# A Simulation Study of Usage-Based Pricing Strategies for Packet-Switched Networks

Leanne P. Breker
Carey L. Williamson
Department of Computer Science
University of Saskatchewan
57 Campus Drive
Saskatoon, SK, CANADA S7N 5A9
e-mail: {lpb133,carey}@cs.usask.ca
phone: (306) 966-8656
FAX: (306) 966-4884

March 12, 1996

**Abstract**

With the Internet experiencing massive growth, new ways to recover ever rising costs and regulate network usage are becoming increasingly important. Several researchers have proposed pricing strategies and econometric models for regulating resource usage in a globally shared network environment. These approaches can take the form of simple "price setting", or can be more closely integrated with the admission control, congestion control, and bandwidth allocation functions of the network.

This paper presents a simulation study of two proposed usage-based pricing strategies (one static, one dynamic) for computer networks. In the static pricing strategy evaluated, a price per packet is assigned and held constant over time. Results of the static strategy simulations show that network utilization decreases as the price per packet increases. Revenue, on the other hand, first increases and then falls as the price per packet increases. In the dynamic pricing strategy simulations, the price varies over time as a result of user demand for bandwidth. As the number of users on the network increases, the price per unit of bandwidth increases. While the behavioural properties of the dynamic pricing scheme make it appear promising as a pricing framework for packet-switched networks, further work is clearly required to better address the tradeoffs between network utilization, revenue, and network efficiency.

**Keywords:** Pricing, packet-switched networks, simulation, performance evaluation

# 1 Introduction

The advent of usage-based pricing for packet-switched computer networks, such as the Internet, is imminent. By usage-based pricing, we refer to pricing schemes that charge network users on a per byte or per packet basis. Such pricing strategies are a significant change from the traditional pricing strategies used by telecommunications companies and network service providers, which up until now have been based solely on connect time (e.g., on a per minute, per hour, or per month basis).

The primary reason alternative usage-based pricing strategies are now under consideration is the rapid growth of the Internet [13]. This growth has been phenomenal, regardless of which metric is used (e.g., number of hosts, number of users, volume of traffic). Current estimates place growth in number of hosts at 85% per year [5], and growth in traffic volume at 10-15% per month [9]. Some researchers [1] suggest that usage-based pricing structures are perhaps the only way to control this exponential growth.

There are several additional reasons why usage-based pricing is particularly attractive. First, the restricted access and abundance of goodwill that the academic community enjoyed on the Internet in the past have changed with the advent of public and commercial access to the Internet [16]. This change creates a need for a *revelation mechanism*, or punishment strategy, as a means to entice malicious users to act in a cooperative way. In this new era of the Internet, user incentives become a very important factor in maintaining network performance. Second, the limited bandwidth of the past has also meant the presence of only low bandwidth applications on the Internet. However, with the advent of high speed Asynchronous Transfer Mode (ATM) networks [6] and increased bandwidth, comes the appearance of new high bandwidth applications [16]. With these elaborate new applications, demand for bandwidth also increases. As this cycle proliferates, it becomes apparent that new methods for congestion control are also becoming increasingly important. Third, new high bandwidth applications and increased demand create the need for more than one service class [16]. No longer will First In First Out (FIFO) methods of service be sufficient to appease the many different types of customers. Multiple service classes will be required to differentiate different types of traffic flows and, since multiple service classes are more complex than a simple FIFO scheme, charges for service classes and changes in network protocols are required.

Clearly, the growth in Internet usage, combined with a revelation mechanism, increasing congestion control and multiple service classes, imply rising costs for network providers. High speed facilities cost more to implement and maintain. New considerations for user incentives, congestion control and flow control all increase the operational costs of networks. Government cutbacks, inflation and increasing taxes make network providers even more concerned about cost recovery.

Several researchers have proposed pricing strategies and econometric models for regulating resource usage and recovering costs on packet-switched networks. These approaches range from simple static schemes, in which one price is set, to dynamic schemes, in which

the price is continually changing based on conditions in the network.

The purpose of this paper is to evaluate usage-based pricing strategies for packet-switched computer networks. There are many such schemes proposed in the literature, and this paper presents simulation studies of two such strategies, one static and one dynamic. The performance studies are intended to illustrate some of the basic issues, behavioural properties, and tradeoffs that arise in the design of network pricing strategies. The paper provides some preliminary data points to network service providers on the feasibility of usage-based pricing, and forms the groundwork for our ongoing work on dynamic pricing strategies for bandwidth allocation in high speed networks.

The remainder of this paper is organized as follows. Section 2 presents a survey and classification of network pricing strategies, along with examples of each type of scheme that appear in the literature. Two of these schemes were selected for detailed evaluation in our performance study. These simulation results are presented in Section 3 and 4 for the static and dynamic pricing schemes, respectively. Finally, Section 5 concludes the paper.

# 2 Network Pricing Strategies

Pricing schemes fall into two broad categories: static and dynamic. Both approaches assume that the network provider has mechanisms to set prices, and to perform accounting and billing for network usage on a per-user basis.

Static pricing strategies generate revenues by setting a price per unit of resource and charging for the amount of resource used. Alternately, static pricing schemes simply set a flat fee that is independent of resource usage. In both cases, the price is not dependent on any network condition and remains unchanged over time.

In dynamic pricing schemes, the price per unit of resource fluctuates based on some network condition. Dynamic pricing strategies are usually more complex and require slightly more accounting than static pricing schemes.

## 2.1 Static Pricing Strategies

A static pricing strategy can be classified into one of several different types. These types range from the very simple flat pricing to more complex strategies such as reservation based pricing. Each of these types is illustrated below with one or more examples from the literature.

### 2.1.1 Flat Pricing

Cocchi et al [3] proposed a pricing scheme in which a price is set and applied on a per byte basis. Users then pay a sum to the network provider based on the number of bytes sent. In Cocchi et al's basic scheme, the network is assumed to operate on a simple FIFO basis with no consideration for different service classes.

### 2.1.2 Priority Pricing

Cocchi et al [3] extended their initial flat pricing scheme to include priority (i.e., service classes). The authors' scheme charged a higher per byte fee for bytes with higher priority, which were then assumed to get better service from the network in comparison with bytes of a lower priority. In this way, users that request a high quality of service pay in monetary terms for the stresses imposed on the network and other users by that requested higher quality of service.

Several interesting conclusions came out of the Cocchi et al simulations. The first is that when high network utilization, or maximum network efficiency, is the goal, the use of a pricing scheme is less important when the network is lightly loaded. In other words, when the network is lightly loaded it is less important to motivate users with monetary incentives. Second, as the network load increases, a wider range of prices is acceptable. In this context, an acceptable price is one in which the selfish behaviour of users results in an efficient usage of network resources [3]. Third, the authors' simulations showed that the range of acceptable priority prices depends on the network topology, the backbone speed and the offered load.

Although significant, the conclusions were based on simulations conducted under very heavily constrained assumptions. For example, the authors assumed that only a user's desired service class varied and not the amount of traffic offered. They also assumed that an administrative body could set prices based on historical evidence of offered load. This assumption is particularly restrictive since demand for network resources is continually increasing making offered load incredibly hard to predict.

### 2.1.3 Reservation Based Pricing

Parris et al [11] proposed a pricing scheme that combined flat pricing, two priority classes, and a reservation model. In this scheme, users choose a class of service and pay a fee per packet based on the class of service chosen. Again, a higher fee is charged for a higher service class. At the time of call setup, users not only determine the class of service they desire, but also how long they need network resources. Based on this user supplied information, a simple accept/reject decision is made by the network. The network accepts the call if sufficient capacity exists on the network and the user can pay for the services requested. If the connection is accepted, a one time setup fee may be charged.

Unfortunately, there are several problems with this proposed scheme, particularly with the setup fee. The first is that including a setup fee is not equitable; users with short conversations subsidize users with long conversations. Second, average utilization is lower when a setup fee is introduced and hence the goal of network efficiency is not fulfilled. Further, a discriminatory effect is introduced with the setup fee in that poorer users may no longer be able to afford even a single connection, while a setup fee hardly affects rich users. In spite of these drawbacks, it was found that the inclusion of a setup fee is always preferable when maximizing revenue. This is a demonstration of the tradeoff between network efficiency and revenue, which will be explored in more detail in Sections 3.

A second example of a reservation based pricing scheme was developed by Ferrari et al [10]. This policy is based on charging users for the *type of service* they request. The type of service is defined by the amount of resources needed for the connection. The simple formula for the type of service requested involves the amount of bandwidth reserved for the connection, the number of buffers reserved for the connection, the CPU time needed for the connection and the delay the call would impose on other callers.

The total charge for a user is defined by the following formula:

$$Total_Charge = TOS * t * tod$$

where $TOS = type_of_service$, $t = duration_of_call$ and $tod = time_of_day$. The *tod* factor is included so that a user making a call during high network utilization is charged more than a user making a call during low network utilization. Because of this, this scheme can also be classified as a time of day pricing scheme. Unfortunately, this policy requires an enormous amount of *a priori* knowledge, which is difficult and not always realistic to obtain.

### 2.1.4 Time Of Day Pricing

*Time of day* pricing refers to schemes that try to accommodate peak and off-peak network periods. Parris et al [11] extended their previous reservation based pricing model to include peak and off-peak periods. The peak period, which is similar to busy business hours, has a much higher call arrival rate. Users are charged a much higher price if they want to use the network during this time. Off-peak periods, on the other hand, are periods where network utilization is low, such as late evening or early morning. Rates are lower during this off-peak time in the hopes of stimulating user demand for resources. Telephone companies often implement this strategy by offering discounts during off-peak hours.

In modelling peak and off-peak periods, Parris et al include an elasticity factor. Elasticity is defined as the percentage of change in quantity demanded due to a small (e.g., 1%) change in price. Elastic users will greatly change the amount they demand when there is a small change in price. In this way, the authors modelled users that may be able to delay their peak call to an off-peak period and hence take advantage of offered savings.

## 2.2 Dynamic Pricing Strategies

Most dynamic pricing strategies rely, in some form or another, on the econometric model of supply and demand. This model assumes an equilibrium price exists given a quantity demanded and a quantity supplied. The equilibrium price is the price where the quantity of resource demanded is equal to the quantity of resource supplied. In general, user demand for a resource decreases as the price of that resource increases. On the other hand, as the price increases, suppliers are willing to produce and sell more of the resource. Figure 1 shows a simple supply and demand curve and the equilibrium price resulting due to their intersection.

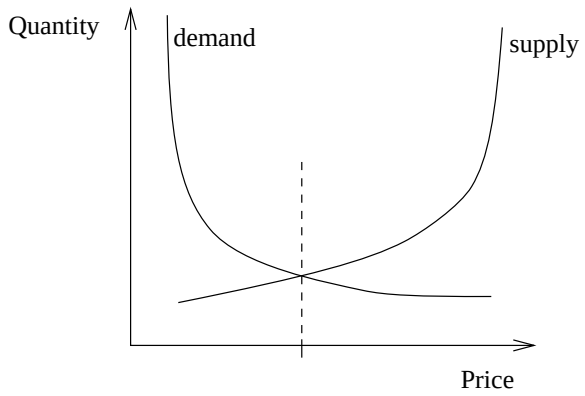



Figure 1: Example Supply and Demand Curves

Pricing strategies based on the supply and demand model attempt to use price to manipulate user demand to either maximize revenue or maximize network efficiency. Several researchers have studied the use of supply and demand in obtaining network efficiency and cost recovery goals. Several are discussed here.

### 2.2.1 Transport Auction

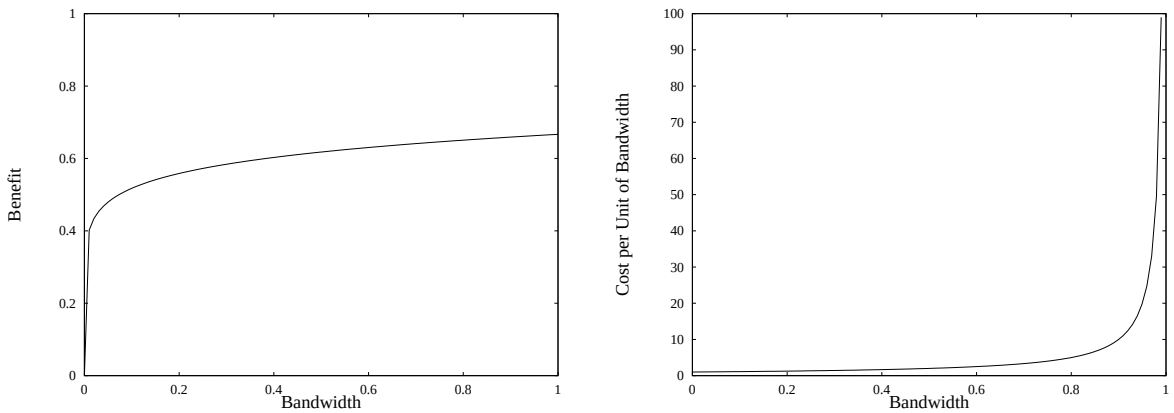
Hayer and Grover [4] propose a unique dynamic pricing strategy in which excess transport capacity on the network is auctioned to users with delay tolerant traffic. The basic idea of Transport Auction is to set the price dynamically in response to the changes in demand so as to stimulate more usage of the telecommunications facilities and thus increase revenues.

In the Transport Auction strategy, the user has an *agent*, or electronic device, which takes as input some amount of traffic and the maximum price the user is willing to pay to send that traffic over the network. When the network price is at or below the user's maximum price, the user's data is offered to the network. The network then accepts or rejects the traffic based on current load and network capacity. A user may also log with the agent the latest possible date that the information can be sent. When that time approaches, the agent may send the traffic at the current price, even if the current price is greater than the logged maximum price.

Although Transport Auction generates a very high revenue, numerous problems still exist. For example, it is a fairly complex strategy and requires users to install an "agent". Second, a revelation mechanism is absent and greedy users may bid down the price or misrepresent their traffic. Third, since Transport Auction occurs on a per trunk basis, there is no ability to reroute traffic, resulting in poor reliability.

### 2.2.2 Dynamic Bandwidth Allocation

Murphy et al [8] take a different but related approach to the use of supply and demand for dynamic pricing in packet-switched networks. In their approach, users are assumed to act in their own best interest. They are also assumed to place benefit on any bandwidth received. This results in each user having a *personal benefit function*. An example benefit function in shown in Figure 2(a). The user is assumed to place the most benefit on the first units of bandwidth received and hence the benefit function increases rapidly at first, when the bandwidth received is low. The benefit function then levels off as more bandwidth is received.[1] This benefit function is used in conjunction with the price set by the network to determine the amount of bandwidth that a user will "purchase" in a given interval. In this context, an interval is simply some regular unit of time.




(a) (b)

Figure 2: Example Curves for Dynamic Pricing Strategy: a) benefit function b) cost function

The network, on the other hand, derives the price per packet in the next interval based on the marginal cost of sending the user's amount of traffic and the price in the current interval. In doing this, the network defines a cost function, an example of which is shown in Figure 2(b), which takes into account the capacity of the link and the sum of all bandwidth used. This cost function could represent queueing delay, for example, or any other reasonable measure of network cost. As the total bandwidth used approaches capacity (i.e., unity, in Figure 2(b)), the price per unit approaches infinity. In using a cost function with this shape, a low price per unit of bandwidth when usage is low encourages users to demand more bandwidth, while a high price per unit of bandwidth when usage is high encourages users to demand less bandwidth.

[1]This behaviour is referred to as *diminishing marginal utility* in the economics literature.

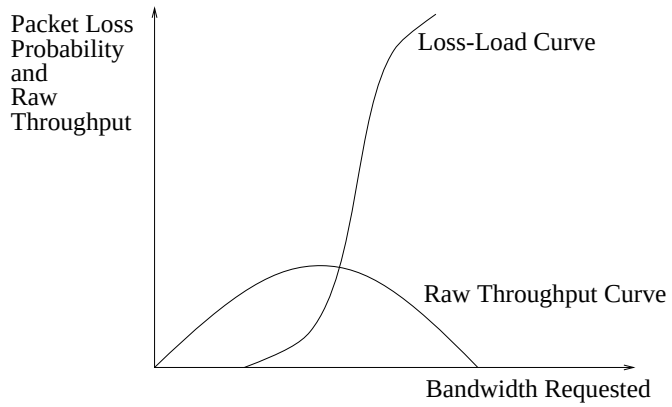



Figure 3: A Sample Loss-Load Curve and Throughput Curve

### 2.2.3 Loss-Load Curves

Another possible dynamic pricing scheme is a congestion control scheme called *loss-load curves*, proposed by Williamson [17]. Loss-load curves [17] allocate bandwidth based on changing network conditions. An example of a loss-load curve is shown in Figure 3. The curve expresses the expected packet loss probability for a user as a function of that user's offered load. The packet loss probability ranges from zero, in which the capacity of the link is not exceeded, to between zero and one, in which the capacity of the link is exceeded and excess traffic must be dropped. Senders request (i.e., bid for) an amount of bandwidth and the loss-load curve is then used to determine the loss the sender will experience. This loss probability in turn determines the bandwidth allocation (called the *raw throughput*) for the user relative to the user's request. An example raw throughput curve is shown in Figure 3.

In choosing where to operate on the loss-load curve, users have the ability to maximize throughput, minimize response time or minimize loss. Users independently choose which strategy to follow. For example, users who wish to minimize loss must request a smaller percentage of bandwidth than users with loss-tolerant traffic. In requesting a decreased percentage of bandwidth, these users operate at the low end of the loss-load curve and hence get minimum loss. Those senders with loss-tolerant traffic may choose to request a higher percentage of bandwidth and hence will experience higher loss by operating at the high end of the loss-load curve. Likewise, users maximizing raw throughput choose where to operate on the raw throughput curve.

Loss-load curves have a concise mathematical representation [17], give the sender control over their loss and incorporate a revelation mechanism to control greedy senders. Unfortunately, the loss-load strategy is a very aggressive approach to congestion control and high levels of packet loss often result.

### 2.2.4 Virtual Loss-Load Curves

Prithviraj [14, 15] proposes a modified version of the loss-load model that reduces its aggressiveness, by having switches advertise a *virtual capacity* rather than their actual capacity to network users. The congestion control strategy incorporates a virtual leaky bucket with loss-load curves. The combination of the virtual leaky bucket and loss-load curves results in increased network throughput, decreased network overload, and decreased sender loss, according to the simulation results reported in [14].

One limitation of the Prithviraj et al strategy is that using a virtual capacity removes the built-in revelation mechanism in the original loss-load model. That is, greedy senders who know about the virtual capacity can cheat and win. As a result, the virtual loss-load model requires sender cooperation in order to function properly.

Pricing may be the added dimension needed in this model to not only recover costs, but also to punish greedy senders. Evaluation of this approach as a framework for dynamic pricing is currently underway.

## 2.3 Summary

There are a number of usage-based pricing strategies proposed in the literature for use on packet-switched networks. These strategies range from very simple static pricing strategies, such as flat fee pricing, to more complicated dynamic pricing strategies, such as those based on the theory of supply and demand from economics.

¿From the foregoing strategies, two have been selected for further evaluation in our performance study: the static pricing scheme proposed by Parris et al [11], and the dynamic pricing scheme proposed by Murphy et al [8]. Simulation results for these pricing strategies are presented in the next two sections.

# 3 Simulation Results: Static Pricing

This section presents the simulation results for the static pricing scheme proposed by Parris, Keshav and Ferrari [11]. This pricing scheme is a reservation based per unit pricing model, with additional features to support peak and off-peak pricing. The simulations conducted are an attempt to recreate the results reported by Parris et al [11], and to understand the general behavioural characteristics of their static pricing strategy.

## 3.1 Experimental Methodology

Three separate simulation experiments were conducted with the static pricing scheme. The first experiment studied simple per packet pricing, in which a sender was charged a fee per packet. The second experiment added a setup cost to the basic per packet model. The third experiment modeled peak and off-peak pricing intervals.

In all three experiments, the users of the network were characterized (as in [11]) with three attributes: class, duration, and budget. Combinations of these three attributes are intended to model users with different network needs and varying financial capabilities.

The class of a user defines the amount of bandwidth that the user desires. Users can request either a large portion or a small portion of the bandwidth per interval. Class A users always request 1% of the bandwidth.[2] This class of user models the typical user with moderate bandwidth needs. 80% of all users are Class A users. Class B users, on the other hand, always request 5% of the bandwidth. Such users have much higher than average bandwidth needs and must pay five times as much for this bandwidth. Only 20% of all users fall into this category.

Each user generates a call with a specific duration. The duration of calls follows a geometric distribution with a mean of 36 intervals for Class A users and a mean of 240 intervals for Class B users. Again, Class B users are assumed to have much more intensive bandwidth needs. Users depart from the network when their calls complete, or when they run out of money to send another packet (whichever comes first).

The last attribute of users is their budget. 20% of all users were considered rich and had a total of \$600 to spend. 80% of the users were poor and only had \$60 to spend. The probability of being rich was independent of the class of a user.

All experiments were conducted on a one link network with a perfect non-blocking switch. Call arrivals followed a Poisson distribution with a mean of one arrival per interval. The network accepted a call if capacity permitted and if the user had enough money to pay for at least one packet.

In all simulations, both utilization and revenue were monitored. Utilization indicates the total average percentage of bandwidth of users over the entire simulation run. A higher utilization is considered advantageous providing capacity is not exceeded. Revenue is the dollar amount collected over all intervals of the simulation run. The simulation run length (as in [11]) was 5000 calls, following an initial warmup period of 200 time units.

## 3.2 Experiment 1: Per-Packet Pricing

The first simulation experiment studied the effect of the per-packet price on network utilization and revenue. In this experiment, the price per packet was varied from \$1 per packet to \$250 per packet in increments of \$1.

Figure 4(a) shows the simulation results obtained for network utilization. As expected, utilization is high when the price per packet is low. As the price per packet increases, utilization steadily decreases. For reference, the simulation results from Table I of the Parris et al paper [11] are also shown in Figure 4(a). Our simulation results match the Parris et al results almost exactly for every price.

Revenue results are shown in Figure 4(b). Simulation results show a general rise in revenue and then a fall as the price per packet increases. There is also a definite saw-tooth

[2]For simplicity, 1% of the bandwidth is equivalent to one packet per interval.

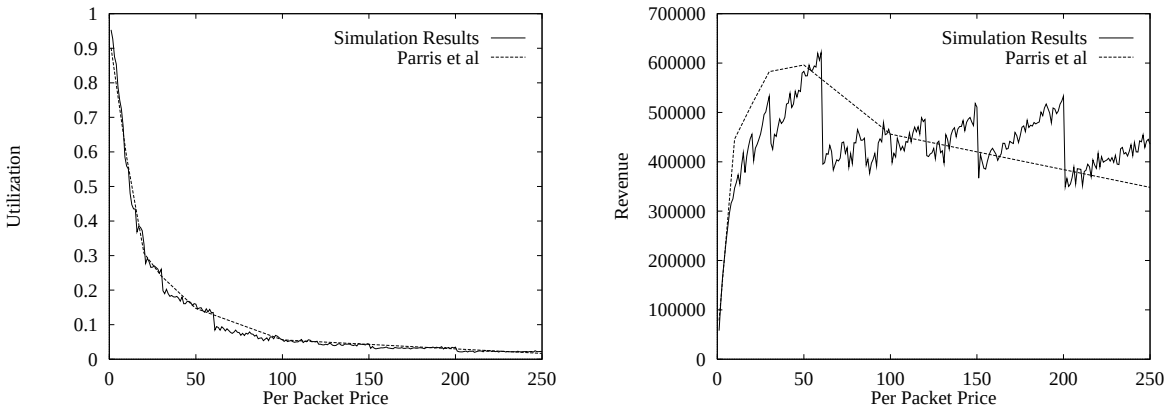



Figure 4: Per Packet Pricing: a) Utilization b) Revenue

pattern with significant periodic falls. These falls stem from the fact that users may not be able to spend all the money they have available. If a user does not have enough money to send one more whole packet, the money remains unspent since sending partial packets is not allowed. For example, changing the per packet price from \$60 to \$61 means that poor users cannot send any packets at all, wasting \$60 each, and changing the per packet price from \$200 to \$201 means that rich users can send only two packets instead of three, wasting \$199 each.

The Parris et al results in Figure 4(b) also show a general increase in revenue and then a decrease. Although the simulation results do not match exactly, they are in the same range and follow the same pattern. Also note that our simulation results present more data points than were reported in [11].

## 3.3 Experiment 2: Effect of Setup Fee

The second experiment differed from the first by the addition of a setup fee. The setup fee is a one-time fee charged to the user upon connection. The setup fee charged was 20% of the per packet price. In these simulations, the network would only accept a call if the user could pay for at least one packet and the entire setup fee.

Figure 5(a) shows the simulation results obtained for utilization for the second set of simulations. Again, a high utilization is achieved when the price is low and utilization steadily falls as the setup price increases. In these simulations, the Parris et al results also closely match the simulation results. In comparing Figures 4(a) and 5(a), the setup fee appears to have made little difference in the utilization.

Figure 5(b) shows the revenue values obtained for the second set of simulations. Again, a rise in revenue first occurs followed by a fall as the price increases. The same saw-tooth pattern as in the first set of simulations is seen and the Parris et al results again are in

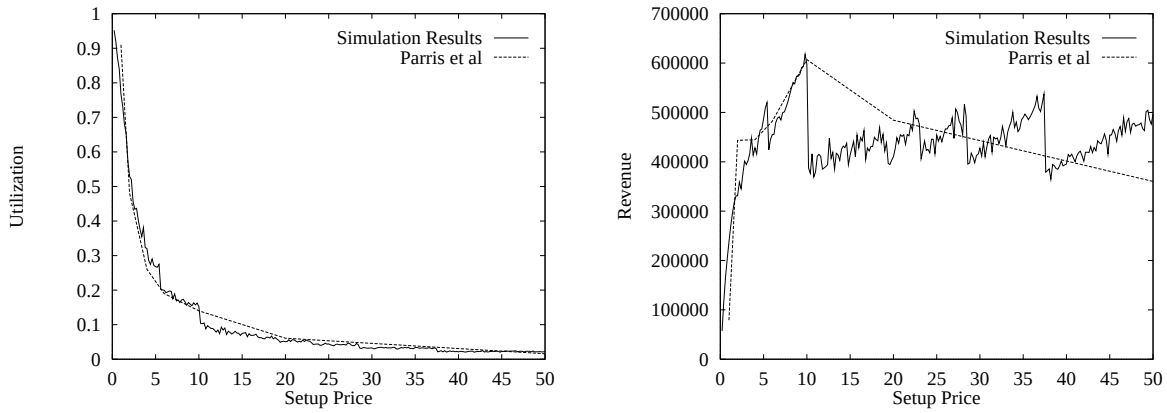



Figure 5: Per Packet Pricing with Setup Cost: a) Utilization b) Revenue

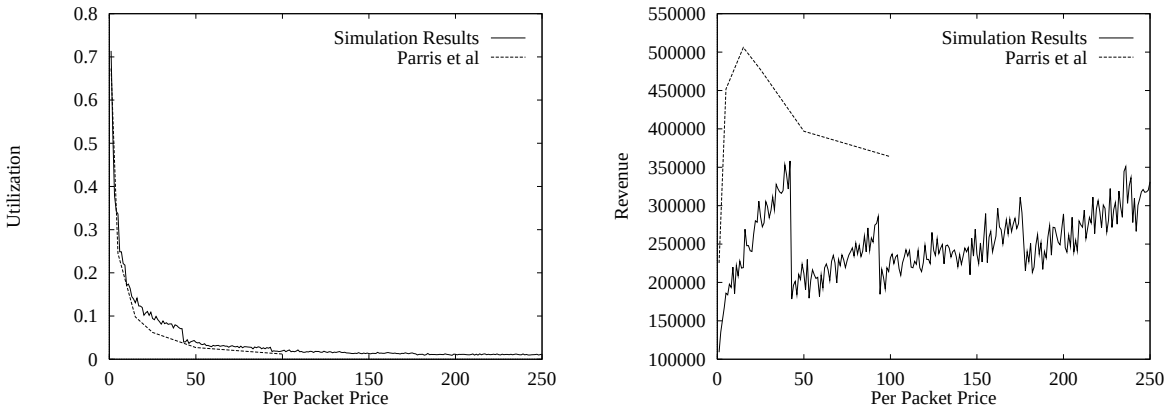
the same region and follow the same general path as the simulation results. In comparing Figure 4(b) with Figure 5(b), the addition of the setup fee appears to have shifted the revenue curve slightly towards the left so that the same revenue is generated with a lower per packet price.

## 3.4 Experiment 3: Peak and Off-Peak Pricing

The third experiment differed from the first two in a number of ways. First, the third simulation divided the time intervals into peak and off-peak periods. During peak periods, the number of call arrivals per interval was drawn from a Poisson distribution with a mean of two arrivals per interval, while the off-peak period (as before) still had a mean of one arrival per interval. Second, the peak price per packet differed from the off-peak price. The peak price was three times the off-peak price and a user whose call started during the peak period was charged the peak price for the duration of the call. Third, during the peak periods, an elasticity factor was associated with each call. Users could either be elastic or inelastic. Elastic users are very price sensitive and hence, even though they may want to make a call during a peak period, they delay their call to an off-peak period to take advantage of the cost savings. Users had a 20% chance of being elastic. Lastly, the setup fee for peak and off-peak calls was set to 40% of the per packet price.

Figure 6(a) shows that these changes had a significant effect on utilization. Although the same general trend is apparent, overall utilization is much lower. The Parris et al results agree with this decline. This decrease in utilization is most likely from the higher setup cost since users will not be able to afford to send as many packets.

Figure 6(b) shows the total revenue obtained for each price in this set of simulations. Again, a rise is seen first followed by a decline as the price per packet increases. The same saw-tooth pattern is also seen but overall generated revenue is considerably lower than in




(a) (b)

Figure 6: Peak Pricing: a) Utilization b) Revenue

the first two scenarios. There is also a marked difference between the Parris et al results and our simulation results. Since the simulated utilization values appear to exactly match Parris et al utilization values, the difference in revenue values most likely arise due to different implementation assumptions in the simulations.

### 3.5 Summary

Several conclusions can be drawn from the above simulations. First, for all three pricing schemes, increasing prices first increases net revenue and then decreases it. Second, as the price increases, the network utilization decreases. Third, the addition of a 20% setup cost to the per packet pricing scheme results in a slightly increased net revenue. Finally, when the setup cost is at 40% of the per packet price and peak and off-peak periods are simulated, a much lower utilization and revenue is achieved.

Overall, the simulations illustrate a definite tradeoff between the goals of maximizing network efficiency, or utilization, and maximizing revenue. In all three experiments, the price setting that achieves the highest network utilization generates very little revenue, while the price setting that generates the maximum revenue results in very low network utilization. In other words, maximizing network efficiency and maximizing revenue are mutually exclusive goals.

## 4 Simulation Results: Dynamic Pricing

The second set of simulation experiments involved the dynamic bandwidth pricing algorithm proposed by Murphy et al [8]. The algorithm is designed to achieve fair bandwidth allocation on ATM networks by using a dynamic usage-based pricing scheme.

In this pricing scheme, the user is characterized with a benefit function. Users are assumed to act in their own best interest and place benefit on any bandwidth they receive. An example benefit function is shown in Figure 2(a). Note that the benefit function increases quickly when bandwidth received is lower and levels off when bandwidth received is higher. As mentioned previously, this shape is considered typical since it is assumed that each user values the first additional units of bandwidth that they receive more than latter additional units.

The network is characterized with a cost function. Figure 2(b) shows a typical cost function. As the total bandwidth of all users increases, the cost per unit of bandwidth also increases. As the total bandwidth utilized approaches the capacity of the network, or the "danger zone" where congestion may occur, the price per additional unit of bandwidth approaches infinity so that no additional units are requested by users.

The algorithm followed in the Murphy et al pricing scheme is fairly simple and involves the following steps:

**Step 0:** network operator chooses initial value for the price

**Step 1:** network operator announces the price to the users at the start of the current pricing interval

**Step 2:** users respond by choosing their requested bandwidth according to

$$\frac{\partial BenefitFunction}{\partial Bandwidth} - CurrentPrice = 0$$

**Step 3:** network operator calculates the marginal cost with respect to bandwidth using:

$$CurrentPrice = \frac{\partial CostFunction}{\partial TotalBandwidth}$$

where

$$CostFunction = \frac{1}{Capacity - TotalBandwidth} - \frac{1}{Capacity}$$

**Step 4:** network operator adjusts price to reduce the difference between terms in step 3. Go to step 1.

In other words, the network sets the price based on the marginal cost of carrying the current load (i.e., the slope of the cost function at the current operating point), and users react to price changes by moving to a point on their benefit function where the marginal benefit (i.e., the slope of the curve) is the same as the current network price. This algorithm is explained in further detail in [8].

## 4.1 Experimental Methodology

In evaluating the Murphy et al algorithm, several different sets of simulations were conducted. The first and second sets evaluate the behaviour of the algorithm when the user community has homogeneous user benefit functions. The third set of simulations allows for heterogeneous user benefit functions. The initial price is then changed in the fourth set. Finally, the last set of simulations evaluates dynamic changes in the number of users on the network.

All experiments assume that the user benefit function is of the form $y = \alpha x^{\beta}$, where $0 \leq \alpha \leq 1$ is a scale component that shifts the curve up or down, and $0 \leq \beta \leq 1$ is a shape component that determines its amount of curvature. $(0 \leq x \leq 1)$ is the fraction of the network bandwidth that a user will request at connection time. With this formulation, the user benefit function is continuous and differentiable (a prerequisite condition for the Murphy et al algorithm), as well as easily parameterizable to reflect a range of different benefit functions.

All three experiments were conducted on a one link network with a perfect non-blocking switch. The simulation run length was 30 time intervals, though our analysis focuses solely on the startup behaviour of the pricing algorithm.

## 4.2 Experiment 1: Homogeneous User Benefit Functions

In the first set of simulations, the user community is defined as having homogeneous bandwidth needs and hence, all users' benefit functions are equivalent. Figure 7(a) shows the results obtained when the homogeneous benefit function is $y = 1.0x^{0.1}$.

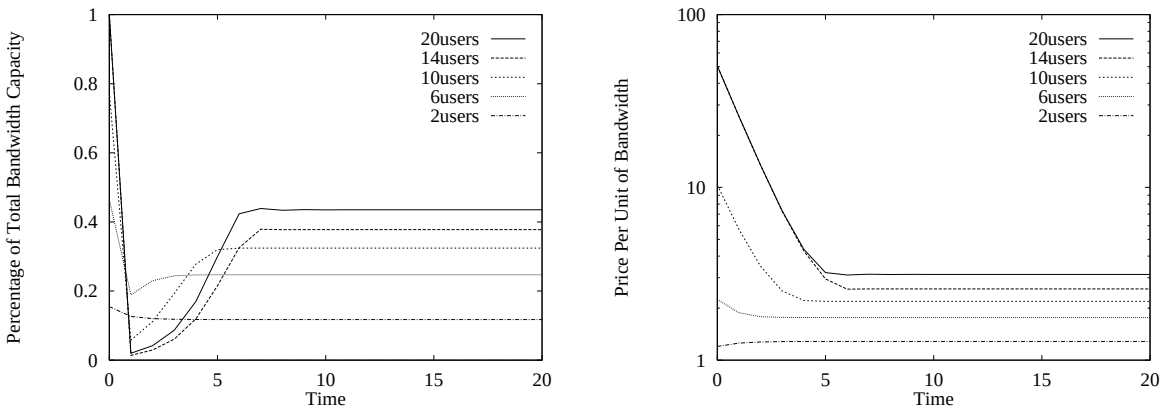



Figure 7: Homogeneous Users ($\alpha = 1.0, \beta = 0.1$): a) total bandwidth b) price

There are two phases apparent in the results. The first is the stabilization phase in which the total bandwidth requested first falls and then rises. The second phase is the equilibrium phase in which the total bandwidth requested does not change over time. Note that it takes

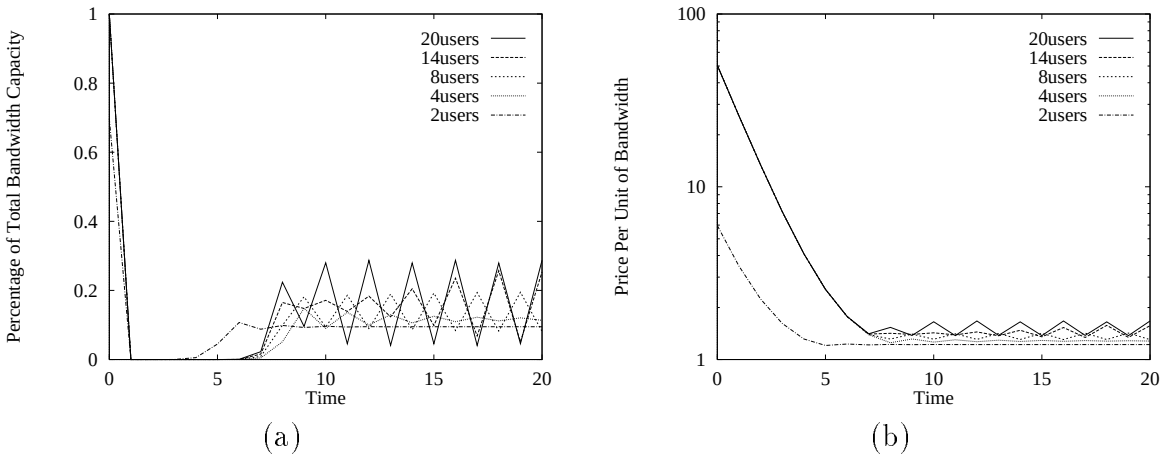
more time to reach equilibrium when there are more users on the network, but equilibrium is always reached. The total bandwidth requested at equilibrium increases as the number of users on the network increases, but is always below 50% of the network capacity in our simulations.

Figure 7(b) shows the price per unit of bandwidth plotted over time for the same experiment. Again, there are two phases present: stabilization and equilibrium. The stabilization phase shows that the price falls first before leveling out. The equilibrium phase for the price begins approximately at 5 time units when there are 20 users on the network and at 2 time units when there are 6 users on the network. This time frame corresponds to the time frame needed for the total bandwidth requested to reach equilibrium. Note that as the number of users on the network increases, so does the equilibrium price.

## 4.3 Experiment 2: Effect of $\beta$

The second set of simulations also involved a homogeneous user group; however, this group had benefit functions defined by $y = 1.0x^{0.9}$ . Figure 8(a) shows vastly different results in comparison with the previous experiment.

In Figure 8(a), there are again two phases. The first phase, the stabilization phase, shows a steep drop in total bandwidth requested by all numbers of users on the network before reaching the second phase. In fact, for a short time the total bandwidth requested appears to be zero. The second phase is not an equilibrium phase but a two-state oscillation phase. The total bandwidth requested appears to oscillate from high to low. In explanation, note that this particular benefit function is nearly linear indicating that users are very elastic. Consequently, a small change in price creates a large adjustment in requested bandwidth. Since all users make this adjustment at the same time, the bandwidth oscillation worsens as the number of users on the network increases.




(a) (b)

Figure 8: Homogeneous Users ($\alpha = 1.0, \beta = 0.9$): a) total bandwidth b) price

Figure 8(b) shows the price variations over time for this same user group. The stabilization phase is similar to Figure 7(b) but with slightly less variation between 4 users on the network to 20 users on the network. The second phase of Figure 7(b) is also a two-state oscillation phase as in Figure 8(a). Again, when there are a greater number of users on the network, the oscillation is more pronounced.

## 4.4 Experiment 3: Heterogeneous User Benefit Functions

The third set of simulations involved users with differing network needs. Users in this heterogeneous user group were given random user benefit functions of the form $y = \alpha x^{\beta}$. The number of users on the network were again varied from 2 users to 20 users. Simulations found that two phases were again present in both the total bandwidth requested and the price. The stabilization phase is similar to the first set of simulations in that the total bandwidth requested and the price first fall and then rise to their final states. The second phase is an equilibrium phase, again similar to the first set of simulations, in which the amount of total bandwidth requested and the price do not change over time. No oscillatory behaviour was found in the results of these experiments. Graphs of these results are available in [2].

## 4.5 Experiment 4: Effect of Initial Price Setting

In the fourth set of simulations, the initial price which the network presented to the users was varied. It was found that the initial price affects the rate at which an equilibrium is reached. As the initial price increases, the equilibrium phase is reached more quickly; however, equilibrium values themselves do not change. These results are also fully explained and graphed in [2].

## 4.6 Experiment 5: Effect of User Arrivals and Departures

Lastly, simulations were conducted to evaluate the impact of users arriving and departing from the network. Users were again assigned random benefit functions to simulate a network where users have varying bandwidth needs.

Figure 9 (a) shows the results obtained for total bandwidth values when arrivals and departures follow a Poisson distribution with a mean of 1.0 per interval. There is again a distinct stabilization phase where the total bandwidth decreases sharply and then rises. The second phase fluctuates slightly as users arrive and depart from the network. Despite the fluctuations, the algorithm behaves well in that the values of total bandwidth are never extreme; they are always within a small range.

Figure 9 (b) shows the price variations for the same simulation set. Very small variations in the price occur despite arrivals and departures from the network. The values for the price are reasonably stable.

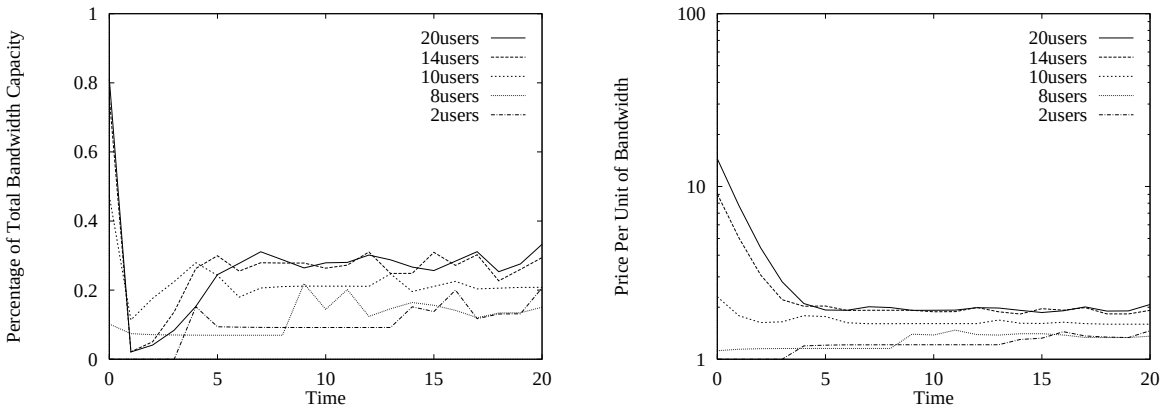



Figure 9: Poisson Arrivals/Departures: a) total bandwidth b) price

### 4.7 Summary

Several conclusions can be made from the simulation results for the dynamic pricing strategy. First, the more users on the network, the longer the stabilization phase (i.e., the longer the period until an equilibrium price is reached). Second, the equilibrium price increases as more users are added to the network (as it should). The sum of all users' bandwidth also increases as more users are added, though the allocated bandwidth per user decreases, as expected. Third, the higher the number of users on the network, the greater the oscillation of both price and bandwidth when oscillation is present. Fourth, varying the initial price changes the stabilization period but does not affect the final state values of both price and total bandwidth requested. Finally, the simulations varying the number of users demonstrates the potential for this algorithm. Despite arrivals and departures of users, the algorithm still maintains a controlled range of values for both total bandwidth requested and price.

Overall, the simulations show that this dynamic pricing algorithm operates very conservatively at a total bandwidth usage below 50% of link capacity. As well, an equilibrium state is not always reached when a homogeneous user community is defined (with the benefit functions of the form assumed here). However, the algorithm does show good potential for operating in a well-behaved manner when user benefit functions are heterogeneous, even when there are dynamic arrivals and departures of users.

## 5 Conclusions

Usage-based pricing will likely become a crucial part of packet-switched computer networks in the very near future. This paper has presented a survey of several proposed pricing strategies for packet-switched networks, and more detailed simulation evaluations of two such pricing strategies, one static and one dynamic.

Several conclusions are evident from our simulation experiments. First, while static pricing strategies are simple and offer a fairly straightforward means for network providers to control network usage and maximize revenue, static pricing schemes are unlikely to offer the generality, flexibility and robustness desired in future packet-switched networks. That is, static schemes, such as the one evaluated in this paper, operate under very restrictive assumptions about user populations and user traffic demands. Extending such schemes to support multiple service classes and more realistic models of user behaviour (if known) will prove cumbersome, if not impossible.

Second, dynamic pricing schemes, if properly designed, can react very well to changing network conditions and easily manipulate user demand by implementing the econometric principles of supply and demand. Clearly, dynamic pricing strategies, which use knowledge of current network conditions rather than just *a priori* price setting, offer greater potential for use on future networks. However, dynamic pricing strategies are usually more complex than static schemes. The tradeoff between complexity and effectiveness must be considered carefully when choosing between pricing schemes.

Finally, there is a distinct tradeoff between the goals of network efficiency (i.e., efficient network utilization) and revenue maximization. This tradeoff was illustrated in both pricing strategies studied in this paper. In the static pricing scheme, the per packet price that maximized revenue for the network provider resulted in a fairly low network utilization. In the dynamic pricing scheme, the equilibrium price achieved by the Murphy et al algorithm resulted in network utilizations from 10-40%, depending on the number of network users. If efficient network utilization is a primary consideration (rather than revenue maximization), then pricing schemes other than those studied in this paper must be explored.

In future work, we plan to study the effectiveness of the loss-load congestion control strategy (and its variants) as a mechanism for dynamic pricing and bandwidth allocation.

# References


[1] J. Adam, "Upgrading the Internet", *IEEE Spectrum*, Vol. 32, No. 9, pp. 24-29, September 1995.

[2] L. Breker, "Evaluation of a Dynamic Bandwidth Pricing Scheme", available at `http://www.cs.usask.ca/grads/lpb133/`, December 1995.

[3] R. Cocchi, S. Shenker, D. Estrin and L. Zhang, "Pricing in Computer Networks: Motivation, Formulation, and Example", *IEEE/ACM Transactions on Networking*, Vol. 1, No. 6, pp. 614-627, December 1993.

[4] J. Hayer, "Transport Auction: A New Service Concept", M.Sc./MBA Thesis, TR-93-05, TR*Labs* Edmonton, 1993.

[5] M. Lottor, data available by anonymous ftp from `ftp://ftp.nw.com/zone` or `http://www.nw.com/`.

[6] S. Minzer, "Broadband ISDN and Asynchronous Transfer Mode (ATM)", *IEEE Communications*, Vol. 27, No. 9, pp. 17-24, September 1989.

[7] J. Murphy and L. Murphy, "Pricing for ATM Network Efficiency", available by anonymous ftp from `ftp.eeng.dcu.ie/pub/tele-communications/murphy`, 1995.

[8] J. Murphy, L. Murphy and E. Posner, "Distributed Pricing for Embedded ATM Networks", available by anonymous ftp from `ftp.eeng.dcu.ie/pub/-tele-communications/murphy`, 1995.

[9] NSFNET Statistics, data available by anonymous ftp from `nic.merit.edu/statistics/nsfnet`.

[10] C. Parris and D. Ferrari, "A Resource Based Pricing Policy for Real-Time Channels in a Packet-Switching Network", Technical Report, Tenet Group, ICSI, UC Berkeley, 1994.

[11] C. Parris, S. Keshav, and D. Ferrari, "A Framework for the Study of Pricing in Integrated Networks", Technical Report, Tenet Group, ICSI, UC Berkeley, 1994.

[12] C. Partridge, *Gigabit Networking*, Addison-Wesley Publishing Company, Reading, Massachusetts, 1994.

[13] V. Paxson, "Growth Trends in Wide Area TCP Connections", *IEEE Network*, Vol. 8, No. 4, pp. 8-17, July/August 1994.

[14] N. Prithviraj, "A Virtual Loss-Load Congestion Control Strategy for High Speed Networks", M.Sc. Thesis, Department of Computer Science, University of Saskatchewan, July 1994.

[15] N. Prithviraj and C. Williamson, "A Virtual Loss-Load Congestion Control Strategy for High Speed Networks", submitted to *ACM Computer Communication Review*, March 1995.

[16] S. Shenker, "Making Greed Work in Networks: A Game Theoretic Analysis of Switch Service Disciplines", Proceedings of ACM SIGCOMM'94, London, UK, pp. 47-57, September 1994.

[17] C. Williamson, "Dynamic Bandwidth Allocation Using Loss-Load Curves", submitted to *IEEE/ACM Transactions on Networking*, January 1994.