Available at www.ComputerScienceWeb.com
POWERED BY SCIENCE @ DIRECT•

ELSEVIER

Computer Communications 26 (2003) 1470–1483

computer communications

www.elsevier.com/locate/comcom

A proposal of new price-based Call Admission Control rules for Guaranteed Performance services multiplexed with Best Effort traffic

Marco Baglietto[a,*], Raffaele Bolla[a], Franco Davoli[a], Mario Marchese[b], Maurizio Mongelli[b]

[a]*DIST—Department of Communications, Computer and Systems Science, University of Genoa, Via Opera Pia 13, 16145 Genova, Italy*
[b]*CNIT—Italian National Consortium for Telecommunications, University of Genoa, Research Unit, Via Opera Pia 13, 16145 Genova, Italy*

Abstract

Pricing for the use of telecommunication services is an issue widely treated in the literature. In the last years it has received a growing attention in order to establish various fairness criteria in the bandwidth allocation for each type of traffic class. A number of pricing models have been proposed and analyzed in the context of Quality of Service (QoS) guaranteed networks (e.g. ATM, IP Integrated Services, IP Differentiated Services) and more recently, also for Best Effort (BE) environments. In the context of QoS networks the pricing scheme can influence the Call Admission Control (CAC) rules. On the contrary, for a BE service, users accept a variable bandwidth allocation, they are not subject to CAC and their pricing policies, according to the Proportional Fairness Pricing, are integrated within the flow control. In this paper we investigate the condition where both BE traffic and traffic explicitly requiring QoS (Guaranteed Performance, GP) are present. We propose three CAC rules for the GP traffic. The aim is to maximize the Internet Service Provider's overall revenue and to establish a bound over the GP traffic prices. Numerical results are presented to show the good performance of the proposed techniques.

Keywords: Call Admission Control; Pricing; Proportional Fairness; Best Effort services; Guaranteed Performance services; IP Quality of Service; Asynchronous Transfer Mode; Flow control

1. Introduction

The exponential growth of the Internet, the pervasive diffusion of the TCP/IP paradigm for the transport of Best Effort (BE) services, and the emergence of Guaranteed Performance (GP) services (as provided by ATM and IP QoS mechanisms), have fostered the development of Internet pricing schemes. Besides differentiating prices according to QoS levels, such schemes should also be capable of achieving 'globally optimal' bandwidth allocations and fairness for the users.

A number of pricing models have been considered and analyzed in the context of Quality of Service (QoS) guaranteed networks, mainly with respect to the Asynchronous Transfer Mode (ATM) world (see, e.g. Refs. [1–8]). In Ref. [6], a dynamic adaptive priority scheme for each Virtual Path (VP), based on the periodic adjustment of prices per unit of bandwidth is proposed: the users decide the bandwidth to be utilized on the corresponding VP and pay accordingly up to the next reallocation. Kelly in Ref. [5] bases the pricing scheme in a QoS context on the concept of effective bandwidth (i.e. the bandwidth that is necessary to satisfy call-level QoS requirements). The user's charge is established in function of the traffic volume and of the call duration. Such function is linear and its coefficients are: a fixed charge, the price per unit time, the price per unit volume; they are fixed at the time of connection acceptance and depend on the user's traffic contract. IP QoS mechanisms (Resource Reservation Protocol [9], Integrated Services [10,11] and Differentiated Services [12]) can be subject to pricing policies very similar to the 'guaranteed bandwidth' environments (circuit-switched or ATM).

On the other hand, in a BE context, where the user does not declare QoS parameters and is not subject to a Call Admission Control (CAC) (i.e. flows are 'elastic', as determined by TCP congestion control, or by TCP-friendly mechanisms at the application level), pricing policies should be different from those adopted in the above mentioned QoS environments. Recent works about

* Corresponding author.
E-mail addresses: mbaglietto@dist.unige.it (M. Baglietto), lelus@dist.unige.it (R. Bolla), franco@dist.unige.it (F. Davoli), mario.marchese@cnit.it (M. Marchese), maurizio.mongelli@cnit.it (M. Mongelli).

0140-3664/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0140-3664(03)00032-X

the design of telecommunication networks have also shown that it is better to serve elastic traffic with variable bandwidth allocation rather than with constant bit rate [30–32]. Moreover, giving the BE traffic only the bandwidth resources unused by the GP services, as proposed in some early models (see for example Ref. [36]), does not seem to be the best way to manage GP and BE service classes together [18]. A new model for BE services, the Proportional Fairness Pricing (PFP), has been proposed in order to manage both congestion control and pricing also for some ‘big users’ that do not subscribe a GP traffic contract. Such big users, in fact, could be interested in receiving a variable bandwidth allocation in order to satisfy their variable QoS requirements (e.g. packet delay constraints), according to the current network traffic congestion condition and to their willingness to pay [34,35]. Low and Varaiya [15] have been among the first to propose a pricing model where prices are periodically adjusted on the basis of a dynamic pricing flow control. The users’ requests are formulated through a ‘utility function’, according to their traffic parameters and QoS requirements and to their ‘willingness to pay’ for a given service. In this context, pricing becomes strictly related with congestion control, as it is capable of determining the optimal rates of the users’ flows that maximize the aggregate source utility [13,14,16,17,20,21,42].

In the literature, models for CAC strategies aimed at maximizing the Internet Service Provider’s (ISP) revenue [27–29,37,38] and models for the integration between congestion control and pricing [13,14,16,17,20,21,42] are usually analyzed separately. Our aim is to establish a unified model, able to manage both the GP and the BE traffic. We propose novel price-based CAC rules, whose main goal is the maximization of the ISP’s revenue. However, at the same time, we obtain a fair bandwidth allocation for BE users (consistently with the PFP scheme). The papers that deal with the CAC in telecommunications networks propose models for the minimization of a weighted sum of the blocking probabilities of different traffic classes, each of which is associated to a particular reward for the ISP’s revenue (see, for example, Refs. [27–29]). In this context, the minimization of the blocking probability corresponds to the maximization of the ISP’s revenue. In a real network, under both the ATM technology [6,7] and the future IPQoS technology [9–12], BE and GP traffic classes are multiplexed together and share the available bandwidth resources. In such a context, a trade-off between revenues coming from GP and BE users must be taken into consideration. The maximization of the ISP’s overall revenue in a heterogeneous network of both GP and BE users would be a quite simple task if the bandwidth allocations and the prices for BE users were much lower than those imposed to GP ones (a situation considered in some models preceding the PFP, e.g. Ref. [36]). In such case, each strategy that minimizes the blocking probability of GP calls achieves the maximum ISP’s revenue, too. On the contrary, if some big users are managed and priced following the BE PFP scheme [34,35], the revenue generated by BE traffic can become very close to the GP revenue. For this reason, the revenue that would be provided by the acceptance of a new incoming GP connection should be compared with the correspondingly possible BE revenue decrease on the links crossed by such new GP connection, and the result should drive the CAC decision. A possible alternative to this is to derive a new tariff for the GP service, to compensate the BE revenue decrease.

The issue of modeling and analyzing the integration of BE and GP traffic pricing is addressed in rather few papers (see e.g. Refs. [18,19]). In Ref. [18] an analytical model for the evaluation of the decay of the BE traffic performance as a function of the GP traffic blocking probability is proposed. Moreover, using a Stackelberg game model, a policy that recalls the paris-Metro pricing scheme [3] is proposed. The aim is to set the prices so that those users who are free to choose between GP or BE service (the so-called mixed users), would be inclined toward the traffic class that results more convenient, so as to reach the best global social welfare. In this paper, we do not decide upon the choice of the traffic class; rather, each incoming flow will be declared to belong either to GP or BE beforehand and we take on-line decisions by using different CAC techniques that are explicitly based on the presence of both traffic categories.

The paper is organized as follows. In Section 2, we introduce the main optimization problem for the BE environment (PFP) following Low and Lapsley [16] and Kelly [13]. Section 3 is devoted to the description of our optimization model. Numerical results are presented in Section 4 and final remarks in Section 5. Conclusions and future work are discussed in Section 6.

2. Proportional Fairness Pricing

The concept of the PFP was motivated by the desire to incorporate the notion of fairness into the allocation of network resources [3]. The PFP is dedicated to the BE service class, because the bandwidth allocation of a PFP user depends on the current network congestion conditions. In the PFP model the willingness to pay for such bandwidth allocation is also taken into account (see also Ref. [8], for an introduction to the congestion price models, or Ref. [42]). The PFP is aimed at maximizing the network social welfare and at guaranteeing a fair bandwidth allocation, either in terms of the Max-min fairness [33], adopted by the ATM forum for ABR services [40], or in terms of the ‘proportional fairness’ proposed in Refs. [13,39,41]. A BE user, in the context of the PFP, can represent a single domestic user, but also a ‘big one’ such as an aggregation of domestic users or a group of LANs (for example a company, one of its branches, or a university campus [34]). Each of these aggregates includes one or more groups of single users that have the same routing path. Their willingness to pay is modeled by a single 'big' utility function that can vary during the day [34,35].

With a notation that slightly differs from that in Refs. [13,16], we consider a telecommunication network composed by a set J of unidirectional links j, each with capacity c_j. We call 'BE user' r a connection established on a specific path consisting of a non-empty subset $J_{BE}(r)$ of J; R_{BE} is the set of active BE users, whose cardinality is denoted by $|R_{BE}|$. We indicate with $R_{BE}(j)$ the subset of BE users that use link j and with $A = \{A_{jr}, j \in J, r \in R_{BE}\}$ the matrix assigning resources to BE users ($A_{jr} = 1$ if link j is used by user r, $A_{jr} = 0$ otherwise). Moreover, let x_r be the rate of user r and $U_r(x_r) : [m_r, M_r] \rightarrow \Re$ the utility function of such user, supposed to be strictly concave, increasing and continuously differentiable over $I_r = [m_r, M_r]$; m_r and M_r are the minimum and maximum transmission rates, respectively, required by user r. Such utility function describes the sensitiveness of user r to changes in x_r. In the context of pricing, one can think of $U_r(x_r)$ as the amount of money user r is willing to pay for a certain x_r. Finally, let $\mathbf{c} = [c_j, j \in J]$, $\mathbf{x} = [x_r, r \in R_{BE}]$, $\mathbf{U}(\mathbf{x}) = [U_r(x_r), r \in R_{BE}]$ be the aggregate vectorial quantities. The main goal of the ISP can now be stated, consisting of the maximization of the sum of all users' utilities, under the link capacity constraints over the given paths [13,16,17,20].

Namely, we have the SYSTEM problem:

$$\mathbf{x}^o = \arg \max_{x_r \in I_r, \forall r \in R_{BE}} \sum_{s \in R_{BE}} U_s(x_s) \tag{1}$$

under the following constraints:

$$\mathbf{A} \cdot \mathbf{x} \le \mathbf{c} \qquad \mathbf{x} \ge 0$$

It is shown in Refs. [13,16,17,20,21] that to formulate a distributed and decentralized solution of the SYSTEM problem it is convenient to look at its dual. Following Ref. [16], the capacity constraints can be incorporated in the maximization by defining the Lagrangian:

$$\begin{aligned} L(\mathbf{x}, \mathbf{p}) &= \sum_{r \in R_{BE}} U_r(x_r) - \sum_{j \in J} p_j \left(\sum_{r \in R_{BE}(j)} x_r - c_j \right) \\ &= \sum_{r \in R_{BE}} \left[U_r(x_r) - x_r \sum_{j \in J_{BE}(r)} p_j \right] + \sum_{j \in J} p_j c_j \end{aligned} \tag{2}$$

where $\mathbf{p} = [p_j, j \in J]$.

The objective function of the dual problem is:

$$D(\mathbf{p}) = \max_{x_r \in I_r, \forall r} L(\mathbf{x}, \mathbf{p}) = \left\{ \sum_{s \in R_{BE}} B_s(p^s) + \sum_{j \in J} p_j c_j \right\}$$

where:

$$B_r(p^r) = \max_{x_r \in [m_r, M_r]} (U_r(x_r) - x_r p^r) \tag{3}$$

$$p^r = \sum_{j \in J_{BE}(r)} p_j \tag{4}$$

Thus, the dual problem for Eq. (1) is:

$$\mathbf{p}^o = \arg \min_{\mathbf{p} \ge 0} D(\mathbf{p}) \tag{5}$$

The first term of the dual objective function $D(\mathbf{p})$ is decomposed into $|R_{BE}|$ separable subproblems (3) and (4). If we interpret p_j as the price per unit bandwidth at link j then p^r is the total price per unit bandwidth for all links in the path of user r. Hence, $x_r p^r$ represents the 'shadow price' for user r [41,42], namely, the bandwidth cost for user r, when it transmits at rate x_r, under the current network conditions in terms of the congestion of the links and of the other users' utility functions. $B_r(p^r)$ is the benefit that can be achieved by user r at the given price p^r under the current network conditions.

Define by $x_r(\mathbf{p})$, $r \in R_{BE}$ the solution to Eq. (3) for a given $\mathbf{p}$. Such $x_r(\mathbf{p})$ may not be the solution of Eq. (1), but if each subproblems (3) takes p^r from the solution of Eq. (5), then the primal optimal source rate $\mathbf{x}^o$ of Eq. (1) can be computed by each individual user r. It is worth noting that, given $\mathbf{p}^o$, individual BE users can solve Eq. (3) separately, without the need to coordinate with other users. In practice, $\mathbf{p}^o$ serves as a coordination signal that aligns the individual optimality of Eq. (3) with the social optimality of Eq. (1). The dual problem can be solved using a gradient projection method, where link prices are adjusted in opposite direction to the gradient $\nabla D(\mathbf{p})$:

$$p_j(t+1) = \left[p_j(t) - \eta \frac{\partial D(\mathbf{p}(t))}{\partial p_j} \right] \tag{6}$$

Noting that $\partial D(\mathbf{p}(t))/\partial p_j = c_j - \sum_{r \in R_{BE}(j)} x_r(\mathbf{p}(t))$, also the solution of the dual problem (5) can be achieved in a decentralized way. Namely, Eq. (6) can be implemented by individual links using only the local information $\sum_{r \in R_{BE}(j)} x_r(\mathbf{p}(t))$ that is the aggregate source rate at link j. At each iteration, user r individually solves Eq. (3) and communicates the results $x_r(\mathbf{p})$ to each link $j \in J_{BE}(r)$ on its path. Each link j then updates its price p_j according to Eq. (6), it communicates the new prices to user r, and the cycle repeats.

An alternative decomposition is proposed in Ref. [13]: the SYSTEM problem (1) can be decomposed, by separately considering an ISP part and a user part. Let $w_r = p^r x_r$, $r \in R_{BE}$, be the shadow price [41,42] for user r, i.e. the price per time unit that user r is willing to pay. Let $\mathbf{w} = [w_r, r \in R_{BE}]$; each BE user solves the following optimization problem.

The USER$_r$ problem:

$$w_r^o = \arg \max_{w_r:\, w_r/p_r \in I_r} \left[U_r\left(\frac{w_r}{p^r}\right) - w_r \right] \tag{7}$$

subject to: $w_r \ge 0$. In practice, a software agent periodically contracts with the network the bandwidth allocation x_r of each user r, it computes w_r in function of its utility, and sends it to the network [35]. The ISP, instead, has to solve the following optimization problem.

The NETWORK problem:

$$\mathbf{x}^o = \arg \max_{x_r \in I_r, \forall r \in R_{BE}} \sum_{s \in R_{BE}} w_s \log x_s \tag{8}$$

under the following constraints:

$$\mathbf{A} \cdot \mathbf{x} \le \mathbf{c} \qquad \mathbf{x} \ge 0$$

Given the vector $\mathbf{w}$ the network computes $\mathbf{x}$, and sends it as a feedback to the flow controller of each user r. Asynchronous distributed approaches to the NETWORK problem have been developed in several works (see e.g. Refs. [13,17]). Malinowski [17] suggests the use of feedback from the real system, while in Ref. [13] it is shown that modeling flow control dynamics through suitable differential equations, can yield to an arbitrarily close approximation to the solution of the problem. More specifically, cost functions are defined for each link j, of the type:

$$\mu_j(t) = \gamma_j \left(\sum_{r \in R_{BE}(j)} x_r(t) \right) \tag{9}$$

where the arguments of the functions $\gamma_j(\cdot), j \in J$, represents the total rates on the link j of the network. Such functions should set a penalty on an excessive use of the resource. The following dynamic system, including pricing and flow control is considered in Refs. [13,17]:

$$\frac{\mathrm{d}}{\mathrm{d}t} x_r(t) = k_r \left(w_r - x_r(t) \sum_{j \in J_{BE}(r)} \mu_j(t) \right) \tag{10}$$

The interpretation in terms of flow control is as follows:

- each link j generates feedback signals according to $\gamma_j(y)$, where y is the flow traversing it;
- the feedback is interpreted as a congestion indicator by the users;
- each user's rate grows with rate w_r and decreases proportionally to the feedback.

It can be shown that, under not too restrictive hypotheses on the form of functions $\gamma_j(\cdot)$, the system of differential equations is globally stable and by adapting the prices w_r according to the solutions of the USER$_r$ problems, the PFP optimum of Eq. (1) can be reached [13].

3. Guaranteed Performance service: Call Admission Control and pricing

In this section we shall consider the presence in the network of both BE traffic and of traffic explicitly requiring guaranteed QoS (GP traffic). In this context, our goal is to influence the BE traffic flow control and to apply a CAC to the GP traffic in order to maximize the ISP's overall revenue. A GP user does not request a variable bandwidth allocation according to a utility function as in the PFP scheme, as it is interested in receiving a fixed bandwidth pipe that cannot vary, in spite of the network traffic conditions. Therefore, the request can be accepted only if there are enough bandwidth resources along the required routing path.

In the literature, when CAC in telecommunications networks is addressed, models for the minimization of a weighted sum of the blocking probability of the incoming traffic classes, each of which is associated to a particular reward for the ISP's revenue, are proposed (see for example Refs. [27–29]). Other works propose models for a static or a dynamic allocation of the prices in order to maximize the ISP's revenue [37,38]. In these papers, the minimization of the blocking probability corresponds to the maximization of the ISP's revenue. In this section we investigate the fact that, if BE PFP users are multiplexed together with GP users, a trade-off between GP and BE revenues should be taken into account. In fact, if also big users can subscribe a PFP traffic contract [34], accepting a new GP connection does not always increase the ISP's revenue, owing to the fact that the BE revenue decreases on the links crossed by such new GP connection. We actually believe that future telecommunication networks will be heterogeneous systems, namely, neither the GP services nor the BE services will be proposed as the unique models for the establishment of a traffic contract between an ISP and the customers. Some customers could be interested in paying a tariff for the establishment of a service with some strict QoS requirements and could be priced according to a GP tariff category. On the other hand, other customers could be interested in paying a tariff for the establishment of a service with a lighter and variable QoS requirements that can also adapt to the network congestion conditions and can be priced according to a BE tariff category, for example according to the PFP scheme. Clearly, as also big users can subscribe a PFP traffic contract, the bandwidth allocations and the prices of the GP users will not necessarily be greater than those of such BE users.

In the following we shall propose three strategies that influence the amount of resources to allocate to the GP traffic, taking into account the bandwidth allocation of the BE one. In our first strategy, a price is given for all the GP traffic and we decide whether it is suitable or not to accept the requests of incoming connections on the basis of a revenue derivative comparison. In the second one, all the requests that pass the CAC bandwidth availability check are accepted, but every GP user is assigned a new price. In the third one, we decide whether it is suitable to accept an incoming GP connection if it increases the estimated overall revenue at the end of its duration. In our model of the network, the total ISP's revenue per unit time G (for example expressed in €/s) is formed by the sum of two terms, concerning the GP and the BE traffic, respectively:

$$G = G_{GP} + G_{BE} \tag{11}$$

The revenue concerning the BE users is given by the $\mathbf{w}^o$ vector as the optimal solution of the corresponding PFP problems (1) and (7) (given the vector $\mathbf{x}^o$ from Eq. (1), the w_r^o of each user is obtained by Eq. (7) as $w_r^o = x_r^o U'(x_r^o)$), while that of the GP traffic is obtained by multiplying the Effective Bandwidth [5] by the assigned charge. The ISP assigns a reserved bandwidth y_r to each user $r \in R_{GP}$, where we denote by R_{GP} the set of active GP users (i.e. all of the GP connections accepted in the network and in progress). Every user r, $r \in R_{GP}$, pays an amount b_r per unit of sent GP traffic data per unit time (e.g. b_r could be €/Mbps per minute). Each time a new GP call asks to enter the network (i.e. a new GP user $\tilde{r}$ wants to start up a connection), the network is asked for a new amount of bandwidth $y_{\tilde{r}}$. The GP's required bandwidth might be computed, for example, by using a technique like the equivalent bandwidth in ATM, and, anyway, the feasibility of the request with respect to the available capacity should be tested [44]. We suppose the BE traffic to be regulated by a flow control mechanism such as in Section 2; so, before the new GP user $\tilde{r}$ enters the network, the rates x_r and the prices per unit time w_r of the current BE users $r \in R_{BE}$, have reached the stationary optimal values x_r^o and w_r^o. If the new bandwidth $y_{\tilde{r}}$ will be reserved for the new GP user $\tilde{r}$, the BE traffic rates x_r^o and price w_r^o, $r \in R_{BE}$, will move to the new optimal values $\tilde{x}_r^o$ and $\tilde{w}_r^o$ according to Eqs. (1) and (7), where the capacity constraints in Eq. (1) become:

$$\mathbf{A}\cdot\mathbf{x} \le \tilde{\mathbf{c}} \tag{12}$$

where $\tilde{\mathbf{c}} = [\tilde{c}_j, j \in J]$ is the residual capacity matrix, with $\tilde{c}_j = c_j - \sum_{r \in R_{GP}, j \in r} y_r$ the residual capacity (capacity not reserved to GP traffic) of link j.

3.1. First CAC control rule: CACPricing1

It is clear that the revenue's rates ($G_{BE} = \sum_{r \in R_{BE}} w_r^o$ for BE traffic and $G_{GP} = \sum_{r \in R_{GP}} b_r y_r$ for GP traffic) change with the traffic change. If a new GP user $\tilde{r}$ enters the network, the GP revenue rate, G_{GP}, increases, while the BE traffic rates decrease (less bandwidth available for BE), so that the BE revenue rate, G_{BE}, decreases, too. In this respect, a possible Acceptance Control Rule for the requests of increasing the GP traffic reserved bandwidth is to accept the new GP bandwidth reservation only if the total revenue rate increases with respect to the current situation. So, in our first proposal, we use the revenue rate to decide whether to accept a new GP request. In particular, let y_r, $r \in R_{GP}$, be the current GP bandwidth reservations, and $y_{\tilde{r}}$ a new bandwidth request for the new GP user $\tilde{r}$ with associated tariff $b_{\tilde{r}}$. The ISP accepts the new bandwidth request if:

$$\sum_{r \in R_{GP}} b_r y_r + b_{\tilde{r}} y_{\tilde{r}} + \sum_{r \in R_{BE}} \tilde{w}_r^o \ge \sum_{r \in R_{GP}} b_r y_r + \sum_{r \in R_{BE}} w_r^o \tag{13}$$

where the $\tilde{w}_r^o$ represent the optimal values of the BE price in the presence of the new GP allocation $y_{\tilde{r}}$.

This results in the following CAC rule.

CACPricing1. Accept any new GP call if it passes the CAC bandwidth availability check and if

$$b_{\tilde{r}} y_{\tilde{r}} \ge \sum_{r \in R_{BE}} (w_r^0 - \tilde{w}_r^0) \tag{14}$$

In this context, the choice of the prices is static [37], namely, the GP prices are chosen according to an off-line planning of the telecommunication network and do not depend on the current utilization of the resources.

3.2. Second CAC control rule: VariableGPPrice

If the GP prices can be freely assigned by the ISP every time a connection is accepted, it is possible to assign them in order to leave the total revenue derivative unchanged:

$$\sum_{r \in R_{GP}} b_r y_r + b_{\tilde{r}} y_{\tilde{r}} + \sum_{r \in R_{BE}} \tilde{w}_r^o = \sum_{r \in R_{GP}} b_r y_r + \sum_{r \in R_{BE}} w_r^o \tag{15}$$

Imposing condition (15) leads to the following pricing scheme for GP users.

VariableGPPrice. Accept any new GP call if it passes the CAC bandwidth availability check and fix its tariff as

$$\bar{b}_{\tilde{r}} = \frac{\sum_{r \in R_{BE}} (w_r^0 - \tilde{w}_r^0)}{y_{\tilde{r}}} \tag{16}$$

In this way, we apply a dynamic pricing policy, where a price is assigned to the incoming GP connection to exactly equal the revenue which will be lost on the BE traffic. As said before, a possible choice for the ISP to maximize its global revenue could consist in imposing very high tariffs to any incoming GP user. Such a policy would be 'unfair' for the GP users. In a sense, the 'VariableGPPrice' strategy can be interpreted as a way to fix a lower bound to the GP prices. Any $b_{\tilde{r}} > \bar{b}_{\tilde{r}}$ would augment the ISP's global revenue.

VariableGPPrice is a dynamic pricing policy [37], because the GP prices are set in function of the current utilization of the resources. It could be interpreted as an extension to the GP traffic of the shadow prices applied to the BE traffic. With this strategy, in fact, we want to assign a price to each new GP connection in function of the decay of performance of the BE traffic when bandwidth $y_{\tilde{r}}$ is no longer available for the latter. In fact, bandwidth for BE traffic decreases along the routing path of the new incoming GP connection; every BE user that has at least one routing link in common with such new GP connection, will receive less bandwidth resources and, consequently, a lower price. This situation remains stable until a new event occurs in the system (such as the termination or an arrival of a BE or GP call). A possible drawback of this strategy stems form the fact that in the PFP scheme, the tariffs are imposed only in function of the current network congestion condition and of the BE users' utility functions. Neither the costs for the building and the maintenance of the network infrastructure, nor any ISP's desirable profit are taken into consideration.

For both strategies presented so far, every time the CAC block acts, it is necessary to foresee the revenue $\tilde{G}_{BE} = \sum_{r\in R_{BE}} \tilde{w}_r^o$ which will be obtained in the future on the BE traffic after the bandwidth reallocation. The basic idea is to use the SYSTEM problem (1) to calculate the new value of the **x** vector after the bandwidth reallocation; then, using the USER_r problem (7), it is possible to calculate the new BE prices and the new BE revenue and evaluate the total revenue after the possible bandwidth reallocation. Eqs. (1) and (7) are mathematically fairly tractable [13], so it seems they can be actually applied on-line by the CAC block during the network evolution. 'CACPricing1' and VariableGPPrice are only based on the current PFP bandwidth and prices allocations. The PFP model does not consider the dynamic nature of the network [37], namely, call arrivals and departures are not taken into account. So, in Eqs. (13) and (15) only the revenues per unit time and not the total revenues are compared. These two strategies act in fact following an 'Open Loop Feedback Control' approach [22]: they exploit information on the current congestion situation on the network links and current utility functions (feedback), but they do not take into consideration what could happen in the future (open loop) in terms of all the possible terminations of connections actually present in the network and in terms of the possible arrivals of new connections.

3.3. Third CAC control rule: accept the incoming GP requests if, at the end of its duration, it increases the total revenue

If the lengths of the connections were explicitly considered, the ISP's decisions would be different from those induced by CACPricing1 and VariableGPPrice. For example, it might happen that, having accepted a new GP connection, the network is forced to refuse other connections, because it does not have any more bandwidth available for the incoming requests; on the other hand, some of these refused connections might contribute to increase the total revenue much more than the accepted one. Taking into account the length of the connections would lead to maximize the total revenue, rather than the revenue rate. In order to accomplish this task, it would be necessary to identify all the possible events that happen during each new incoming GP connection lifetime and to solve the SYSTEM problem (1) for each time interval between them. In fact, every time a (BE or GP) connection ends, more bandwidth becomes available for the BE traffic, so the system converges to a new balance of the total revenue. In the following, we propose a heuristic technique that considers this additional information and takes into account the possible future opening of new BE and GP connections.

Again, let $\tilde{r}$ be the new incoming user asking, at time $\tilde{t}$, for a GP connection, expected to terminate at $\tilde{T}$. Suppose $\tilde{r}$ to be subject to the CAC based on the total revenue comparison and let $G_{TOT}(\tilde{t},\tilde{T})$ and $\tilde{G}_{TOT}(\tilde{t},\tilde{T})$ be the total revenues the ISP expects to obtain in $[\tilde{t},\tilde{T}]$ in the case $\tilde{r}$ is refused or accepted, respectively. To calculate the two terms $G_{TOT}(\tilde{t},\tilde{T})$ and $\tilde{G}_{TOT}(\tilde{t},\tilde{T})$ it is necessary to break $[\tilde{t},\tilde{T}]$ in all of the sub-intervals where there are no changes in the **w** and **x** vectors. To take into account also the arrival of new requests from GP and BE traffic, we have used a heuristic approach based on Montecarlo simulation: when a new request of GP connection occurs, we generate n different simulation runs of length $[\tilde{t},\tilde{T}]$. At the end of each simulation we calculate the overall revenue, considering the terminations of the BE and GP connections within $[\tilde{t},\tilde{T}]$ and the arrivals of new BE and GP connections within the same time interval. In this way, all of the bandwidth reallocations occurring during the new connection are considered. In the n different simulations, as to the new GP connections starting after $\tilde{t}$, a CAC strategy based only on the bandwidth availability is applied and an estimate of the expectation of the revenue is computed. Two situations are considered: the case in which the new GP connection is refused (G_{TOT}), and the case in which the new GP connection is accepted ($\tilde{G}_{TOT}$). The final choice is to accept the incoming GP user $\tilde{r}$ only if it increases the estimate of the mean value of the total revenue obtained at the end of the recursive procedure ($\tilde{G}_{TOT} \geq G_{TOT}$). In the following, we present the pseudocode of the procedure used to decide whether to accept or not the incoming GP connection by applying this strategy.

This strategy (called 'CACPricing2' (Fig. 1) in the following), as the CACPricing1, is a static pricing policy and it is related to the family of the so-called 'Receding Horizon' techniques. A performance index (the revenue) that is referred to a finite temporal window (the duration of the new GP connection) is maximized. The perfect information on the termination instants of all the connections that will end during the new GP connection is exploited and averaging is performed on the arrivals of the new GP and BE calls. All of this additional information leads to an estimation of the expectation of the overall revenue that could be obtained accepting or refusing the new GP call. It is clear that this approach is very time consuming and cannot be applied in a real scenario where the CAC block acts on-line, but it could be very useful to test the performance of the previous techniques, where only an optimization on the revenue rate is applied and the events that can occur during the new GP connection life time are ignored.

4. Numerical results

We have developed a simulation tool that describes the behaviour of the network at call level to verify the performance of the proposed price-based CAC mechanisms. The simulator does not model the packet level. The COST

CACPricing2.

$\tilde{t}$ = *arrival time of GP user* $\tilde{r}$;

$\tilde{T}$ = *termination time of the new GP connection;*

n : number of simulations of the network evolution;

T_EV: set of the time instants of all of the GP and BE connections' terminations inside $[\tilde{t}, \tilde{T}]$;

for i=1 to n

> *B_EV*^{i}*: i-th event list of BE and GP births inside* $[\tilde{t}, \tilde{T}]$ *and corresponding death times generated according to distributions used by the main simulation;*
>
> *EV*^{i}*= T_EV* $\cup$ *B_EV*^{i}*;*
>
> *On the basis of EV*^{i}*:*
>
> - *Simulate the evolution of the network in the case* $\tilde{r}$ *is refused;*
> - G^i_{TOT} = *resulting overall revenue obtained in* $[\tilde{t}, \tilde{T}]$;
> - *Simulate the evolution of the network in the case* $\tilde{r}$ *is accepted;*
> - $\tilde{G}^i_{TOT}$ = *resulting overall revenue obtained in* $[\tilde{t}, \tilde{T}]$;

$E[G_{TOT}] = \sum_{i=1}^{n} G^i_{TOT} / n;$ *//Estimation of the expectation of the revenue at the end of* $\tilde{T}$ *in the case* $\tilde{r}$ *is refused:*

$E[\tilde{G}_{TOT}] = \sum_{i=1}^{n} \tilde{G}^i_{TOT} / n;$ *//Estimation of the expectation of the revenue at the end of* $\tilde{T}$ *in the case* $\tilde{r}$ *is accepted*

if $E[\tilde{G}_{TOT}] \geq E[G_{TOT}]$ *then "accept* $\tilde{r}$ *" else "refuse* $\tilde{r}$ *";*

Fig. 1. Pseudocode of the procedure used by the CACPricing2 strategy.

239 experimental network [23] (depicted in Fig. 2) is utilized for the tests; it is composed by 20 links and by 11 nodes. We consider a subset of 10 active routes, where each active route can generate both BE and GP traffic connections:

Route 1: {0, 1, 5, 7, 10}
Route 2: {4, 8, 12, 19, 15}
Route 3: {14, 17, 16}
Route 4: {2, 5, 6, 12}
Route 5: {3, 8, 9, 10}
Route 6: {5, 7, 9, 11}
Route 7: {3, 4, 13}
Route 8: {1, 19, 12}
Route 9: {17, 18, 11}
Route 10: {3, 4, 5, 7, 10, 16}

Our test scenarios are very close to the simulations tests of some among the most relevant papers on CAC and Pricing in telecommunication networks (see for example Refs. [27,28] for the pure CAC and Refs. [34,37–39] for the Pricing). We have imposed a probability distribution over all of the significant variables of the problem: interarrival times of the BE and GP users, required bandwidth and utility functions, to produce variable traffic conditions.

The simulations performed fall in the category of the so-called 'finite time horizon' or 'terminating' simulations [24]. For computational time reasons (in particular for the CACPricing2 strategy) the Independent Replications technique for the analysis of stochastic simulation systems [24] (i.e. the repetition of the same simulation with different pseudorandom number generators until a confidence interval is reached for the performance parameter) could not be applied. The pricing strategies are compared in terms of the total revenue and blocking probability of GP traffic. A fixed sequence of realizations of the stochastic processes involved in the problem has been used (i.e. the comparison is done over a sample path of events). In this way it is guaranteed that the characteristics of the requests of all the traffic classes are

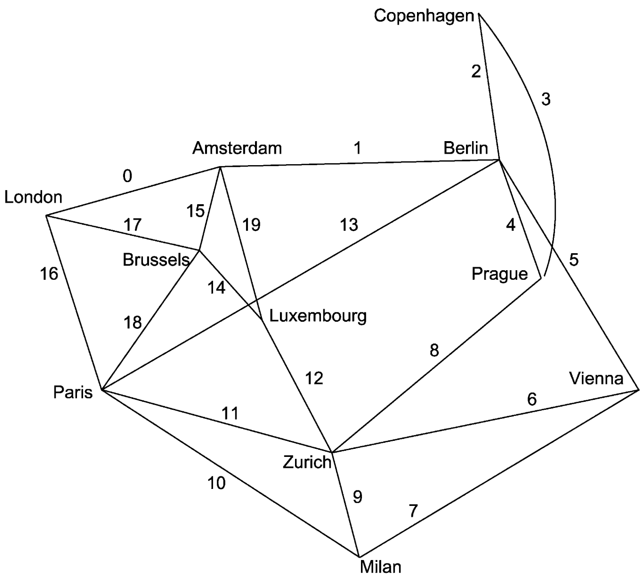

Fig. 2. Topology of the test network.

identical in each simulation where a different Pricing technique is applied.

We have defined as 'static' a scenario where each parameter follows a probability distribution with a fixed mean value. This corresponds to a real situation where the users' behavior, for each traffic class, remains the same during the simulation (for example in terms of mean arrival rate or mean duration time of the calls). On the contrary, we have defined as 'dynamic' a scenario where each parameter follows a probability distribution with different mean values, namely, the users' behaviour can change during the simulation. Both scenarios have been used for the tests.

4.1. The static scenario

Connections are generated following independent Poisson distributions with mean rate $\lambda_r^{(BE)}$ and $\lambda_r^{(GP)}$ for each route r, for BE and GP traffic, respectively. The call durations follow independent exponential distributions with mean values $1/\mu_r^{(BE)}$ and $1/\mu_r^{(GP)}$. The bit rate of the BE traffic is controlled according to the PFP scheme using Eqs. (1) and (7). Each GP call requires an amount of bandwidth generated from an exponential distribution with mean value γ. We use the following utility function for the BE traffic:

$$U(x) = \alpha\sqrt{x} \tag{17}$$

where the parameter α is generated with an exponential distribution with mean value $\bar{\alpha} = 1$.

The simulation data are summarized in the following:

- $\lambda_r^{(BE)} = \lambda_r^{(GP)} = \lambda = 10$ calls per minute $\forall r \in \{1, \ldots, 10\}$
- $1/\mu_r^{(BE)} = 1/\mu_r^{(GP)} = 1/\mu = 1$ minute $\forall r \in \{1, \ldots, 10\}$
- $c_j = c = 5$ Mbps (link capacity), $\forall j \in \{0, \ldots, 19\}$
- $\gamma = 1$ Mbps (average bandwidth required by a GP call)
- Time of simulation: 100 min
- n-number of simulations of the procedure used by the CACPricing2 strategy (Fig. 1): 5.

The results obtained with the proposed CAC rules are compared with fixed CAC rules, which accept a constant percentage p of calls that do not violate the bandwidth constraints. This means that a CAC checking bandwidth availability is applied when a call enters the network. Among the calls that respect this rule, only the percentage p is really accepted. We have considered three different percentages: $p = 100\%$ ('AlwaysAccept' strategy), $p = 50\%$ ('HalfAccept strategy') and $p = 0\%$ ('NeverAccept' strategy). Clearly, also the NeverAccept strategy generates revenue: all of the GP calls are refused and the overall ISP's revenue comes from the BE traffic. These fixed CAC rules are aimed at representing two extreme conditions, and an average one, as well, concerning the acceptance of the incoming GP requests; in this way the contribution of the GP traffic to the total revenue is highlighted. Fig. 3 shows the total revenue for two different choices of the price b (money per unit of sent traffic data the GP user pays for:

- Case 1 $b = 0.1$ €/Mbps per min.
- Case 2 $b = 1$ €/Mbps per min.

These two price choices are aimed at evaluating, in a static scenario, the two typical revenue situations of the system. As we can see from Fig. 3, the best fixed strategies in terms of achieved revenue are: the AlwaysAccept strategy for $b = 1.0$ and the NeverAccept strategies for $b = 0.1$. In fact, in the case $b = 0.1$ the BE traffic contributes for the largest part of the total revenue; on the other hand, GP traffic represents the largest part of the total revenue in the case $b = 1.0$. When b is very low (e.g. $b = 0.1$), the maximum revenue is obtained when the GP traffic is not accepted (NeverAccept technique); if the value of b is higher (e.g. $b = 1.0$), the best performance is obtained using the AlwaysAccept technique. This behaviour is due to the fact that, if the price paid by the GP users is low, it is convenient to refuse all of the GP calls, leaving all the bandwidth to the BE traffic. On the other hand, in the opposite situation, giving all the bandwidth to the GP traffic is the most convenient choice in terms of the ISP's revenue. The CACPricing1 technique offers a good level of performance in both cases, showing a good adaptation to the GP price changes. The behaviour of the Variable-GPPrice technique does not depend on b; the tariff of each connection is dynamically decided on the basis of formula (16). The obtained performance turns out to be the best one in absolute in the $b = 0.1$ situation, while it appears quite poor if $b = 1.0$. This shows that the $b = 0.1$ value is too low (GP users pay less than they would have to), while $b = 1.0$ is slightly too high.

If the revenue performance of CACPricing1 is compared with that obtained by CACPricing2, the latter seems not to be able to produce significant improvements, despite the greater computational complexity. This might be due

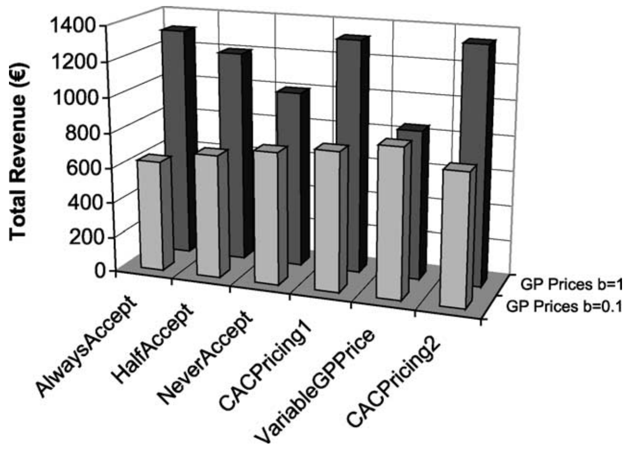

Fig. 3. Static simulation scenario #1. Total revenue [€] with two values of GP prices.

to two reasons:

- The CACPricing1 technique (which is based only on the maximization of the actual revenue per unit time) already works well and the inaccuracy that is made ignoring the opening and the termination of new connections in the future is not so significant.
- The CACPricing2 technique should work better if the number of recursive simulations would increase, but this is a very time consuming approach, so it is quite infeasible, owing to the great computational time required. We have tried to do it anyway up to $n = 10$, but without any significant improvement.

Concerning the Blocking Probability (depicted in Fig. 4), CACPricing1 offers good performance if $b = 1.0$. In the $b = 0.1$ case, instead, more than 67% of the GP connections are refused. However, it is important to remind that from the point of view of the revenue, in this situation, the best solution was NeverAccept: CACPricing1 obtains similar revenue, but with a much lower blocking probability. In terms of the GP blocking probability CACPricing2 offers lower values than the CACPricing1 ones. So, we can see that the greater computational complexity of the CACPricing2 has more impact on the blocking probability than directly on the revenue. As obvious, the technique VariableGPPrice has the same blocking probability as the AlwaysAccept strategy, because it accepts all the connections that pass the first CAC level based on the bandwidth availability.

Let us consider another static situation very close to the previous one, but where the willingness to pay of the BE users is increased by an order of magnitude; namely, the parameter α of Eq. (17) is generated with an exponential distribution with average value $\bar{\alpha} = 10$. The results are summarized in Figs. 5 and 6. In this situation, for both the $b = 0.1$ and the $b = 1.0$ cases, not accepting the GP calls results to be the best choice, because the tariffs and the associate revenue regarding the GP traffic are very low. The BE users' availability to pay dominates. Again the proposed CAC strategies determine the maximum revenue and a lower blocking probability, too.

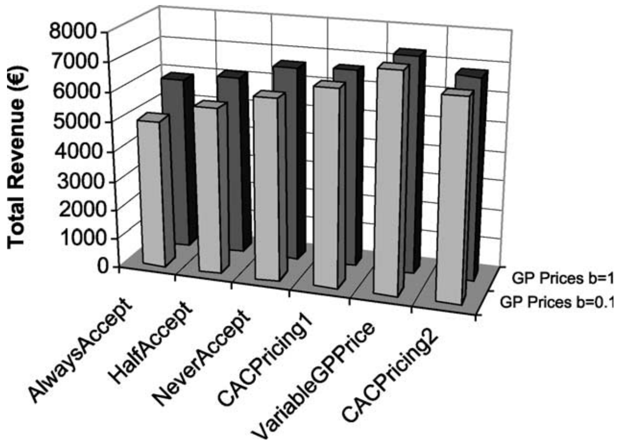

Fig. 5. Static simulation scenario #2. Total revenue [€] with two values of GP prices.

On the basis of these results we can note that all of the static scenarios are characterized by one of these possible situations:

- *predominance of BE revenue* (Static simulation scenario #2 and Static simulation scenario #1, $b = 0.1$ case): in this case CACPricing1 and CACPricing2 offer good performance, but with high blocking probability, while VariableGPPrice maximizes the total revenue and minimizes the blocking probability.
- *predominance of GP revenue* (Static simulation scenario #1, $b = 1.0$ case): VariableGPPrice decreases its revenue performance because the prices (just coming from the change imposed on the revenue of the BE traffic) are much lower than the imposed fixed price. CACPricing1 and CACPricing2 provide good results without increasing too much the blocking probability.

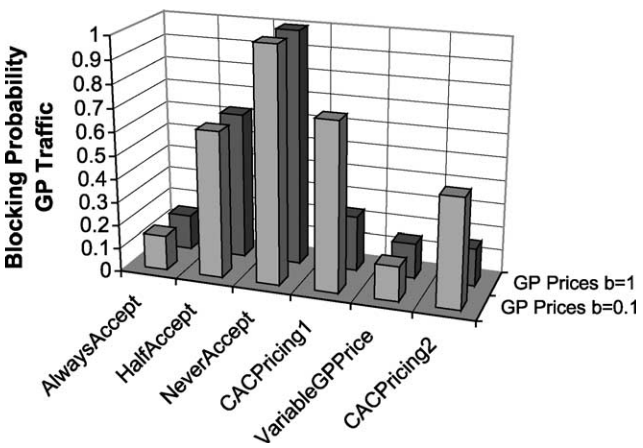

Fig. 4. Static simulation scenario #1. GP traffic Blocking Probability, with two values of GP prices.

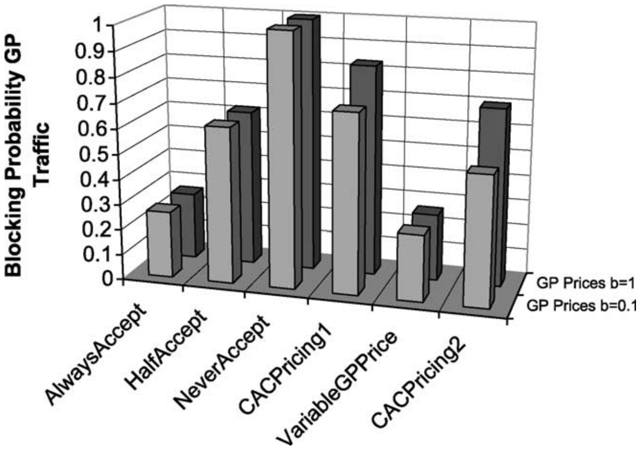

Fig. 6. Static simulation scenario #2. Blocking Probability for GP traffic, with two values of GP prices.

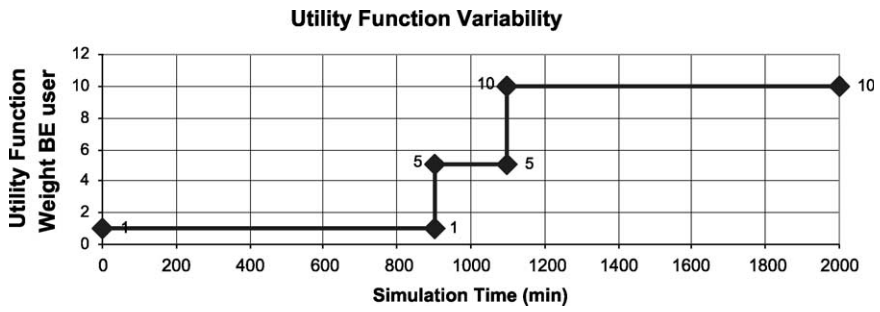

Fig. 7. Dynamic simulation scenario #1. Utility function variability.

4.2. The dynamic scenario

Now we consider a situation in which the volume of the traffic and the users' behaviour can change within the same simulation. In fact, it is well known that the traffic profile depends on the period of the day (see e.g. Refs. [25,26]) most of the traffic carried during the day is professional traffic (e.g. between companies), consisting prevalently of GP traffic, while BE traffic (for example, residential traffic) dominates in the evening. Moreover, if also big users can subscribe a PFP traffic contract and vary the form of their utility functions, in order to comply with variable QoS requirements [34,35], it is reasonable to expect that any 'good' CAC strategy can optimally react to such variable traffic conditions. We consider two different dynamic scenarios: in the first one the utility functions are changed in order to increase the willingness to pay of the BE users, in the second one the mean interarrival time of the BE calls is increased. Each of these scenarios is aimed at simulating a strong change over the BE traffic behaviour.

From Section 4.1, we can note that the fixed strategy that maximizes the revenue is AlwaysAccept in the first static scenario when $b = 1$ and NeverAccept in the second static scenario, again when $b = 1$. In both circumstances, the strategies proposed in this paper provide the maximum revenue, so it is reasonable to expect that, in dynamic conditions, they are able to provide better CAC choices than the fixed techniques.

Consider now a situation in which the willingness to pay of BE users changes a number of times within the same period of simulation, e.g. the utility functions for the BE traffic are generated according to Eq. (17), where the parameter α is generated with an exponential distribution with mean value increasing from 1 (first static scenario) to 10 (second static scenario) (see Fig. 7).

The other simulation data are the same as in the previous static scenario, but the simulation time is increased to 2000 min. The results are summarized in Figs. 8 and 9. Observing the values of the revenue obtained at the end of the simulation period it is clear that, in dynamic conditions, the fixed strategies do not succeed in optimizing the overall revenue. The proposed strategies can better suit dynamic traffic conditions in terms of utility functions variability.

We now consider a different dynamic situation, where the willingness to pay remains the same, but there is a strong increase in the interarrival frequency of the BE users from 1 call per minute in the first 1000 min to 20 calls per minute in the last 800 min (Fig. 10). The utility functions for the BE traffic are generated according to Eq. (17), where the parameter α is generated with an exponential distribution with mean value $\bar{\alpha} = 1$ and the other simulation data are the same as in the previous static scenario, but again with an increase in the simulation period to 2000 min.

As in the previous dynamic scenario, we can see from Figs. 11 and 12 that the proposed strategies guarantee the maximization of the revenue. Moreover, we can also see that the CACPricing2 strategy maintains a higher revenue and a lower blocking probability than those of the CACPricing1 strategy. In dynamic traffic conditions, the differences in performance obtained by these two strategies are more evident than in the static case. Despite the fact that

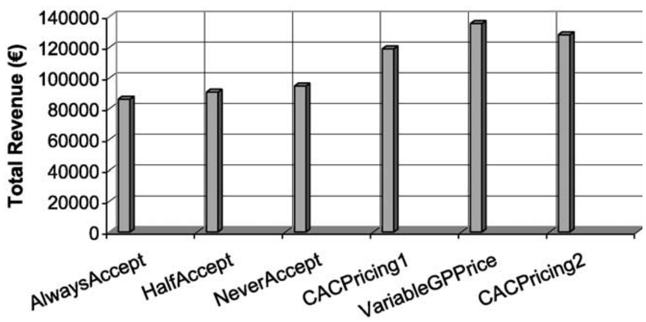

Fig. 8. Dynamic simulation scenario #1. Total revenue [€].

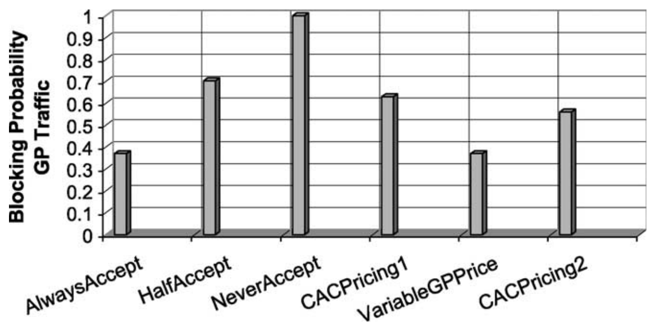

Fig. 9. Dynamic simulation scenario #1. GP traffic Blocking Probability.

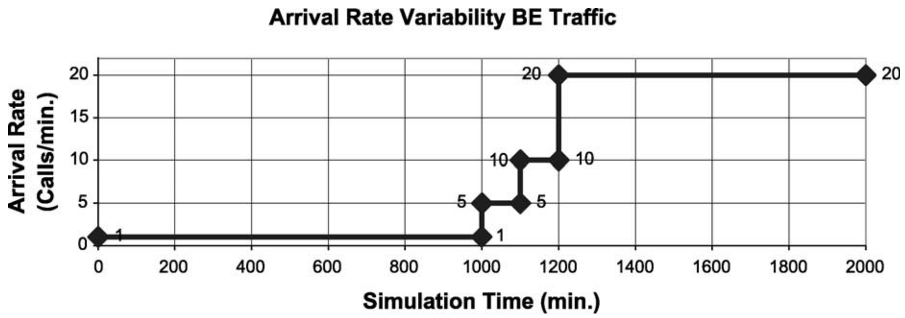

Fig. 10. Dynamic simulation scenario #2. Arrival rate variability (BE traffic).

such differences exist, we can say that the approximation applied by the CACPricing1 is acceptable in terms of the revenue and blocking probability performance.

4.3. The case of imperfect information on the utility functions

Until now, we have supposed the ISP to know the utility functions of BE users perfectly. When we apply a CAC rule, using Eqs. (1) and (7), we foresee the BE revenue ($\tilde{G}_{BE} = \sum_{r \in R_{BE}} \tilde{w}_r^o$) exploiting all of the information available from the network traffic conditions (i.e. routing and utility functions of the active BE users). While it is reasonable to perfectly know the routing of each active connection, the same does not hold for the knowledge of the users' utility functions. It is more likely to only have partial information about them. When a BE user enters the network, it is automatically subject to the flow control and there is no way, for the ISP, to know perfectly its utility function. For this reason, it is necessary to test our CAC rules in the case that Eqs. (1) and (7) are updated only with an 'average' form of the utility functions.

We consider again the first dynamic scenario, where there is a strong variability in the form of utility functions. They are generated according to Eq. (17):

$$U(x) = \alpha\sqrt{x}$$

where the parameter α is a random variable with mean value $\bar{\alpha}$ increasing from 1 to 10 (see Fig. 7).

We call 'CACPricing1_MeanU' and 'VariableGPPrice_MeanU' the CACPricing1, VariableGPPrice techniques where only partial information about the utility functions of the active BE users is employed, namely Eqs. (1) and (7) are updated with the average utility function:

$$U(x) = \bar{\alpha}\sqrt{x}$$

where $\bar{\alpha}$ is the current mean value of α according to the current period of simulation. We consider two different probability distributions over α : an exponential distribution in the first case and a uniform distribution between 0 and $2\bar{\alpha}$ in the second one. We have chosen these probability distributions in order to evaluate the error of the CACPricing1_MeanU and VariableGPPrice_MeanU strategies in function of the increase in the variance of α around $\bar{\alpha}$. In fact, the exponential distribution guarantees a larger variance, i.e. a larger deviation, in the BE users' willingness to pay.

We considered the following strategies: NeverAccept (that is the best fixed CAC rule in the first dynamic scenario), CACPricing1, CACPricing1_MeanU, Variable-GPPrice and VariableGPPrice_MeanU. We can see from Figs. 13 and 14 and Tables 1 and 2 (Total Revenue and GP Blocking Probability) that the CACPricing1_MeanU increases the GP Blocking Probability of the CACPricing1 technique. This means that the imperfect information about the current BE traffic utility functions leads to the wrong choice in front of the incoming new GP requests, and, for this reason, it guarantees a smaller total revenue performance than CACPricing1. VariableGPPrice_MeanU, due to the estimation, affects the charge applied to the incoming GP connection, leading to an increase of the revenue. It is important to note that the larger variance of the exponential probability distribution leads to a larger difference between

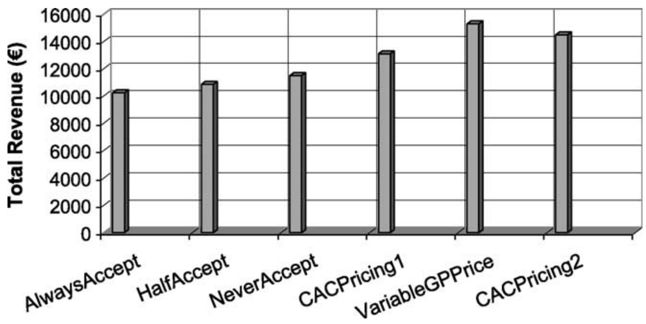

Fig. 11. Dynamic simulation scenario #2. Total revenue [€].

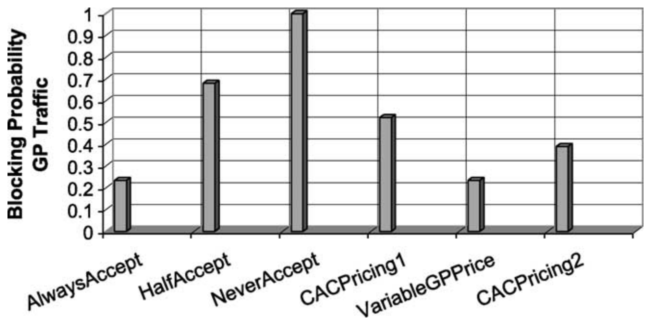

Fig. 12. Dynamic simulation scenario #2. Blocking Probability (GP traffic).

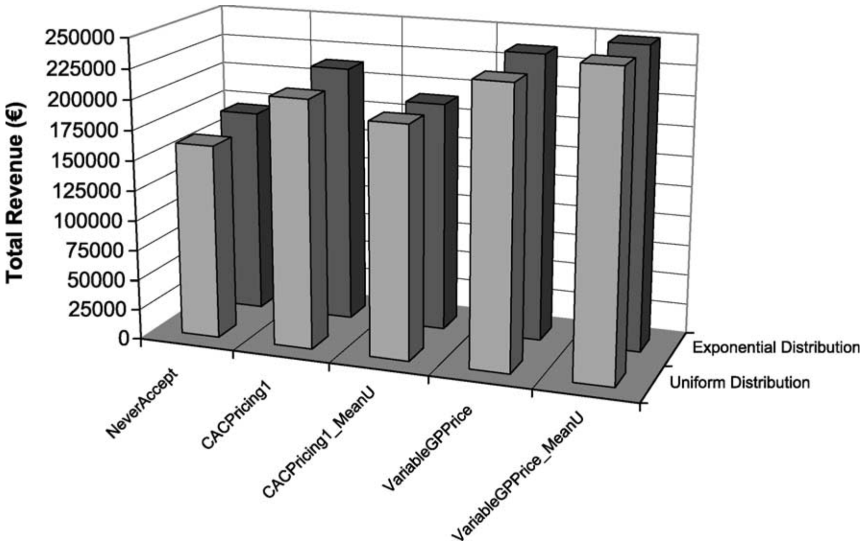

Fig. 13. Dynamic simulation scenario #1, the case of unknown utility functions: Total revenue [€].

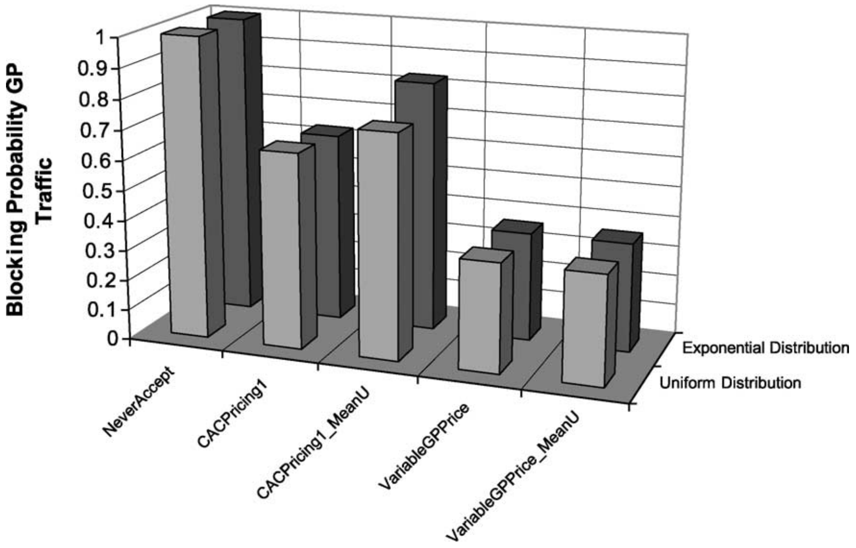

Fig. 14. Dynamic simulation scenario #1, the case of unknown utility functions: GP Blocking Probability.

the CACPricing1 and VariableGPPrice performance and the CACPricing1_MeanU and VariableGPPrice_MeanU performances. For example, the CACPricing1_MeanU performance decay is larger in the presence of the exponential probability distribution: the GP blocking probability differs from the CACPricing1 of 0.093 in the uniform distribution case and of 0.201 in the exponential distribution case. As to the total revenue, the difference is 13,352.4092 in the uniform distribution case and 23,425.286 € in the exponential distribution case.

5. Final remarks

In this work, we have proposed three Price-Based CAC rules for a heterogeneous environment of GP and BE connections. The first CAC rule is aimed at maximizing the ISP's revenue given the price imposed to the GP calls. It uses a revenue rate comparison based on the current network condition, so it can be applied in a real scenario. The optimization problems (1) and (7) used in our model to foresee the next PFP equilibrium are in fact mathematically

Table 1
Dynamic simulation scenario #1, the case of unknown utility functions: total revenue [€]

Total revenue [€]	Uniform distribution	Exponential distribution
NeverAccept	161,303.9668	170,212.1668
CACPricing1	205,956.3458	214,200.3458
CACPricing1_MeanU	192,603.9366	190,775.0598
VariableGPPrice	231,083.8	237,574.2
VariableGPPrice_MeanU	249,176.2765	269,351.2752

Table 2
Dynamic simulation scenario #1, the case of unknown utility functions: GP Blocking Probability

Blocking Probability	Uniform distribution	Exponential distribution
NeverAccept	1	1
CACPricing1	0.65	0.63
CACPricing1_MeanU	0.743	0.831
VariableGPPrice	0.362	0.362
VariableGPPrice_MeanU	0.362	0.362

fairly tractable [13]. The third CAC rule has the same aim and verifies, with a heuristic simulation approach, the performance of the first one. It is based on a comparison between the revenue obtained at the end of two different simulations regarding the network evolution in the cases when the incoming GP connection is accepted or refused, respectively. Clearly, the third CAC rule is more precise, because it considers the overall revenue at the end of each simulation, also taking into account the events that can occur in the system during the incoming GP connections' life time. From all of the presented simulations results, it is clear that the performance of the first CAC rule is quite close to that of the third one, so the method proposed with the first CAC rule effectively guarantees a good level of optimality for both the ISP's revenue and the GP blocking probability at a smaller computational effort. The second CAC rule is aimed at evaluating a bound over the prices imposed to the GP calls, if they are accepted after the bandwidth availability check. It is based on a comparison between the current shadow prices imposed to the BE users and it guarantees that the ISP, accepting the incoming GP call, does not decrease its overall revenue.

As regards the simulation scenarios, in a static situation of the system parameters, the proposed pricing mechanisms achieve a lower blocking probability than that of the fixed CAC strategy that maximizes the overall revenue; in a dynamic scenario, with time varying system parameters, we can see that the fixed strategies are not sufficient to maintain the highest revenue, while the proposed CAC rules optimize the overall revenue. In a dynamic scenario, an ISP, adopting only fixed strategies, can reach the best revenue only with a perfect estimate of the traffic variability and by calculating off-line the best fixed strategy that has to be used in every time interval where the traffic conditions reach an overall stability. On the contrary, adopting the proposed strategies, and keeping them updated with an estimate of the current traffic conditions, is sufficient to have the CAC block always supply the best choice. As we have seen in Section 4.3, our CAC strategies are also able to maintain a good level of optimality in the presence of an imperfect information about BE users' utility functions.

Finally, we can say that it is reasonable to expect that an Authority for the control of the market could force the ISP to fix the price for the GP traffic, in order to manage the establishment and maintenance of the network infrastructure, plus a tariff able to guarantee a further revenue, quite similar to the one achieved by the VariableGPPrice strategy. We know from Ref. [37] that, in large telecommunication networks, the revenue guaranteed by every optimal dynamic pricing strategy can be always reached by an optimal static pricing strategy if the statistics of the sources are quite 'regular' (i.e. stationarity and ergodicity of the interarrival times of the calls in function of the prices). We have found out something similar in this work, too. Given the revenue obtained by the VariableGPPrice strategy, it is always possible to find a static GP price that can ensure the same revenue applying the CACPricing1 strategy. We have also found that the dynamic strategy (the VariableGPPrice) has a much lower blocking probability than the static ones (CACPricing1 and CACPricing2). A lower blocking probability allows, at the same revenue, to satisfy a greater number of users. This is a also a useful property in terms of the ISP's revenue, if, for instance, a flat tariff is added to the price imposed to every user for the establishment of a new connection (e.g. a tariff in function of the cost of the ISP to signal the establishment of a new GP connection by using a resource reservation protocol [9,43]).

6. Conclusions and future work

In this paper we have proposed three price optimization mechanisms that operate in networks where there are both BE and GP traffics.

They are based on a decentralized flow control method (PFP) for the BE traffic and on three original CAC rules for GP calls. The simulation results presented show that all of the proposed mechanisms adapt well to traffic changes in order to maintain the best global revenue for the ISP.

Future work could include the analysis of the dynamics of the PFP optimum in the presence of fluctuations in the bandwidth allocation of the BE traffic and how this can influence the revenue forecast applied in our CAC rules. Moreover, it could be very interesting to add some mechanisms to both the PFP scheme and to CAC rules, which could take into account the costs for the building and the maintenance of the network infrastructure and guarantee a desiderable profit for the ISP. A model for the maximization of a unified social welfare for both the GP and the BE users is under investigation, too.

References

[1] C. Courcoubetis, V.A. Siris, G. Stamoulis, Integration of pricing and flow control for Available Bit Rate services in ATM networks, Proc. IEEE Globecom, London, UK (1996) 644–648.
[2] L.A. DaSilva, Pricing for QoS-enabled networks: a survey, IEEE Commun. Surv. 2nd quart. (2000) 2–8.

[3] M. Falkner, M. Devetsikiotis, I. Lambadaris, An overview of pricing concepts for broadband IP networks, IEEE Commun. Surv. 2nd quart. (2000) 2–8.
[4] F. Kelly, D. Songhurst, Charging schemes for multiservice networks, in: V. Ramaswami, P. Wirth (Eds.), Teletraffic Contributions for the Information Age, Elsevier, Amsterdam, 1997.
[5] F.P. Kelly, Charging and accounting for bursty connections, in: L.W. McKnight, J.P. Bailey (Eds.), Internet Economics, MIT Press, Cambridge, MA, 1996.
[6] J. Murphy, L. Murphy, Bandwidth allocation by pricing in ATM networks, Technical Report, Dublin City University, July 1994.
[7] J. Murphy, L. Murphy, E. Posner, Distributed pricing for embedded ATM networks, Proc. IFIP Conf. Broadband Commun. (BB-94), Paris, France, March (1994).
[8] J. Walrand, P. Varaiya, High-performance Communication Networks Ch. 10, Network Economics, 2nd ed., Morgan-Kaufmann, San Francisco, CA, 2000.
[9] R. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin, Resource ReserVation Protocol (RSVP), Version 1 Functional Specification, IETF, RFC 2205, September 1997.
[10] Integrated Services Working Group, http://www.ietf.org/html.charters/intserv-charter.html.
[11] J. Wroclawski, The Use of RSVP with IETF Integrated Services, IETF, RFC 2210, September 1997.
[12] Differentiated Services Working Group, http://www.ietf.org/html.charters/diffserv-charter.html.
[13] F.P. Kelly, A.K. Maulloo, D.K.H. Tan, Rate control for communication networks: shadow prices, proportional fairness and stability, J. Oper. Res. Soc. 49 (3) (1998) 237–252.
[14] F.P. Kelly, Mathematical modelling of the Internet, in: B. Engquist, W. Schmid (Eds.), Mathematics Unlimited—2001 and Beyond, Springer, Berlin, 2001, pp. 685–702.
[15] S.H. Low, P.P. Varaiya, A new approach to service provisioning in ATM networks, IEEE/ACM Trans. Networking 1 (5) (1993) 547–553.
[16] S.H. Low, D.E. Lapsley, Optimization flow control, I: basic algorithm and convergence, IEEE/ACM Trans. Networking 7 (6) (1999) 861–874.
[17] K. Malinowski, Optimization network flow control and price coordination with feedback; proposal of a new distributed algorithm, Comput. Commun., Spec. Issue Adv. Perform. Eval. Comput. Telecommun. Networking 25 (11–12) (2002) 1028–1036.
[18] E. Altman, D. Artiges, K. Traore, On the integration of Best-Effort and Guaranteed Performance services, Eur. Trans. Telecommun., Spec. Issue Architectures, Protocols and Quality of Service for the Internet of the Future (2) (1999).
[19] M. Baglietto, R. Bolla, F. Davoli, M. Marchese, A. Mainero, M. Mongelli, A unified model for a pricing scheme in a heterogeneous environment of QoS-controlled and Best Effort connections, Proc. Int. Symp. Perform. Eval. Comput. Telecommun. Syst. (SPECTS 2002), San Diego, CA, July (2002) 414–421.
[20] S. Low, Optimization flow control with on-line measurement, Proc. 16th Int. Teletraffic Cong., Edinburgh, UK June (1999) 237–249.
[21] S. Low, Multipath optimization flow control, Proc. ICON, Singapore 2000.
[22] D. Bertsekas, Dynamic Programming and Optimal Control, 2nd ed., Athena Scientific, Belmont, MA, 2001.
[23] Cost 239 Ultra high capacity optical transmission networks, Final report of COST project 239, 1999.
[24] K. Pawlikowski, Steady state simulation of queueing processes: a survey of basic problems and solutions, ACM Comput. Surv. 22 (2) (1990) 123–170.
[25] W. Ben-Ameur, H. Kerivin, New economical virtual private networks, Commun. ACM (2001).
[26] W. Ben-Ameur, E. Gourdin, B. Liau, N. Michel, Dimensioning of Internet networks, Proc. DRCN2000, Munich, Germany (2000) 56–61.
[27] P. Marbach, O. Milhatsch, J. Tsitsiklis, Call admission control and routing in integrated services networks using neuro-dynamic programming, IEEE J. Select. Areas Commun. 18(2) (2000) 197–208.
[28] K. Gokbayrak, C. Cassandras, Adaptive call admission control in circuit-switched networks, IEEE Trans. Automat. Contr. 47 (6) (2002) 1234–1248.
[29] K. Ross, Multiservice Loss models for Broadband Telecommunication Networks, Springer, Berlin, 1995.
[30] S. Shenker, Fundamental design issues for the future Internet, IEEE J. Selected Areas Commun. 13 (7) (1995) 1176–1188.
[31] S.H. Low, Equilibrium allocation and pricing of variable resources among user-suppliers, Perform. Eval. 34 (4) (1998).
[32] S.H. Low, Equilibrium allocation of variable resources for elastic traffics, Proc. INFOCOM'98, San Francisco, CA, March (1998).
[33] D. Bertsekas, R. Gallager, Data Networks, Prentice-Hall, Englewood Cliffs, NJ, 1987.
[34] M. Malowidzki, K. Malinowski, Optimization flow control in IP networks: distributed pricing algorithms and reality oriented simulation, Proc. Int. Symp. Perform. Eval. Comput. Telecommun. Syst. (SPECTS 2002), San Diego, CA, July (2002) 403–413.
[35] C. Courcoubetis, G.D. Stamoulis, C. Manolakis, F.P. Kelly, An intelligent agent for optimizing QoS-for-money in priced ABR connections, Telecommun. Syst., Spec. Issue Network Econ. (1998).
[36] F.P. Kelly, D. Walker, J. Solomon, Tariffing in the new IP/ATM environment, Telecommun. Policy 21 (4) (1997) 283–295.
[37] X. Lin, B. Shroff, Pricing-based control of large networks, In: Evolutionary Trends of the Internet, Proc. 2001 Tyrrhenian International Workshop on Digital Communications, IWDC 2001, Taormina, Italy.
[38] I.C. Paschlidis, J.N. Tsitsiklis, Congestion-dependent pricing of network services, IEEE/ACM Trans. Networking 8 (2000) 171–184.
[39] C. Touati, E. Altman, J. Galtier, On fairness in bandwidth allocation, Rapport de recherche de l'INRIA Sophia Antipolis, Septembre 2001.
[40] Traffic management specification, version 4.0 Technical Report AF-TM-0056.000, ATM Forum Traffic Management Working Group, April 1996.
[41] F.P. Kelly, Charging and rate control for elastic traffic, Eur. Trans. Telecommun. 8 (1997) 33–37.
[42] H. Yaiche, R.R. Mazumdar, C. Rosenberg, A game theoretic framework for bandwidth allocation and pricing in broadband networks, IEEE/ACM Trans. Networking 8 (5) (2000).
[43] D.O. Awduche, L. Berger, D.-H. Gan, T. Li, V. Srinivasan, G. Swallow, RSVP-TE: extensions to RSVP for LSP Tunnels, Internet draft, February 2001, draft-ietf-mpls-rsvp-lsp-tunnel-08.txt.
[44] J.M. Pitts, J.A. Schormans, Introduction to IP and ATM Design and Performance, 2nd ed., Wiley, New York, 2000.