Network Pricing Schemes: A Survey & Simulation Results

Leanne P. Breker
Department of Computer Science
University of Saskatchewan
57 Campus Drive
Saskatoon, SK S7N 5A9
CANADA
e-mail: lpb133@cs.usask.ca

Supervisor: Carey Williamson

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 summary of proposed pricing strategies for computer networks, as well as an experimental evaluation of both a static strategy and a dynamic strategy. Simulations of the static strategy 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. Further results and a summary of the pricing schemes that appear in the literature are outlined below.

Introduction

The arrival of the Information Age has brought sweeping changes in everything from the way we access our money to the way we reserve a hotel room. Internet usage, in particular, is rapidly escalating, bringing with it new problems that cannot be ignored.

First, the restricted access and abundance of goodwill that the academic community enjoyed in the past has disappeared with the advent of public access (Shenker 1994). This absence of goodwill has meant an awareness of the need for a revelation mechanism, or punishment strategy, to entice malicious users to act in a cooperative way. Never before have user incentives been such an important factor in maintaining network performance.

Second, the limited bandwidth of the past has also meant the absence of high bandwidth applications. However, with the appearance of high speed Asynchronous Transfer Mode (ATM) networks and increased bandwidth, comes the appearance of new high bandwidth applications (Shenker 1994). With these flashy new applications, demand for bandwidth also increases. As this cycle proliferates, it becomes apparent that new methods for congestion control are also increasingly important.

Third, new high bandwidth applications and increased demand justifies the need for more than one service class (Shenker 1994). No longer will First In First Out (FIFO) methods of service be sufficient to appease the many different types of customers. Multiple service classes are again more complex than a simple FIFO scheme and require more changes in network protocols.

Finally, the growth of Internet usage has brought about rising costs. High speed facilities cost more to implement and maintain. New considerations for user incentives, congestion control and flow control all increase the operation 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 data networks. These approaches range from simple static schemes, in which one price is set, to dynamic schemes, in which the price is continually changing.

This paper is organized as follows. First, a basic classification of pricing schemes is presented followed by several examples of each type of scheme that appear in the literature. Two of these schemes have been studied in more detail and specific results can be found in the Simulations section. Finally, conclusions are drawn.

Classification

Pricing schemes fall into two broad categories: static pricing schemes and dynamic pricing schemes. 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 may 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.

The second category of pricing schemes is dynamic pricing schemes. In dynamic pricing schemes, the price is usually set per unit of resource and fluctuates based on some network condition. Because of this, dynamic pricing schemes are usually more complex.

Static Pricing Strategies

Static pricing strategies can be divided into several different types. These types range from the very simple flat pricing to more complex strategies such as reservation based pricing or time of day strategies. Each type is illustrated below with one or more examples from the literature.

Flat Pricing

Cocchi et al (Cocchi et al. 1993) proposed a pricing scheme in which a price is set and applied on a per byte basis. Users pay a sum to the network provider based on the number of bytes sent. In Cocchi et al's basic scheme, no consideration was given to service classes and the network is assumed to operate on a FIFO basis. This pricing scheme formed the base of an extended scheme in which priority classes were implemented. The priority pricing scheme is outlined below.

Priority Pricing

Cocchi et al (Cocchi et al. 1993) expanded their initial strategy that used flat pricing and developed a new strategy that includes priority or service classes. In this new scheme, a higher per byte fee was charged for bytes with higher priority. Bytes with higher priority 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 stress imposed on the network and other users that is the result of that 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 (Cocchi et al. 1993). Third, 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 network utilization is continually increasing making offered load incredibly hard to predict.

Reservation Based Pricing

Parris et al (Parris, Keshav & Ferrari 1994) proposed a pricing scheme which 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 priority class. At the time of call setup, users not only determine the class of service they desire, but also how long they will be using 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.

It was found in the Parris et al simulations that as the per packet price and setup price increase, the total revenue increases to a maximum and then begins to decrease. As well, the total revenue generated is consistently higher when a setup fee is included. These conclusions are discussed in detail in the Simulations section below.

Unfortunately, there are several problems with this proposed scheme, particularly with the setup cost. The first is that including a setup cost is not equitable; users with short conversations subsidize users with long conversations. Second, a discriminatory effect is introduced with the setup cost in that poorer users may no longer be able to afford a connection. Further, average utilization is lower when setup costs are introduced and hence the goal of network efficiency is not fulfilled. However, inclusion of a setup cost is always preferable when maximizing revenue. This is an instance where the tradeoff between network efficiency and revenue is demonstrated.

A second example of a reservation based pricing scheme was developed by Ferrari et al (Parris & Ferrari 1994). 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 amount charged to 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 both difficult and not always realistic to obtain.

TOD Pricing

TOD or *time of day* pricing refers to schemes that try to accommodate peak and off-peak network periods. Parris et al (Parris, Keshav & Ferrari 1994) 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 (eg. 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.

The main conclusion drawn from TOD simulations is the fact that user calls are dispersed over a longer period resulting in a more even distribution of utilization. Unfortunately, this results in a lower maximum utilization over all time periods. In addition, no revelation mechanism is implemented making these TOD pricing schemes susceptible to misbehaving users.

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. 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. The equilibrium price is the price where the quantity of resource demanded is equal to the quantity of resource supplied. Figure 1 shows simple supply and demand curves and the equilibrium price resulting due to their intersection.

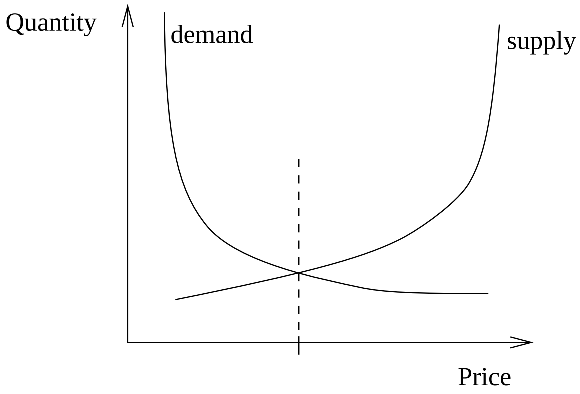


Figure 1: Example Supply and Demand Curves

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 cost recovery and network efficiency goals. Three are cited here.

Transport Auction

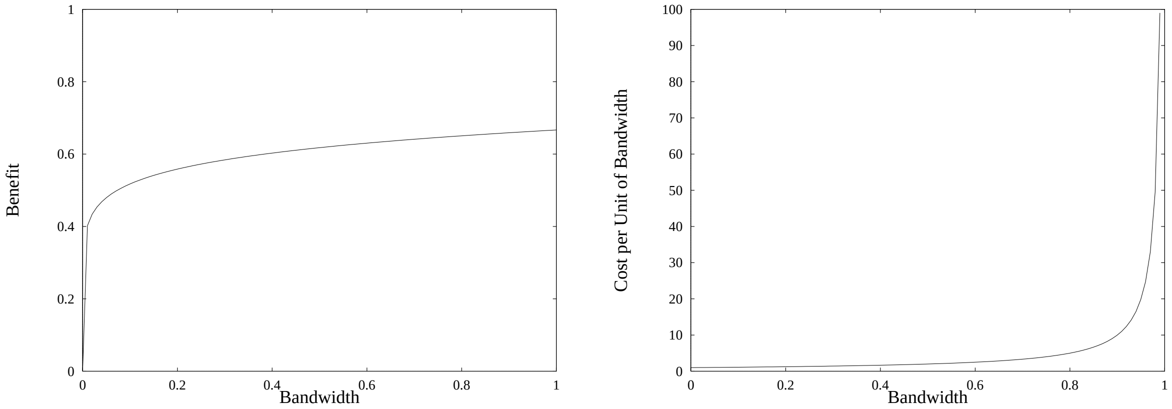
The first dynamic pricing strategy is a unique idea by Hayer (Hayer 1993) in which excess capacity on the network is auctioned to users with delay tolerant traffic. The basic idea of Transport Auction is to set the price dynamically in response to the changes in demand so as to stimulate usage and thus 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, since Transport Auction occurs on a per trunk basis, there is no ability to reroute traffic resulting in poor reliability. The third and perhaps most significant problem is that a revelation mechanism is absent and greedy or misbehaving users may bid down the price or misrepresent their traffic.

Dynamic Bandwidth Allocation

In the second dynamic pricing strategy, Murphy et al (Murphy, Murphy & Posner 1995) take a different



(a) (b)

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

but related approach to the use of supply and demand in data 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 curve. An example benefit curve 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 curve is increasing when bandwidth received is low. The benefit function then levels off as the amount of bandwidth received is increased. This curve 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 setting, an interval is simply some regular unit of time.

The network, on the other hand, derives the price of bandwidth in the next interval based on the marginal cost of sending the user's traffic and the current price. In doing this, the network defines a cost function, as shown in Figure 2(b), which takes into account the capacity of the link and the sum of all bandwidth used. As the total bandwidth used approaches the link's capacity, the cost per unit approaches infinity. In using a cost function with this shape, a low cost per unit of bandwidth when capacity is low should encourage users to demand more bandwidth, whereas a high cost per unit of bandwidth when capacity is high should encourage users to demand less bandwidth.

It was found through experimentation that an equilibrium state is not always reached when the user community has the same benefit function. Instead, utilization and price oscillate between a high and low state. In particular, simulation results indicate that the more elastic the users are, the more severe the oscillation. However, it was also found that an equilibrium state is reached when the user community has heterogeneous, or varying, benefit functions. The Murphy et al algorithm was studied in detail and is presented in the Simulations section.

Virtual Leaky Bucket and Loss-Load Curves

The third dynamic pricing scheme is actually a congestion control scheme by Prithviraj and Williamson (Prithviraj & Williamson 1995) that may extend itself to a framework for pricing. The congestion control strategy incorporates two basic components: an augmented virtual leaky bucket and loss-load curves.

The first component, an augmented virtual leaky bucket, acts as a basic leaky bucket (Partridge 1994), but instead of dropping offending cells, the cells are tagged. If congestion later occurs, the tagged cells are the first to be dropped. Users may also vary the token generation rate in this type of leaky bucket, unlike in the basic leaky bucket. The virtual leaky bucket is simple but extremely conservative.

The second component, loss-load curves (Williamson 1994), allocate bandwidth based on changing network conditions. An example of a loss-load curve is shown in Figure 3. A loss-load curve exists for each of the switch's outgoing links. The curve expresses the sender's cell loss probability as a function of his or her load. Cell 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 an amount of bandwidth and the loss-load curve determines the loss the sender will experience. Users have the ability to maximize throughput, minimize response time or minimize loss when choosing where they operate on the loss-load curve. For example, users who wish to minimize loss must request a smaller percentage of bandwidth than users with loss-tolerant traffic. On the other hand, those senders with loss-tolerant traffic who wish to maximize throughput may choose to request a higher percentage of bandwidth and hence experience greater loss. An example raw throughput curve is also shown in Figure 3.

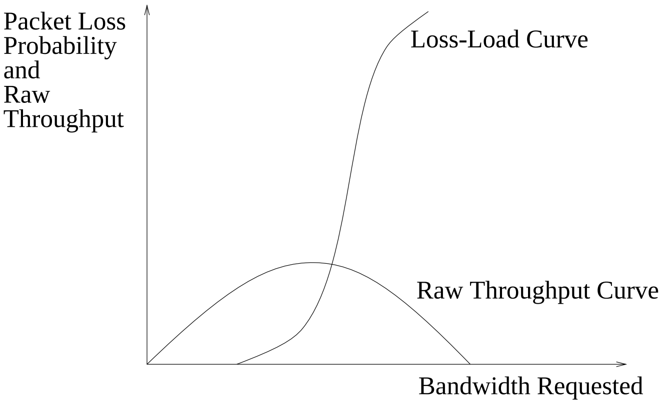


Figure 3: A Sample Loss-Load Curve and Throughput Curve

Loss-load curves have a concise mathematical representation, (see (Williamson 1994)), 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 cell loss often results.

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 simulations in (Prithviraj & Williamson 1995). Unfortunately, the switches are given a virtual capacity in the combined model in order to reduce loss, removing the loss-load revelation mechanism and making sender cooperation again necessary. Pricing may be the added dimension needed in this model to not only recover costs but also to punish greedy senders.

Simulations

Presented in this section are simulations of two pricing schemes. The first simulation involves a static pricing scheme by Parris, Keshav and Ferrari (Parris, Keshav & Ferrari 1994) in which a reservation based strategy is implemented. The simulations conducted are an attempt to recreate the results reported by Parris et al.

The second simulation involves Murphy, Murphy and Posner's dynamic bandwidth allocation pricing scheme (Murphy, Murphy & Posner 1995). This scheme was simulated to evaluate its potential for use in future strategies. Several interesting results are reported.

Static Pricing Scheme

The first experiment involves three separate simulations using Parris et al's (Parris, Keshav & Ferrari 1994) reservation based per unit pricing model. The first simulation implemented simple per packet pricing in which a sender was charged a fee per packet. The second added a setup cost to the basic per packet model and the last experiment modelled peak and off-peak pricing intervals.

The users of the network were characterized in all simulation sets with three attributes including 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 could either request a large portion of the bandwidth per interval connected or a small portion. A Class A user always requests only 1% of the bandwidth. This class of user models a typical user with moderate bandwidth needs. For simplicity, 1% of the bandwidth is equivalent to one packet. 80% of all users in the following simulations are Class A users.

Class B users, on the other hand, always request 5% of the bandwidth. This user has much higher than average bandwidth needs and must pay 5 times as much for this bandwidth. Only 20% of all users fall into this category.

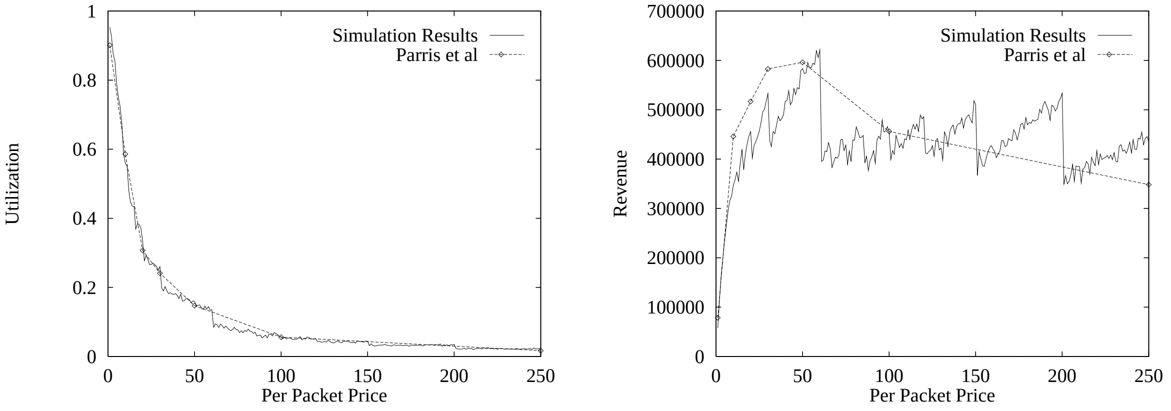
Each user will be connected to the network for a number of time intervals. This number constitutes the duration attribute. The duration of calls follows an geometric distribution with a mean of 36 intervals for Class A users and a mean of 240 intervals for Class B users. Again, Class B users are assumed to have much more intensive bandwidth needs.

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 all users were poor and only had \$60 to spend. The probability of being rich was independent of the class of a user.

The first set of simulations were done 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 this set of simulations, the price per packet was varied from \$1 per packet to \$250 per packet in increments of \$1. The network dropped a connection if the sender ran out of money or their requested duration had elapsed.

In these simulations, both utilization and revenue were monitored. Utilization indicates the total percentage of bandwidth used over the entire simulation run. A higher utilization is considered advantageous as long as capacity is not exceeded. Revenue is the dollar amount collected over all intervals of the simulation run. Again, higher revenue values are more advantageous.

Figure 4(a) shows the simulation results obtained for utilization as well as the Parris et al results reported in (Parris, Keshav & Ferrari 1994). As expected, utilization is high when the price per packet is low. As the price per packet increases, utilization steadily decreases until it remains steady around \$150 per packet. Simuation results match Parris et al results almost ex-



(a) (b)

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

actly for every price.

Revenue results are shown in Figure 4(b). Simulation results show a general rise and then fall as the price per packet increases. There is also a definite saw-tooth pattern with significant periodic falls. These falls may involve the fact that a user may not have the opportunity to spend all money allotted him or her. If the user does not have enough to send one last whole packet, the money must sit idle since sending partial packets is not allowed.

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 not all price level results were reported in (Parris, Keshav & Ferrari 1994).

The second set of simulations were identical to the first set with the addition of a setup cost. The setup cost is a one-time fee charged to the sender upon connection. The setup cost 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 cost.

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 cost 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. As in the first set of simulations, the same saw-tooth pattern is seen and the Parris et al results again are in 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 cost appears to have shifted revenue slightly towards the left so that the same revenue is generated with a lower per packet price.

The third set of simulations was different from the first two in a number of ways. First, the third simulation divided the number of intervals into peak and off-peak periods. During peak periods, call arrivals followed a Poisson distribution with a mean of two arrivals per interval. The off-peak period remained unchanged and still followed a Poisson distribution with 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 arrived 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 were 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 cost for peak and off-peak packets was increased 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 for simulation set three. 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 the first two scenarios. There is also a marked difference between the Parris et al results and the simulation results. Since the simulated utilization values

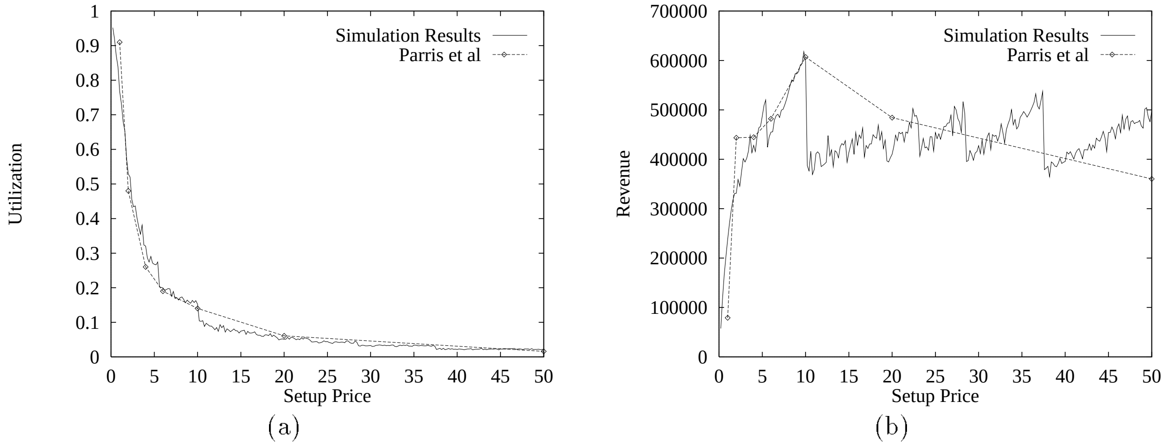


Figure 5: Per Packet Pricing with Setup Cost: a) Utilization b) Revenue

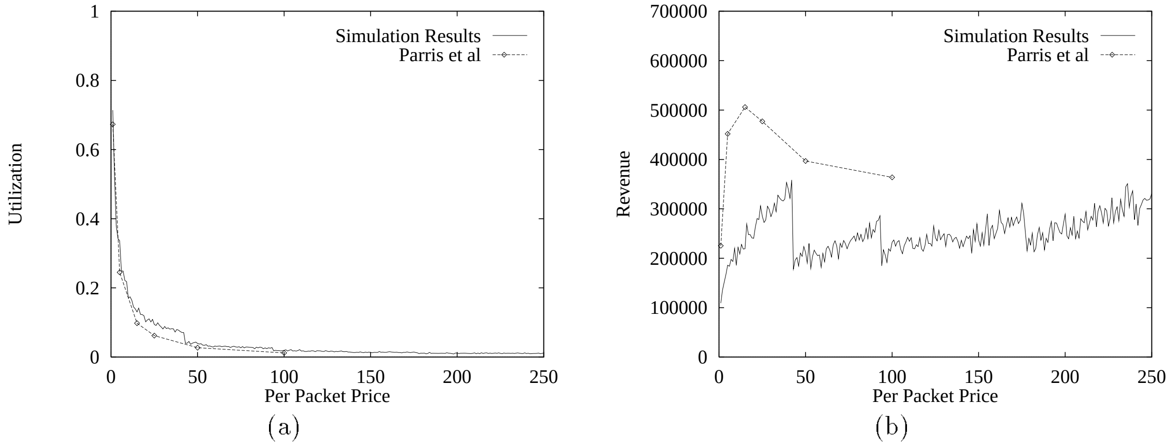


Figure 6: Peak Pricing: a) Utilization b) Revenue

appear to exactly match Parris et al utilization values, the variation in revenue values most likely arise due to different implementation assumptions.

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.

Dynamic Pricing Scheme

The second experiment involved implementing a dynamic bandwidth pricing algorithm developed by Murphy et al (Murphy, Murphy & Posner 1995). 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 curve. Users are assumed to act in their own best interest and place benefit on any bandwidth they receive. An example benefit curve is shown in Figure 2(a). Note that the curve is increasing 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 benefit curve is defined as having two components, a scale component, which shifts the curve up or down, and a shape component, which determines its amount of curvature. The function for the benefit curve is defined by: $y = \alpha x^{\beta}$ where $0 \leq \alpha \leq 1$ is the scale component, and $0 \leq \beta \leq 1$ is the shape component. $0 \leq x \leq 1$ is the fraction of the network bandwidth that a user will request at connection time.

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.

This algorithm is further explained in detail in (Murphy, Murphy & Posner 1995).

The cost function used in the algorithm characterizes the network. 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 overflow may occur, the cost per additional unit of bandwidth approaches infinity such that no additional units are requested by users.

In evaluating the Murphy et al algorithm, several different sets of simulations were conducted. Each simulation was conducted on a one link network with a perfect non-blocking switch. If the total bandwidth requested by all users exceeded the capacity of the link, cells were sent in a round robin fashion from user 1 to user n with excess cells discarded. In these simulations, the number of users on the network varies from 2 to 20.

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

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 greatest range of values for total bandwidth requested occurs when there are twenty users on the network. The total bandwidth requested also increases as the number of users on the network increases. The second phase is the equilibrium phase in which the total bandwidth requested does not change over time. Note that it takes less time to reach this state when there are fewer users on the network.

Figure 7(b) shows the price per unit of bandwidth plotted over time for the same user group. Again, there are two phases present: stabilization and equilibrium. The stabilization phase shows that the price falls first before levelling out. The range of values covered by the price decreases as the number of users decreases. In fact, when there are two users on the network, the price appears to rise before equilibrium, not fall. 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.

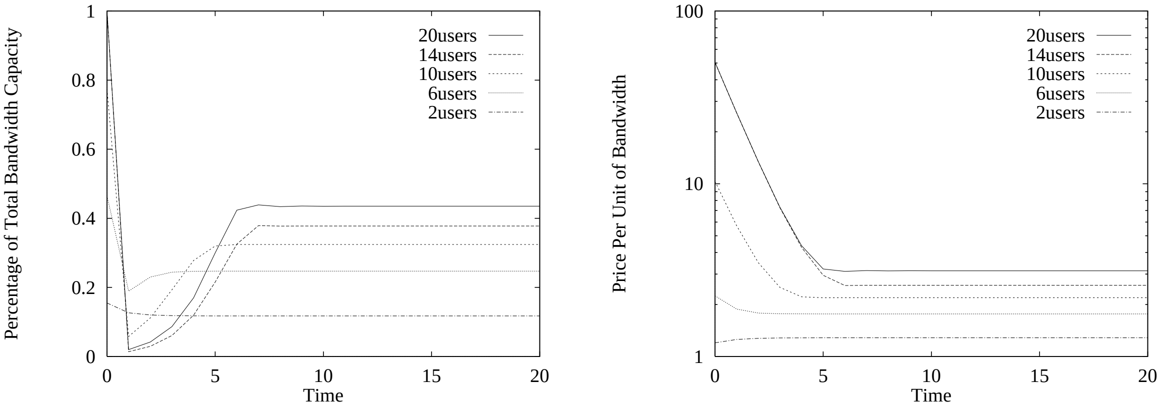
The second set of simulations also involved a homogeneous user group; however, this group had a different benefit curve defined by $y = 1.0x^{0.9}$. Figure 8(a) shows vastly different results in comparison to the previous homogeneous user group.

In Figure 8(a), there are again two phases. The first phase, the stabilization phase, shows a vast fall in total bandwidth requested by all numbers of users on the network before reaching the second phase. In fact, for a short time 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. This oscillation appears to be worse as the number of users on the network increases.

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 and 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.

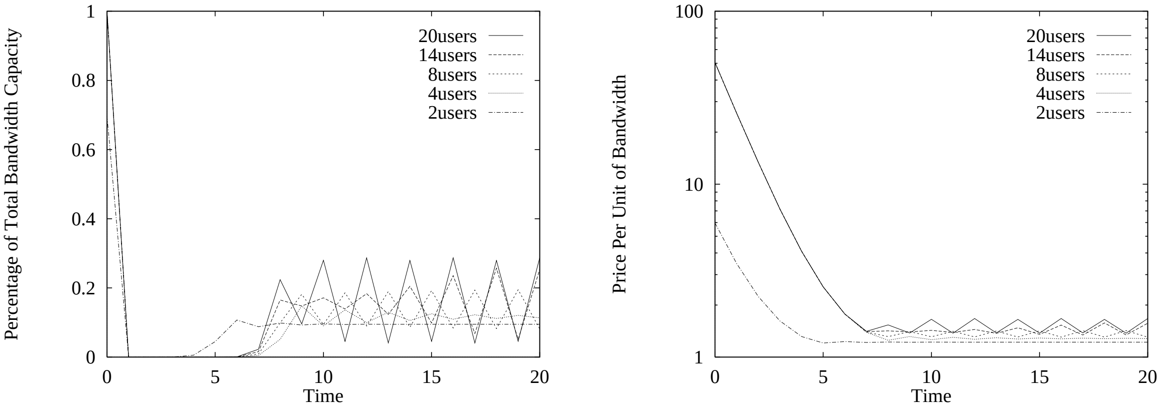
The third set of simulations involved users with differing network needs. Users in this heterogeneous user group were given random user benefit curves of the form $y = \alpha x^{\beta}$. The number of users on the network were again varied from 2 users to 20 users. Simulations found that two phases were 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. Graphs of these results are found in (Breker 1995).

Finally, the initial price which the network presented to the users was varied in a final set of simulations. 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 and others are also fully explained in (Breker 1995).



(a) (b)

Figure 7: Homogeneous Users Alpha = 1.0, Beta = 0.1: a) total bandwidth b) price



(a) (b)

Figure 8: Homogeneous Users Alpha = 1.0, Beta = 0.9: a) total bandwidth b) price

Several conclusions can be made from the simulation results with respect to the number of users on the network. First, the more users on the network, the longer the stabilization phase (ie. the longer the period until an equilibrium price is reached). As well, the more users on the network, the greater the range of values spanned by the utilization and total bandwidth requested before this second state is reached. Second, the equilibrium price increases as more users are added to the network. The sum of all users bandwidth also increases as more users are added. Third, the higher the number of users on the network, the greater the oscillation of both price and bandwidth when oscillation is present.

The simulations in which the initial price was changed also produced several interesting results. First, varying the initial price changes the stabilization period but does not affect the final state values of both price and total bandwidth requested. Second, as the initial price increases, the more quickly the final two-state oscillation or equilibrium state is reached.

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. Lastly, the initial price affects the rate at which the second phase is achieved, but not the values at the second phase.

Conclusions

New ways to recover ever rising costs and regulate network usage are becoming increasingly important. Usage-based pricing is one viable option to both generate revenue for the service provider and promote efficient network usage. This paper has presented a survey of several proposed pricing schemes, and a more detailed simulation evaluation of two such pricing strategies, one static and one dynamic.

Several interesting results were obtained from the simulations. First, static pricing schemes are simple and offer a fairly straight forward method of raising revenue. However, static schemes simply do not adapt to changing network conditions and hence do not encourage efficient network usage. Further, in current simulations, they operate under very heavy restrictions making it nearly impossible to extend them easily to meet the needs of future networks.

Second, dynamic pricing schemes react very well to changing network conditions and easily manipulate user demand by implementing the econometric principles of supply and demand. They offer the flexibility and generality that static schemes lack. Unfortunately, dynamic pricing schemes are slightly more complex and have not been fully examined.

Finally, it was apparent in both pricing schemes studied that a trade-off between the goals of network efficiency (i.e., network utilization) and generated revenue exists. For example, in the static scheme, the price per packet at which revenue was maximized resulted in a low utilization value. In the dynamic scheme, when the equilibrium price was reached, utilization was again below 50%.

Future Directions

The Murphy et al dynamic bandwidth allocation scheme operates fairly conservatively but shows potential in generating revenue for service providers. It also demonstrates an achievable state of equilibrium in network usage. However, the Murphy et al algorithm must be further tested with dynamic arrivals and departures of network users to ensure continued equilibrium. If this proves successful, the aggressive loss-load curves in (Williamson 1994) and its variants will be explored in combination with the Murphy et al approach.

References

Breker, L. 1995. Evaluation of a Dynamic Bandwidth Pricing Scheme. Available at tt http://www.cs.usask.ca/grads/lpb133/.

Cocchi, R.; Shenker, S.; Estrin, D. and Zhang,L. 1993. Pricing in Computer Networks: Motivation, Formulation, and Example. *IEEE/ACM Transactions on Networking* 1(6):614-627.

Groves, T. 1995. Financing the UCSD Network. Available at tt http://gopher.econ.lsa.umich.edu/-EconInternet.html.

Hayer, J. 1993. Transport Auction: A New Service Concept. M.Sc./MBA Thesis, TR-93-05, TRLabs Edmonton.

Murphy, J., and Murphy, L. 1995. Pricing for ATM Network Efficiency. Available by anonymous ftp from tt ftp.eeng.dcu.ie/pub/tele-communications/murphy.

Murphy, J.; Murphy, L., and Posner, E. 1995. Distributed Pricing For Embedded ATM Networks. Available by anonymous ftp from tt ftp.eeng.dcu.ie/pub/tele-communications/murphy.

Partridge, C. 1994. *Gigabit Networking.* Addison-Wesley Publishing Company, Reading, MA.

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

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

Prithviraj, N., and Williamson, C. 1995. A Virtual Loss-Load Congestion Control Strategy. Technical Report, Department of Computer Science, University of Saskatchewan.

Shenker, S. 1994. Making Greed Work in Networks: A Game Theoretic Analysis of Switch Service Disciplines: In *Proceedings of ACM SIGCOMM'94*, 47-57. London, UK.

Williamson, C. 1994. Dynamic Bandwidth Allocation Using Loss-Load Curves. Department of Computer Science, University of Saskatchewan.