

Congestion Control and Charging for the ABR Service Category in ATM Networks

O. Casals^a, C. Blondia^b, L. Cerdà^a and B. Van Houdt^{b*}

^a Polytechnic University of Catalunya, Computer Architecture Department,
c/ Jordi Girona, 1-3, 08034 Barcelona, Spain

^b University of Antwerp, Department of Computer Science and Mathematics,
Universiteitsplein 1, B-2610 Antwerpen, Belgium

ABSTRACT

The Available Bit Service (ABR) is a promising “best effort” service designed to achieve in ATM networks high efficiency and low cell loss. Since the ATM Forum approved a first standard, intensive research has been done about ABR. The aim of this paper is to analyse the main research topics involved in ABR, namely: the evaluation of traffic and congestion control schemes, conformance definition and policing and charging.

Keywords: ATM, ABR, flow control, switch mechanism, conformance definition, charging.

1. INTRODUCTION

The Available Bit Rate Service Category (ABR) has been introduced to support ATM connections originating from users which are willing to accept unreserved bandwidth and which are able to adapt their cell rate to changing network conditions and available resources. Information about the state of the network (e.g. with respect to congestion) and the availability of resources is sent to the source as feedback information through special control cells, called Resource Management cells (RM-Cells). Services that are compliant to this feedback control information experience a low cell loss ratio and obtain a fair share of the available bandwidth. There is no guarantee with respect to the delay or delay variation. As this control scheme operates at the time scale of a complete round trip delay, the ABR Service Category requires large buffers to be present in the network. The traffic parameters used for this category are the Peak Cell Rate and a minimal usable bandwidth, called the Minimum Cell Rate (MCR). The only QoS guarantee is the cell loss ratio (CLR). The available bandwidth may vary in time, but shall never be lower than the MCR. Typical applications using this category are Remote Procedure Calls, Distributed File Transfer, Computer Process Swapping, etc.

In this paper, we describe the behaviour of the end stations, as specified by the ATM Forum, and compare some methods that the switch can implement to detect and control congestion. Three approaches to evaluate the performance of these schemes have been used: a fluid flow approach, a method based on control theory and a discrete-time Markov chain approach. The first two are briefly described while the third one is applied to evaluate the throughput and buffer requirements in an Explicit Rate (ER) switch. A conformance algorithm used by the network to monitor whether the source transmits according to the traffic contract based on a dynamic version of the Generic Cell Rate Algorithm (DCGRA) is discussed. Finally different approaches for a usage-based pricing of ABR are analyzed.

2. FLOW CONTROL SCHEMES: DESCRIPTION

The congestion scheme described in the ATM Forum specifications¹ for the ABR service category is a rate-based, closed-loop, per-connection control that uses the feedback information from the network to regulate the rate at which the sources transmit cells. In this scheme, the transmission rate of each connection is controlled by means of special control cells called Resource Management Cells (RM-Cells). RM-Cells are transmitted embedded in the Data-Cell flow from the Source End System (SES) to the Destination End System (DES). The DES “turns around” the RM-Cells, which return to the SES along

* Further author information

O.C.: E-mail: olga@ac.upc.es; Telephone: +34 93 401 6985

C.B.: E-mail: blondia@uia.ua.ac.be; Telephone: +32 3 820 2404

L.C.: E-mail: llorenc@ac.upc.es; Telephone: +34 93 401 6798

B.H.: E-mail: vanhoudt@uia.ua.ac.be; Telephone: +32 3 820 2408

the same path carrying congestion information. Depending on the congestion information received in the RM-Cell, the SES increases or decreases its transmission rate. An ABR connection is always bi-directional and the sources located at both termination points must implement the SES and DES behavior. The standard specifies the SES and DES behavior and several methods that a switch can implement to control congestion that we describe next.

2.1. SES and DES Behavior

At the connection set up the source negotiates the maximum and minimum rate at which it may transmit (PCR and MCR); the Initial Cell Rate (ICR) at which it may start transmitting, the number of cells per RM-Cells (Nrm), the Rate Increase Factor (RIF) and the Rate Decrease Factor (RDF). After Nrm-1 Data-Cell transmissions, the SES sends a RM-Cell. The main fields of this cell and the initial values set by source are the following:

- Explicit Rate (ER), set to PCR,
- Current Cell Rate (CCR), set to the Allowed Cell Rate (ACR) of the source. The ACR is a parameter maintained by the source which fixes the maximum rate at which cells may be scheduled for transmission,
- Congestion Indication (CI) Bit, set to 0 (no congestion),
- No Increase (NI) bit, set to 0 (increase rate),
- Direction (DIR) bit, set to forward.

Note that the RM-Cell rate is the source rate divided by Nrm. Thus, if the source rate is low, the time intervals between the RM-Cells would be large, decreasing the opportunities of the source to change the transmission rate. To solve this problem the standard defines the Trm parameter. At each scheduled transmission time, the source checks if the time since the last RM-Cell transmission is higher than Trm. If it is, a RM-Cell is transmitted regardless of the Nrm count.

The cell flow is received by the DES, which must store the Explicit Forward Congestion Indication Bit (EFCI) of the last Data-Cell received. On receiving a forward RM-Cell it must change the CI bit to a congested state depending on the EFCI bit stored, change the DIR to backward and send (“turn-around”) the RM-Cell back to the SES along the same path.

On receiving a backward RM-Cell the SES adjusts the ACR. Just in case that a backward RM-Cell is received with CI = 0 and NI = 0, the SES is allowed to increase its rate (ACR) by no more than RIF·PCR. On receiving an RM-Cell with CI = 1, the SES has to decrease the ACR by at least RDF·ACR. Finally the ACR must be set at most to the ER field. In any case the ACR may not be reduced below the MCR or increased over the PCR. The source may perform a “Rescheduling option” which consists of rescheduling a transmission time of a cell in order to take advantage of an increase in the ACR.

Other source rules are defined to guarantee the efficiency of the flow control in special situations, e.g. in case of heavy congestion. Refer to the standard¹ for details.

2.2. ABR Switch Mechanisms

The ATM Forum establishes that a switch shall implement at least one of the following methods to control congestion: set the EFCI bit of the data cells; set CI or NI in forward and/or backward RM-Cells; reduce the ER of forward and/or backward RM-Cells. Switches that set the EFCI or CI bit to indicate a congestion state are known as binary switches. Switches that modify the ER field are called ER switches. Moreover, the ATM Forum establishes that ABR switches should allocate the bandwidth among the sources following fairly. The Max-Min fairness criterion has been used as a goal for switch mechanisms. This algorithm consists of computing the fair rate of the link as the available bandwidth minus the bandwidth of constrained sources over the number of unconstrained sources. A source is said to be constrained if it cannot transmit at the fair rate because it is internally rate limited, or because it is limited to a lower rate by another switch. In the following we briefly describe several switch mechanisms ordered by their complexity degree. A further comparison may be found².

The EFCI switch³, the simplest one, was described in the early proposal of ABR. It monitors the queue length and if it is larger than a threshold it sets CI = 1 in backward RM-Cells. A main problem of this switch is the lack of fairness. For example, the CI bit of backward RM-Cells will be set more often in those VC going through a higher number of congested links.

The EPRCA switch⁴ tries to achieve fairness computing a heuristic approximation of the fair rate exponentially averaging the rate of unconstrained VCs. When the queue length is larger than a threshold, the switch reduces the ER of unconstrained VCs down to a fraction of the fair rate.

The ERICA switch⁵ has become popular because of its robustness and simplicity. This switch measures the ABR input rate in order to compute the overload factor $z = ABR \text{ input rate} / TCR$, and the $Fairshare = TCR / \text{Number of active ABR Sources}$, where the TCR is a fraction of the available bandwidth. This is done at each interval given by N cell arrivals. When receiving a backward RM-Cell the switch computes the $VCShare = VC \text{ rate} / z$. Finally, the ER field of the backward RM-Cell is reduced, if greater, down to $\max(Fairshare, VCShare)$.

Figures 2 and 3 give a pictorial view of the evolution of source rates and queue length for an EFCI and an ERICA switch obtained by simulation. The network topology is shown in Fig. 1. The switch is fed by 5 ABR greedy sources staggered 30 ms. All links have a delay of 1 ms and a capacity of 365 Cells/ms (≈ 155 Mbps). Table 1 shows the source parameters. In the EFCI we have set a queue threshold of 100 cells and in the ERICA $N = 100$ cells and $TCR = 0.9 \cdot \text{Link cell rate}$. The figures show that the ERICA switch has lower rate oscillations and lower queue length. The EPRCA has a performance in between the EFCI and the ERICA.

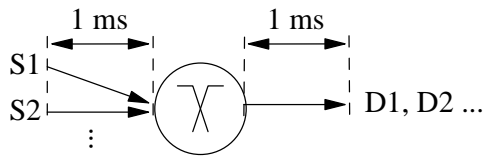


Figure 1. Network Topology.

Table 1. Source Parameters.

PCR [Cells/ms]	MCR [Cells/ms]	ICR [Cells/ms]	RIF	RDF	Nrm [Cells]
365	0	2	1/64	1/16	32

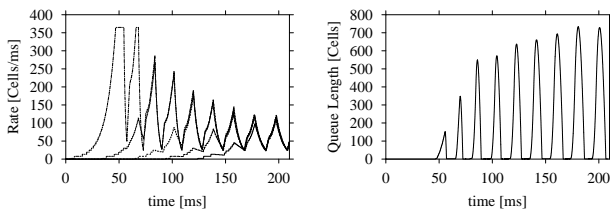


Figure 2. EFCI Switch.

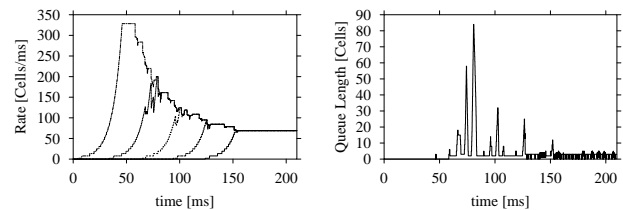


Figure 3. ERICA Switch.

3. FLOW CONTROL SCHEMES: EVALUATION

In this Section we first briefly describe two well-known approaches that have been used by several authors to evaluate the performance of the ABR flow control scheme: a fluid flow model and a model based on control theory. We then present a discrete-time Markov model we have used to derive some guidelines to engineer the ABR service category.

3.1. Fluid Flow Model

Consider a system consisting of a number of greedy homogeneous ABR end station and a single bottleneck link. Each of these ABR traffic sources has the same parameters. The propagation delay from the ABR sources to the switch and from the switch to the destination end station are considered in the model. Congestion in the switch is detected by means of two threshold values. When the queue length exceeds the higher threshold value, the node detects congestion, until the queue length goes below the lower threshold value. Considering the traffic flows as fluid, the ABR system dynamics are described by means of a set of coupled differential equations with two real value continuous variables, the ACR and the ABR buffer occupation. Due to the delay of the feedback information, the system oscillates and control cycles can be isolated. Each cycle consists of a number of phases that are governed by the set of coupled differential equations. Upper bounds for the buffer occupancy in case of an EFCI Marking mechanism can be found in⁶. The evaluation of the EPRCA mechanism can be found in⁷, while a more complete analysis is given in⁸.

3.2. Control Theory Model

Control theory seems to be an appropriate approach to the design of closed loop rate based flow control. Several studies have been made for a system consisting of a link with limited capacity whose input consists of controllable ABR traffic and CBR/VBR traffic, assumed to be non-controllable. There is a delay between the control action of rate modification in the

switch and the instant this information reaches the SES. The goal of this control action is to avoid cell loss and to maximize the link utilization. The control loop should be stable over a wide range of traffic patterns and propagation delays. We mention a few approaches. In ⁹ a control theoretical approach is followed to study a Relative Rate marking scheme, while ¹⁰ presents a linear dynamic model for the design of an Explicit Rate mechanism. A dual proportional-plus-derivative controller is used in ¹¹ for an Explicit Rate control loop.

3.3. Discrete-Time Markov Model

Consider a system consisting of two source end stations and a switch (see Figure 4). Time is supposed to be discrete with unit the time needed to process a cell in the switch, referred to as a slot. One end station generates CBR/VBR traffic and the other generates ABR traffic according to an explicit rate congestion control scheme. Both the ABR and CBR/VBR traffic are input to the switch and compete for the bandwidth of the same output port in the switch. The switch acts as a virtual destination station for the ABR traffic. The CBR/VBR traffic is assumed to be the superposition of on/off sources with geometrically distributed on and off period. For the ABR traffic two cases are considered: a first case where the ABR source is persistent (or greedy) and a second case where the ABR source is an on/off source. In this paper we restrict the study to greedy ABR sources. The model takes into account the distance between the ABR SES and the switch, denoted by τ and expressed in number of slots. The round trip delay equals 2τ . The algorithm used to compute the ER is based on the ERICA scheme. During a time interval of fixed length, called observation period P , the switch counts the number of arrivals from both the CBR/VBR traffic and from the ABR traffic. Denote these numbers by N_c , resp. N_a . The total input rate during this observation period is then given by $i=(N_c+N_a)/P$. The overload factor is computed as $o=i/TCR$. The Explicit Rate communicated to the ABR-SES is now given by $r=\min[PCR \max[MCR \ Na/(o.P)]]$. We assume that at the end of each observation period, the switch sends an RM cell with the newly computed ER to the SES. Upon arrival of this RM cell, the ABR-SES immediately applies this new ER as ACR. For this simple model, an embedded Markov chain technique leads to two important measures: the throughput of the ABR traffic¹² and the ABR buffer capacity required to ensure a given cell loss¹³ (e.g. $CLR \leq 10^{-9}$).

In this paper we use this model to show the influence of the following parameters on the ABR throughput and the ABR buffer occupancy :

- Length of the observation period P .
- The distance between the ABR-SES and the switch τ .
- The burstiness of the CBR/VBR traffic p : the CBR/VBR traffic is modeled as a single on/off source (or a superposition of a finite number of such sources); the burstiness is expressed by means of the mean duration p of the on period.

In what follows, a few examples are given which illustrate the type of results that may be derived using this discrete-time Markov model.

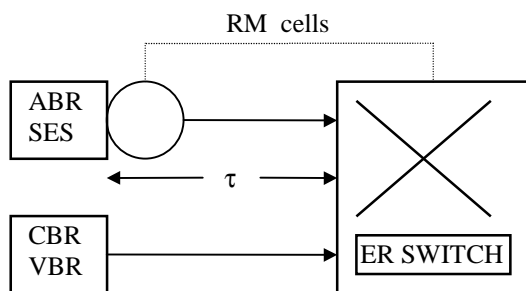


Figure 4. System Configuration

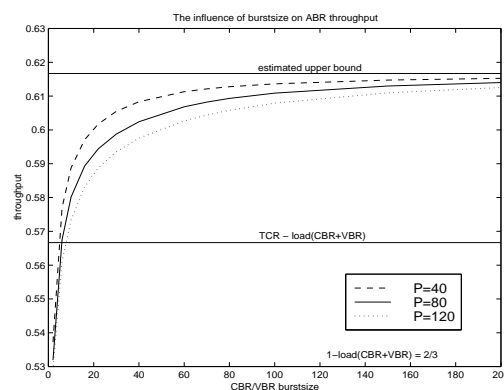


Figure 5. ABR Throughput

3.3.1. Impact of CBR/VBR Burstiness and Observation Period Length on the ABR Performance

Consider a system consisting of CBR/VBR traffic modeled as a single on/off source with geometrically distributed on and off period with respective means p and q . We let $q=2 \times p$. Furthermore, consider a greedy ABR source with parameters $PCR=0.8$, $MCR=0.25$ and let $TCR=0.9$. The round trip delay is supposed to be negligible, i.e. $\tau=0$. The throughput as a function of the burstiness p of the CBR/VBR traffic for different values of the observation period P (namely $P=40, 80$ and 120) is shown in Figure 5, while the corresponding results for the required buffer space to ensure $CLR \leq 10^{-9}$, are depicted in Figure 6. For a fixed observation period length, these figures show that, due to the guarantee of a minimum cell rate $MCR \neq 0$, higher throughputs are obtained for increasing burstiness p of the CBR/VBR traffic, at the expense of a considerable increase of the required buffer capacity. This fact is illustrated in Figure 7, where the influence of the MCR on the required buffer space is depicted. From both the throughput and buffer occupancy curves, we observe that for p much larger or much smaller than P , the length of the observation period P has no impact on the system performance.

3.3.2. Impact of the CBR/VBR Burstiness and the Round Trip Delay on the ABR Performance

Consider a system as described above, with $\tau \neq 0$ and $P=80$. From Figure 8 we see that when p and $2\tau+P$ have about the same magnitude, higher throughput is obtained for larger values of 2τ , but as can be seen from Figure 9, at the expense of larger buffers. A maximum throughput is obtained for p around $2\tau+P$. Moreover, the round trip delay is insignificant if p is much larger or much smaller than 2τ . Remark however that compared to the influence of P on the throughput (Figure 5), the impact of τ lasts longer (in Figure 8 there are still significant differences for $p=3000$).

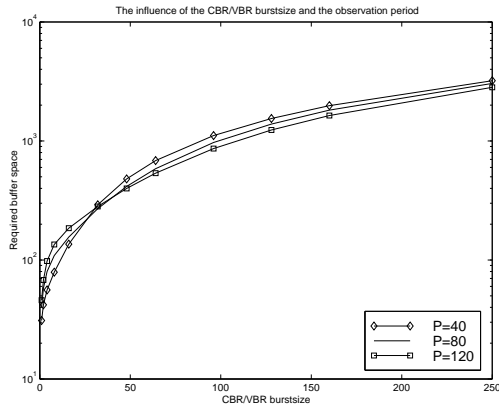


Figure 6. ABR Buffer Occupancy

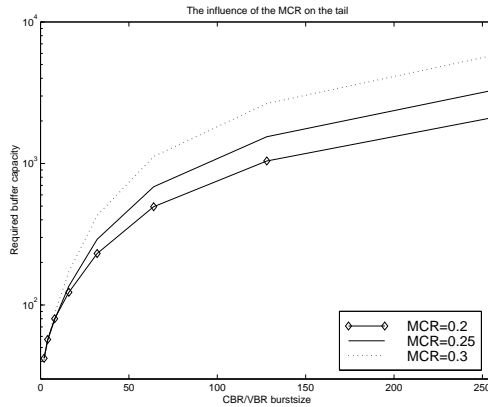


Figure 7. MCR and ABR Throughput

3.3.3. Engineering Rules

From these numerical results, and other investigations¹⁴, we may draw the following conclusions and derive some engineering rules.

- (i) *The length of the observation period P* : The value of P is unimportant if not in the same order of magnitude as the mean burst size p of the CBR/VBR traffic. In case that it has about the same magnitude we get higher throughput results for smaller values of P . Longer buffers for larger values of P are needed if the length P is smaller than p . Otherwise larger values of P result in a smaller buffer occupation.
- (ii) *The Round Trip Time τ* : for bursts in the same order of magnitude as the round-trip delay added with P we get a (slightly) higher throughput and need larger buffers for more distant sources. The round trip delay 2τ becomes less relevant as the difference between the mean CBR/VBR burst size and $P+2\tau$ increases. As with the observation period P , the distance of the ABR-SES is irrelevant if it has not the same order of magnitude as p . Also the effect of the round-trip time 2τ is of no importance if the length of the observation period is much larger.
- (iii) *The Burstiness of the CBR/VBR Traffic p* : The burstiness of the CBR/VBR traffic has a major impact on both the throughput and the required buffer space. Clearly longer bursts require larger buffers. The growth caused by increasing p depends upon the value of the MCR. For the throughput we get better performance as p is increased at least as long as p is smaller than $2\tau+P$. Then when further increasing p a minor decrease might occur depending on the distance between the ABR-SES and the switch where a stronger decrease is observed for more distant sources.

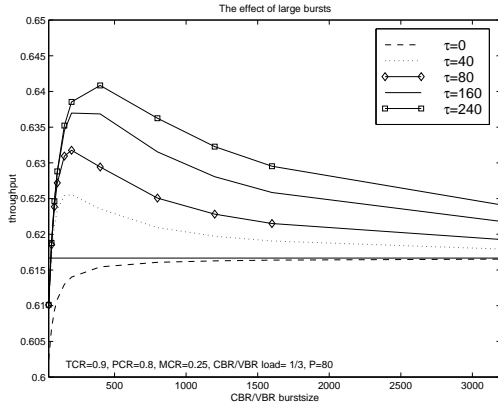


Figure 8. ABR Throughput

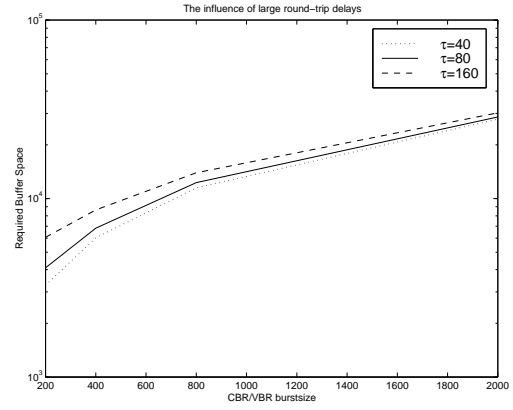


Figure 9. ABR Buffer Occupancy

4. CONFORMANCE DEFINITION AND POLICING

A conformance algorithm may be defined as the formalism used by the network to monitor whether the source transmits according to the traffic contract. ITU-T and the ATM Forum have defined the Dynamic Generic Cell Rate Algorithm (DGCRA) as an example conformance definition for an ABR connection. In this section the DGCRA is briefly described, its drawbacks are shown and possible improvements are suggested on basis of simulation results.

4.1. Overview of the Dynamic Generic Cell Rate Algorithm

The DGCRA decides the conformance in a cell by cell basis according to the following algorithm. At arrival time of cell n , the value $y_n = c_n - a_n$ is measured, where a_n is the arrival epoch and c_n is a theoretical arrival time. The cell n is non-conforming if y_n is greater than the cell delay variation tolerance (τ_l) and is conforming otherwise. The theoretical arrival time is computed at each cell arrival by the following algorithm:

After the initialisations $c_0 = a_0$, $LVST_0 = a_0$, $I_0^{old} = I_0$, the set of theoretical arrival times $\{c_n\}_{n>0}$ is computed at the arrival epochs a_n as:

$$c_n = LVST_n + \min(I_n, I_n^{old}), \quad LVST_n = \begin{cases} c_n, & \text{if } 0 \leq y_n \leq \tau_l \\ a_n, & \text{if } y_n < 0 \\ LVST_n, & \text{if } \tau_l < y_n \end{cases}, \quad I_n^{old} = \begin{cases} I_n, & \text{if } y_n \leq \tau_l \\ I_n^{old}, & \text{if } \tau_l < y_n \end{cases} \quad (1)$$

$LVST_n$ stands for the Last Virtual Scheduled Time at cell n arrival. The theoretical arrival time c_n is given by $LVST_n$ plus an increment $\min(I_n, I_n^{old})$ equal to the inverse of the expected source rate at the interface. I_n^{old} is introduced because the first cell received after a new increase I_n may be received at this increase, because of the rescheduling option of the source or at the previous increment I_n^{old} . Taking the minimum of I_n and I_n^{old} the algorithm stay on the safe side.

The sequence I_n is computed based on the rate changes conveyed by the backward RM-Cell stream. A rate change received at the measuring point at a given time may be applied to the forward cell stream after a delay equal to the round trip feedback from the measuring point to the source. Since the round trip feedback is variable, two time constants τ_2 and τ_3 have been specified to be negotiated together with τ_l for the conformance definition. These τ_2 and τ_3 are respectively an upper bound and a lower bound of the round trip feedback delay. To be on the safe side, the conformance algorithm schedules rate increases conveyed in the backward RM-Cells after a delay τ_3 and rate decreases after a delay τ_2 .

According to τ_2 and τ_3 , two algorithms have been proposed to compute I_n . Algorithm "A" provides the tightest conformance but is too complex to implement. Therefore a simpler algorithm "B" has been defined. The main complexity of algorithm "A" consists of keeping track of all the rate changes conveyed by the backward RM-Cells up to the time they are applied to the forward stream. Algorithm "B" simplifies this problem by storing at most two rate changes. The consequence is that algorithm "B" is only able to follow an approximate envelope of the rate changes.

4.2. Improving algorithm “B”

As stated in the previous section, the tightness of the rate conformance of the DGCRA may be reduced due to the difficulty of following the rate changes at the measuring point. Moreover, the algorithm does not perform a CCR conformance while switches usually use the CCR to estimate the VC rates. Therefore, the absence of a CCR conformance may lead to misbehaviour of the feedback control of switches that make use of the CCR if any source does not properly sets the CCR to the ACR.

In order to solve these drawbacks, we propose a UPC based on the CCR. Remember that sources set the CCR of forward RM-Cells to the ACR. The idea is thus using the CCR conveyed by the forward RM-Cells to perform the rate conformance. The proposed UPC based on the CCR would perform a conformance test at two levels. At a first level it would check that cells are conforming to the rate conveyed by the CCR. At the second level it would check that the source correctly changes the ACR by means of the expected rate at the measuring point computed by one of the algorithms “A” or “B”. Such UPC would guarantee that the source rate will never exceed the value conveyed by the CCR, and thus the switches could safely use it. A CCR non-conformance condition would occur in case the UPC receives a forward RM-Cell with the CCR exceeding the expected rate. In this case the UPC could take several actions, e.g. the CCR could be reduced down to the expected rate or the RM-Cell could simply be discarded.

With the current source behaviour given by the ATM Forum a source may transmit cells following a forward RM-Cell at a rate up to the CCR until a backward RM-Cell conveying a different rate is received. Consequently, with such source behaviour, it is not possible to use the CCR conveyed by the forward RM-Cells to perform the rate conformance because at any time cells may be received at a higher rate. To solve this problem we propose to constrain sources to delay rate increases just after a forward RM-Cell transmission. However delaying rate increases could slow down the access to the unused network bandwidth since the RM-Cells are transmitted after each $(N_{rm}-1)$ th data-cell. To solve this drawback we propose allowing the sources to advance a RM-Cell transmission before its N_{rm} -th turn. Advancing RM-Cell transmission would increase the RM-Cell rate and thus the overhead introduced by the RM-Cell transmission. We therefore introduce the concept of an “RM-Cell advance algorithm” which would decide whether rate increases are worth to advance a RM-Cell transmission or not.

4.3. CI and NI Conformance

The algorithms “A” and “B” defined by the ATM Forum just take into account the feedback conveyed by the ER of the backward RM-Cells (thus define a so called ER-conformance algorithm). However to specify a conformance definition to the SES, the feedback conveyed by the CI and NI bits of backward RM-Cells should also be considered. For example, an ER-conformance algorithm will be inappropriate for a binary switch that conveys the congestion information by means of the CI bit. In ¹⁵ we have proposed modifications to the algorithm “B” in order to take into account the rate changes conveyed by CI and NI.

However, if the rate changes are conveyed exclusively by the CI and NI bits, the rate at a given instant may depend on all previously computed rates. This makes the policing of sources controlled by binary switches a rather difficult task. Assume for example a source routed through binary switches and a UPC keeping track of CI and NI bits for conformance. If the source computes a lower rate than the UPC when receiving a backward RM-Cell, the following rate changes computed at the UPC would overestimate the ACR value that should be policed. Note that using the ER-feedback, each conveyed ER might carry the absolute value at which the source sets the ACR and thus “synchronises” the source and the UPC and avoids this problem. Several reasons may lead a source to compute a lower rate than the conformance algorithm. For example, a congested switch located between the source and the UPC would introduce rate decreases that could not be taken into account by the UPC.

We conclude that in order to perform a CI and NI conformance it would not be enough to keep track of rate changes conveyed by these bits. Such a UPC should also check that the source and the conformance algorithm perform the same rate changes and “re-synchronise” both in case the expected rate is overestimated.

4.4. Simulation Analysis

In this section we show the tightness problem of algorithm “B” and the improvements previously described. The results have been obtained by simulation using the network topology of Fig. 1, but now one ABR greedy source is multiplexed with background ABR sources that randomly become active and silent. We assume the UPC located at the switch input and policing the greedy source. In our model no CDV is introduced in the cell flow up to the UPC. Therefore, the feedback delay is equal to the round trip propagation delay (2 ms), and thus $\tau_2 = \tau_3 = 2$ ms. We are just interested in the ability of the algorithm to compute an accurate value of the expected rate, therefore the parameter τ_1 has no influence on our results.

Figures 10 and 11 superimpose the following curves:

- The expected rate at the interface (PACR) using the modified algorithm “B” which keeps track of the NI and CI bits,
- The rate policed by a UPC based on the CCR (CCR-PACR),
- The rate at which the forward cell flow arrives at the interface, which will be the transmission rate (ACR) delayed by the propagation delay (1 ms), thus we have plotted $ACR(t-1)$. The expected rate would be coincident with $ACR(t-1)$ in a UPC performing the tightest rate conformance.

The figures show that the PACR is a step envelope of the $ACR(t-1)$ with step duration nearly equal to the feedback delay. This inaccuracy is a consequence of the simplification of algorithm “B” keeping at most two-scheduled rate changes. The figures also show this problem can be avoided by using the CCR (note that the CCR-PACR curve is nearly coincident with $ACR(t-1)$).

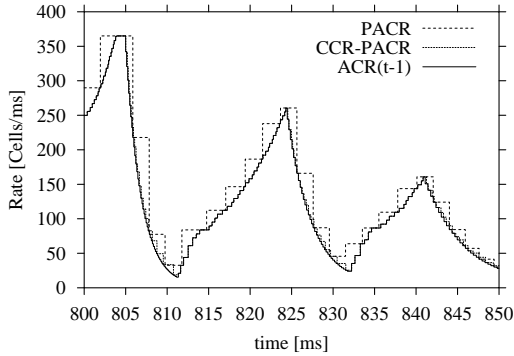


Figure 10. ACR and PACR with a binary switch.

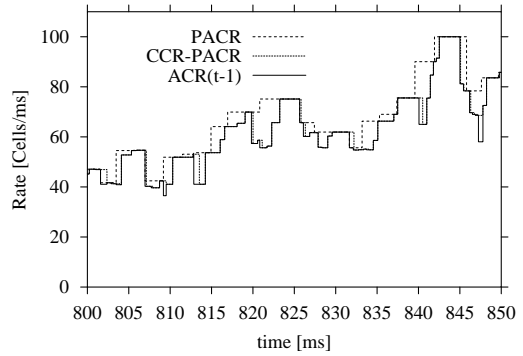


Figure 11. ACR and PACR with an ERICA switch.

5. CHARGING

Charging may be an essential condition for ABR sources to adapt their traffic demand. We analyze different approaches for a usage-based pricing of ABR. We can classify these approaches as “static” or “dynamic”. A charging scheme is static if the charging parameters are established at the connection set up and do not change afterwards. If they change, the scheme is dynamic. We analyze a static and a dynamic charging scheme already suggested and propose new alternatives that solve some of their drawbacks¹⁶.

A usage-based pricing seems adequate for an ATM network. For sources with guaranteed bandwidth as CBR and VBR, the allocated bandwidth or the generalized concept of the effective bandwidth, and the duration of the connection is a good characterization of their resource usage. Charging of ABR may be more complex as bandwidth is not requested by the sources but assigned by the network. Therefore, it is difficult for the network to know the user appraisal of resources for charging.

5.1. Static Charging Schemes

Songhurst and Kelly¹⁷ have proposed a pricing model where the charge of a connection is given by the expression:

$$\text{Total charge of a connection} = a(x) \cdot T + b(x) \cdot V + c(x) \quad (2)$$

where T is the duration of the connection, V the volume submitted by the connection, x the tariff choice and $c(x)$ is a fixed subscription fee. An expression for $a(x)$ and $b(x)$ based on the effective bandwidth can be obtained for VBR connections. For ABR sources the authors propose to use the above expression with $a(x)$ proportional to the MCR, i.e. $a(x) = \gamma \cdot \text{MCR}$, and $b(x)$ is assumed to be much lower than γ or even zero. In order for the sources to select an MCR according to their bandwidth appraisal, the authors assume that the network divides the free bandwidth proportionally to the MCR.

The static model previously described has the drawback that many data sources may not be able to choose an adequate MCR due to their bursty nature. For such sources a pricing model which charges the transmitted volume rather than the duration of the connections would fit better. We propose a static model for the ABR Service, which consists of several prices per cell tariffs p_i . Users select a tariff p_i at connection set up in order to charge the cells transmitted at the shared bandwidth.

We assume that the network allocation algorithm divides the free bandwidth proportionally to the chosen tariff p_i . Some sources may need a guaranteed MCR, which is likely to be charged based on the duration of the connection. According to these considerations, the following charging equation has been deduced:

$$\text{Total charge of a connection} = \gamma \cdot \text{MCR} \cdot T + p_i \cdot \max(V - \text{MCR} \cdot T, 0) \quad (3)$$

Note that if the connection transmits always at a rate $r \geq \text{MCR}$, then $\text{MCR} \cdot T$ is the volume of traffic submitted at the guaranteed MCR. Therefore, the first term of the right side of equation (3) charges this traffic at γ [unit of price/cell]. The second part of equation (3) is intended to charge the volume of traffic transmitted above the MCR at the price p_i chosen by the user, $p_i < \gamma$. Clearly, when the source rate $r \geq \text{MCR}$, the cells given by $\max(V - \text{MCR} \cdot T, 0)$ are the cells transmitted above the MCR. This is not a drawback, however, because it would penalize users that choose a guaranteed MCR higher than their needs.

5.2. Dynamic Charging Schemes

Based on the principle of social welfare optimization a dynamic charging for the ABR service has been proposed¹⁸. The method consists of using the forward RM-Cells to convey the source demand to the switches, and the backward RM-Cells to convey the prices to the sources. An iterative algorithm that in equilibrium satisfies the source demand and maximises the network revenue adjusts both demand and prices. In what follows we briefly describe this method.

Let C_l be the capacity available for ABR traffic traversing link l and a_l the price per unit of bandwidth (per unit of time) charged to the connection traversing link l . Denote by R_c the route of connection c and by w_c the charge of connection c . Clearly $w_c = \sum_{l \in R_c} a_l$. Assume that each connection c has a bandwidth demand equal to D_c . Finally denote x_c the actual rate of connection c . It can be shown that, in equilibrium, the welfare optimization leads to the following equations:

$$x_c = D_c(w_c), \text{ for all } c \quad (4)$$

$$\sum_{c: l \in R_c} x_c \leq C_l, \text{ for all } l \quad (5)$$

$$a_l \left(C_l - \sum_{c: l \in R_c} x_c \right) = 0, \text{ for all } l \quad (6)$$

Equation (4) says that the social welfare is maximized when the connection rates equal their demands. Note that equations (5) and (6) implies that $a_l = 0$ for the non-congested links ($\sum x_c < C_l$) and $a_l \neq 0$ for the congested links ($\sum x_c = C_l$). The algorithm consists of the links updating the price a_l^n at fixed interval n of duration δ (referred as charging intervals). Prices at interval n are decreased or increased if $\sum D_c^{n-1} > C_l$ or $\sum D_c^{n-1} < C_l$ respectively.

In order to integrate this charging scheme in the ABR flow control, two new fields have to be added to the RM-Cells: a request bandwidth (RB) and a price per unit of bandwidth (PB) field. The switches increase the PB field by the computed price a_l^n . Based on the prices w_c conveyed by the PB field, the sources set the RB field with a demand function $D_c(w_c)$ characteristic of the source bandwidth appraisal. Then, the switches use the RB values to estimate $\sum D_c^n$ and compute the price a_l^{n+1} , and so on. Finally, a billing unit located at the network edge keeping track of the PB field of the backward RM-Cells should be used to compute charges.

5.2.1. A Dynamic Pricing Model Based on the Switch Loads

A problem of the pricing model previously described is that users do not have an incentive to specify their true demand into the RB field, and they could set a misleading value in order to modify the prices (for example, the demand field could be set always to zero). We propose an alternative algorithm¹⁶ that uses the offered load ($\sum x_c^n$) instead of the source demand ($\sum D_c^n$) to decide when to increase or decrease prices. Note that in this scheme the RB field described in the previous section is not needed anymore. In the following we describe and analytically analyse the algorithm.

Assume that the control algorithm of the switch adjusts the source rates such that the offered load of link l converges to a certain target cell rate TCR_l . We define the overload as $O_l^{n-1} = \sum_{c: l \in R_c} x_c^{n-1} / \text{TCR}_l$. Based on the overload we propose the following algorithm to compute the prices:

$$a_l^n = \begin{cases} \max\left\{\left(1 + h \cdot \text{sgn}\left[O_l^{n-1} - \alpha\right]\right) a_l^{n-1}, 0\right\}, & \text{if } a_l^{n-1} \neq 0 \\ a_0, & \text{if } a_l^{n-1} = 0 \end{cases} \quad (7)$$

where $\text{sgn}[x] = \begin{cases} +1, & \text{if } x \geq 0 \\ -1, & \text{if } x < 0 \end{cases}$ and $\alpha \lesssim 1$ is a constant parameter. Note that now prices are adjusted such that the link load converges to $\alpha \cdot \text{TCR}_l$.

To analyse the algorithm (7), assume for simplicity a network with only one switch and all the sources located at the same distance from the switch. Let τ be the round trip delay from the sources to the switch. First note that in equilibrium, if $a_l^n \geq 0$, equation (7) can be rewritten as $a(t) = a(t - \delta) + h \cdot \text{sgn}[O(t - \delta) - \alpha] \cdot a(t - \delta)$. With the approximation $dw(t)/dt \approx w(t) - w(t - \delta)/\delta$ we obtain $da(t)/dt = h \cdot \text{sgn}[O(t - \delta) - \alpha] \cdot a(t)/\delta$. The solutions of this equation are of the form

$$a(t) = \begin{cases} K_1 e^{\beta t}, & \text{if } O(t) \geq \alpha \\ K_2 e^{-\beta t}, & \text{if } O(t) < \alpha \end{cases}. \text{ Assume that at time } t_0^+, O(t_0^+) > \alpha \text{ and equation } K_1 e^{\beta t} \text{ applies (cfr. Fig. 12). The price}$$

increases until the time instant t_1 , at which the overload at the switch falls below α . In equilibrium the source rate is equal to the demand (equation (4)). We have $O(t) = \Sigma D_c(a(t - \tau))/\text{TCR}$. Let a_α be the price at which $D_c(a_\alpha)/\text{TCR} = \alpha$. At time $t_1 - \tau$ the price $a(t)$ reaches a_α at the switch. This value is received by the sources at $t_1 - \tau/2$. The corresponding rate reduction is received by the switch at time t_1 . We have the following relations (cfr. Fig. 12): $K_2 = K_1 e^{\beta(t_1 - t_0)}$, $K_1 = K_2 e^{-\beta(t_2 - t_1)}$, $K_2 e^{-\beta(t_2 - \tau - t_1)} = a_\alpha$, $K_1 e^{\beta(t_1 - \tau - t_0)} = a_\alpha$, which imply: $t_1 - t_0 = t_2 - t_1 = 2\tau$, $K_1 = a_\alpha e^{-\beta\tau}$, $K_2 = a_\alpha e^{\beta\tau}$, and thus the period (T) and the amplitude (Δ) of the oscillation are given by:

$$T = 4\tau \quad (8)$$

$$\Delta = a_\alpha (e^{\beta\tau} - e^{-\beta\tau}) \quad (9)$$

Note that the previous algorithm is always stable and easy to tune. The only parameter h could be fixed with the following reasoning. Assume that the design criterion is that the oscillation relative to the equilibrium price has to be lower than a certain bound B_{max} (i.e. $\Delta/a_\alpha \leq B_{max}$). If the amplitude of the oscillation is small, from equation (9) we have $\Delta \approx a_\alpha 2\beta\tau$. Remember that $b = h/\delta$, therefore, the previous bound is equivalent to:

$$h \leq \frac{\delta}{2\tau} a_\alpha B_{max} \quad (10)$$

Given the maximum round trip delay from a source to the switch, and a desired B_{max} , equation (10) could be used to assess the value of h . Note however that the time constant of the algorithm is $1/\beta = \delta/h$. Therefore, reducing the value of h reduces also the convergence speed of the algorithm.

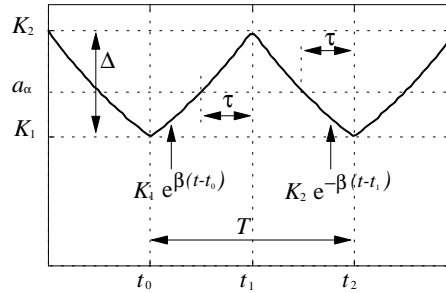


Figure 12. Evolution of the prices.

5.2.2. Integration of the Dynamic Pricing in the ERICA Switch

In this section we describe how the Dynamic Pricing scheme previously described could be integrated in the ERICA switch (see Section 2.2.). In order to compute the fair rates, the ERICA switch computes the switch overload at each measuring interval given by N cell arrivals. The switch could therefore apply algorithm (7) to compute the prices at each measuring interval, just after computing the overload. Then the PB field of the backward RM-Cells could be increased with the computed price at the same time as the ER field is set to the fair rate. In this case, the charging interval δ would correspond to the measuring interval. Assuming that cells arrive at target cell rate (TCR), this interval would be given by $\delta = N/\text{TCR}$. Substituting in (12), the parameter h of the charging algorithm would be given by:

$$h \leq \frac{N}{2 \text{TCR } \tau} B_{max} \quad (11)$$

5.2.3. Simulation Analysis

In this section we give a pictorial view of the source rates and prices that would be obtained applying the dynamic model. We consider the network topology of Fig. 1 with an ERICA switch implementing the charging scheme described in the previous section. The switch parameters are TCR= 0.9 Link Rate and Counting Interval= 100 Cells. The switch initialises the pricing parameters $a_i^0 = 0$, $\alpha = 0.95$ and $a_0 = 0.01$. Parameter h has been adjusted applying equation (11) choosing $B_{max} = 0.02$. The propagation delay is 1 ms, thus $\tau = 2$ ms and after substitution we have $h = 0.038$.

We assume that the switch is fed by two greedy sources S1 and S2 which adjust the transmission rate to $\min\{D_c(w), \text{ER}\}$ at each backward RM-Cell arrival, where w is the price per unit of bandwidth and ER the explicit rate carried by the RM-Cell. Source parameters are RIF = 1, Nrm = 32, ICR = 2 Cells/ms and PCR = Link Rate. We choose the source demand $D_c(w)$ assuming that the user criterion is to fix a maximum charge per unit of time p_c . Remember that w is the price per transmitted cell posted by the network and D is the source rate, the charge after a time period T becomes: charge = $w D \cdot T$. Therefore, to upper bound the charge per unit of time by p_c (i.e. charge/ $T \leq p_c$), the source demand has to be $D \leq p_c/w$. The source is upper bounded by the PCR, and hence such a source demand function is given by: $D_c(w) = \min\{p_c/w, \text{PCR}\}$.

We suppose that source S2 is ready to pay 5 times more than source S1, and thus we have set $p_c = 5$ for source S2 and $p_c = 1$ for source S1. Figures 13 and 14 show the evolution of the prices computed at the switch and the source rates. Source S2 is staggered 400 ms from source S1. Thus, initially the price is stabilised such that the source rate of S1 reaches the TCR of the switch. When the source S2 becomes active the prices increase until source S2 rate is 5 times the rate of S1, according to the source demands.

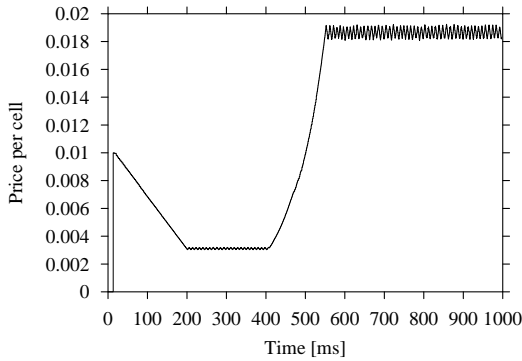


Figure 13. Prices Computed at the Switch.

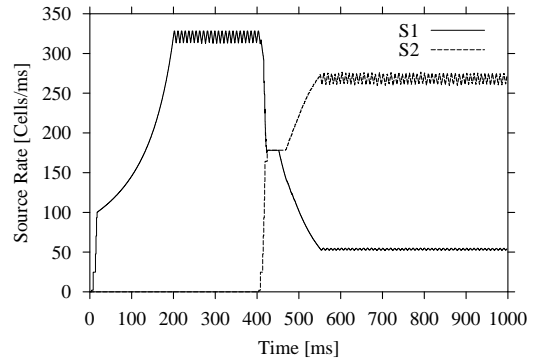


Figure 14. Source Rates.

6. CONCLUSIONS

In this paper we have described the flow control scheme used for the ABR service category according to the specifications defined by the ATM Forum. Several switch mechanisms are described and compared. A detailed analysis is made for the ERICA switch. A discrete-time Markov model leads to the assessment of the influence of the following parameters on the ABR throughput and the required ABR buffer capacity : the observation period, the round trip delay and the burstiness of the CBR/VBR traffic. The analysis leads to some simple engineering rules for the ABR service category.

We have described the DGRCA proposed as the Conformance Definition for ABR. We have shown that the DGRCA may have a lack of tightness of the policed rate, moreover, it does not perform a conformance of all the ABR parameters (e.g. the CCR, CI and NI). We have proposed improvements to solve these drawbacks.

Charging of ABR is an open issue. We have analysed some existing proposals and suggest new alternatives. We classify charging approaches as "static" when prices are established at the connection set up and do not change afterwards, and "dynamic" otherwise. Static charging schemes are simpler, although switches are required to divide the available bandwidth according to a weight related to the applied tariff (weighted allocation). Dynamic charging has the advantage of allowing an

optimisation of tariffs from an economically point of view. Dynamic schemes are more complex since switches have to update the prices continuously according to the link occupancy however, weighted allocation is not needed.

7. ACKNOWLEDGEMENTS

This work was supported in part by the Commission of the European Union, under project ACTS AC094 "EXPERT". The first and third authors were also supported by the Ministry of Education of Spain under grant TIC96-2042-CE. The second and fourth authors were also supported by the Vlaams Actieprogramma Informatietechnologie under project ITA/950214/INTEC "Design and Control of Broadband Networks for Multimedia Applications".

8. REFERENCES

1. ATM Forum Technical Committee TMWG, ATM Forum Traffic Management Specification Version 4.0, 1996.
2. A. Arulambalam, X. Chen and N. Anasri. "Allocating Fair Rates for Available Bit Rate Service in ATM Networks". *IEEE Communications Magazine*, pp 92-100, 1996.
3. N. Yin and M. G. Hluchyj. "On Closed-Loop Rate Control for ATM Cell Relay Networks". *IEEE INFOCOM'94*, 1c.4, pp 99-108, 1994.
4. L. Roberts. "Enhanced Proportional Rate Control Algorithm (EPRCA)". ATM Forum contribution number 94-0735R1, 1994.
5. R. Jain "ERICA Switch Algorithm: A Complete Description", ATM Forum contribution number 96-1172, 1996.
6. M. Ritter, "Network Buffer Requirements of the Rate-based Control mechanism for ABR Services", in *Proc. of the IEEE INFOCOM'96*, pp. 1190-1197, San Francisco, 1996.
7. H. Ohsaki, M. Murata, H. Suzuki, C. Ikeda and H. Miyahara, "Rate-Based Congestion Control for ATM Networks", *ACM SIGCOM Computer Communication Review*, pp. 60-72, 1995.
8. M. Ritter, "Modeling of Flow Control Mechanisms for the Available Bit Rate Services", *Ph.D. Thesis*, University of Wuerzburg, Germany, 1998.
9. M.A. Marsan, A. Bianco, R. Lo Cigno and M. Munafo, "Four Standard Control Theory Approaches for the Implementation of RRM ABR Services", *ATM Networks, Performance Modelling and Analysis*, Volume 3, pp. 259-279, Chapman and Hall, 1997.
10. Y. Zhao and S.Q. Li, "Feedback Control of Multiloop ABR Traffic in Presence of CBR/VBR Traffic Transmission", in *Proc. of the IEEE ICC'96*, pp. 1717-1721, Dallas, 1996.
11. G. Ramamurthy and A. Kolarov, "Application of Control Theory for the Design of Closed Loop Rate Control for ABR Service", in *Proc. of the ITC 15*, pp. 751-760, Washington, 1997.
12. C. Blondia and O. Casals, "Throughput analysis of the Explicit Rate Congestion Control Mechanism", in *Proc. of the 10th ITC Specialist Seminar on Control in Communications*, pp. 89-101, Lund, Sweden, 1996.
13. C. Blondia, O. Casals and B. Van Houdt, "Buffer Analysis of the Explicit Rate Congestion Control Mechanisms for the ABR Service Category in ATM Networks", to appear in the *Proc. of PICS '98*, Lund, May 1998.
14. C. Blondia, O. Casals and B. Van Houdt, "Buffer and Throughput Analysis of the Explicit Rate Congestion Control Mechanisms for the ABR Service Category in ATM Networks", submitted for publication, 1998.
15. L. Cerda and O. Casals, "Improvements and Performance Study of the Conformance Definition for the ABR Service in ATM Networks", in *Proc. of the 10th ITC Specialist Seminar on Control in Communications*, pp. 323-334, Lund, Sweden, 1996.
16. L. Cerda, O. Casals, Dynamic and Static Charging of the ABR Service in ATM Networks, to appear in *GLOBECOM'98*.
17. D. Songhurst, F. Kelly, "Charging Schemes for Multiservice Networks", in *Proc. of the ITC 15*, pp. 751-760, Washington, 1997.
18. V.A. Siris, C. Courcoubetis, G.D. Stamoulis, "Integration of Pricing and Flow Control for Available Bit Rate Services in ATM Networks", in *Proc. of the GLOBECOM'96*, London, UK, 1996.