QoS Rewards and Risks: A Multi-market Approach to Resource Allocation*

Errin W. Fulp and Douglas S. Reeves

Departments of Computer Science & Electrical and Computer Engineering
North Carolina State University, Raleigh NC 27695, USA

Abstract. A large number of network applications require a particular Quality of Service (QoS), that can be provided through proper network resource allocation. Furthermore, certain applications (multimedia oriented) may require guarantees of resource availability for predictable QoS. This paper introduces a distributed multi-market approach to network resource allocation. In this approach link bandwidth is bought and sold in two types of markets: the reservation market and the spot market. Together, these markets provide bandwidth guarantees and immediate availability. In addition, users have more flexibility when purchasing bandwidth that will maximize their individual QoS. Experimental results, using actual MPEG-compressed traffic, will also demonstrate the rewards and risks associated with purchasing various amounts in the reservation and spot markets.

1 Introduction

Many network applications require a certain Quality of Service (QoS) from the network for their proper operation. QoS may include bounds on packet delay, loss and jitter. The network can provide QoS by properly allocating its resources, such as link bandwidth, processor time, and buffer space. Furthermore, some applications may require a guarantee of resource availability for a duration of time. For example, the QoS of an application may be sensitive not only to the amount of link bandwidth allocated, but also any changes in the allocated amount that may occur. Due to the finite supply of resources and the various demands, fairly and efficiently allocating resources to provide QoS is a challenging problem.

Recently, microeconomics has been applied to network resource allocation and flow control. Congestion pricing is a well-known microeconomic approach that charges users for their consumption of resources, and prices are set based on supply and demand [1,4,6,11]. Alternatively, prices can be set with respect to

* This work was supported by AFOSR (grants F49620-96-1-0061, F49620-97-1-0351 and F49620-99-1-0264) and DARPA Tolerant Networks Program of ITO. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the AFOSR, DARPA or the U.S. Government.

G. Pujolle et al. (Eds.): NETWORKING 2000, LNCS 1815, pp. 945–956, 2000.

marginal costs [7]. With such a model, prices can be calculated in a distributed fashion to encourage high utilization of network resources as well as a fair distribution. However in many cases, the transient behavior and the method of distributing intermediate prices (and/or allocations) during convergence is generally ignored. Furthermore, many have not been validated in a detailed way using realistic network configurations and actual traffic.

In [3] we introduced a congestion pricing technique based on the competitive market model. Unlike other congestion pricing methods, this distributed technique allows changing traffic demands by dynamically adjusting the link bandwidth price in response to current supply and demand. It has been proven and demonstrated that this method can achieve Pareto-optimal and fair allocations as well as high link utilization [2]. Although able to achieve optimal allocations under dynamic conditions, a limitation of this method (shared with other methods) is that no guarantee of future bandwidth availability can be made. This may be acceptable for "elastic-traffic" that can easily adjust demands based on network conditions [12]. For other applications (such as high definition video), changing bandwidth amounts may result in a sudden reduction of QoS or in the worst case, being forced out of the economy due to high prices. Therefore, it is advantageous to shield these users from unpredictable price fluctuations.

In this paper, a *multi-market* approach for bandwidth allocation is presented. Multi-market methods have been proposed for allocating link bandwidth and buffer space [5,11]; however the primary focus was not to provide price or QoS stability. In our approach link bandwidth is bought and sold in two types of markets: the reservation market and the spot market. In the reservation market, bandwidth is bought and sold in amounts for a duration of time. Bandwidth purchased in the reservation market may be used to provide some required (minimal) QoS. In the spot market, a user can immediately use whatever amount of bandwidth they find affordable, with no reservation overhead. Therefore, this multi-market approach combines the unique advantages of the single spot market [2] (such as immediate availability) with the price stability offered in the reservation market. This is done in a distributed fashion with a state-less implementation. The flexibility of the multi-market approach is well suited to accommodate the diverse QoS needs of various elastic and non-elastic network applications.

The remainder of this paper is structured as follows. Section 2 describes the new multi-market technique in detail. Section 3 provides a demonstration of multi-market allocation method. Finally, section 4 reviews the multi-market technique, summarizes the results and discusses some open questions.

2 A Multi-market Approach

Similar to the spot market economy [3], the multi-market economy is based on a competitive market model, where pricing is done to promote Pareto-optimal and fair distributions, as well as high utilization [8]. However, unlike the spot market approach the multi-market provides guarantees of resource availability

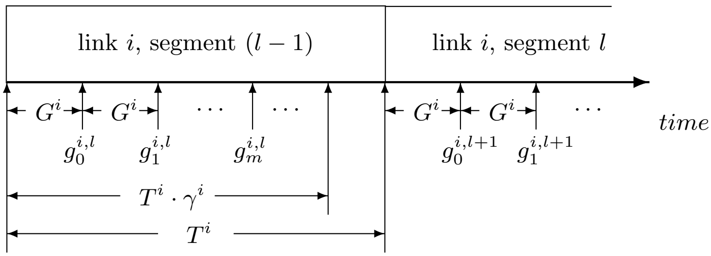


Fig. 1. Example segments and price calculation points for link i reserved bandwidth.

as well as price and QoS stability. This results in an allocation method that is appropriate for elastic and non-elastic traffic.

There are three entities in this network economy: users (those who execute network applications), Network Brokers (NB) and switches. Using the competitive market nomenclature, users are consumers, switches are producers and network brokers are used to assist the exchange of resources in the market. While there are many resources in a computer network, this paper focuses on the pricing of link bandwidth.

2.1 Switch

The network consists of several switches interconnected with links. For a unidirectional link between two switches, we consider the sending switch as owner of the bandwidth of that link. For link i denote the total capacity of link bandwidth as S^i. The capacity is then divided into two types, *reserved* and *spot* bandwidth. Reserved bandwidth is sold as an amount for a duration of time, while spot bandwidth is sold as a non-storable resource. With this distinction, reserved and spot bandwidth are considered separate resources. Reserved bandwidth has the unique advantage of ownership over a period of time, while the advantage of spot bandwidth is its immediate availability. Each resource is sold in its own local (link) market, therefore the switch will associate two markets per output port (thus the "multi-market" designation). These markets operate independently and asynchronously since there is no need for market communication (for example, price comparisons) or synchronization from switch to switch. Since the link capacity is divided into reserved and spot bandwidth, the switch must differentiate the traffic using either type. We assume that a bit will be set in the header of the packet, indicating the packet is using reserved bandwidth.

Reserved Bandwidth Market Link i will sell a maximum fraction β^i of S^i as reserved bandwidth. The reserved bandwidth is divided into equal non-overlapping intervals of time called segments, where the length of each segment is denoted as T^i. Portions of the segment are then sold to users with an auction procedure. Users are only able to bid for an amount of the next segment; therefore, the reserved bandwidth of segment l is auctioned during segment $(l-1)$. At the beginning of the auction for segment l of link i, users forward a bid (for an amount of reserved bandwidth they wish to purchase) to the switch. The sum of these bids, denoted as $h_m^{i,l}$, is recorded by the switch and is used to update the price. During the auction, the price of reserved bandwidth for segment l is adjusted at regular intervals G^i, as seen in figure 1. The mth auction price for reserved bandwidth of link i, segment l is denoted as $g_m^{i,l}$. The price for the next interval is calculated using the following tâtonnement process,

$$g_{m+1}^{i,l} = g_m^{i,l} \cdot \frac{h_m^{i,l}}{\beta^i \cdot S^i} \tag{1}$$

The tâtonnement process adjusts the price at regular intervals, based on the demand ($h_m^{i,l}$) and the supply ($\beta^i \cdot S^i$). If the demand is greater than the supply, then the price increases (and vice versa). After a new auction price is calculated, it is distributed to NB's, who may submit updated bids. It is important to note the switch does not need to store individual bids. Users can initially submit a bid amount, then send only changes (differences) to the switch. This process repeats until an *equilibrium price* $g_*^{i,l}$ for segment l is determined. An equilibrium price causes demand to equal the supply. At this price bandwidth is sold and the resulting allocation is Pareto-optimal and fair [2]. New bidders (users who have not yet participated in bidding for the segment) are not allowed to participate after $T^i \cdot \gamma$ (where $0 < \gamma < 1$) has passed. This provides time for the auction to converge to the equilibrium price before the segment begins. At the end of auction the switch notifies the users that a new segment has begun[1]. Users are then able to use the amount of reserved bandwidth they defined in their last bid (explicit notification is not necessary). Since only aggregate information (not individual) is used and it is not necessary for the switch to store individual bids, the auction process for the reservation market can be considered a state-less implementation.

Bandwidth Spot Market As described in [3], bandwidth in the spot market is considered a non-storable resource (similar to electricity); therefore users/NB's are unable to purchase spot bandwidth in hopes of using it a later time. Once a NB has determined an affordable amount of spot bandwidth, the user can send immediately (no reservation overhead is required)[2].

The spot market price for link i is calculated at the switch, at discrete intervals of time. The price during the nth interval is constant and is denoted as p_n^i. At the end of the nth spot market price interval, denote the demand for spot market bandwidth as d_n^i and the amount of reserved bandwidth currently used as r_n^i where $r_n^i \leq \beta^i \cdot S^i$. The total spot market supply for link i is $S^i - r_n^i$; therefore any reserved bandwidth that is not used can be sold in the spot market. At the

[1] Methods of price distribution and user notification, as well as queueing delay issues are discussed in [2].

[2] Without reservations, selling the same spot bandwidth to multiple users may occur. How this is avoided in described in [2].

end of the nth interval, the switch updates the spot market price of link i using a *modified tâtonnement process* [3]. A limitation of the tâtonnement process, in its original form (equation 1), is the inability to dynamically adapt to changing demands[3]. To handle such dynamics, the spot market price is determined using the following modified tâtonnement process.

$$p^i_{n+1} = p^i_n \cdot \frac{d^i_n}{\alpha^i \cdot S^i - r^i_n} \tag{2}$$

The modified tâtonnement process adjusts the price at regular intervals, based on the demand (current spot traffic) and the supply. The bandwidth supply is the total bandwidth times a constant α^i (where $0 < \alpha^i \leq 1$) minus the amount of reserved bandwidth currently used (r^i_n). This modification causes the price to increase after some percentage (α^i) of the available spot bandwidth has been sold. This is evident from the equation, since the price will only increase if the numerator is greater than the denominator ($d^i_n > \alpha^i \cdot S^i - r^i_n$). Once the new price p^i_{n+1} is calculated it is distributed to the users/NB's using the link. Upon receiving the new price users/NB's adjust their transmission rate. As demand changes, the modified tâtonnement process dynamically adjusts the price seeking the new equilibrium price p^i_*. Similar to the reservation market, the spot market is state-less since the bandwidth price is calculated using only the local aggregate demand, supply and current price.

2.2 User

A user, executing a network application, requires link bandwidth for transmission. The amount of bandwidth desired (or the desired rate) is determined from the application and is denoted as a. Based on prices and wealth, the user can afford a range of bandwidth (less than or equal to a), and some amounts will be preferred over others. In economics these preferences are represented with a utility function. The utility function maps a resource amount to a real number, that corresponds to a satisfaction level. The utility curve can be used to compare resource amounts based on the satisfaction the user will receive. For this economy we will use *QoS profiles* [9] for the utility curves. The profile can be approximated by a piece-wise linear curve with three different slopes, as seen in figure 3(b). The horizontal axis measures the bandwidth ratio of allocated bandwidth to desired bandwidth. The vertical axis measures the satisfaction and is referred to as a QoS score. Our QoS scores range from one to five, with five representing an excellent perceived quality and one representing very poor quality.

Since there are two different types of resources in the economy (reserved and spot bandwidth), the user must identify how the reserved and spot bandwidth may be substituted for one another. In microeconomics these preferences

[3] In the reservation market the equilibrium price must be determined before bandwidth is sold; therefore, if demands change a new equilibrium price must be determined.

are represented with an indifference curve. An indifference curve indicates the combinations of resources that result in the same utility [8]. For our economy, indifference curves indicate the combination of reserved and spot bandwidth that result in the same utility. These curves are normalized to the current desired amount of bandwidth a, as seen in figure 2(a). For example in figure 2(a), the indifference curve labeled "prefer-reserved" was generated from the equation $y = k \cdot (1 - x)^2$, where y is the amount of spot bandwidth and x is the amount of reserved bandwidth. The preference for reserved bandwidth increases as k increases. The other indifference curve given in figure 2(a) represents a user who has no preference for reserved or spot bandwidth. In this case, spot and reserved bandwidth are considered "perfect substitutes" and the user will always prefer the *cheaper* (lower cost) bandwidth. In the case where the two types of bandwidth have the same price, we assume the user will prefer equal amounts of either type of bandwidth. The only assumptions required for the indifference curve is that it must be continuously differentiable and convex to the origin (required for determining the amount of bandwidth to purchase).

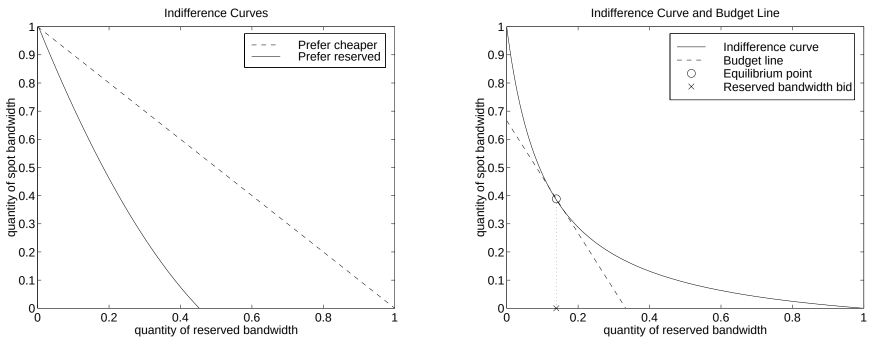
Finally, the user is charged continuously for the duration of the session (analogous to a meter). To pay for the expenses, we will assume the user provides an equal amount of money over regular periods of time [2]. We will refer to this as the budget rate of the user, W (\$/sec).

2.3 Network Broker

Users can only participate in the network economy through a network broker (NB). This entity is an agent for the user and is located between the user and the edge of the network. Representing the user in the economy the NB performs the following tasks: connection admission control, policing, packet marking, and purchasing/bidding decisions. Although the NB works as an agent for the user (making purchasing decisions), we assume that the NB operates honestly in regards to both the switches and the user.

The NB monitors the user and the prices by gathering and storing information about each. From the user, the NB collects and stores; the QoS profile, indifference curve, a and W. The NB also stores the route, R, that connects source to the destination, where R consists of v links, $\{l^i, i = 1 \ldots v\}$. For each link on R, the NB collects current reserved and spot bandwidth prices[4]. The NB will divide the budget rate, W, into a vector of v budget rates $\boldsymbol{w}$, where $\boldsymbol{w} = \{w^i, i = 1 \ldots v\}$ and w^i corresponds to link i. Separate budgets are used to localize the effect of prices to each link. Using this information the NB controls network admission by initially requiring the user to have enough wealth to afford at least an *acceptable* QoS; otherwise, the user is denied access. The NB also levies the user for their consumption. In addition, the NB polices the user, ensuring only the bandwidth purchased is used, and marks the packets (assigning which are to use reserved bandwidth). Finally, the NB determines the reserved bandwidth bid and the amount of spot bandwidth to purchase.

[4] The requirement that the NB must know the entire route, and store a distinct price per link, can be relaxed [2].



(a) Example indifference curves, prefer-reserved and prefer-cheaper bandwidth.

(b) The equilibrium point, where slopes of the indifference curve and budget line are equal.

Fig. 2. Example indifference curves and how they are used.

Bidding and Purchasing Bandwidth Since reserved bandwidth is sold in an auction format, the NB must bid for reserved bandwidth at each link on the route. The bid (the amount the user will purchase) for link i is based on the budget w^i, a statistic of the spot bandwidth price, and the reserved bandwidth auction price for this link $g^{i,l}$. A statistic is required for the spot bandwidth price due to its volatility. For our discussion (and simulation) we will use the maximum spot price, $\hat{p}^i$, measured during the current segment. Using this information the NB maximizes the utility of the user $u(x, y)$,

$$\max\{u(x,y)\}, \quad g^{i,l} \cdot x + \hat{p}^i \cdot y \leq w^i \tag{3}$$

where x is the amount of reserved bandwidth and y is the amount of spot bandwidth. As defined in [8], the first order condition of this constrained maximization problem is,

$$\frac{\partial u/\partial x}{\partial u/\partial y} = \frac{g^{i,l}}{\hat{p}^i} \tag{4}$$

The NB must spend the budget to equalize the ratio of marginal utility to the price of each resource. Plotting the budget line with the indifference curve, as seen in figure 2(b), the slope of the budget line is $g^{i,l}/\hat{p}^i$. Therefore, the point where the slope of the indifference curve and the budget line are equal is where the utility of the user is maximized. The second order condition is not required due to the convexity assumption of the indifference curve [8]. The reserved bandwidth bid is the x component of this point and is forwarded to link i.

The NB keeps a table storing the amount and price of reserved bandwidth purchased at each link in the route. When the segment for link i is sold, the amount purchased c^i and the price f^i is updated in the table. The maximum amount of reserved bandwidth that can be used is,

$$e = \min_{i=1\ldots v} \{c^i\} \tag{5}$$

which is the minimum amount of reserved bandwidth purchased at any link. If the desired bandwidth is greater than the purchased reserved bandwidth, then spot bandwidth is used for transmitting the remaining portion $(a - e)$. The amount of spot bandwidth to use y is,

$$y = \min \left\{ \min_{i=1\ldots v} \left\{ \frac{w^i - e \cdot f^i}{p_n^i} \right\}, a - e \right\} \tag{6}$$

which is the maximum amount of spot bandwidth that is affordable, but no more that what is required $(a - e)$.

2.4 Optimality and Network Dynamics

For any allocation method it is important to address the optimality and fairness that is achievable. For an economy consisting of multiple competitive markets, once the tâtonnement processes (one per market) reach equilibrium, the resulting allocation is Pareto-optimal and fair [2]. Due to the nature of multimedia traffic, it is also important to determine if the allocation technique can adequately handle network dynamics (users entering/exiting and variable bit rate sources). Due to the complexity of actual network dynamics, simulations have been used to demonstrate the performance of the competitive market approach. Experimental results have shown the spot market achieves optimal allocations over 90% of the time under realistic network conditions [3]. In the next section, simulation is used to demonstrate the performance of the multi-market economy under similar conditions.

3 A Demonstration of the Multi-market Economy

In this section the performance of the multi-market network economy is demonstrated via simulation. Experiments performed will consist of a realistic network configuration, allowing users to randomly enter the network and use actual MPEG-compressed traffic. Simulation results will show that users who prefer reservations will experience less of a QoS impact than those who do not. Results also show that the preference for reservations is at the expense of lower average QoS. This represents the *QoS rewards and risks* of the multi-market.

The network simulated consisted of 160 users and their associated NB's, four switches and seven primary links, as seen in figure 3(a). Each output port carried traffic from 40 users and connected to a 45 Mbps link. Links interconnecting switches were 1000 km in length, while links connecting sources to their first switch were 25 km in length. Users had routes consisting of one, two or three hops. The network can be described as a "parking lot" configuration, where multiple sources use one primary path. The multi-market economy had the following initial values. The spot market parameter α (targeted utilization) was 90%. Switches sold equal amounts of reserved and spot bandwidth; therefore β was 45%. Reserved bandwidth prices were initialized to 55 and segments were 15 minutes in duration. Longer segments could have been selected; however, we were interested in observing the transition effects from one segment to another and the smaller segment size reduced the simulation time. Spot market prices were initialized to 50 and the update interval was 20 times the longest propagation delay. Assume no propagation delay between the user and their NB.

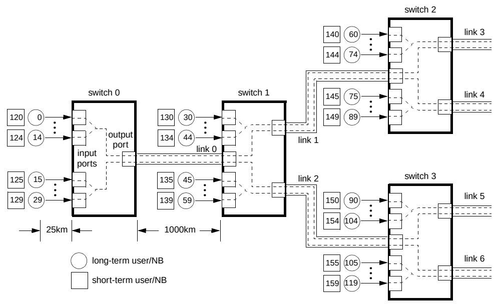
Users had budget rates[5], $\boldsymbol{w}$, of 3×10^8/sec and used the QoS profile given in figure 3(b). Since all users have the same budget rate, they are considered equal (purchasing power) in the economy [2]. The source for each user was one of 15 MPEG-compressed traces obtained from Oliver Rose at the University of Würzburg, Germany [10][6]. Although users have the same wealth and QoS profile, for this demonstration users are considered either *long-term* or *short-term*. Long-term users measure, over the duration of the simulation, the different QoS obtained from preferring different amounts of reserved and spot bandwidth. Short-term users are introduced in the economy to cause sudden demand shifts, which may occur in actual networks (peak load times). Together, we are interested in the QoS achieved by the various users in the multi-market economy.

A total of 120 users are considered long-term and have sessions that last the duration of the simulation. Half of the long-term users prefer reserved bandwidth, the remaining prefer cheaper bandwidth (indifference curves given in figure 2(a)). The long-term users enter the network at random times uniformly distributed between 0 and 600 seconds. The remaining 40 users are considered short-term. These users transmit a short segment of an MPEG video (under 3 minutes, randomly determined). Due to the relative shortness of their session, these users will only purchase bandwidth from the spot market. Starting at 3000 seconds the short-term users enter the network with a Poisson distribution of mean 120 seconds. We are interested in the link bandwidth utilization and the QoS provided to each type of long-term user. Allocation graphs are provided to measure the utilization of link bandwidth, while QoS graphs measure the average QoS observed by long-term users.

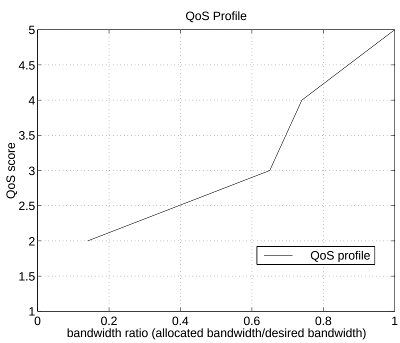
For this simulation, the example bandwidth allocations, prices and average QoS are given in figures 3(c) - 3(e). The results from 1800 to 6500 seconds are displayed since we are only interested in the effect the short-term users have in the economy. As seen in figure 3(c), all of the available reserved bandwidth for link 3 was sold, while the total bandwidth allocated stayed within the vicinity of α (targeted utilization). Similar results were noted for the remaining links. As seen in figure 3(e), before the short-term users entered the network (time less than 3000 seconds) prefer-cheaper users enjoyed a higher QoS. During this time, prefer-cheaper users only purchased bandwidth from the spot markets,

[5] The denomination is based on bps, if based on Mbps the budget would be 300/sec.

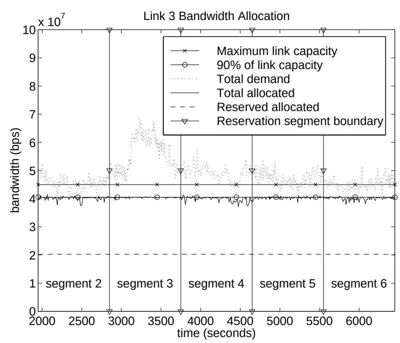
[6] Traces can be obtained from the ftp site `ftp-info3.informatik.uni-wuerzburg.de` in the directory `/pub/MPEG`



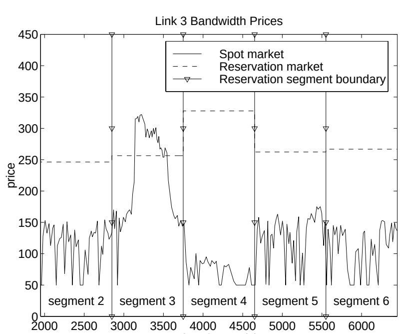
(a) Network configuration used for simulations.



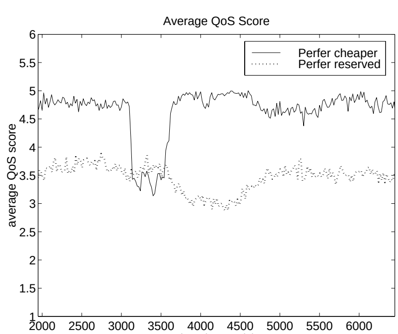
(b) QoS profile.



(c) Link 3 allocation.



(d) Link 3 bandwidth prices.



(e) Average QoS scores for all long-term users.

Fig. 3. Multi-market simulation set-up and results.

since prices were lower (as seen in figure 3(d)). In contrast, prefer-reserved users spent their entire budget in the reservation markets. Purchasing from the spot markets yielded higher QoS for the prefer-cheaper users, because these users only purchased what they needed at any time. This allows users to efficiently share the spot market bandwidth. In the case where the total demand of the prefer-cheaper users exceeded the spot market supply, each prefer-cheaper user received an equal-share of the spot market bandwidth. The prefer-reserved users were allocated an equal-share of the reserved bandwidth supply for the duration of the simulation (no more). For this reason, prefer-reserved users observed a lower QoS, until the short-term users arrived. When the short-term users arrived (during segment 3 of figure 3(c)), the spot market price increased in response to the increase in demand. During this time prefer-cheaper users received a lower QoS, since they had to compete with the new arrivals. The standard deviation in QoS observed by the prefer-cheaper users during this segment was 0.65, which was six times higher than the prefer-reserved users. The prefer-reserved users continued to receive approximately the same level of QoS since they purchased reserved bandwidth for segment 3. The spot market price increase, during segment 3, did cause a higher reservation market price for segment 4, as seen in figure 3(d). The prefer-reserved users could afford less reserved bandwidth during this segment, resulting in a slightly lower QoS. Afterwards, prices and QoS observed by the users returns to the previous values.

This simulation provides some insight to the rewards and risks of purchasing various amounts of bandwidth in the spot and reservation markets. In the example, prefer-cheaper users enjoyed the reward of a higher QoS (average of 4.65), but were susceptible to the risk of spot market price fluctuations. Applications that can gracefully handle such unpredictable bandwidth changes (such as teleconferencing) would perform better by accepting the risk associated with the spot market (which was typically cheaper). In contrast, prefer-reserved users opted for the more stable reservation market, but generally received a lower QoS (average of 3.46). Applications, such as high definition video, that can not adapt well to sudden bandwidth fluctuations would perform better with the stability provided from the reservation market.

4 Conclusions

This paper introduced a decentralized bandwidth allocation method based on a multi-market economy. A computer network can be viewed as an economy consisting of three entities (users, Network Brokers and switches) and two different markets/resources (reserved and spot bandwidth). Switches own the bandwidth sought by users, which is sold in the reservation and spot markets. Reserved bandwidth has the advantage of ownership over a period of time, providing the user with some predictability of their expected QoS. In contrast, spot bandwidth has the advantage of immediate availability without the reservation overhead. Therefore, the multi-market approach integrates the benefits of the spot market and the reservation market in one allocation technique. Both market types are modeled as competitive markets; therefore, Pareto-optimal and fair allocations are possible. The flexibility of the multi-market approach is well suited to accommodate the diverse QoS needs of various network applications (from elastic-traffic to applications that prefer predictability). Experimental results showed that users who preferred reservations experienced a more stable QoS than those who do not. However, the preference for reservations came at the expense of lower average QoS. Therefore this represents the QoS rewards and risks of the multi-market approach. Some areas for future work include the association of indifference curves with QoS profiles and dynamically adjusting market parameters in response to market trends.

Acknowledgements. The authors wish to thank Maximilian Ott and Daniel Reininger of C & C Research Laboratories, NEC USA for their significant contributions to this research.

References

1. D. F. Ferguson, C. Nikolaou, J. Sairamesh, and Y. Yemini. Economic Models for Allocating Resources in Computer Systems. In S. Clearwater, editor, *Market Based Control of Distributed Systems*. World Scientific Press, 1996.
2. E. W. Fulp. *Resource Allocation and Pricing for QoS Management in Computer Networks*. PhD thesis, North Carolina State University, 1999.
3. E. W. Fulp, M. Ott, D. Reininger, and D. S. Reeves. Paying for QoS: An Optimal Distributed Algorithm for Pricing Network Resources. In *Proceedings of the IEEE Sixth International Workshop on Quality of Service*, pages 75 – 84, 1998.
4. F. Kelly, A. K. Maulloo, and D. K. H. Tan. Rate Control for Communication Networks: Shadow Prices, Proportional Fairness and Stability. *Journal of the Operational Research Society*, 49:237 – 252, 1998.
5. S. H. Low. Equilibrium Allocation of Variable Resources for Elastic Traffics. In *Proceedings of the IEEE INFOCOM*, pages 858 – 864, 1998.
6. J. K. MacKie-Mason and H. R. Varian. Pricing Congestible Network Resources. *IEEE Journal on Selected Areas in Communications*, 13(7):1141 – 1149, Sept 1995.
7. J. Murphy, L. Murphy, and E. C. Posner. Distributed Pricing for ATM Networks. *ITC-14*, pages 1053 – 1063, 1994.
8. W. Nicholson. *Microeconomic Theory, Basic Principles and Extensions*. The Dryden Press, 1989.
9. D. Reininger and R. Izmailov. Soft Quality-of-Service for VBR+ Video. In *Proceedings of the International Workshop on Audio-Visual Services over Packet Networks, AVSPN'97*, Sept. 1997.
10. O. Rose. Statistical Properties of MPEG Video Traffic and Their Impact on Traffic Modeling in ATM Systems. Technical Report 101, University of Würzburg Institute of Computer Science, Feb. 1995.
11. J. Sairamesh, D. F. Ferguson, and Y. Yemini. An Approach to Pricing, Optimal Allocation and Quality of Service Provisioning in High-speed Packet Networks. In *Proceedings of the IEEE INFOCOM*, pages 1111 – 1119, 1995.
12. S. Shenker. Fundamental Design Issues for the Future Internet. *IEEE Journal on Selected Areas in Communications*, 13(7):1176 – 1188, 1995.