In International Conference on Telecommunications, Melbourne, Australia, April 2-4, 1997

ABR Pricing Experiments in a Real Network*

C. Courcoubetis†, F. P. Kelly‡, V. A. Siris†, G. D. Stamoulis†, and R. Weber‡,

† Institute of Computer Science, FO.R.T.H.
and Dept. of Computer Science, Univ. of Crete
P.O. Box 1385, GR 711 10 Heraklion, Crete, Greece
{courcou,vsiris,gstamoul}@ics.forth.gr

‡ Statistical Laboratory
University of Cambridge
16 Mill Lane, Cambridge CB2 1SB, U.K.
{fpk,rrw1}@statslab.cam.ac.uk

Abstract: *The success of the Available Bit Rate (ABR) service will depend not only on pure technological issues, but also on whether its pricing structure provides the right incentives for users to efficiently use network resources, thus minimizing the negative effects of congestion. We describe a testbed for experimenting with various pricing schemes. The testbed attempts to be as realistic as possible, and allows real end-users to visually experience the effects of pricing and network congestion.*

1 Introduction

Available Bit Rate[1] (ABR) is one of the five service categories identified by the ATM Forum for ATM-based integrated services networks, [1, 3]. It is intended for sources that have the ability to reduce or increase their information rate in response to congestion control signals from the network. Thus, the ABR service category provides a mechanism for conveying congestion information from the network to the source, which reacts appropriately. In addition, users can specify a minimum throughput requirement (Minimum Cell Rate - MCR), which is committed to them for the duration of their connection. ABR connections will use the bandwidth that is left over from Constant Bit Rate (CBR) and Variable Bit Rate (VBR) connections. It is also possible to reserve some percentage of the total capacity for ABR services.

There has been considerable excitement around the ABR service due to its potential to satisfy the performance requirement of many applications while efficiently utilizing link capacity. However, its success will also depend on whether it provides the right incentives for users to efficiently use network resources, thus minimizing the negative effects of congestion. The latter is currently one of Internet's most intense problems, and it is attributed to its ineffective pricing structure (cf. [2]), namely *flat rate* pricing where prices depend solely on the link rate of the access pipe from the customer to the Internet service provider. Such a scheme offers no incentives for users to send traffic at a rate less than the rate of their access pipe. Considering aggregate user benefit, welfare economics suggests pricing schemes where users will be charged according to the amount of bandwidth they occupy and whose price will be determined by the total demand (hence by the actual need for bandwidth). While economic theory can provide the necessary guidelines and justification of various pricing schemes, low implementation overhead should be balanced against theoretical soundness of the underlying model. The development of successful pricing schemes will require an understanding of the interaction of pricing with other network controls (e.g., flow control), and the effects of conditions that exist in real networks (e.g., variability and unpredictability of the bandwidth available for ABR connections).

In this paper, we describe a testbed for conducting ABR pricing experiments, in order to demonstrate, evaluate, and compare various pricing models in terms of their implementation requirements and dynamic behavior. This testbed attempts to be as realistic as possible: it captures the essential issues of a large network providing ABR services, and supports most of the technological features anticipated in such a network. Since there is no such fully ABR-compliant equipment currently available, we had to provide the mechanisms to realistically emulate (in real-time) the missing parts. An important feature of the testbed is that it enables real users to visually experience the effects of pricing and network congestion, through the degradation of video quality, and to use pricing to control the performance of the services they receive. Hence, using the testbed, one is able to demonstrate and explain the difference (and, in fact, the advantages) of usage-based pricing compared to non-usage based pricing.

The remainder of the paper is organized as follows: In Section 2, we overview the ABR pricing schemes of

*This work was partly supported by the European Commission under ACTS Project CA$hMAN (AC-039).

[1] ABR essentially coincides with one of the ATCs (ATM Transfer Capabilities) identified by ITU, which is also referred to as the ABR ATC.

[6, 8, 5]. In Section 3, we describe the testbed in detail. In Section 4, we discuss and assess its main features, including the objectives of experiments that can be conducted. Finally in Section 5, we present some concluding remarks.

2 ABR Pricing Schemes

The pricing structure for best-effort services should be *incentive compatible*, i.e., it should lead the user to select the service offering that best fits his (the users) needs. A key characteristic of any incentive compatible pricing scheme for best-effort service is that the price for a certain amount of service increases when congestion in the network increases. Although, an ABR service does not coincide with best-effort in the sense of Internet, it has no strict performance guarantees either. Thus, ABR pricing can be based on similar principles as pricing for best-effort services in the Internet. (The main differences between the two kinds of services lies in the fact that ABR can have some minimum performance guarantees through the MCR.)

We consider two pricing schemes for ABR services which have both been developed within the E.C. funded ACTS Project CA$hMAN (Charging and Accounting in Multiservice ATM Networks), [5, 4]. We briefly describe these schemes hereafter, in order to make the issues to be faced by the testbed more tangible .

In the first scheme, [8, 5], prices are based on the Minimum Cell Rate (MCR) and involve measurements of the duration of a connection (Time, T) and the total number of cells transferred that are above MCR (Volume, V). The total charge of the connection is

$$a \cdot MCR \cdot T + b \cdot V, \tag{1}$$

where a is the charge per unit of time and unit of MCR, and b is the charge per unit of volume transferred in excess to the MCR. Both of these values are set by the network. The value of b is typically small and may be equal to zero. Hence, according to this scheme users are priced based largely on their MCRs. The rationale behind this pricing scheme is the following. Consider a user who wishes to transfer a file. The user is free to choose any MCR. The duration of the file transfer depends both on the chosen MCR and the level of congestion. In the presence of congestion, the rate with which the connection can send is equal to MCR plus some share of the bandwidth that remains if we subtract the sum of the MCRs from the bandwidth available for ABR: $C_{\mathrm{ABR}} - \sum MCR$. The scheme assumes that this left-over bandwidth is shared among connections *in proportion* to the their specified MCR. This discourages users from splitting a single connection into multiple connections [8]. As congestion increases, the available bandwidth for the connection decreases, hence the duration of the transfer increases. From (1) we see that the charge also increases. Thus the price a user pays is higher during congestion, while his performance decreases. This is in accordance with incentive compatible pricing.

The second pricing scheme, [6, 5], is justified by the theory of social welfare maximization. Prices per unit of volume vary dynamically in response to varying conditions of network load in such a way that, in equilibrium, the demand for bandwidth will be equal to the supply. It turns out that, with appropriately defined user demand functions, in the state of equilibrium, both network revenue and social welfare (i.e., the aggregate benefit or satisfaction of the network and its users) are maximized. According to this approach, charges are computed in time intervals of constant duration. The price $w(k)$ per unit of volume during interval k is the sum of the prices on all links along the connections path. The charge for each interval is $w(k)V(k)$, where $V(k)$ is the number of cells sent in interval k. Hence, the total charge is

$$\sum_{k=1}^{K} w(k)V(k), \tag{2}$$

where K is the total number of intervals a call lasts. Prices for each link are adjusted locally at each switch based on the total demand for bandwidth on that link: prices decrease (resp. increase), in an iterative fashion, when the aggregate demand exceeds (resp. is less than) the available capacity. How frequently prices need to change depends on how frequently the aggregate user demand changes. We anticipate that for networks which multiplex a large number of connections, as will be the case for ATM, the aggregate user demand does not change very often.

3 Testbed for ABR Pricing Experiments

The testbed for conducting ABR pricing experiments is shown in Figure 1. In its initial stage it comprises two ATM switches connected by a single ATM link. The *Charging and Accounting Unit* (CAU) is transparent to the end stations and ATM switches, and is responsible for measuring quantities such as the total duration and the total volume (i.e., total number of cells) for a connection. These measurements are sent to the Network Management Station (NMS) which derives the final charge for the connection. The Charging and Accounting Unit also implements flow control and price related functions.

Regarding the user applications accessing this link, three configurations have been defined:

1. The ATM link is accessed by a number ($= n$) of user applications, each served by the ABR service. Each

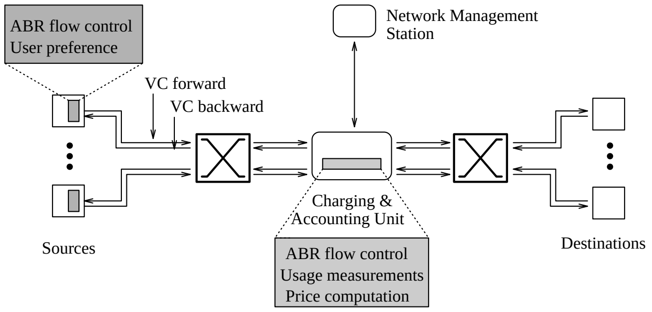

Figure 1: Testbed for ABR pricing experiments.

application resides on an end station and has a VC connection to a destination which passes through the priced link.

2. The ATM link is accessed by a single user application served by the ABR service, as well as by background traffic generated according to a model that approximates traffic of multiple ABR applications as in a large system. Again, the application resides on an end station and has a VC connection to a destination which passes through the priced link; the same applies to background traffic.

3. A combination of multiple user applications and background traffic.

The first configuration is appropriate for investigating a system with a small number of users, since their number n is bounded by the number of connections that can be supported by the Charging and Accounting Unit. On the other hand, subject to appropriate selection of the model for background traffc, the second configuration can be employed in analyzing a larger system with many users. (The model of background traffic is discussed later.) The fact that the testbed has only one link is not as restrictive as it can initially appear, because similarly with Internet even a large network may have a single (or a few) bottleneck link(s). For example, a transatlantic link is often the bottleneck of a network connecting users from Europe and U.S.A.

3.1 Software modules

The experimental testbed configuration with n applications includes $n+1$ software modules. One is located at the Charging and Accounting Unit which implements ABR's flow control and pricing related functions. The other n modules are identical in functionality, and are located at the n source applications. Each such module is responsible for the user specific flow control functions and user preferences. The testbed configuration for large systems includes 3 software modules, Namely, that of the Charging and Accounting Unit, that of the user application (both are the same with the ones in the configuration with n users), and that generating background traffic. Similarly, the third configuration includes $n+2$ software modules, where n is the number of applications.

Next we discuss in more detail each of the aforementioned modules.

Source flow control and user preference: The flow control functions performed at the user side involve reacting to congestion control signals from the network, and thus follow relevant specifications [1]. According to specifications, a source (i.e., a user) periodically sends special control cells, called Resource Management (RM) cells. Once they reach the destination, RM cells are sent back to the source. Thus, the flow of RM cells creates a feedback loop which is used by the network to send congestion related information to the source. Due to the lack of ATM adapters which support ABR flow control, we had to implement the functionality at the application level. If for some reason (e.g., performance) RM cells cannot be sent in-line (i.e., in the same VC that is used for the data transfer), a remedy is the following: For each VC used to transfer data, we open a second VC, the "control VC", which is solely used to transfer the equivalent control information between the end applications and the Charging and Accounting Unit.

In addition to flow control functions, the user module also includes user preference functions. These involve selecting the amount of bandwidth or MCR to request. It is this capability that allows users to control the performance they receive by increasing or decreasing the price they are willing to pay.

On top of the flow control and user preference functions, we have included a software module which reads MPEG (Motion Picture Expert Group) compressed video from a file and sends it over an ABR connection to the destination; this plays back the MPEG file as it is received. The sending module adjusts its sending rate in response to congestion. There are two possibilities. First, the transmission of the whole MPEG file can be delayed, hence the deterioration of the playback quality at the destination is experienced through a decrease in the number of frames played per second. A second alternative is for the sender to preferentially drop B frames[2]during periods of congestion. In this case, the deterioration of the playback quality at the destination is experienced through "jumps" in the frame sequence.

CAU flow control and pricing: The Charging and Accounting Unit support for flow control includes the implementation of a sharing policy for the bandwidth re-

[2]MPEG compressed video contains three frame types: I, P, and B. Among the three, B frames can be dropped without affecting the decoding process of other frames.

maining after subtracting the sum of the MCRs. For example, excess bandwidth can be shared among ABR connections in proportion to their Minimum Cell Rate (MCR) or according to a fair share policy. This functionality enables us to experiment with various bandwidth sharing policies, and investigate how these affect pricing. When ABR flow control becomes fully available in commercial switches, the software module of the Charging and Accounting Unit will be simplified accordingly.

In addition to flow control functions, the Charging and Accounting Unit performs pricing and accounting functions. Specifically, the CAU measures the duration and total number of cells sent through a connection. These are sent to the Network Management Station (NMS) which computes the final charges. Furthermore, for the dynamic pricing scheme described in Section 2, the CAU implements the price update function; this involves increasing or decreasing the price per unit of bandwidth depending on whether the aggregate demand for bandwidth is more or less than the capacity available for ABR connections.

Backround traffic generation: This should be based on a model for multiplexed ABR traffic. Similar to the source module described preciously, it should also react to the posted price and network congestion. The precise definition of such a model is still an open issue. A simple first approach is to use a model for multiplexed Internet traffic, since such traffic shares several similarities with ABR. The demand for Internet traffic (and its dependence on prices) is currently investigated in the Internet Demand Experiment (INDEX) Project ([7]), which can prove a valuable source of input for defining the model of interest.

The bandwidth available for ABR connections depends on the bandwidth used by higher priority traffic (i.e., CBR and VBR). The fluctuation of bandwidth for ABR connections can be emulated by means of a call generator running on the Network Management Station, which randomly generates CBR and VBR calls, each with its own duration and bandwidth requirement. The Charging and Accounting Unit, at regular intervals, can obtain from the Network Management Station the amount of bandwidth of the link available for the ABR connections. Alternatively, we can introduce connections at random times which carry VBR and CBR traffic over the priced link.

4 Features of the Testbed

The main features of the testbed described above are as follows:

- The testbed can accommodate a wide variety both of ABR pricing schemes and of bandwidth allocation policies. Thus, it can demonstrate the feasibility of such a pricing scheme (and whether it matches with a certain bandwidth allocation policy), and identify possible implementation difficulties or bottlenecks.

- The testbed can be used to investigate a system with a large number of connections, by emulating multiple user applications as backround traffic.

- In the testbed, we can introduce different delays to the user side software modules or to different "constituents" of the modeled background traffic. This is important, because connections in a real network will have various round trip delays, which may differ by orders of magnitude. Such a wide range of round trip delays will effect the dynamic behavior of both flow control and pricing.

- The testbed can demonstrate visually the dependence of prices on the level of congestion. This can be visualized by sending MPEG compressed video from the source, and having it played back, as it is received, at the destination. The quality of the playback will decrease as congestion increases. The user can increase the quality by increasing the price he is willing to pay for transmission, or by requesting a higher MCR.

4.1 Objectives of the Experiments

Apart from demonstrating feasibility of a pricing scheme, and the dependence of prices on the level of congestion, the objectives of the pricing experiments include the following:

- Compare various pricing schemes, and investigate their effects in a real network. Important issues include stability and convergence properties, and their relation with flow control. In addition, one can evaluate the gains in using load measurements for determining prices.

- Investigate the impact of the time scale of the fluctuation of the amount of available bandwidth for ABR traffic; this is an important factor that affects the behavior of both the pricing and flow control schemes.

5 Concluding Remarks

In this paper, we have described a real network that can serve as a testbed for ABR pricing schemes. The testbed attempts to be as realistic as possible by capturing the important issues of a large network providing ABR services. An important feature of the testbed is that it enables real users to visually experience the effects of pricing and congestion, and allows us to investigate the effects of issues such as stability and convergence, on the end user applications.

An important issue deserving further attention is the modeling of the background ABR traffic, as well as the evaluation and exploitation of experimental results on the pricing schemes already defined.

References

[1] ATM Forum. *ATM Forum Traffic Management Specification Version 4.0.* April 1996.

[2] J. P. Bailey and L. Mcknight (editors). *Internet Economics.* MIT Press, Massachusetts, 1996.

[3] F. Bonomi and K. W. Fendick. The Rate-Based Flow Control Framework for the Available Bit Rate ATM Service. *IEEE Network*, pages 25–39, March/April 1995.

[4] CA$hMAN. Charging and Accounting Schemes in Multi-
service ATM Networks (CA$hMAN), ACTS Project AC-039. http://www.isoft.intranet.gr/cashman/.

[5] CA$hMAN Consortium. *First Year Results on Pricing Models.* Deliverable D06, CA$hMAN, ACTS Project AC-039, 1996.

[6] C. Courcoubetis, V. A. Siris, and G. D. Stamoulis. Integration of Pricing and Flow Control for Available Bit Rate Services in ATM Networks. In *Proceedings IEEE Globecom'96*, pages 644–648, London, UK, November 1996.

[7] INDEX. Internet Demand Experiment (INDEX) Project. University of California at Berkeley. http://www.INDEX.Berkeley.EDU/public/.

[8] F.P. Kelly. Charging and Accounting for Bursty Connections. In J. P. Bailey and L. Mcknight, editors, *Internet Economics*, Massachusetts, 1996. MIT Press.