mean t_{on} (t_{off}).

This model has successfully been used to describe voice traffic with silence removal, see [11] and [12], and in [13] a video source has been modelled as a number of independent on/off sources. Si and Mark has in [14] suggested the corresponding discrete on/off model as the basis for source modelling in an integrated network.

2.1.2 Multiple concatenated links

Models for evaluation of the quality of service exist in the case of a single link. An example is the on/off model (see section 3). However, the original characteristics of the cell stream will change as the cells of a given connection passes through a number of links due to multiplexing of cells.

The difficult issue of describing this change in characteristics is beyond the scope of this paper. Our analysis therefore concerns a single link performance evaluation, and is based on a single cell multiplexer with a queue served according to the FCFS discipline.

Furthermore it is expected that the preventive load control approach will ensure small delays in the buffers. This implies that the cell streams will only experience minor changes in characteristics when passing a link.

2.2 Quality of service (QOS)

In an integrated broadband network, different services will have different requirements of QOS. Voice traffic has restrictive delay requirement but can accept a moderate cell loss ratio. High speed data might be very sensitive to cell loss but very delay-tolerant. For services using coding techniques in which cell loss induces loss of synchronisation, the absolute time between cell losses might be more important than the loss ratio itself.

To take into account all kinds of different QOS requirements is beyond the scope of this paper. We have chosen to use the cell loss ratio as a measure for the QOS. The models applied in later sections can also estimate delay fragiles such that the approach taken is applicable in this case too.

Full integration of services implies that the most restrictive QOS requirements must be applied to all services. In our calculations we have therefore chosen a restrictive cell loss ratio of 10^{-9} .

2.3 Transfer efficiency

Primarily, the cost of a network and the cost of using it is due to several causes. One parent of cost is managing the network - that is management of connection requests, maintenance, managed flexibility etc. Another parent is the needed network capacity: the networks must be dimensioned in order to ensure a certain QOS. The needed network capacity is determined by the user demand and the bandwidth utilisation which the network can demonstrate. The transfer efficiency is the average utilisation with respect to bandwidth of a single ATM link - the maximum of this is a guide-line of what may be expected from an ATM network. The link transfer efficiency will be different on different links, but this section only refers to a qualitative average for describing the issue.

First, there is a number of well known contributions to transfer efficiency reduction:

- ATM cell overhead (approximately: 10%)
- synchronisation cells for providing framed or periodic ATM. For example, this amounts to a 1% reduction if each 135th cell is used for synchronisation

In this paper the effective payload is chosen to 135 Mbit/s in order to compensate for the efficiency deficiencies mentioned above.

Now some others aspects of transfer efficiency related to ATM are:

- admission control. The AC only allocates a certain percentage of the link bandwidth to ensure a limitation of the

cell loss probability at ATM cell multiplexers. Examples of achieved transfer efficiency are given in section 4

- network architecture options like service segregation, virtual paths (networks) tend to decrease the total available bandwidth due to a reduced effect from statistical multiplexing
- in future customer premises networks (CPN), some sources may not utilise their allocated bandwidth provided through their declared FEPs. A margin may be needed in order to limit cell loss in the FE function

When only single link performance has been evaluated for AC there will be a diminished fidelity to the FEP's of a connection after passage of multiple ATM links. For CBO services this requires a margin between declared FEP's and the FEP's used for AC at all links on the route. For bursty services the declared FEP may be used.

3. LINK PERFORMANCE EVALUATION

In this section four models to evaluate the cell loss ratio are presented. First a fluid flow queueing model which calculates loss rates for a number of different on/off sources is presented. This model serves as a reference model. Then the arrival process is approximated by a number of identical on/off sources and loss formulas are found by the fluid flow model. Finally two approximate models are presented. They both approximate the actual queueing system with a corresponding multi-server loss system. The multi-server model which agrees best with the reference model is then used to obtain a simple AC algorithm.

The traffic input is in all cases modelled as independent on/off sources as described in section 2.1.1. The sources are partitioned into a number of classes, all sources in a given class being statistically identical.

Input parameters are:

c: the number of classes

The traffic of class i ($i=1, \dots, c$) is described by

N_i : the number of sources in class i

p_i : the constant arrival rate in the on-state

$t_{on}(i)$: the average duration of the on-state

a_i : the fraction of time in which a source is in the on-state

B: the buffersize

p_{out} : the transfer rate of the output link

3.1 Reference Model: Heterogeneous Fluid Flow

In [15] and [16] a fluid flow queueing model for heterogeneous on/off traffic is presented and in [16] formulas for time congestion and loss rates are given. A short description is given below.

The input to the buffer is modelled as a fluid running into a reservoir (the buffer) and the departure from the buffer is the fluid running out of a hole in the bottom of the reservoir.

The assumption about exponentially distributed on- and off durations implies the equilibrium buffer distribution can be found as solution to a set of first order differential equations. With the equilibrium buffer distribution known, the loss ratio is easily derived.

The number of equations in the set of differential equations is:

$$N := (N_1 + 1)(N_2 + 1) \dots (N_c + 1) \quad (3.1)$$

As the expression for N indicates, then this model requires considerable computation time, and it is not feasible to use it directly in the AC algorithm.

3.2 One-step Approximation: Homogeneous Fluid Flow

In [7] it is suggested to approximate an arbitrary arrival process by a number of identical on/off sources by matching the following three statistical parameters

- 1) The mean arrival rate (λ)
- 2) The variance of the arrival rate (σ^2), and
- 3) An asymptotic variance parameter v defined by:

$$v = \lim_{t \rightarrow \infty} \frac{\text{Var}(N(0,t))}{t} \quad (3.2)$$

where $N(0,t)$ is the number of cells arriving in the time interval $(0,t)$, and $\text{Var}(N(0,t))$ is the variance of this random variable.

The advantage of using these parameters is that they are additive, that is $\lambda_{1+2} = \lambda_1 + \lambda_2$, $\sigma_{1+2}^2 = \sigma_1^2 + \sigma_2^2$, and $v_{1+2} = v_1 + v_2$.

Since a homogeneous on/off stream is described by the four parameters N, a, p , and t_{on} , there is one degree of freedom when matching λ, σ^2, v . We have chosen the following additional constraint:

N = the number of established connections on the link

When the arrival stream is on/off traffic appearing in c classes with parameter set as given in the start of this section, this yields

$$N = \sum_{i=1}^c N_i \quad (3.3)$$

Matching λ, σ^2, v for the homogeneous and heterogeneous stream yields:

$$N a p = \lambda = \sum_{i=1}^c N_i a_i p_i \quad (3.4)$$

$$N a (1-a) p^2 = \sigma^2 = \sum_{i=1}^c N_i a_i (1-a_i) p_i^2 \quad (3.5)$$

$$2N a (1-a)^2 p^2 t_{on} = v = \sum_{i=1}^c 2N_i a_i (1-a_i) p_i^2 t_{on}(i) \quad (3.6)$$

These formulas are proven in [7] and references in there.

The cell loss ratio for the homogeneous approximation is obtained by applying the fluid flow model. Section 3.5 evaluates this model.

3.3 Two-Step Approximation: Binomial Multi-Server Approximation

This approximation is to be viewed as an approximation of the homogeneous on/off stream, and therefore it is a two-step approximation of the original heterogeneous stream.

Consider a homogeneous on/off stream with parameter set (N, a, p, t_{on}) which is offered to the ATM multiplexer.

Let X denote the number of sources which at an arbitrary time is in the on-state. X is binomially distributed with parameters N and a , that is:

$$P(X=i) = \binom{N}{i} a^i (1-a)^{N-i} \quad (3.7)$$

where P denotes the probability and $K(N,a)$ is the usual binomial coefficients.

To obtain a simple formula for the loss ratio we approximate the queuing system with a corresponding multi server system. The idea is that each source in the on-state will occupy a server. First we introduce the following assumption:

The average number of cells arriving from a specific source during an on-state is much larger than the buffer capacity, that is $p_{on} \gg B$.

Put

$$n = \left\lfloor \frac{p_{out}}{p} \right\rfloor \quad \text{where } \lfloor x \rfloor \text{ denotes the integer part of } x \quad (3.8)$$

$n = p_{out}/p$ where n is the integer part of p_{out}/p .

that is n is the largest number of sources which can be in the on-state at the same time without introducing congestion.

The system can now be viewed as a multi server system with n servers. With the above assumption cell losses will occur as soon as the number of sources in the on-state exceeds n . If the actual on-duration is small or the peak rate is small such that the buffer is able to absorb several bursts this model will overestimate the cell loss.

In traditional loss systems, the traditional Engset formula will give the blocking probability. However, since we are approximating a queuing system, we use the formula

$$P_{loss} = \frac{\sum_{i=n+1}^N p(i-n) \binom{N}{i} a^i (1-a)^{N-i}}{N a p} \quad (3.9)$$

The term $p(i-n)$ is the rate at which cells are lost in state i , and the numerator therefore expresses the rate at which cells are lost. The denominator is the offered traffic. [7] uses a similar approach to find the loss ratio.

In section 3.5 numerical examples are used to evaluate this model.

3.4 Pascal Multi-Server Approximation

Heterogeneous traffic offered to a multi-server loss system has been investigated intensively in the literature, see [8], [10], [17], [18] and [19]. In this section we apply a technique, which has been used for calculating the probability distribution of the number of occupied servers in multi-server systems, to obtain approximate loss rates for the queuing problem.

The link capacity p_{out} is divided into n basic bandwidth units (BBU) each of size b , such that $p_{out} = nb$. The link is now considered as a n server loss system, implying that the buffer is of negligible size. Each source in class i demands $d_i = p_i/b$ BBU's in the on-state. It is assumed that p_i is an integer multiple of b .

In our case with heterogeneous on/off traffic the mean M and variance V of the arrival intensities is:

$$M = \sum_{i=1}^K N_i a_i d_i \quad (3.10)$$

$$V = \sum_{i=1}^K N_i a_i (1-a_i) d_i^2 \quad (3.11)$$

According to [8], [10], [17] and [19] it has been noticed that the steady state distribution of the number of occupied servers $Q(i)$ is approximately invariant for different offered traffic with fixed mean and variance. An accurate candidate for $Q(i)$ can be obtained from a Pascal (negative binomial) arrival process implying the following recursion formula for $Q(i)$

$$iQ(i) = (\alpha + (i-1)\beta)Q(i-1) \quad (3.12)$$

in which the relation between α , β , M , and V is:

$$\alpha = \frac{M^2}{V} \quad (3.13)$$

and

$$\beta = 1 - \frac{M}{V} \quad (3.14)$$

Using this way of matching the original queuing system with the above loss system, we obtain the following formula for loss ratio of class i traffic,

$$P_{loss}(i) = \frac{\sum_{k=i+1}^n (i-k)Q(i)}{M} \quad (3.15)$$

where $k=n-d_i$.

The overall loss ratio is then the weighted mean of the individual loss ratios.

Since this model also uses an approximation to a loss system, it is expected that it will overestimate the cell loss when the on-duration or peak rate is small. Section 3.5 compares this model with those given in section 3.1, 3.2 and 3.3.

3.5 Numerical Examples

Several calculations have been carried out to compare the above models. The following parameters have been chosen fixed:

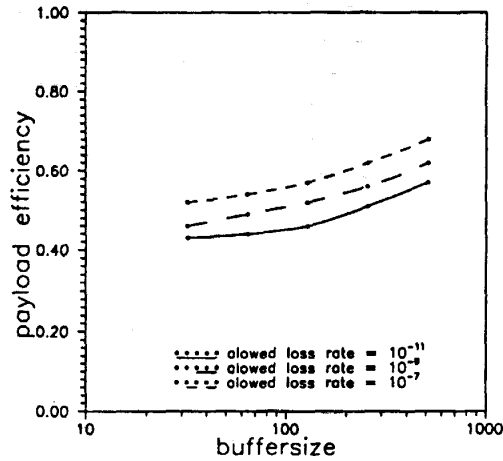
$p_{out}=135$ Mbit/s (motivated in section 2.3)

Buffersize $B=88$

Small buffers has the advantage of introducing only small delay, small delay jitter, and small changes in the characteristics of the cell streams. The disadvantage is a potentially small network efficiency due to potential cell loss. However, our calculations shows that the increase in network efficiency which can be obtained by increasing the buffersize is small. Fig. 3 shows the obtained transfer efficiency (based on the reference model) as a function of the buffersize.

In fig. 4 the cell loss probability is illustrated as a function of the traffic mix. The traffic appears in two classes, a class with 2 Mbit/s sources, and a class with 10 Mbit/s sources. As the figure illustrates then all four models shows a significant increase in loss ratio when the proportion of 10 Mbit/s traffic is increased with a constant load of 0.52. The homogeneous approximation described in section 3.2 estimates the loss rate accurately, while the multi server approximations both overestimates the loss rate.

In fig. 5 the loss ratio from fifty 2 Mbit/s and ten 10 Mbit/s sources offered to the multiplexer is illustrated as a function of the duration of the on-state for 10 Mbit/s sources. When the on-duration increases there is a dramatic increase in loss ratio. The homogeneous approximation only captures a minor part of this behaviour but for moderate and large durations it is in agreement with the reference model. The multi server approximations are

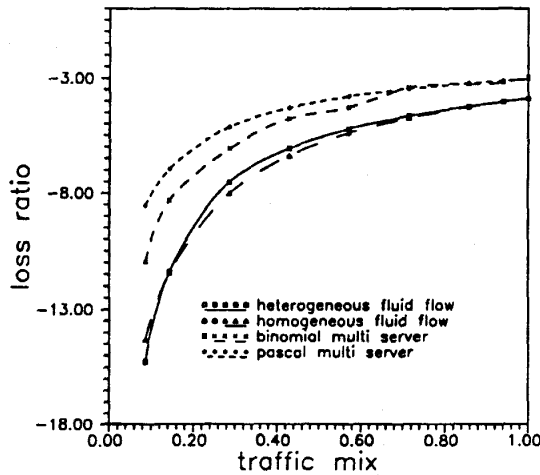


Class 1
 Peak rate = 2 Mbit/s
 Mean rate = 1 Mbit/s
 $t_{on} = 25$ ms
 Class 2
 Peak rate = 10 Mbit/s
 Mean rate = 2 Mbit/s
 $t_{on} = 5$ ms

$P_{tot} = 135$ Mbit/s
 = class 1 sources = $5x$
 = class 2 sources = x

Fig 3 The maximum payload efficiency as a function of the buffersize for three acceptable call loss ratios (based on the reference model)

Figure 3



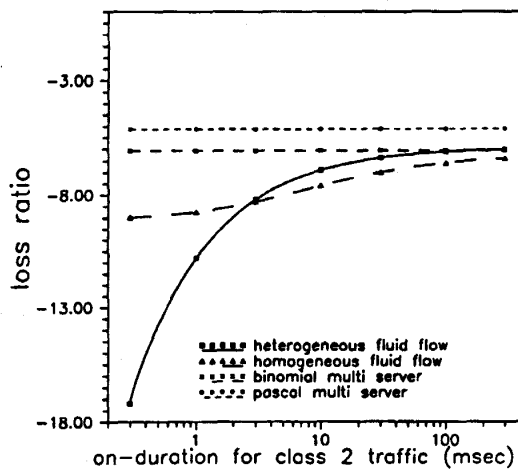
Class 1
 Peak rate = 2 Mbit/s
 Mean rate = 1 Mbit/s
 $t_{on} = 25$ ms

Class 2
 Peak rate = 10 Mbit/s
 Mean rate = 2 Mbit/s
 $t_{on} = 5$ ms

$P_{tot} = 135$ Mbit/s
 buffersize = 88
 load = 0.52

Fig. 4 The loss ratio as function of the traffic mix t where t is the fraction class 2 traffic to overall traffic

Figure 4



Class 1
 Peak rate = 2 Mbit/s
 Mean rate = 1 Mbit/s
 $t_{on} = 25$ ms

Class 2
 Peak rate = 10 Mbit/s
 Mean rate = 2 Mbit/s

$P_{tot} = 135$ Mbit/s
 buffersize = 88
 load = 0.52

class 1 sources = 50
 # class 2 sources = 10

Fig. 5 The loss ratio as a function of the on-duration for class 2 traffic

Figure 5

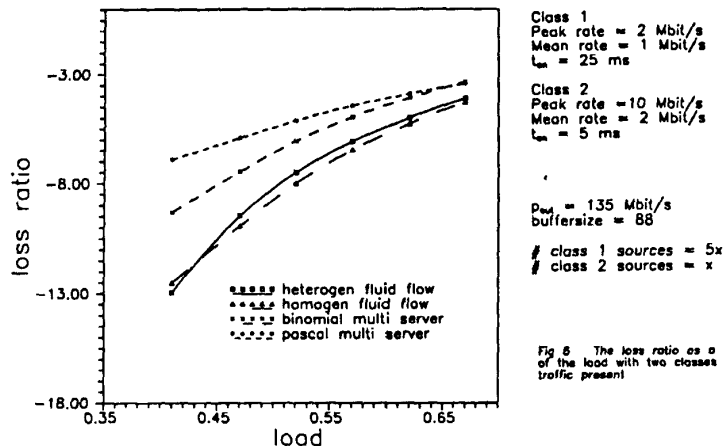


Figure 6

due to the approximation of a loss system independent of the on-state duration. As expected, then the loss rate of the homogeneous model converges to the binomial loss rate for large on-durations.

In fig. 6 the loss ratio from a fixed mix of 2 Mbit/s and 10 Mbit/s sources offered to the multiplexer is illustrated as a function of the load. The increase in loss ratio based on the fluid flow models is dramatic in the range 0.35 to 0.45 and for higher load the curves flattens out. The homogeneous approximation accurately estimates the loss ratio except for low load while the multi-server approximations overestimate the loss ratio, the binomial approximation being the most accurate, as in the previous cases.

Fig. 7 illustrates the same as fig. 6 with the exception that one 30 Mbit/s bursty source is included in the traffic stream.

From the figures the following conclusions can be made:

- Bursty sources with high bitrates introduces a much higher cell loss ratio than less bursty sources with moderate bitrates.

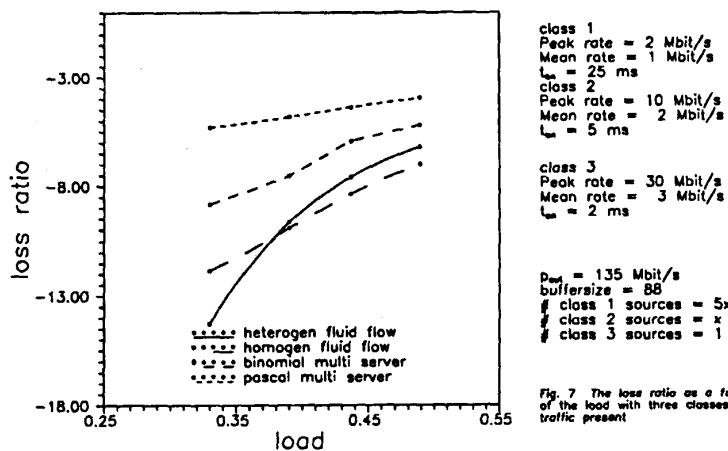


Figure 7

- The on-duration is, especially for bursty sources with high bitrates, a very important parameter
- The homogeneous approximation is over a wide range of system parameters sufficiently accurate.
- The binomial and pascal multi server approximation both overestimates the loss ratio, the binomial multi server model being the most accurate.

4. ADMISSION CONTROL AND FLOW ENFORCEMENT

The term admission control refers to the local decision whether or not to accept a new connection on a specific link. The criterion is based on:

- 1) A set of parameters for each of the existing connections on the link.
- 2) A corresponding set of parameters for the new connection.

This parameter set must be in close accordance with the FEP's and the parameters used to model the traffic sources, see fig. 1. In this paper where all sources are modelled as on/off sources it is natural to use the following three parameters:

mean rate m , peak rate p , and peak duration t_{on} .

The mean and peak rates are simple and accepted parameters while the peak duration, which contains information of the correlation in the source is more complex

If one of the multi-server models are to be used in the AC algorithm, only the mean and peak rate are to be used.

There are two requirements to the admission control criterion:

- 1) It must be simple to reduce call set-up time.
- 2) It should be accurate or at least not underestimate the actual load on the link.

In practice there are at least two different ways of adopting a AC criterion.

- 1) If one of the performance model is an analytical model in which approximations can be made to obtain a simple formula for the cell loss ratio, this is one way of adopting a AC criterion.
- 2) On the basis of numerous evaluation examples it might be possible empirically to obtain a simple and accurate AC criterion.

In the next two sections the first approach is used and evaluated.

4.1 An Admission Control Criterion Based on the Appr. Model

Since the AC criterion have to be simple it is not feasible to use any of the fluid flow models, and therefore the most accurate of the multi-server approximations must be used. From section 3.5 it is seen to be the binomial approximation. We therefore base the AC criterion on this model.

Consider a link serving n virtual connections $1, \dots, n$, where the parameters for connection i is (p_i, m_i) m_i being the mean bitrate. Put

$$a = \frac{\lambda}{p}, \quad \lambda = \sum_{i=1}^n N_i a_i p_i, \quad \sigma^2 = \sum_{i=1}^n N_i a_i (1 - a_i) p_i^2 \quad (4.1)$$

The parameter set (N, a, p) for the aggregate stream is:

$$N = n \quad (4.2)$$

$$a = \frac{\lambda}{\lambda + N \sigma^2} \quad (4.3)$$

$$p = \frac{\lambda}{N a} \quad (4.4)$$

(4.3) and (4.4) due to (3.5) and (3.6).

4.1.1 An algorithm

A request for a new connection with parameter set (p_{n+1}, m_{n+1}) is made. Then the admission control algorithm is as follows:

(The algorithm)

```

BEGIN
   $a_{n+1} := \frac{m_{n+1}}{p_{n+1}}$ ;
   $\lambda' := \lambda + m_{n+1}$ ;
   $\sigma'^2 := \sigma^2 + m_{n+1}(1 - a_{n+1})p_{n+1}^2$ ;
   $N' := N + 1$ ;
   $a' := \frac{\lambda'}{\lambda' + N'\sigma'^2}$ ;
   $p' := \frac{\lambda'}{N'a'}$ ;
   $n' := \text{trunc}\left(\frac{p_{max}}{p'}\right)$ ;
   $P_{loss} := \frac{\sum_{i=n'+1}^n (i-n) \binom{N}{i} a^i (1-a)^{N-i}}{Na}$ ;
  IF  $P_{loss} < \epsilon$  THEN
  BEGIN
    Accept connection;
     $\lambda := \lambda'$ ;
     $\sigma^2 := \sigma'^2$ ;
     $a := a'$ ;
     $N := N'$ ;
     $p := p'$ ;
  END
  ELSE reject connection;
END
  
```

ϵ denotes the maximum acceptable cell loss ratio.

4.2 Evaluation Examples

The maximum link utilisation depends on the traffic on the link and the required QOS. In this section we have chosen a cell loss ratio $\epsilon = 10^{-9}$, motivated in section 2.2.

Two examples of traffic have been considered. In example one two classes of traffic are present, a class of sources with mean and peak rate of respectively 1 Mbit/s and 2 Mbit/s, and a class of sources with mean and peak rate of 2 Mbit/s, 10 Mbit/s respectively. The number of sources in class one is kept 5 times the number of sources in class 2. Table 1 states the maximum obtained transfer efficiency when peak allocation of bandwidth is applied, when the admission criterion given in section 4.1 is applied, and with the heterogeneous fluid flow model.

	Peak allocation of bandwidth	Admission control algorithm based on section 4.2	Admission control based on the reference model
Payload efficiency	0.31	0.42	0.49
Gross rate efficiency	0.27	0.36	0.43

Two classes of traffic is present:

class 1: 5x sources, $p = 2$ Mbit/s, $a = 0.5$, $t_{on} = 25$ ms
 class 2: x sources, $p = 10$ Mbit/s, $a = 0.2$, $t_{on} = 5$ ms

Table 1 Comparison of the transfer efficiency obtained for three different admission control algorithm with a mix of two different classes of traffic

Example two differs from example one in only one respect, one bursty 30 Mbit/s (mean rate 3 Mbit/s) source is added to the traffic stream. Table two states the results obtained.

With the above mentioned traffic mixes the obtained transfer efficiency is rather low, and the gain obtained from statistical multiplexing, based on algorithm 4.1.1, is not remarkable. With the considerations made in section 2.3 the real efficiency might be even lower. However, with other traffic mixes the figures will be different.

4.3 Flow Enforcement Parameters (FEP)

As it was seen in the previous two sections then the actions taken by the AC algorithm are based on the peak bitrate and the mean bitrate of each of the connections. From the evaluation

	Peak allocation of bandwidth	Admission control algorithm based on section 4.2	Admission control based on the reference model
Payload efficiency	0.28	0.33	0.41
Gross rate efficiency	0.24	0.29	0.36

Three classes of traffic is present:

class 1: 5x sources, $p = 2$ Mbit/s, $a=0.5$, $t_{on}=25$ ms
class 2: x sources, $p = 10$ Mbit/s, $a=0.2$, $t_{on}=5$ ms
class 3: 1 source, $p = 30$ Mbit/s, $a=0.1$, $t_{on}=2$ ms

Table 2

examples in chapter 3 it is seen that violation might lead to unacceptable QOS for all connections.

It is therefore mandatory to introduce a flow enforcement function which supervises the peak bitrate and mean bitrate of the connections.

A flow enforcement function to perform this could be implemented as two leaky buckets, one supervising the peak bitrate and one supervising the mean bitrate.

Three parameters, the increment factor I, the decrement factor D, and the threshold value T, characterizes the leaky bucket algorithm. The relation between decrement factor D_p (D_m), increment factor I, and peak rate p (mean rate m) is:

$$D_p = \frac{pI}{P_{out}}, \quad D_m = \frac{mI}{P_{out}} \quad (4.5)$$

where P_{out} is the payload of the link.

The threshold value T_p (T_m) can be viewed as a time constant. In case of peak rate monitoring it is chosen low, and in case of mean rate monitoring it is chosen high such that the source is allowed to exceed the mean bitrate for a considerable time.

5. CONCLUSION

In this paper preventive load control in an ATM network has been investigated. An admission control criterion, based solely on the parameter set for each established connection, declared during connection set up has been presented. The criterion has been evaluated by a fluid flow queueing model, and the obtained transfer efficiency is derived. A flow enforcement algorithm to supervise each individual connection is presented.

ACKNOWLEDGEMENT

The paper is a result of the work carried out in the performance group within the RACE ATMOSPHERIC project, and we would like to thank J.H.Dejean for valuable comments, and the project for financial support.

REFERENCES

- [1] G.M. Woodruff, R.G.H. Rogers, P.S. Richards; *A Congestion Control Framework for High-Speed Integrated Packetized Transport*, Globecom '88, Florida, Session 7.1, pp. 203-207
- [2] J.S. Turner; *New Directions in Communications (or Which Way to the Information Age?)*, IEEE Communications Magazine, Vol 24, no.10, Oct. 86.

[3] K. Sällberg, B. Stawenow, I. Andersen; *A Resource Allocation Framework for B-ISDN*, ISS '90, Stockholm, Sweden.

[4] J.Y. Hui; *Resource Allocation for Broadband Networks*, ICC 88, session 31.6, pp 1004-1011

[5] B. Eklundh, K. Sällberg, B. Stavenow; *Asynchronous Transfer Modes - Options and Characteristics*, 12th International Teletraffic Congress, Torino, Italy, June 1988.

[6] J. W. Roberts, J.T. Virtamo; *The Superposition of Periodic Cell Arrival Streams in an ATM Multiplexer*, submitted to the IEEE Trans. Commun.

[7] K. Lindberger; *Traffic Analysis of Statistical Multiplexing in ATM-networks*, Contribution to COST 224 and 8th Nordic Teletraffic Seminar, Otnas, Finland, Aug 1989.

[8] J.L. Lutton, J.W. Roberts; *Traffic Performance of Multi-slot Call Routing Strategies in an Integrated Service Digital Network*, ISS '84 Florence, Session 22.6

[9] V.B. Iversen; *The Exact Evaluation of Multi-Service Loss Systems with Access Control*, Teleteknik, English ed. 1987, no. 2, pp 56-61

[10] L.E.N. Delbrouck; *A unified Approximate Evaluation of Congestion Functions for Smooth and Peak Traffic*, IEEE Trans. Commun., Vol. Com-29, pp 85-91, Feb. 81

[11] J.N. Daigle, J.D. Langford; *Models for Analysis of Packet Voice Communications Systems*, IEEE JSAC, Vol SAC-4, no. 6, Sept. 86, pp 847-855

[12] R.C.F. Tucker; *Accurate Method for Analysis of a Packet-Speech Multiplexer with Limited Delay*, IEEE Trans. Commun., Vol Com-36, no. 4, April 88, pp 479-483

[13] B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson, J.D. Robbins; *Performance Models of Statistical Multiplexing in Packet Video Communications*, IEEE Trans. Commun., Vol Com-36, no. 7, July 88, pp 834-844

[14] S. Li, J.W. Mark; *Traffic Characterization for Integrated Services*, INFOCOM 88, Session 8A.3, pp 781-790

[15] L. Kosten; *Stochastic Theory of Data-Handling Systems with Groups of Multiple Sources*, Int. Symp. on Performance on Computer Comm. Systems, IFIP, Zürich, March 1984

[16] S. Jacobsen, L. Dittmann, K. Moth; *A Fluid Flow Queueing Model for Heterogeneous on/off Traffic*, 8th Nordic Teletraffic Seminar, Otnas, Finland, Aug 1989.

[17] J.W. Roberts; *Traffic Models for the Telecom 1 Integrated Services Network*, ITC-10 Montreal, session 1.1

[18] M.E. Bershai; *The Poissonian Spectrum Method for Treating a Loss System Serving Non-Poissonian Multi-Bit-Rate Traffic*, INFOCOM 89, pp. 1010-1019

[19] R.I. Wilkinson; *Theories for Toll Traffic Engineering in the USA*, Bell Syst. Tech. J., Vol 35, no 2, pp.421-514, 1956