The Economic Impact of Network Pricing Intervals⋆

Errin W. Fulp[1] and Douglas S. Reeves[2]

[1] Department of Computer Science, Wake Forest University,
P.O. Box 7311, Winston-Salem N.C. 27109-7311 USA
`fulp@wfu.edu`
[2] Department of Computer Science *and* Department of Electrical and Computer
Engineering, N.C. State University, Box 7534, Raleigh N.C. 27695 USA
`reeves@eos.ncsu.edu`

Abstract. Interval pricing can provide an effective means of congestion control as well as revenue generation. Using this method, prices are fixed over intervals of time, providing adaptability and predictability. An important issue is the interval duration associated with price updates. While previous research has discussed the effect of interval lengths on congestion control, this paper investigates the economic impact of price interval duration. Smaller intervals yield higher profits since prices are more responsive to changing demands. However, experimental results indicate only a modest profit gain (no more than 5%) is achieved when smaller intervals are used as opposed to larger intervals (for example 100 times longer). Given users preferences toward fewer price changes, smaller price intervals may hold few economic benefits.

1 Introduction

A growing number of applications (e.g., multimedia-oriented) require certain network resources for their proper operation. Network resources such as bandwidth of each physical link, buffer space and processing time at each node, are finite and must be allocated in a fair and cost-effective manner. It has been demonstrated that pricing is an effective method for achieving fair allocations as well as revenue generation [1–7]. Furthermore, pricing resources prevents users from over allocating resources or requesting service classes more stringent than necessary.

Most pricing-based methods rely on the basic microeconomic principle of supply and demand to determine the appropriate resource price. For example, price increases can be used to lower demand as it approaches supply (capacity). Conversely, decreasing prices can be used to encourage usage. The objective is to

---

price resources so demand equals supply, at this point allocations are provably fair and profit is maximized [8]. Given a microeconomic-model of user demands, the appropriate price and supply can be determined. However, given the dynamic nature of computer networks (changing user demands), determining when prices should change (time-scale) remains an important issue.

Users prefer a price that is predictable, while service providers prefer a price that can change based on current congestion [9, 10, 7]. For example, spot market prices are updated over short periods of time to reflect congestion [11]. While this method does provide fair allocations under dynamic conditions, users can not accurately predict the cost of their sessions due to possible price fluctuations. In general, most users dislike the uncertainty and perceived complexity associated with a dynamic or spot price, even though it may result in lower costs [10]. In contrast, fixed prices provide predictable costs; however, there is no incentive for the user to curtail consumption during peak (congested) periods. Service providers typically prefer a price structure that reflects actual usage; thus, providing more control. As a compromise, prices can be based on slowly varying parameters such as Time of Day (ToD) statistics. As noted in [12, 13, 5, 9], the aggregate demand for bandwidth changes considerably during certain periods of the day. This is depicted in figure 1, which shows the aggregate traffic arriving and departing Wake Forest University campus during a week. Peak demand is evident during business hours, while network usage decreases at night. Given a ToD based pricing structure, an important question is determining the appropriate interval length. The effect of ToD duration (pricing interval) on congestion control has been discussed (smaller intervals provide better congestion control [5, 7]); however, the specific impact on profit has not.

This paper investigates the effect of pricing intervals on profit. For example, given a service provider and a set of customers, this paper seeks to determine the economic advantages of different sized pricing intervals. Having many smaller pricing intervals can result in higher profits. However, experiments indicate the profit increase is less than 5% on average as opposed to implementing longer intervals. As a result, the economic benefit of having smaller duration intervals is limited considering users prefer flat and less complex pricing structures.

The remainder of this paper is structured as follows. Section 2 describes the network pricing model used, consisting of individual users and a service provider. Optimal strategies for bandwidth provisioning and allocation (pricing) are then presented in section 3. In section 4, the economic impact on different sized pricing intervals is discussed and investigated experimentally. Finally, section 5 provides a summary of network pricing intervals and discusses some areas of future research.

2 Network Pricing Model

The pricing model consists of a network service provider and a set of users (customers). Users require bandwidth for their network applications. A user may start a session at any time, request different levels of QoS, and have varying
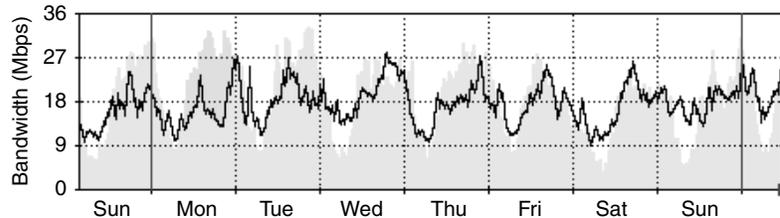
Fig. 1. Aggregate bandwidth usage for Wake Forest University. The dark line represents the incoming traffic, while the shaded area is traffic departing campus.

session lengths. Furthermore, users desire immediate network access (minimal reservation delay). In contrast, service providers own large amounts of bandwidth (or rights to bandwidth) and create connections across their network [14]. A connection would serve multiple users, and persists over a long period of time[3]. These large aggregate connections provide scalability; however, single-user connections (e.g., using the ReSource reserVation Protocol, RSVP [16]) are also possible. Once a connection is established (provisioned), portions of the connection are sold (allocated) to individual users at a price.

The price of bandwidth (charged to users) will be usage-based, where the user cost depends on the current price and the amount consumed. As previously mentioned, an important issue is the time-scale associated with the price. For example, prices could remain fixed for long periods of time or continually change based on current congestion levels. As a compromise, this model will use interval pricing. A day will be divided into T equal length periods of time, where $t = 1, ..., T$. To provide predictability, these prices (next day) are known a priori by the users (prior pricing [7]) via a price-schedule $\{p_t\}$, where p_t is the price of bandwidth during the t ToD period. Given the price schedule, the interval durations, and the required bandwidth, the user can predict the cost of a session. The bandwidth of a connection is sold on a first come first serve basis; no reservations are allowed. If the amount is not available at the beginning of the session, the user is blocked. However, users who can not afford p_t are **not** considered blocked.

The service provider is responsible for establishing connections and allocating portions of the connection to individual users. Within this model, acquiring resources for a connection is provisioning, while selling portions of the connection will be referred to as allocation. Primary goals of the service provider will be profit maximization and minimizing the blocking probability.

3 Optimal Resource Provisioning and Allocation

A service provider must manage several connections simultaneously, where each connection could represent a unique pair of ingress and egress routers, a QoS

---

[3] Analogous to an Internet DiffServ connection across a network domain [15].

3

class, or both. As described in [8], the service provider is interested in maximizing the profit of each connection, which is achieved when the difference between the revenue generated minus the cost is maximized. This is given in the following formula,

$$\max \left\{ \sum_{t=1}^{T} [r(x_t) - c(s)] \right\} \tag{1}$$

The revenue generated by the connection during ToD period t is $r(x_t)$ and is based on x_t which is the aggregate user demand. Note the profit maximization is over T consecutive ToD periods (which could represent a DiffServ SLA term). Viewing this as an optimization problem, the first order conditions are

$$\sum_{t=1}^{T} \frac{\partial r(x_t)}{\partial x_t} = T \cdot \frac{\partial c(s)}{\partial s} \tag{2}$$

Note the connection supply, s, is constant for each ToD period. The left-hand side of equation 2 is also referred to as the marginal revenue, which is the additional revenue obtained if the service provider is able to sell one more unit of bandwidth. The right-hand side of equation 2 is referred to as the marginal cost, which is the additional cost incurred. A solution to the optimization problem exists, if the cost and revenue functions are continuous and convex. Therefore, to determine the appropriate provisioning amounts and prices, these functions must identified.

The Cobb-Douglas demand function will be used to model aggregate user demand (multiple users using the same connection). The Cobb-Douglas demand function is commonly used in economics because it is continuous, convex, and has a constant elasticity. A constant elasticity assumes users respond to proportional instead of absolute changes in price, which is more realistic. Therefore, this demand function is popular for empirical work. For example, the Cobb-Douglas demand function has been used for describing Internet demand in the INDEX Project [17]. The Cobb-Douglas function has the following form,

$$x_t(p_t) = \beta_t \cdot p_t^{-\alpha_t} \tag{3}$$

where p_t is the bandwidth price during ToD t and the approximate aggregate wealth of users is denoted by β_t. The elasticity (own-price) during ToD t is α_t, which represents the percent change in demand in response to a percent change in the price. Due to the dynamic nature users, demand is expected to change over time. These changes may reflect variable application demands, ToD trends, pricing, or the introduction of new technology. For this reason, demand prediction and estimation must be employed [4, 12], where the demand curve parameters (α_t and β_t) are estimated using previous ToD measurements. During a ToD the price is set and the demands are recorded (demand during a ToD will vary based on the number of users and their applications). This results in a set of demands for each ToD price. For this paper, the highest demand observed for a ToD price was used for parameter estimation, which resulted in a conservative

demand curve. If the estimation is correct, demand should equal supply, which results in a zero blocking probability.

Given the estimated aggregate demand function, the revenue earned is the price multiplied by the demand,

$$p_t \cdot x_t(p_t) = \left(\frac{\beta_t}{x_t(p_t)}\right)^{\frac{1}{\alpha_t}} \cdot x_t(p_t) = \beta_t^{\frac{1}{\alpha_t}} \cdot [x_t(p_t)]^{1-\frac{1}{\alpha_t}} \qquad (4)$$

Taking the derivative of equation 4 with respect to demand yields the marginal revenue for ToD period t. Similarly, taking the derivative of the cost function yields the marginal cost. Substituting these values into equation 2 results in a system of equations that can be solved for s which is the appropriate amount to provision. Since the marginal equations (revenue and possibly costs) are non-linear, a direct solution can not be found [8]. For this reason, gradient methods (e.g. Newton-Raphson) can be used to determine the optimal provisioning amounts [18]. Due to the time typically associated with negotiating (or establishing) a connection (SLA) [8], calculations can be performed off-line; therefore, convergence time is not critical.

Although the technique presented in this section does determine the optimal prices and connection supply, it does not (like many others) address the appropriate price-interval length. Since price-interval length can impact profit, its economic effect must be considered when managing network resources.

4 Pricing Intervals

As described in the introduction and depicted in figure 2, the duration of a price interval can range from extremely large to very small. Smaller intervals provide greater congestion control since prices can quickly adjust based on user demand [10, 7]. This flexibility also increases profits, since prices can be set to encourage usage. Consider the allocated bandwidth for a single price interval given in figure 2(a). Near the end of this graph, the profit decreases due to a reduction in demand. Using shorter intervals, prices are lowered during this time as seen in figure 2(c). This price reduction encourages more usage and increases profits. Even using seven intervals during a day, seen in figure 2(b), results in higher profits than a single price. Of course, once users have purchased their desired maximum bandwidth, further price reductions will not increase demand. Ideally, a service provider wants to update the price whenever the aggregate user demand changes. Changing the price at a rate faster than the change in demand would not increase profits. Although smaller price intervals are advantageous to the service provider, users prefer a simpler price structure. Generally, users are adverse to the possibility of multiple price changes during a session, even if it could result in a lower cost [7]. As a result, a service provider could lose customers, which would result in lower profits, if rival service providers offer a simpler pricing structure (fewer intervals).

Determining the appropriate number of intervals in a day can be viewed as an optimization problem. Let $g(T)$ represent the profit (revenue minus the
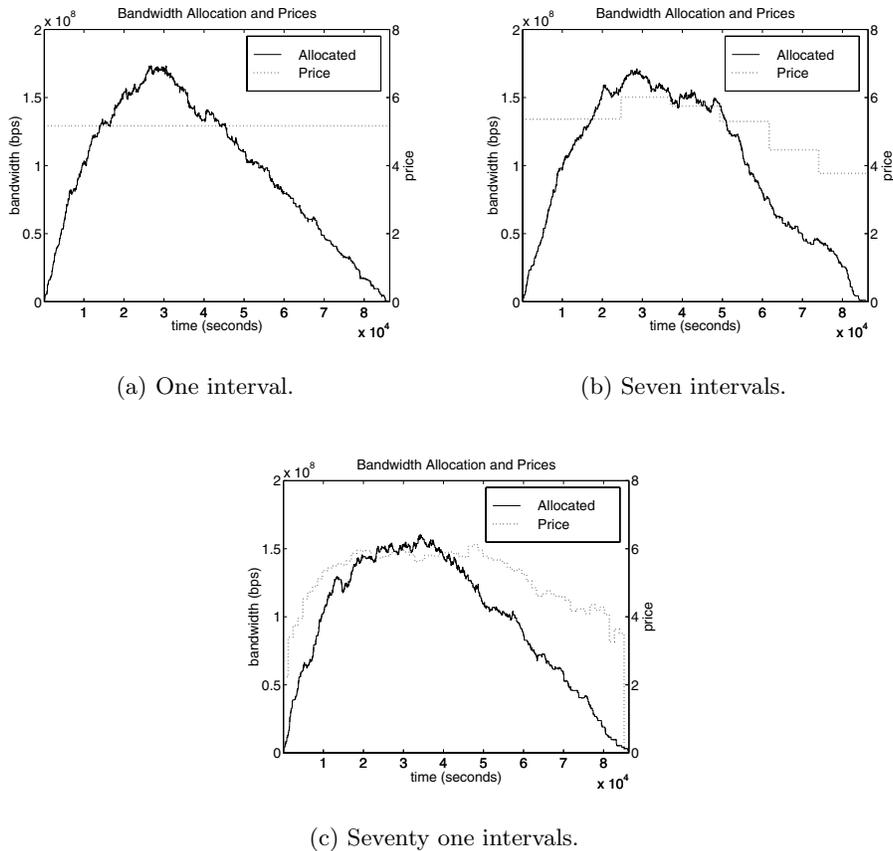
(a) One interval.

(b) Seven intervals.

(c) Seventy one intervals.

Fig. 2. Effect of interval duration on aggregate user demand.

actual connection cost) obtained using T intervals, which can be determined using equation 1. In addition, let $l(T)$ represent the loss in revenue due to the number of intervals in a day[4]. We will assume $l(T)$ increases as the number of intervals increases, which represents a higher inconvenience or complexity value users associate with more pricing intervals. Furthermore, let $l(1) = 0$ indicating users prefer a flat price. Under these circumstances the service provider seeks to determine T such that,

$$\max_{T \geq 1} \{g(T) - l(T)\} \tag{5}$$

A solution for T can be found if these functions are continuous and convex. The value of $g(T)$ can be determined using equation 1 and demand estimation

---

[4] Again, equation 1 only considers the cost of the connection in terms of resources required to support it.

techniques. Determining $l(T)$ is difficult, since it relies on user feedback on pricing intervals preferences. Although we can assume the function increases as T increases (users prefer flat prices), realistic data is required to determine the actual properties of this function. We can gain insight on the economic effect using simulation by measuring the profit associated with different price-interval lengths. If there is relatively little monetary gain using smaller intervals, the service provider should implement a simpler user-friendly price structure.

4.1   Experimental Results

In this section, the effect of price interval duration on service provider profit is investigated using simulation. For each simulation, a random number of users, uniformly distributed between 200 and 500, interacted with a single service provider during a day. Users started their sessions at random times using a Poisson distribution with mean equal to 9:00am. The duration of a session was then uniformly distributed between 0 and 12 hours. This resulted in a high aggregate demand during the midday, simulating peak hours [13]. Users had an elasticity α uniformly distributed between 1.75 and 2.5 (consistent with the IN-DEX project [17]), and a wealth β uniformly distributed between 1×10^8 and 2×10^8. The minimum demand of each user was uniformly distributed between 0.5 Mbps and 2 Mbps to represent a variety of traffic. Furthermore, users were considered independent and did not share connections. The service provider used the optimization techniques described in the previous section (including demand estimation) to provision and price bandwidth.

For each experiment, we are interested in measuring the percent change in profit as the number of intervals increases during a day[5]. A single experiment consisted of six simulations, each simulating a different number of pricing intervals. User arrivals and demands were randomly generated at the beginning of each experiment and were used for the six simulations. For each experiment, the first simulation was performed to estimate demand curve parameters and to determine the smallest aggregate demand interval. An aggregate demand interval is the amount of time between successive demand changes. Changing prices more frequently than the shortest aggregate demand interval should not increase profits, as described in the beginning of this section. Let T_* represent the number of pricing intervals based on the shortest aggregate demand interval. Five additional simulations were then performed, where the number of intervals were equal to 1, $T_*/100$, $T_*/10$, T_*, and $2 \cdot T_*$. Note T_* was greater than 100 for every experiment. Finally, 5000 independent experiments were conducted (30000 simulations total) to provide averages and 95% confidence intervals.

The experimental results are shown in figure 3, which depicts the percent change in profit as the number of intervals increases. Note, the percent change in profit is compared to the profit obtained from a single interval. As expected, profit increased as the number of intervals increased. Initially, the profit increased rapidly. At $T_*/100$ the change was 0.081, while the profit increase was 0.116 at

---

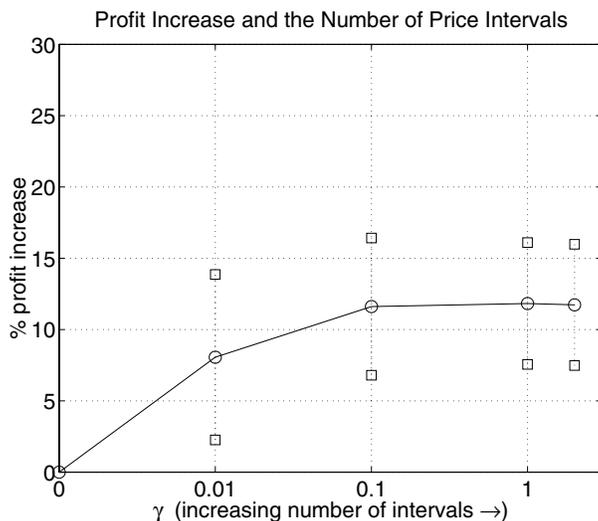[5] For a given day, we assume all intervals are equal in duration.

Fig. 3. Simulation results showing the profit change as the number of pricing intervals in a day increases. T_* is the number of pricing intervals based on the shortest aggregate demand interval. Percent change in profit is based on the profit of a single interval. For each data point, the 95% confidence intervals are shown.

$T_*/10$ (a difference of 0.035) . However, the profit increase for T_* intervals was 0.118, which is negligible as compared to the profit obtained from $T_*/10$ intervals. Using more than T_* intervals did not increase profits as compared to T_* intervals, which was expected. Furthermore, the standard deviation of the average-profit-change reduced as the number of intervals increased, from 0.058 at $T_*/100$ to 0.043 at $2 \cdot T_*$. No users were blocked in any of the simulations performed. In summary, a service provider can increase profits by using more intervals; however, there is limited monetary benefit associated with using intervals lengths that approach the smallest change in aggregate user demand.

5   Conclusions

Pricing has been demonstrated as an effective method for allocating network resources as well as revenue generation. Pricing resources also prevents users from over allocating resources or requesting service classes more stringent than necessary. The objective is to price resources so demand equals supply, at this point allocations are provably fair and profit is maximized. However, given the dynamic nature of computer networks (changing user demands), determining when prices should change (time-scale) is problematic. Prices could remain fixed over long periods of time providing predictable network usage costs, but limited congestion control. In contrast, prices could continually change based on demand yielding effective congestion control, but no cost predictability. Interval pricing

is a compromise between these two extremes, where prices remain fixed over a certain time span. This method can provide predictability as well as congestion control; however, determining the appropriate interval length is difficult. The impact of interval length on congestion control has been addressed, indicating shorter intervals provide better control. This paper investigated the economic impact of price interval duration. Shorter intervals can provide higher profits; however, experimental results indicated these gains are modest. For example, experiments conducted indicate a profit increase no larger than 5% was achieved when smaller intervals were used as opposed to larger intervals (for example 100 times longer). Therefore, given users preferences toward fewer price changes, smaller price intervals hold few economic benefits.

Future areas of research include determining user price-interval preferences and investigating variable length price-intervals. Users are adverse to multiple price changes during a session; however their preferences must be quantified to solve the optimization problem introduced in this paper. Furthermore, most interval pricing techniques require all intervals to be equal length. Using different sized intervals during the day may increase profits; yet, still provide a user friendly price structure. Both items are integral in the development of interval pricing.

References

1. Anerousis, N., Lazar, A.A.: A framework for pricing virtual circuit and virtual path services in atm networks. ITC-15 (1997) 791 – 802
2. Courcoubetis, C., Siris, V.A., Stamoulis, G.D.: Integration of pricing and flow control for available bit rate services in atm networks. In: Proceedings of the IEEE GLOBECOM. (1996) 644 – 648
3. Ferguson, D.F., Nikolaou, C., Sairamesh, J., Yemini, Y.: Economic models for allocating resources in computer systems. In Clearwater, S., ed.: Market Based Control of Distributed Systems. World Scientific Press (1996)
4. Kelly, F., Maulloo, A.K., Tan, D.K.H.: Rate control for communication networks: Shadow prices, proportional fairness and stability. Journal of the Operational Research Society **49** (1998) 237 – 252
5. Paschalidis, I.C., Tsitsiklis, J.N.: Congestion-Dependent pricing of network services. IEEE/ACM Transactions on Networking **8** (2000) 171–184
6. Wang, X., Schulzrinne, H.: Pricing network resources for adaptive applications in a differentiated services network. In: Proceedings of the IEEE INFOCOM. (2001)
7. Yuksel, M., Kalyanaraman, S.: Effect of pricing intervals on the congestion-sensitivity of network service prices. In: Proceedings of the IEEE INFOCOM. (2001)
8. Fulp, E.W., Reeves, D.S.: Optimal provisioning and pricing of internet differentiated services in hierarchical markets. In: Proceedings of the IEEE International Conference on Networking. (2001)
9. Reichl, P., Flury, P., Gerke, J., Stiller, B.: How to overcome the feasibility problem for tariffing internet services: The cumulus pricing scheme. In: Proceedings of the ICC. (2001)
10. Shih, J.S., Katz, R.H., Joseph, A.D.: Pricing experiments for a computer-telephony-service usage allocation. In: Proceedings of the IEEE Globecom. (2001)

11. Fulp, E.W., Ott, M., Reininger, D., Reeves, D.S.: Paying for qos: An optimal distributed algorithm for pricing network resources. In: Proceedings of the IEEE Sixth International Workshop on Quality of Service. (1998) 75 – 84
12. Morris, R., Lin, D.: Variance of aggregated web traffic. In: Proceedings of the IEEE INFOCOM. (2000)
13. Odlyzko, A.: The economics of the internet: Utility, utilization, pricing, and quality of service. Technical Report 99-08, DIMACS (1999)
14. Nichols, K., Jacobson, V., Zhang, L.: A two-bit differentiated services architecture for the internet. Internet Draft `http://ds.internic.net/internet-drafts/draft-nichols-diff-svc-arch-00.txt` (1997)
15. Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., Weiss, W.: An architecture for differentiated services. IETF RFC 2475 (1998)
16. Braden, R., Zhang, L., Berson, S., Herzog, S., Jamin, S.: Resource reservation protocol (rsvp) - version 1 functional specifications. IETF RFC 2205 (1997)
17. Varian, H.R.: Estimating the demand for bandwidth. Available through `http://www.INDEX.Berkeley.EDU/public/index.phtml` (1999)
18. Yakowitz, S., Szidarovszky, F.: An Introduction to Numerical Computations. second edn. Macmillan (1989)