Exploring Traffic Pricing for the Virtual Private Network[1]

Zhangxi Lin
Center for Research in Electronic Commerce
The University of Texas at Austin
zlin@uts.cc.utexas.edu
&
College of Business Administration
Texas Tech University
zlin@ba.ttu.edu

Dale O. Stahl
Department of Economics
The University of Texas at Austin
stahl@eco.utexas.edu

Peng Si Ow, Andrew B. Whinston
Center for Research in Electronic Commerce
The University of Texas at Austin
phighnam@aol.com, abw@uts.cc.utexas.edu

Abstract: This paper presents a virtual private network traffic pricing model with first-in-first-out and round-robin bandwidth scheduling. A transaction-level pricing architecture based on proxy server technology is proposed for the implementation. The experiment using real-time test data shows that the pricing mechanism can effectively improve a VPN's transmission efficiency measured in the service welfare rate.

1. Introduction

Evolved from the private network that originally uses phone lines with dedicated connections, the virtual private network (VPN) is a value-added network built upon all types of network clouds, such as ATM and frame relay, with secured virtual paths for data transmissions. The explosive growth of the Internet is also spilling over to VPNs, so that VPNs are emerging as a dominant computing network for corporations. There are numerous VPN products on the market now, such as e-Network by IBM, MultiVPN by Ascend, VTCP/Secure by InfoExpress, SmartGate by V-ONE, VPN-1 by Check Point, etc. Although the Internet-based VPN has brought users cost-effective solutions to support enterprise resource planning projects, Internet traffic is becoming more congested as a result of exponentially-increasing traffic loads, which substantially diminish the net benefits of users and service providers. Since 1991, we have researched the use of network resource pricing which was initialized by Stahl and Whinston [SW94] for managing traffic on the general computer network. We explored the use of an incentive-compatible resource price structure that can provide better quality of service for higher valued data flows by reducing those lower valued data flows, and as a result, improve network service welfare [GJPGW97]. We derived a general-equilibrium economic model, denoted as GSW priority pricing model, for network resource allocation [GSW97, LGJSW99]. We tested the model under various scenarios by a series of experiments. A simulation of a public network such as the Internet was developed for studying the model feasibility and to investigate equilibria.

Encouraged by these initial results, we have moved on to focus on developing a prototype VPN traffic pricing system architecture and technology, and addressing issues related to the implementation of a practical network traffic pricing system targeted at industrial application demands [LOSW99]. This paper presents these latest efforts. First we introduce a traffic pricing model for the VPN. It is a customized GSW model for the VPN with round-robin scheduling extensions. Then we describe a pilot prototype system for VPN traffic pricing. This system is different from the software-based simulation system for the GSW model in that it is built on a configurable network composed of several computers and it uses real-time data flows for experiments. It is able to carry out various kinds of experiments by

[1] We are grateful to Intel Corporation for its donation of the network computing equipment to the Center for Research in Electronic Commerce, with which all the experiments referred in this paper are conducted.

changing its testbed infrastructure and experimental parameters. Some of the results will be shown in this paper.

2. An M/G/1 Traffic Pricing Model for the VPN

We propose a *transaction-level pricing* architecture [LZ99] for the VPN traffic pricing system to cope with implementation issues which include digital contracting, VPN traffic pricing efficiency, logistic systems such as payment and accounting, integration of traffic pricing system and existing traffic control techniques, and user acceptability. By definition, a *transaction* is a series of services from one or more servers in response to a user request arriving within the local area network behind a VPN gateway. The cost to the user is the fee for the bandwidth usage to all data flows incurred by this request plus the devaluation caused by a response delay. Users can initiate *digital contracts* with a bandwidth broker without looking into details of data flow structure. In this context, the digital contract is an electronic agreement between the user and the bandwidth broker for services that are charged when requested. Once a user submits a request for service to an application server, it is automatically acknowledged that the user has electronically accepted the billing terms. Pricing and billing are exercised on per-request basis. With transaction level implementation, the VPN gateway schedules routing tasks in regard to the application needs and priorities. This process allows the pricing mechanism to be incorporated within enterprise resource planning applications.

A proxy server based VPN traffic pricing (Figure 1) is the underpinning infrastructure for transaction-level VPN pricing architecture. A proxy server, employed as the bandwidth broker to schedule the data flows with a pricing mechanism for an affiliated VPN gateway, is denoted as a *traffic proxy server* (TPS). A TPS is an extension of the VPN gateway responsible for traffic control functions. Our proposed VPN traffic pricing system consists of a minimum of five conceptual components in addition to regular VPN functional modules. These components include: 1) a set of utilities facilitating job submission at the client end; 2) user registration and authentication; 3) user account management; 4) job scheduling and queue management; and 5) queue monitoring and service pricing. A TPS operates all the components except the set of job submission utilities at the client end component.

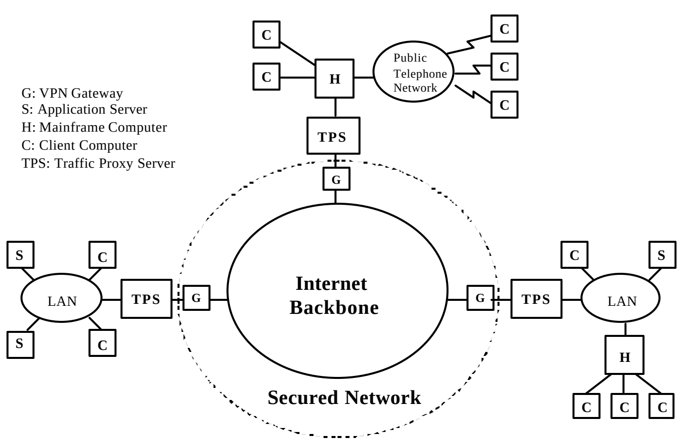


Figure 1: Infrastructure of proxy server based VPN traffic pricing system

We assume that the job request arrival at a VPN gateway is a Poisson process with a general size distribution. Consequently, the process of TSP bandwidth allocation can be modeled as an M/G/1 queueing system. Applying different bandwidth scheduling approaches to this queueing system, such as first-in-first-out (FIFO) and round-robin (RR), results in different pricing formulas.

We denote I as the set of users, J as the set of job types, K as the set of priority classes, and Q as the set of job sizes. Following GSW model's approach [GSW97], we define the expected service welfare from a type-j job submitted by user i as $V_{ij} - \boldsymbol{d}_{ij}\boldsymbol{t}_{jk}$, where V_{ij} is the job's gross value, $\boldsymbol{d}_{ij}$ is the delay cost coefficient and t_{jk} is the expected throughput time for the job entering priority-k queue. We assume both V_{ij} and $\boldsymbol{d}_{ij}$ are random variables with a joint density function $g(V_{ij},\boldsymbol{d}_{ij})$. User i submits a type-j job to priority class k if a job's net value $V_{ij} - \boldsymbol{d}_{ij}\boldsymbol{t}_{jk} - r_{jk}$ is greater than zero, where r_{jk} is the price for type-j job at priority k, and k is chosen to maximize this net value among all feasible priority choices. Denote the traffic rate of type-j jobs submitted by user i to priority class k by x_{ijk}. It is a function of exogenous traffic rate, users' values, and the time and monetary costs of the job. Define $\boldsymbol{j}_{jk} = \sum_{i\in I} x_{ijk}$ as the total traffic rate of type j jobs submitted to the priority class k. The expected throughput time for a type-j job submitted to the priority-k service can be defined as a function of the distribution of arrivals by job type and priority ($\boldsymbol{j}$ $= \{\boldsymbol{j}_{jk} \mid j\ \hat{I}\ J, k\ \hat{I}\ K\}$), and the bandwidth B: $t_{jk} = W_{jk}(\boldsymbol{j}\,; B)$. Then we can derive the optimal price for job j at priority k as:

$$r^*_{jk} = \sum_{i\in I} \sum_{m\in J} \sum_{h\in K} x_{imh} d_{im} \frac{\partial \Omega_{mh}}{\partial \boldsymbol{j}_{jk}}, \qquad (*)$$

In the above pricing formula, the $x_{imh}\boldsymbol{d}_{im}$ term is the flow cost of delay for each user of the VPN gateway, and the ¶W_{mh}/¶$\boldsymbol{j}_{jk}$ term is the marginal time delay caused by an extra type j job at priority k. Thus, the optimal resource allocation is achieved when a type j job at priority k pays for the marginal increase of the aggregate flow cost of delay for all users. The price will be high when the traffic is heavy and/or the cost of time is high.

With first-in-first-out (FIFO) scheduling, the expected throughput time depends on the job type only through the job size. Therefore, the queue waiting time (w_k) and the pricing formula can be expressed in terms of job size (q) instead of job type j. By using the M/G/1 queue waiting time formula, we can obtain a pricing formula for FIFO scheduling that is quadratic in job size:

$$r^*_{qk} = \sum_{h\in K} \boldsymbol{j}_h \bar{\boldsymbol{d}}^{(h)} (a_{1h} q + a_{2h} q^2) \quad "q\ \hat{I}\ Q, k\ \hat{I}\ K$$

where $a_{1k} = \dfrac{w_k(2 - \sum_{l<k} \boldsymbol{r}_l - \sum_{l\le k} \boldsymbol{r}_l)}{B(1 - \sum_{l<k} \boldsymbol{r}_l)(1 - \sum_{l\le k} \boldsymbol{r}_l)}$ when $h < k$, $a_{1k} = \dfrac{w_k}{B(1 - \sum_{l\le k} \boldsymbol{r}_l)}$ when $h = k$,

$$a_{2k} = \frac{1}{2B^2(1 - \sum_{l<k} \boldsymbol{r}_l)(1 - \sum_{l\le k} \boldsymbol{r}_l)} = w_k \sum_{q\in Q} \boldsymbol{j}_{qk} q^2 ,$$

and $\bar{\boldsymbol{d}}^{(k)} = \sum_{l\in I} \sum_{m\in J} \dfrac{x_{lmk}}{\boldsymbol{j}_k} \boldsymbol{d}_{lm}$ is the mean of d_{ij} over i and j, weighted by the flow in priority k.

We now consider the pricing formula for a non-prioritized RR scheduling scheme. RR scheduling has different features from FIFO scheduling. In a RR scheduling system job's throughput time—the time that a job stays in a VPN gateway—is proportional to job size and the average number of jobs in the queue during servicing. The throughput time of a size q_j job in a non-prioritized queue is $t(q_j) = [(L^*+1)$ q_j - $\boldsymbol{r}$ /2]/B, where $L^* = \dfrac{\boldsymbol{r}^2 E[q^2]}{(1-\boldsymbol{r})(E[q]+E[q^2])}$ is the average number of jobs being serviced, $\boldsymbol{r}$ is the bandwidth utilization ratio of the gateway, and $E[q]$ and $E[q^2]$ are the expected size and size-squared values for the gateway. The type-j job submission criterion for user i is $q_j(v_{ij} - \delta_{ij}\tau) \geq r_j$, or $v_{ij} - \delta_{ij}\tau \geq r_0$, where $v_{ij} = V_{ij}/q_j$ is job's unit value, $\tau = t_j/q_j$ is the unit throughput time, and $r_0 = r_j/q_j$ is the unit price. We can derive the optimal unit price:

$$r_0 \approx \frac{\bar{\delta}}{B}(L^* + \frac{L^* - \rho}{1 - \rho}),$$

where $\bar{\delta} = \sum_{l \in I} \sum_{m \in J} x_{lm} \delta_{lm} q_m / \sum_{l \in I} \sum_{m \in J} x_{lm} q_m$ is the mean of δ_{ij} over i and j, weighted by data volume rates.

The approximate form for the pricing with RR scheduling indicates that the expected number of jobs in the queue is a critical factor in a job's price and the price is proportional to the size of a job.

The accurate expression for r_0 includes two more terms. One of them is in a quadratic form and hence is inconvenient in use. We ran a set of experiments to compare the outcomes from the approximate formula with the exact expression. The results show that the difference in price level is within 0.1%. This level of error has in fact no effect on the accuracy of the outcomes because another group of experiments show that the prototype VPN pricing system can tolerate a price bias up to 5% without a significant impact on the welfare rate.

3. Experimenting VPN traffic pricing

We have set up an environment that allows us the flexibility to configure different VPN infrastructures. Current experimental study focuses on a simple VPN system comprising of a single TSP unit shown in Figure 2. This VPN has two subsystems - the VPN traffic pricing kernel and the system control interface. The kernel subsystem is composed of a client computer, a VPN gateway, a traffic proxy server, and an application server. Controlling and monitoring this VPN are accomplished via a control panel and a real-time graphic interface that displays performance statistics. The system allows testing different network resource scheduling techniques, such as FIFO and RR, and pricing strategies on a real-time basis.

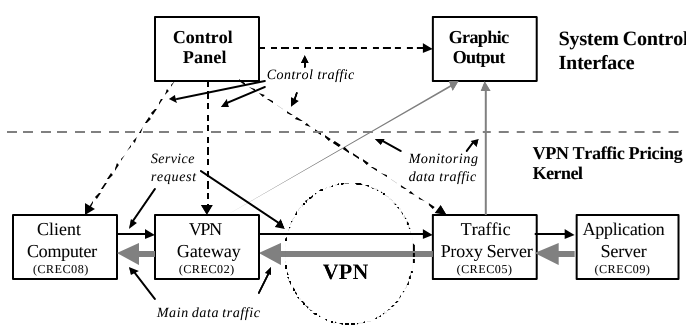
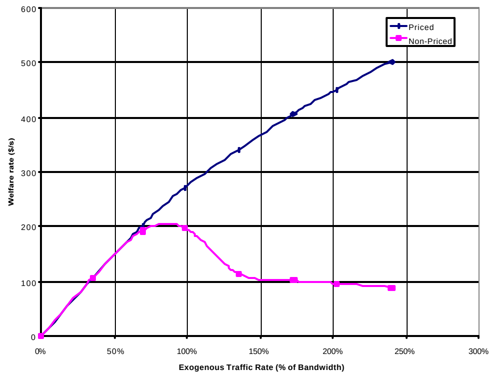


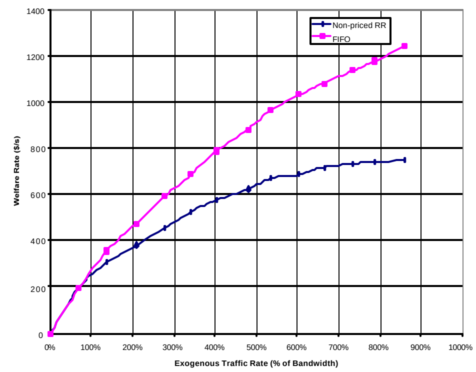
Figure 2: An autonomous traffic pricing unit for experimentation

In order to test if the performance from the hardware-based simulation is the same as that from the GSW [GSW97] software-based simulation, we used the same parameter values used in the GSW simulation. These include the distributions of the job size, the job value and the delay cost coefficient. Experiments turned out the following results:

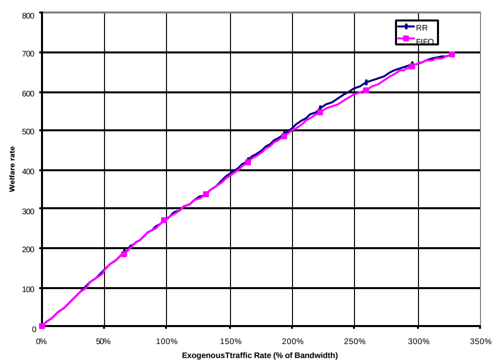
1) The outcome from the simulation based on the above distributed system using the FIFO scheduling matches that from the GSW model simulation very well. VPN traffic pricing significantly improves network welfare rate in FIFO scheduling (Figure 3a). The curve of the welfare rate from the non-priced FIFO scheme starts to decline when exogenous traffic rate increases and approaches the capacity, while that from the priced FIFO scheme keeps going up as traffic rate increases. This is exactly the same result as the GSW model simulation revealed.



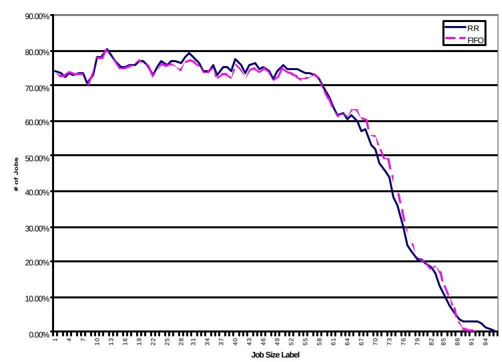
(a) Priced FIFO versus non-priced FIFO



(b) Non-priced round-robin (denoted as RR in the figure) versus priced FIFO



(c) Welfare rate comparison between priced round-robin (denoted as RR) scheduling and priced FIFO scheduling



(d) Job arrival distribution comparison between priced round-robin (denoted as RR) and priced FIFO schemes at an exogenous traffic rate of 4 Mbps.

Figure 3: Effectiveness of VPN traffic pricing

2) Priced FIFO scheduling performs better than non-priced RR scheduling does (Figure 3b). Welfare rate yielded from a non-priced RR scheme increases with the augmented traffic rate. This is because the expected throughput time for a job in RR scheduling is proportional to the job's size. This provides the incentive for users not to submit the jobs with lower unit values. Since the submission decision is based on the comparison between the unit job value and the unit waiting cost, the efficiency of the network is better than that of non-priced FIFO. However, the increase rate of the welfare rate from non-priced RR scheduling scheme is much lower that that from a priced FIFO scheme. The gap widens when traffic rate goes up.

3) Pricing is also effective in the RR case with the same welfare rate as that from priced FIFO routing (Figure 3c).

4) Pricing on FIFO and RR scheduling schemes results in almost the same job arrival rate distributions (Figure 3d).

4. Future research plan

We are currently upgrading our experimental system to incorporate features such as a user-friendly operational interface, job submission utilities, etc. Our intention is to explore the effects of such a pricing system on the users with budgets and vice versa. This will be in conjunction with other on-going instructional projects in which students utilize a VPN to set up and run electronic businesses. By running this project on our prototype VPN with and without TPS pricing, we will be able to assess not only the network performance under real loads but also the user satisfaction under the various experimental treatments. In another experimentation plan, we will provide students with fundable resources that can be used for accessing our VPN or cashed-in should the student opt to use an alternative non-priced VPN. The objective would be to ascertain if the users perceive a quality of service on our priced VPN sufficient to warrant paying the fee charged. We will also interview users to ascertain their satisfaction and solicit suggestions for improvements in all aspects of the user interface. This experimental work will guarantee a versatile prototype VPN traffic pricing system that well meets the requirements of practical industrial applications.

References:

[GJPSW97] Gupta, Alok, Boris Jukic, Manoj Parameswaran, Dale O. Stahl, and Andrew B. Whinston, "Streaming the Digital Economy: How to Avert a Tragedy of the Commons", *IEEE Internet Computing*, vol. 1, no. 6, Nov/Dec 1997, pp38-46.

[GSW97] Gupta, A., Dale O. Stahl, and Andrew B. Whinston., "A Stochastic Equilibrium Model of Internet Pricing", *Journal of Economic Dynamics and Control*, 21, 1997, pp697-722.

[LGJSW99] Li, Mingzhi, Alok Gupta, Boris Jukic, Dale O. Stahl and Andrew B. Whinston, "Estimating Internet Users' Demand Characteristics", working paper, 1999, http://crec.bus.utexas.edu/works/articles/demand.pdf.

[LOSW99] Lin, Zhangxi, Peng Si Ow, Dale O. Stahl and Andrew B. Whinston, "Improving the Performance of Virtual Private Network by Pricing Traffic", working paper, March 1999.

[LZ99] Zhangxi Lin, "Exploring Traffic Pricing for the Virtual Private Network", Ph.D. dissertation, to be completed in November 1999.

[SW94] Stahl, Dale O. and Andrew B. Whinston, "A General Economic Equilibrium Model of Distributed Computing", in W. W. Cooper and A. B. Whinston (eds.), *New Decisions in Computational Economics*, pp175-189, 1994.