Contribution to : 7th International Teletraffic Congress Seminar on
"Broadband Technologies: Architectures, Applications, Control, and Performance"
Morristown, NJ, October 9 – 11, 1990

Policing Mechanisms for ATM Networks – Modelling and Performance Comparison

Erwin P. Rathgeb[1]

Institute of Communications Switching and Data Technics, University of Stuttgart

Abstract

ATM networks, as proposed by CCITT as the solution for the future BISDN, provide a high flexibility with respect to varying bandwidth requirements for different services as well as with respect to the momentary bitrate within a connection. For the network operator this results in a need to control the individual connections by means of a so called "policing" or "usage parameter control" function to ensure an acceptable quality of service for all existing connections. In this paper basic design objectives and requirements for such a function are described. These objectives serve as a basis for the comparison of some of the mechanisms proposed so far, namely the "Leaky Bucket", the "Jumping Window", the "Moving Window" and the "Exponentially Weighted Moving Average" mechanisms.

1 Introduction

ATM networks have been proposed by CCITT as the solution for the future BISDN [1] and a complete set of recommendations is currently under study [2]. In an ATM network all information is packetized and transferred in small, fixed size blocks called cells using the virtual connection concept. The statistical multiplexing of cells belonging to different virtual connections will provide the users with the possibility to have a bitrate varying in a wide range during a connection according to their needs (bitrate on demand). Flexible User/Network Interfaces (UNIs) in the range of 150 and 600 Mbit/s will allow connections ranging from low to very high bitrates to support services like voice, interactive video and high speed data.

The increased flexibility with respect to the bandwidth requirements of the calls makes necessary extended signalling procedures for the call establishment. A "traffic contract" including the relevant service attributes characterizing the connection, like ,e.g., maximum and mean cell rate, and the quality of service (QOS) requirements of the connection with respect to cell loss, delay and delay jitter, has to be negotiated with the network at call setup. A connection acceptance control (CAC) function [12, 6, 19] then has to decide whether the new connection can be accepted or not. This decision has to be made based on the knowledge about the traffic characteristics and QOS requirements of the new connection and of the already existing connections sharing the same network resources. The CAC function relies on the traffic parameters negotiated during call establishment and therefore these parameters have to be enforced to ensure proper functioning.

2 The Policing Function in ATM Networks

Basic characteristics of ATM networks are the provision of broadband user access interfaces, packet-oriented information transfer without flow control between the user and the network and the use of the statistical multiplexing principle. Due to these properties, a virtual connection can in principle exceed the negotiated traffic parameters up to the maximum capacity of the UNI and a new network function called "usage parameter control" or "policing" function as defined in CCITT recommendation I.311 [3] is needed. This function has to restrict the behaviour of the traffic source to the characteristics negotiated in the contract in order to protect the network against excessive congestion resulting in a degradation of the quality of service for all connections sharing the same network resources.

In order to protect all network resources, the policing function has to be located as close as possible to the actual traffic source. Nevertheless, it has to remain under the direct control of the network

---

[1] The author is now with Bellcore, Morristown N.J.

provider. Depending on the customer access network configuration, the policing function may be performed on virtual circuits (VCs), on virtual paths (VPs, logical grouping of several VCs) or on the total traffic volume on an access link within components like concentrators, local exchanges and ATM cross-connects (c.f. [3]).

Violations of the traffic contract may result from equipment malfunctioning, from attempts to intentionally degrade the quality of service or – more likely – may be motivated by possible economical or operational advantages.

To protect the network and the coexisting connections, action has to be taken by the policing function after detecting a violation of the contract parameters. The most obvious action is to discard all cells for which a violation of the traffic contract has been detected [3]. However, it has to be recognized that the set of policing parameters proposed by CCITT in recommendation I.311, namely average cell rate, peak cell rate and duration of the peak, is not sufficient to completely describe ATM traffic sources. Furthermore, not all of these characteristics may be known at call setup with the required accuracy and some of them may be modified before the cells reach the policing function, e.g., due to jitter introduced in the CPN. This results in a certain probability that the policing function makes a wrong decision and discards cells from a source which is in accordance with the traffic contract. As the policing function is a network function, the resulting cell loss contributes to the network performance and has to be kept in the range which is also accepted for the other network components like multiplexers and switching nodes. To achieve this objective, a margin has to be foreseen between the negotiated and the actually policed traffic parameters. This margin has to be taken into account when dimensioning the CAC function, which should in this case be based on the worst case traffic allowed by the policing function rather than on the negotiated traffic parameters.

Another policing action is to mark the violating cells for preferred deletion within the network in case of a congestion [3, 7, 9].

An additional task of the policing function is to verify that valid address fields are used. However, this aspect of the policing function will not be considered in the following.

3 Description and Modelling of Policing Mechanisms

Requirements for Policing Mechanisms The policing function must be available during the whole connection and has to operate in real-time. This requires that the mechanism used is fast, simple and cost effective to implement in hardware.

Ideally, the policing mechanism should be transparent for connections which respect the parameters defined in the traffic contract and take no policing actions on their cells. It should, on the other hand, act on all cells violating the contract to limit the behaviour of the source in the intended way. From these requirements an ideal limit curve for the policing function can be derived (see Section 4).

In addition to these requirements, the dynamic reaction time of the mechanism to parameter violations has to be short to avoid flooding of the relatively small buffers in the network. To meet these rather conflicting requirements, several policing mechanisms have been proposed so far. Besides the "Leaky Bucket" mechanism [7, 10, 13, 16, 23] several other mechanisms which are described ,e.g., in [5, 4, 17] will be discussed and compared in the following sections.

Traffic Source Models Especially if the policing function is performed on a single VC, the assumptions about the characteristics of the traffic source have a significant influence on the results. Therefore the 2-phase Burst/Silence model has been used for the comparison of the mechanisms. This source is a fairly realistic model, e.g., for packetized voice [11], still picture [8] and interactive data services. It allows to vary the relevant parameters, namely maximum cell rate, mean cell rate and mean peak duration, independent of each other while still allowing an analytical evaluation for some of the mechanisms. The model can be characterized by the number of cells per burst (geometric, mean $E[X]$), the minimum inter cell time Δ and the silence duration (negative-exponential, mean $E[S]$). With $\alpha^{-1} = E[X] \times \Delta$ and $\beta^{-1} = E[S]$ the mean cell rate λ is defined as $\lambda = 1/\Delta + \frac{\alpha\Delta}{\beta}$.

To show the sensitivity of the results to the assumptions about the traffic sources, the widely used Poisson input process as well as the autoregressive Markovian model used in [14] to describe the cell stream of a variable bit rate video codec will be used. Since the model does not take into account the variation of the bitrate during a single frame – which is an important factor for a network with finite buffers – two extreme cases will be used in Section 4:

1. The cells are sent equidistantly according to the actual bitrate of the frame. This is the smoothest traffic that can be generated (coefficient of variation of about 2).
2. The cells are sent equidistantly according to the maximum bitrate of the codec, which implies a pattern with a burst and a silence period during one frame. This variant yields a highly variable output stream with a coefficient of variation of about 10.6 for the cell interarrival times.

The actual intra-frame behaviour will be somewhere in between both extremes, but the first, desirable variant can be realized by providing sufficient buffering in the codec.

The Leaky Bucket Mechanism (LB) The LB mechanism consists of a counter which is incremented by one each time a cell is generated by the source and decremented in fixed intervals as long as the counter value is positive. If the momentary cell arrival rate exceeds the decrementation rate, the counter value starts to increase. It is assumed that the source has exceeded the admissible parameter range, if the counter reaches a predefined limit and suitable actions (e.g. discard or mark cells) are taken on all subsequent cells until the counter has fallen below its limit again. This mechanism is easy to implement using one counter for the system state, one counter for the decrementation interval and a variable for the counter limit.

To evaluate the probability for the detection of a parameter violation – which is identical to the cell loss probability if violating cells are discarded – the G/D/1-s delay loss system can be used as an exact model for the Leaky Bucket mechanism. The service time of the model is chosen to be equal to the decrementation interval and the number of customers in the system (including server and queue) directly represents the state of the counter. Therefore, the counter limit N is equal to $(s+1)$.

For the evaluation of the system with a general, discrete time renewal input process, the iterative analysis approach based on the workload in the system at consecutive arrival instants [21] can be used in principle. For the stationary case and moderate system sizes, a direct solution of the linear equation system for the workload distribution at arrival instants ($u(k)$) is more efficient because the iterative algorithm converges relatively slow in the region of interest. The linear equation system is defined by

$$u(k) = \sum_{i=0}^{sD} u(i)a(D-k+i) + \sum_{i=sD+1}^{(s+1)D} u(i)a(i-k) \quad 0 < k \leq (s+1)D \tag{1}$$

The last, additional equation is given by the normalization condition. All system characteristics can be computed from this distribution of unfinished work at arrival instants, which is identical to the waiting time distribution for FIFO queues. In particular, the probability for a violation p_{viol} can be computed as

$$p_{viol} = \sum_{i=sD+1}^{(s+1)D} u(i) \tag{2}$$

It should be noted that the G/D/1-s queue is only a valid model for the violation probability of the LB mechanism. No cell is actually queued and the resulting cell stream entering the network is not the output process of the queueing system.

To apply the solution described above, which is valid for general discrete time renewal processes, the 2-phase source model has to be adapted by discretizing the negative-exponential probability density function of the silence phase. The discretization error has no significant influence on the results if the discretization steps are dimensioned properly.

An approximate solution for the loss probability of the stationary system with the 2-phase source model can also be obtained using the "Uniform Arrival and Service" model [10, 22].

The Jumping Window Mechanism (JW) With the JW mechanism, the maximum number of cells accepted from a source within a fixed time interval (window) is limited to a maximum number (N). The next interval starts immediately at the end of the preceeding interval (jumping window) and the associated counter is started again with an initial value of zero. Therefore, the time interval during which a specific cell is influencing the counter value varies from zero to the window width. The implementation complexity of this mechanism is comparable to the complexity of the LB mechanism.

The probability that policing actions have to be taken on a cell can be computed using the counting process, which is characterizing the number of arriving cells in an arbitrary time interval.

For general renewal arrival processes, the probability distribution for the counting process can in principle be obtained using renewal theory. The probability distribution $x_n(t)$ for n arrivals in an interval of length t results in

$$x_n(t) = A^{(n)}(t) - A^{(n+1)}(t) = \begin{cases} 1 - A^*(t) & n = 0 \\ \frac{1}{\lambda}\left(a^*(t) \star a^*(t)\right) & n = 1 \\ x_{n-1}(t) \star a(t) & n \geq 2 \end{cases} \tag{3}$$

where A^* and $a^*(t)$ denote the probability distribution function (PDF) and the probability density function (pdf) of the forward recurrence time for the arrival process, respectively. $A^{(i)}$ denotes the PDF for the time until the n-th cell arrival starting at an arbitrary point in time and $a(t)$ denotes the pdf for the interarrival times. The probability that a violation is detected by the JW mechanism is given by

$$p_{viol} = \frac{\sum_{i=1}^{\infty} i\, x_{N+i}(T)}{\sum_{i=1}^{\infty} i\, x_i(T)} = \frac{\lambda T - N + \sum_{i=0}^{N-1} (N-i)\, x_i(T)}{\lambda T} \tag{4}$$

The Laplace transform for the distribution of the counting process for the 2-phase source is given by

$$x_{LT,n}(s) = \frac{\lambda}{s^2}\left[a_{LT}(s)^{n-1} - 2a_{LT}(s)^n + a_{LT}(s)^{n+1}\right] \qquad \text{with} \qquad a_{LT}(s) = \left((1-\alpha\Delta) + \frac{\alpha\Delta\beta}{s+\beta}\right) e^{-s\Delta} \tag{5}$$

and the distribution in the time domain can be evaluated applying the inverse transform ($\mathcal{LT}^{-1}$)

$$\begin{aligned} \mathcal{LT}^{-1}\left\{\frac{a_{LT}(s)^n}{s^2}\right\} = & \left[\sum_{k=0}^{n} \binom{n}{k} (1-\alpha\Delta)^{n-k} (\alpha\Delta)^k \times \right. \\ & \left. \left(t - n\Delta - \frac{1}{\beta}\sum_{i=1}^{k}\left(1 - e^{-\beta(t-n\Delta)}\left(\sum_{j=0}^{i-1}\frac{(\beta(t-n\Delta))^j}{j!}\right)\right)\right)\right] \sigma(t-n\Delta) \end{aligned} \tag{6}$$

Using the correspondence equation

$$\mathcal{LT}\left\{\sum_{i=1}^{\infty} i\, x_{N+i}(t)\right\} = \frac{\lambda}{s^2}\left[a_{LT}(s)\right]^N \tag{7}$$

where $\mathcal{LT}$ denotes the Laplace transform, equation 4 can be expressed as

$$p_{viol} = \frac{1}{T}\mathcal{LT}^{-1}\left\{\frac{a_{LT}(s)^N}{s^2}\right\}_{t=T} \tag{8}$$

With the analogon of equation 3 in the discrete time case, the distribution for the counting process can be evaluated for general distributions. The convolutions can be substituted by Fast Fourier Transforms

(FFTs) and a complex multiplication of the distribution vectors in the frequency domain, which makes the algorithm more effective.

The Exponentially Weighted Moving Average Mechanism (EWMA) The EWMA mechanism uses fixed, consecutive time windows like the JW mechanism. The difference is, that the maximum number of accepted cells in the i-th window (N_i) is a function of the allowed mean number of cells per interval N and an exponentially weighted sum of the number of accepted cells in the preceeding intervals (X_i) according to the rule

$$N_i = \frac{N - \gamma S_{i-1}}{1-\gamma}, \qquad 0 \leq \gamma < 1 \qquad \text{with} \qquad S_{i-1} = (1-\gamma)X_{i-1} + \gamma S_{i-2} \tag{9}$$

which can also be expressed as

$$N_i = \frac{N - (1-\gamma)(\gamma X_{i-1} + \cdots + \gamma^{i-1} X_1) - \gamma^{i+1} S_0}{1-\gamma} \tag{10}$$

S_0 is the initial value of the EWMA measurement. The factor γ controls the flexibility of the algorithm with respect to the burstiness of the traffic. If $\gamma = 0$, N_i is constant and the algorithm is identical to a JW mechanism. A value of $\gamma > 0$ allows more bursty traffic. Although the computation of N_i can be made rather effective for special values of γ, the implementation complexity of this mechanism is slightly higher than for the previous mechanisms. The EWMA mechanism will be evaluated using event-by-event simulation techniques.

The Moving Window Mechanism (MW) Similar to the JW mechanism, the maximum number of cell arrivals within a given time interval T is limited by this mechanism. The difference is, that each cell is memorized for exactly one window width thus strictly limiting the number of accepted cells in any possible interval of length T to a maximum of N. Therefore, the arrival time of each cell is stored and a counter is incremented by one for each arrival. Exactly T time units after the arrival of an accepted cell the counter is decremented by one again. This mechanism can be interpreted as a window, which is steadily moving along the time axis. This mechanism requires that the arrival times of up to N cells are stored. The complexity is therefore directly related to the counter limit and is considerably higher than for the other mechanisms even for relatively low counter limits. Furthermore, the other mechanisms where the counter limit only influences the required range for the counter allow a much more flexible dimensioning.

The MW mechanism can be modelled by a G/D/n–0 system, where the deterministic service times reflect the window width T and the number of servers is defined by the maximum allowed number N of cells in the interval. For the 2-phase source model, the mechanism will be evaluated by simulation.

4 Comparison of the Mechanisms

Definitions for the Comparison For the Jumping Window, the EWMA and the Moving Window mechanism, the ratio of the maximum accepted number of cells per interval N and the window width T gives the long term average cell rate that is controlled by the mechanisms. For the Leaky Bucket mechanism, the controlled average cell rate is given by the reciprocal of the decrementation interval D. With a choice of

$$T = \frac{N}{\lambda C} = \frac{N}{\lambda_p} \qquad \text{and} \qquad D = \frac{1}{\lambda C} = \frac{1}{\lambda_p} \tag{11}$$

all mechanisms are dimensioned to control the same mean cell rate λ_p. C is an overdimensioning factor describing the ratio of λ_p to the mean cell rate of the source (λ).

Non-deterministic traffic sources will violate the policing criterion with a certain probability because of their short term statistical fluctuations even if they respect the long term average. This probability can be decreased for a given policed cell rate λ_p by increasing the counter limit N for the LB mechanism.

The same applies to the window mechanisms if the ratio of N and the window width T is kept constant. For the EWMA mechanism also the factor γ can be increased for this purpose. On the other hand, the incrementation of N (and γ) also increases the reaction time of the mechanisms.

Another method to decrease the violation probability is to overdimension the mechanisms ($C > 1$). This will, however, decrease the ability of the mechanisms to detect real, long term parameter violations.

Statistical Sources It is obvious that for counter limits in the range of the buffer sizes within the network or above, the peak cell rate has to be controlled in any case to prevent flooding of these buffers. This can be achieved using a separate policing mechanism with appropriate dimensioning (see e.g. [10, 18]). Therefore, in this section the dimensioning aspects for mean cell rate policing are discussed under the assumption that the sources do not violate the negotiated peak cell rate.

Figure 1 shows the influence of the counter limit N on the violation probability assuming arrivals according to the 2-phase source with the parameters used in [11] for packetized voice ($E[X] = 22$, $E[S] = 650\ ms$, $\Delta = 16\ ms$). The mechanisms have been dimensioned to control the exact mean cell rate of the source ($C = 1$, $\lambda = \lambda_p$), 0.8 has been chosen for the factor γ of the EWMA algorithm. It is obvious, that N has to be prohibitively high for all mechanisms to limit the violation probability for a well behaving source to values in the range of 10^{-9}, which would be required if violating cells would be discarded. Even if service individual cell loss probabilities up to 10^{-5} would be acceptable for some loss insensitive services, it would not make sense to operate the mechanisms with mean rate dimensioning (the LB mechanism e.g. would require a counter limit well beyond 10^5).

Figure 2 shows the influence of N for the same source if a factor $C = 2.277$ is chosen resulting in a λ_p relatively close to the maximum cell rate of the source ($1/\Delta$). Violation probabilities in the range below 10^{-9} can be achieved using a counter limit of 100 for the LB mechanism. The EWMA mechanism shows a similar behaviour. To get the same violation probability for the other window mechanisms with a comparable N, λ_p has practically to be set equal to the peak cell rate of the source. For a λ_p close to the mean cell rate ($C = 1.1$), N would have to be about 1800 for the LB mechanism to achieve the same violation probability. This would allow the source to send a burst up to 2933 cells (equivalent to a burst duration of 47 seconds) at the peak cell rate when starting with a counter value of 0.

The dynamic behaviour of the mechanisms is shown in Figure 3. The mechanisms have been dimensioned for the same violation probability of 10^{-2} at the negotiated cell arrival rate using a factor $C = 2.277$. The actual cell rate of the source has been increased to 1.5 times the nominal rate by increasing $E[X]$ to 45.25. The violation probability for the MW mechanism increases rapidly starting from the point where the number of emitted cells (counted from the start of the connection) is higher than the counter limit. This is because the time window is started by the first arriving cell and the number of accepted cells within the first interval of length T after the emission of the first cell is strictly limited to N. The other three mechanisms allow more than N cells during the first T time units after the emission of the first cell, which results in a more gradually increasing violation probability. Despite the small mean counter limit of 10, the EWMA mechanism shows the worst dynamic behaviour, because the actual counter limit starts with a value of 50 in the first interval and it takes some window periods until the overload can be detected.

The effects caused by a variation of the long term actual cell rate are shown in Figure 4 for the LB and for the JW mechanism. To be able to dimension both mechanisms to a violation probability of about 3×10^{-10} at the nominal cell rate, the source characteristics have been changed to $E[X] = 5$ and $E[S] = 147.72\ ms$ yielding a smoother traffic while keeping the mean and peak cell rate the same. N has been set to 100, the cell rates have been normalized to the nominal cell rate.

The ideal limit curve for the policing mechanisms in Figure 4 is characterized by zero violation probability up to the nominal arrival rate and a violation probability according to the percentage of excess cells beyond that point.

As expected, the LB mechanism with $C = 1.42$ approaches the ideal limit considerably faster than the JW mechanism with $C = 2.71$ indicating a significantly better ability to detect static parameter violations. Since the violation probability of the JW mechanism is always lower than for the MW mechanism assuming the same dimensioning, the MW mechanism is even less sensitive. A source exploiting the full range given by the mechanisms, e.g., a deterministic source (see below), could even exceed the negotiated cell rate up to a factor equal to C without any cell loss.

To achieve the same violation probability as for the LB mechanism using a comparable factor C, the counter limit had to be about 750 for the JW mechanism resulting in very long policing intervals and a bad dynamic behaviour. A MW mechanism with such a counter limit would furthermore be very expensive to implement, because the arrival times of all 750 cells would have to be remembered.

Figure 5 shows the comparison with the LB ($C = 1.42$) and the EWMA mechanism using the same source, $N = 45$ and a nominal violation probability of 2×10^{-5} (due to the limitations of the simulation). The EWMA mechanism with the same factor C and $\gamma = 0.8$ shows a slightly better behaviour for moderate overload situations, whereas the EWMA mechanism with $C = 1.27$ and $\gamma = 0.91$ detects static overload significantly better. The drawback of these relatively high γ values is that the mechanisms react considerably slower to parameter violations.

The sensitivity of the policing mechanisms to the source assumptions is depicted in Figure 6 using the LB mechanism with $C = 1.42$ as an example. It is obvious that the frequently used Poisson – and also the geometric – arrival processes are far too optimistic. Only peak rate policing is possible for the 2-phase source model with the parameters for a still picture service ($E[X] = 1953.125$, $E[S] = 11s$, $\Delta = 256\mu s$ [10]) which has a peak bit rate of 2 Mbit/s and a mean bit rate of 86.96 kbit/s. This results mainly from the large number of cells in a burst and from the fact, that the counter limit has to be significantly higher than $E[X]$ to allow policing close to the mean rate. With a factor $C = 1.1$ for example, N would have to be about 400,000 to obtain a violation probability in the range of 10^{-10}, which is not realistic.

The sampling time of the LB mechanism is also not sufficient for the video sources and a significantly higher N would be required to be able to police close to the mean cell rate. Figure 6 also shows the influence of the intra frame behaviour of the source. The more bursty type is more difficult to police, but the difference decreases with an increasing counter limit.

Worst Case Sources The first requirement that a worst case traffic source has to satisfy is, that the policed cell rate λ_p is always fully exploited. This requirement can be fulfilled by a deterministic source. However, the resulting traffic is very smooth and can therefore be multiplexed rather efficiently. To decrease the gain from statistical multiplexing, the cells have to be clustered in bursts as much as possible and the resulting worst case traffic will be a periodic burst/silence pattern.

For the MW mechanism this pattern is given by a burst of length N followed by a silence period of length $T - (N-1)\Delta$, where Δ denotes the cell interarrival time during the burst.

If the source synchronizes to the time window in the JW mechanism, it can concatenate two bursts of N cells at the end of one interval and at the beginning of the next. Then a silence period of $2T - (2N-1)\Delta$ is required yielding a period of the pattern of $2T$. The same pattern characterizes the long term worst case traffic of the EWMA mechanism. However, due to the variable counter limit even longer bursts are possible for a limited time.

The worst case pattern for the LB mechanism has a maximum burst length N_{max} which is given by the smallest integer number that satisfies the inequality

$$N_{max} > \frac{N-1}{1-\Delta\lambda_p}, \qquad \Delta < \frac{1}{\lambda_p} \tag{12}$$

The required silence period has a length of $N_{max}(1/\lambda_p - \Delta) + \Delta$. The period of the pattern is N_{max}/λ_p. For the example in Figure 4 with a counter limit of 100 the maximum burst length for the LB mechanism would be 198 cells and thus approximately the same as for the JW mechanism (200 cells). The required silence period following such a burst, however, would be 3184 ms for the LB and only 172 ms for the JW mechanism. In the case of Figure 5 with a counter limit of 45 the burst length would be 89 for the LB and 90 for the EWMA mechanism. This burst has to be followed by a silence period of 1440 ms for the LB versus 1456 ms for the EWMA mechanisms, which indicates that these mechanisms show a comparable static worst case behaviour.

5 Conclusions

The fundamental problem with respect to mean cell rate policing is, that the characteristics of the source have to be estimated on the basis of a relatively short sample, which gives rise to a certain probability for incorrect policing decisions. An extension of the sampling period (or more flexibility due to a large factor γ for the EWMA), on the other hand, also increases the reaction time of the mechanisms. A larger margin between the policed cell rate and the negotiated cell rate reduces the ability of the mechanisms to detect actual long term overload and leads also to a more unfavourable worst case behaviour.

Another problem is caused by the inaccuracies and uncertainties in the knowledge about relevant parameters, like the mean bitrate, in the establishment phase of the call. These effects imply increased margins between the policed bitrate and the actual bitrate and reduce the effectivity of the policing as well as of the call acceptance control function.

The comparison of the policing mechanisms and their dimensioning in this paper taking into account the remaining violation probability for a source in agreement with the contract parameters, the sensitivity to static overload, the dynamic reaction time, the worst case traffic admitted into the network and the implementation complexity shows, that the LB and the EWMA are the most promising mechanisms. The other window mechanisms are not flexible enough to cope with the short term statistical fluctuations of the source traffic. Moreover, the MW mechanism is rather expensive to implement for a realistic parameter dimensioning.

The optimum dimensioning and the effectivity of the mechanisms depends heavily on the characteristics of the traffic sources and their QOS requirements. Smooth, continuous bitstream like traffic streams can be policed much more effective than highly variable and bursty ones. For real-time services with very long bursts, like, e.g., the still picture service, policing close to the mean cell rate requires unrealistically long sampling periods. For this type of service dynamic bandwidth allocation schemes (burst switching protocols) might be useful to avoid peak rate policing and consequently peak rate bandwidth allocation. Video services with variable bitrate show both short term variations within a frame or between consecutive frames and long term variations, e.g., due to scene changes, with time constants of several seconds. Both effects can not be captured sufficiently by a simple policing mechanism and the long term variations would make necessary a very long sampling period. One possible approach for these services might be to design the codecs to restrict their output in a way that it is better suited for policing [15] without significant degradation of the picture quality. For services with less stringent real-time requirements the traffic could also be shaped within the terminal to allow effective policing and multiplexing.

The worst case traffic patterns accepted by the policing mechanisms may differ significantly from the patterns expected on the basis of the traffic contract due to large counter limits and overdimensioning factors C which may be required. This fact has to be reflected in the dimensioning of the connection acceptance control function in order to be able to ensure a sufficient quality of service.

The performance of the policing function can probably be improved by using a combination of several policing entities with different dimensioning for a single traffic stream or by using more complex mechanisms like, e.g., the Gabarit [4]. It has to be evaluated, however, whether the additional complexity yields an adequate performance improvement.

Acknowledgement This work has been supported by Philips Communications Industry, Nuremberg, FRG and by the Commission of the European Communities under the RACE contract 1022 "Technology for ATD".

References

[1] CCITT: Rec. I.121: "Broadband Aspects of ISDN". *Blue Book, Vol. III.7,* Geneva 1989.

[2] CCITT: Temporary Documents of the Study Group XVIII Meeting in Geneva, January 1990.

[3] CCITT: Draft Recommendation I.311: "B-ISDN General Network Aspects". In [2].

[4] F. DENISSEN, E. DESMET, G. H. PETIT: The Policing Function in an ATM Network. *Proc. of the 1990 Intl. Zurich Seminar on Digital Communications,* pp. 131–144, Zurich, March 1990.

[5] F. VAN DEN DOOL: Policing and Load Control and related Functions. *RACE R1022, Document DNL_311&321_008_CD_CC,* June 1988.

[6] Z. DZIONG, J. CHOQUETTE, K.-Q. LIAO, L. MASON: Admission Control and Routing in ATM Networks. *Proc. of the ITC Specialist Seminar Adelaide 89,* Paper 15.3, Adelaide, September 1989.

[7] A. ECKBERG, D. LUAN, D. LUCANTONI: Bandwidth Management: A Congestion Control Strategy for Broadband Packet Networks – Characterizing the Throughput-Burstiness Filter. *Proc. of the ITC Specialist Seminar Adelaide 89,* Paper 4.4, Adelaide, September 1989.

[8] G. GALLASSI, G. RIGOLIO, P. VACCARI, L. VERRI: Resource Allocation in ATM Networks. *Proc. of the 3rd RACE 1022 Workshop,* Paper 22.6, Paris, October 1989.

[9] G. GALLASSI, G. RIGOLIO, L. FRATTA: ATM: Bandwidth Assignment and Bandwidth Enforcement Policies. *Proceedings of the IEEE Globecom '89,* Paper 49.6, Dallas, November 1989.

[10] E. GARETTI, R. MELEN, A. TONIETTI: Efficiency of the Leaky-Bucket Mechanism in the Framework of ATM Resource Allocation. *Proc. of the 3rd RACE 1022 Workshop,* Paper 21.2, Paris, October 1989.

[11] H. HEFFES, D. LUCANTONI: A Markov Modulated Characterization of Packetized Voice and Data Traffic and Related Statistical Multiplexer Performance. *IEEE Journal on Selected Areas in Communications*, Vol. SAC-4, No. 6, pp. 856-868, September 1986.

[12] J. Y. HUI: Resource Allocation for Broadband Networks. *IEEE Journal on Selected Areas in Communications, Vol. 6, No. 9,* pp. 1598–1608, December 1988.

[13] W. KOWALK, R. LEHNERT: The 'Policing Function' to Control User Access in ATM Networks - Definition and Implementation. *Proceedings of the ISSLS'88*, Boston, September 1988.

[14] B. MAGLARIS, D. ANASTASSIOU, P. SEN, G. KARLSSON, J. D. ROBBINS: Performance Models of Statistical Multiplexing in Packet Video Communications. *IEEE Transactions on Communications, Vol. 36, No. 7,* pp. 834–844, July 1988.

[15] F. PEREIRA, L. MASERA: 'Constant Quality' Coding in ATM Networks: From Videophone to HDTV. *Proceedings of the 3rd RACE 1022 Workshop,* Paper 31.1, Paris, October 1989.

[16] E. P. RATHGEB, T. H. THEIMER: The Policing Function in ATM Networks. *Proceedings of the XIII International Switching Symposium,* Stockholm, June 1990.

[17] E. P. RATHGEB: Comparison of Policing Mechanisms for ATM Networks. *Proceedings of the 3rd RACE 1022 Workshop,* Paper 21.1, Paris, October 1989.

[18] J. ROBERTS: Jitter due to an ATM Multiplex – Application to Peak Rate Policing. *COST 224 – TD(89)035 – August 1989.*

[19] F. C. SCHOUTE: Simple Decision Rules for Acceptance of Mixed Traffic Streams. *Proceedings of the 12th International Teletraffic Congress (ITC),* Paper 4.2A.5, Turino, June 1988.

[20] M. SIDI, W.-Z. LIU, I. CIDON, I. GOPAL: Congestion Control Through Input Rate Regulation. *Proceedings of the IEEE Globecom '89,* Paper 49.2, Dallas, November 1989.

[21] P. TRAN-GIA, H. AHMADI: Analysis of a Discrete-Time $G^{[X]}/D/1-s$ Queueing System with Applications in Packet-Switching Systems. *Proceedings of the INFOCOM'88*, New Orleans, 1988.

[22] R. C. F. TUCKER: Accurate Method for Analysis of a Packet-Speech Multiplexer with Limited Delay. *IEEE Transactions on Communications, Vol. 36, No. 4,* pp. 479–483, April 1988.

[23] J. S. TURNER: New Directions in Communications (or Which Way in the Information Age ?). *Proceedings of the Zurich Seminar on Digital Communications*, pp. 25–32, Zurich, March 1986.

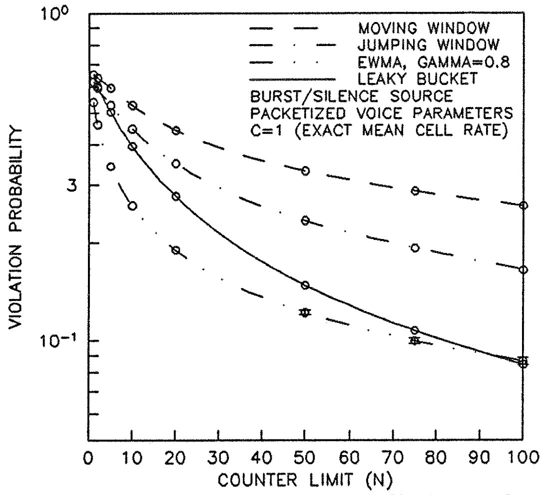

Figure 1: Influence of the counter limit on the violation probability ($C = 1$)

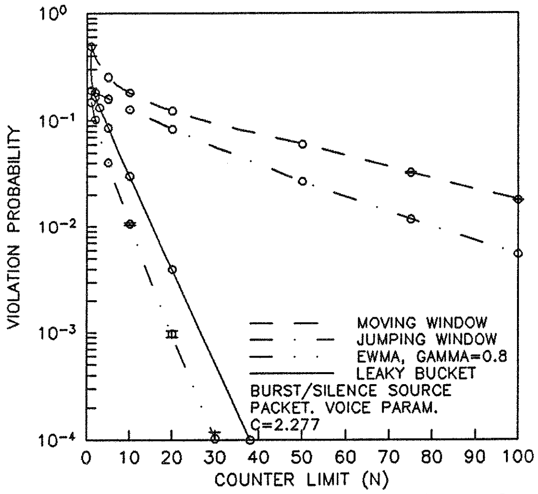

Figure 2: Influence of the counter limit on the violation probability ($C = 2.277$)

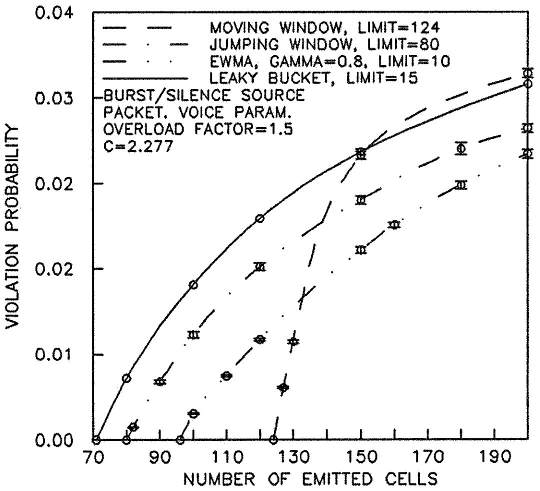

Figure 3: Instationary behaviour of the mechanisms

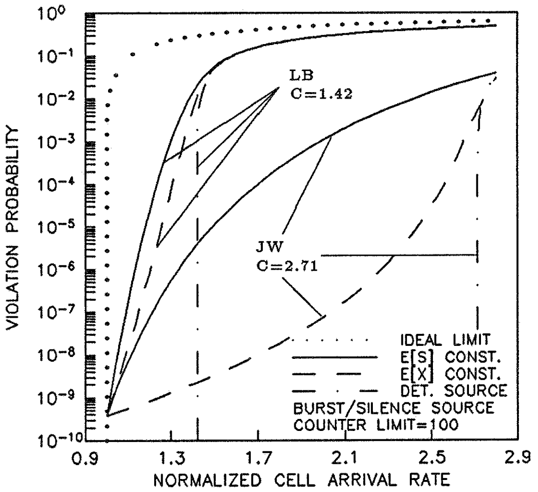

Figure 4: Influence of Cell Rate Variations on violation probability (LB, JW)

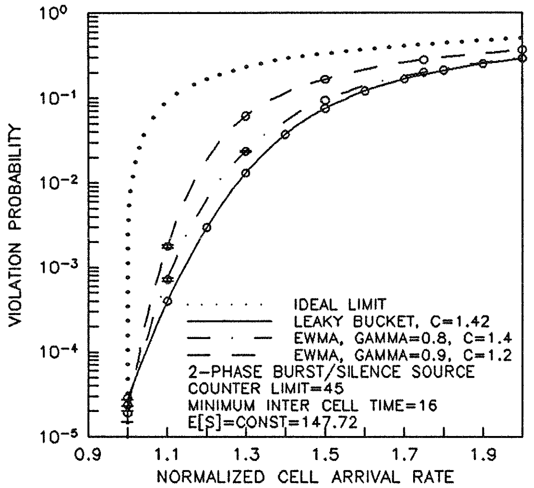

Figure 5: Influence of Cell Rate Variations on violation probability (LB, EWMA)

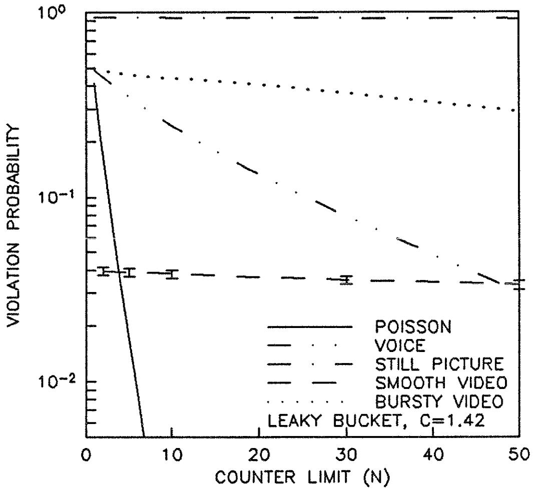

Figure 6: Influence of the source characteristics on the violation probability