2000

# Market-Based QoS Interconnection Economy in the Next-Generation Internet

Junseok Hwang
*Syracuse University*

Martin B.H. Weiss
*University of Pittsburgh*

# Market-Based QoS Interconnection Economy in the Next-Generation Internet

Junseok Hwang [a,1] Martin B. H. Weiss [b]

[a] *School of Information Studies, Syracuse University, Syracuse, NY 13244, USA*

[b] *Telecommunications Program, University of Pittsburgh, Pittsburgh, PA 15260, USA*

**Abstract**

Technical and regulatory changes motivate the economics of the interconnections to be re-examined for the next-generation Internet. This paper presents a framework for the economics of market-based QoS interconnection networks. The economic model presented captures the characteristics of opportunity costs and demands and suggests strategies for QoS pricing and allocation. We use the statistical results of network simulation to present numerical examples of those market opportunity costs with respect to the contracted QoS network services. Using the model, we show the network economics of market-based QoS interconnections.

*Key words:* Network Economics, Internet QoS, Interconnection Pricing, Next-Generation Internet.

## 1 Introduction

The convergence of various telecommunication networks relies critically on the interconnection of various other networks, especially the Internet. Currently, unlike telephony, interconnections of the Internet are not regulated, but managed by market mechanisms [1]. Recent regulatory research agency studies (FCC (USA), IIAA (Internet Industry Association of Austria), and Telus (Canada)) suggest market-based solutions that are concerned with the regulation of future Internet backbone interconnections which are different from the historical interconnection policy. The interconnections of the next-generation Internet are hindered by various new problems such as interconnection settlements and pricing that are different from the issues of conventional telephone interconnections and the best-effort Internet interconnections.

[1] Corresponding author. Tel.: +1-315-4434473; fax: +1-315-4435806.
*E-mail address:* jshwang@syr.edu (J. Hwang)

Changes in both the technical and the regulatory nature of the Internet evolution motivate a re-examination of the economics of network infrastructure interconnections encompassing those new issues arising from the deployment of next-generation Internet services.

The use of the Internet itself has changed in character and grown in both scale and scope. In 1999 there were more than forty national Internet backbone providers and about five thousand ISPs (Internet Service Providers) in the United States.[2] The size and capacity of the ISPs (Internet Service Providers) vary remarkably from access providers to backbone providers. The increased traffic of various network service applications raises several issues, including resource contention, efficient use of bandwidth, and the enhanced level of service quality or, at least, protection from service degradation. The importance of supporting various Quality of Service (QoS) in the Internet has recently intensified. The Internet is evolving into a next-generation network that supports various QoS applications in addition to best-effort services.

QoS in the Internet is the network service performance measure, thanks to the differential treatment of packets in the network that range from queueing and service discipline to service contracts and protocols. Levels of QoS are concerned with different QoS measures such as average delay, delay variation, and packet loss rates. Various techniques to support QoS in the Internet have recently been proposed. This is a fundamental departure from the traditional best-effort nature of the current Internet; however, there are still many issues to be addressed in order to implement the appropriate QoS mechanisms, which include IntServ (Integrated Services) and DiffServ (Differentiated Services). IntServ seeks to support connection-oriented[3] QoS service on top of IP networks for specific connections and flow-based behavior. DiffServ is the approach of keeping the philosophy of the current connectionless[4] Internet and of supporting QoS at the aggregate level of applications based on the network service classification. DiffServ tends, in practice, to be used more for core network interconnection services than for end-to-end services due to the scalability of the technology's service architecture.

The economics of QoS network interconnections for the future Internet will differ from those of telephony and the best-effort Internet with various new aspects such as connection settlements, QoS transport costs, QoS pricing strategy, and QoS allocation. QoS-Interconnection is about connectivity for multi-service network services. Telephony uses interconnections for a single-service application (voice) and associated enhanced services. On the best-effort Internet, interconnection refers to the rate of simple data aggregates connection

[2] Source: Boardwatch from ISPWORLD.

[3] Taking a service approach similar to the telephone network, where the connection is set up before its service.

[4] Network service is offered without individual connection setup.

among different networks. In addition, the basis for QoS interconnection pricing may need to be different from that used in the traditional approach. The interconnection price of telephony is mostly based on the proxy cost model measuring forward-looking incremental costs, and charging is mainly focused on which network terminates a call. Furthermore, different QoS services may have different cost functions to provide, and additional services can be both beneficial and costly to the networks. The cost functions of different QoS networks may make interconnection agreements difficult, and the capital cost-based pricing may not reflect the real value of packet-switched network interconnection. In addition, the value of network service may not be easily and accurately measured for multi-service networks in the QoS Internet, because there might be different opportunity costs among different service connections. As the Internet becomes the network that provides differentiated QoS services, the positive network externalities may not always drive the economics of interconnection for differentiated services without proper QoS pricing mechanisms among networks.

These different economic aspects of QoS interconnections of the next-generation Internet have raised the following general, but important, questions. How will different QoS network services affect the resource and quality allocations over the interconnections of the future Internet? What will be the consequences on the economics of interconnection and its settlements?

This article addresses the economic problems of the market-based network resource management for QoS interconnection networks. This study proposes that interconnecting QoS networks will manage both QoS pricing and QoS bandwidth allocations for optimal interconnection, and different QoS networks will have different optimal QoS allocation policies due to the different marginal costs (marginal bandwidth opportunity cost) associated with each QoS mechanism and policy. Different QoS networks have different cost characteristics for different levels of QoS service. This study investigates how the cost of quality of different QoS networks will characterize the optimal interconnection strategies of individual network service providers and motivate them to interconnect with each other. The economics of various hybrid interconnections depends not only on the bandwidth opportunity costs, but also on other service-related factors such as pricing, service classification, and demand characteristics. We deal with those relevant network economic factors in the scope of QoS interconnections.

We develop network economic models to characterize the different performance properties of various QoS services in the next-generation Internet. We then introduce an optimization problem to address the optimal interconnection behavior of the network service providers. This problem is motivated by the fact that network service providers and operators will act reasonably to maximize their payoffs in an economic sense for QoS Interconnection. To assess the values of the parameters of those economic models for various QoS networks, we use the method of simulation to dimension networks with various size and capacity. We use the cost data of best-effort network interconnections to estimate the cost functions of different networks. Using the cost functions of various QoS networks and selected QoS pricing strategies, we capture service strategies on the pricing and resource allocation of QoS networks.

## 2 Models

The economic models in this article encompass the models for cost, demand, pricing, and interconnection. Empirical network economic models are developed to characterize the different economic variables and properties of different QoS networks. One motivation for the development of the network economic model is to capture the economic behavior of QoS interconnection networks. The network economic model is structured with the following components:

- Demand Assessment of QoS Service Interconnection
- Market Opportunity Cost of QoS Service Interconnection
- QoS Interconnection Service Edge Pricing
- QoS Settlements and Interconnections

This section deals with the detailed tasks of each process of the BMP operation. We formulate this economic model for Differentiated Service (DiffServ) QoS networks. The DiffServ network economic model is implemented in the form of simulation (Section 3). We start with some related works on network economic modeling.

### *2.1 Related Works*

Economic modeling studies for resource allocations and market management have been studied by several researchers [2–4]. Mackie-Mason presented a general optimization model of the resource allocation as a network planning problem [2]. The study proposed a "smart market" mechanism for solving the problem. A smart market requires simultaneous efficient routing and bandwidth allocation for reservations made in advance. Kumanran et al. [3] presented a general economic equilibrium model of users and service providers in a multi-class, multi-QoS network. The maximization of the users' utility and the revenues of service providers are set as the objective functions of the model. The model investigated the analytic results for various market conditions and demand assumptions. Hwang et al. in [4] modeled the profit maximization model for interconnecting networks under various settlement conditions and showed simple numerical results of different conditions.

### *2.2 Demand Assessment*

Demand functions can be modeled by capturing utility characteristics of different network services. The demand function captures price utility for differentiated network services. Different types of utility functions can be modeled for different types of QoS mechanism networks. The network service demand for each service class can be represented as aggregate average data rates for the given price and quality distribution. As shown in the following, the demand function can be modeled by the parameters of the maximum aggregate data rate for the differentiated network service of an average network interconnection, and the highest price ($ per kbps per time) the interconnecting networks are willing to pay for the service. The following general-demand function is an aggregate demand data rate for the given capacity for the specific services:

$$R_i(t) = F[R_{i-max}, p_i(t), p_{i-max}(t), \alpha_i(t)] \tag{1}$$

where:
$p_i(t)$: price per unit of time and bandwidth for service $i$
$p_{i-max}$: maximum willingness to pay for the service $i$
$\alpha_i(t)$: factor for demand elasticity and convexity on price for service $i$
$R_{i-max}$: maximum aggregate data rate for service $i$ at zero price and best available service
$R_i(t)$: aggregate demand data rate for service $i$.

Various forms of demand function can be assumed, but any form of the demand model requires the above listed assessment information.

### *2.3 Market Opportunity Cost*

In general terms, the opportunity cost of the current use of a particular input is its worth in other particular uses. Opportunity costs can be represented in terms of both dollar values and other uses of resources. In a market-based economy of network interconnection, capacity and network quality are valued based not on the capital costs of production but on the market opportunity costs of market value and scarcity of the bandwidth.

The market opportunity costs of the network service can be assessed using information based on the estimation of network usage and the market value of the bandwidth. An estimate of network usage can be assessed through measurement at the network interconnection points, and the market value of the bandwidth might be assessed through price information of the bandwidth market. Since the economic interpretation of the opportunity cost can be clearly explained by a brief discussion of a dual production program in the form of linear programming [5], the principles of opportunity costs are seen in this way.

In the form of a production problem involving an $m$-resource and an $n$-product, linear programming can be represented as:

$$\max \quad \pi = \sum_{i=1}^{n} F_i R_i \tag{2}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} a_{ij} R_i \leq K_j, \qquad \forall j = 1, ..., m \tag{3}$$

$$R_i \geq 0, \qquad \forall i = 1, ..., n \tag{4}$$

Accordingly, the dual of the above problem can be written as:

$$\min \quad \pi^* = \sum_{j=1}^{m} K_j f_j \tag{5}$$

$$\text{s.t.} \quad \sum_{j=1}^{m} a_{ji} f_j \geq F_i, \qquad \forall i = 1, ..., n \tag{6}$$

$$f_j \geq 0, \qquad \forall j = 1, ..., m \tag{7}$$

The objective function of the primal form is the maximization of the total gross profit in dollars, and $F_i$ and $R_i$ refer to the marginal profit in dollars per unit of product and the allocation amount (also demand) for product $i$, respectively. The coefficient $a_{ji}$ denotes the amount of the $j$th resource used in producing a unit of the $i$th product. In view of the fact that $\overline{\pi}^* = \overline{\pi}$, the objective value of the dual program can be in dollar values. $K_j$ refers to the total available quantity of the $j$th resource in the fixed firm's capacity, and $f_j$ denotes the valuation of the resource $j$ in dollars per unit to the $j$th resource. Notice that $f_j$ is not a market price, since the resource in question is already in the firm's possession as part of its fixed production capacity. Rather, $f_j$ is a value to be imputed to the resource.

In the objective functioning of the dual program, recalling that the sum of $K_j$ is the total fixed capacity of available resources in the firm, the expression

$$\pi^* = \sum_{j=1}^{m} K_j f_j \tag{8}$$

denotes the total value to be imputed to the resources. The objective is to minimize the total imputed value (opportunity cost) of the resources in the production capacity.

This study examines the economic interpretation of the following constraint of the above dual program:

$$\sum_{j=1}^{m} a_{ji} f_j \geq F_i, \qquad \forall i = 1, ..., n. \tag{9}$$

Since the coefficient $a_{ji}$ denotes the amount of the $j$th resource used in producing a unit of the $i$th product, the left side of the constraint represents the total opportunity cost of producing a unit of the product $i$. Thus the total opportunity cost of a production should be imputed at a level no less than the gross profit from the product. The last non-negativity constraint $f_j \geq 0$ is economically sensible, because a positive value should be imputed to a resource unless the resource is not fully utilized, so that null opportunity cost is incurred in putting it to productive use. This means that a positive opportunity cost for a resource $j$ ($f_j > 0$) is always to be associated with the full utilization of the resource in the optimal solution.

These observations provide an economic interpretation of the opportunity cost such that, in the optimal allocation (dual property of $\overline{\pi}^* = \overline{\pi}$), the total gross profit must be allocated into the resources of the production capacity entirely through the opportunity cost. Under such conditions, $f_j$ is referred to as an opportunity cost for the $j$th resource.

In retaining the principles of the opportunity cost discussed above, the market opportunity of different QoS interconnection services can be assessed in the following ways. Opportunity cost in the network is the real sacrifice (internal to the ISP) of providing additional units of a commodity (service) when a network (firm) ought to choose between or allocate scarce goods such as bandwidth (as for different services or commodities) for fixed capacity networks. Therefore, the opportunity cost of resource allocation decisions in a multiservice network is the value of the best available service alternative. Consequently, understanding the opportunity costs for different types of services that an ISP provides is important in terms of optimal resource allocation economic behavior. By using market information, the opportunity costs of the QoS differentiated network interconnection services can be practically assessed by measuring usable bandwidth (idle bandwidth) for specific interconnection services in a fixed capacity network.

To assess the opportunity costs for different network services, the estimates of network usage (available bandwidth) and the value of the raw bandwidth are required. Network usage can be assessed through measurement at the edge of the interconnecting networks. The bandwidth value can be assessed through the dynamic information available from the bandwidth commodity market (for example, the dynamic price information of an online bandwidth commodity market such as www.ratexchange.com or www.band-x.com).

For the measured offered load and capacity of the network, the opportunity costs of the different QoS interconnection service classes are assessed as the unit of dollars per kbps per time in the following ways. For any fixed capacity interconnection for a specific bandwidth management period, the maximum usable bandwidth for a QoS interconnection service class will be bounded by at least the usable bandwidth of an overengineered network of the QoS requirements for that service if the network supports only that type of traffic load. Therefore, the available bandwidth for a high-priority interconnection will be approximately the same as that of the entire capacity of the interconnection if no other constraints are enforced. Consequently, in such a situation the opportunity cost of the highest priority will be the function of the total capacity interconnection market commodity value, divided by the QoS service's achievable average utilization in kbps. Opportunity costs of different quality achieved by alternate treatments of packets at different levels can be calculated in such a way when they are of the highest priority in the service interconnection.

The opportunity costs of other priorities will also be dependent on the available capacity of the interconnection. The available capacity for a certain priority can be stated as total capacity times $(1 - \mathit{utilization})$ of the higher priority service classes. Therefore, in a similar way, the opportunity cost of this second priority service can be a function of the available interconnection capacity price (total interconnection price * available utilization for the second priority QoS service) divided by the total achievable average utilization in kbps. This algorithm of opportunity assessment means that the opportunity cost for lowest or best-effort PHB opportunity costs will approach zero when the interconnection capacity is almost fully utilized by the higher PHB classes. This means that resources for best-effort traffic cannot be entirely utilized so that a zero opportunity cost is incurred in putting it to productive use. This property is consistent with the principle of opportunity costs presented previously with the dual model.

### *2.4 Edge Interconnection Pricing*

Interconnection pricing is the decision on how much the QoS networks would charge the interconnecting networks for QoS services. Various approaches to price control of multiservice network resources include priority (service performance differentiation) pricing [6,7], congestion spot-pricing [8,9], optimal control pricing [10,11], and measurement-based edge-pricing [12–15]. Edge pricing, especially, is motivated by a series of approximations to true congestion costs or ideal marginal costs. Approximations are used because it is unrealistically difficult, or impossible, to exact each user's usage based on one user's behavior. Therefore, edge-pricing seeks a reasonable proxy for those costs with a feasible approximation method, and using the approximately captured costs schedules the price at the edge of the networks. We used edge-pricing in our model by approximating market opportunity cost through measuring at the interconnection. Other motivations for this can be found in [12,15].

The pricing of different class network services plays a key role in determining resource allocation among different network service classes. The price schedule acts as a control mechanism for resource allocation (offered service selection and provisioning) among different revenue-generation strategies (different service classes and bandwidth sales) within a fixed-capacity, constrained network. It is particularly important to the allocation of different classes of services for the interconnection of the given market prices of bandwidth and availability. Here, the model presents the price scheduling mechanism for aggregate differentiated services based on the opportunity cost (aggregate pricing uses $ per kbps per time, while flow-based pricing has the unit of $ per call or flow per time). The dynamic pricing of interconnection is edge pricing [12], since it is charged at the interconnection point using proxy opportunity costs.

Revenue-maximization price schedules for the individual service classes can be derived from demand function analysis; however, optimal profit-maximization pricing for different network service classes requires the observation of the opportunity costs of each service class. The optimal price and service schedule classes can be determined as aspects of network profit maximization. The estimated market opportunity costs and demand for different service classes will be used to define price scheduling that maximizes network profit. A fully competitive market will force the networks to set their service prices at the levels of market opportunity costs.

If we consider a network domain producing differentiated service $i$, a price schedule interval can be selected where a potential demand for $R_i$ is the equilibrium demand if the price $p_i$ stays fixed over a given interval of time. Assume that the aggregate service $i$ perceives a consumer surplus $\nu_i$ from the service and $\nu_i$ is randomly distributed across the aggregate service $i$ with a probability density function $r_i(\nu)$. The demand function is then an integral of consumer surplus over any given possible price. If the density function is independent for each service $i$, then for a given price $p_i$ in a time interval $\delta t$, the normalized potential demand can be expressed by

$$R_i(p) = \int_p^{\infty} r_i \delta\nu. \tag{10}$$

In a case where the services $i$ and $j$ have a joint random density function $r_{i,j}(\nu_i, \nu_j)$ for the benefits $\nu_i$ and $\nu_j$, the normalized expected demand should be expressed as:

$$R_i(p_i, p_j) = \int_{p_i}^{\infty} \int_0^{\nu_i + p_j - p_i} r_{i,j}(\nu_i, \nu_j)\delta\nu_j\nu_i; \tag{11}$$

$$R_j(p_i, p_j) = \int_{p_j}^{\infty} \int_0^{\nu_j + p_i - p_j} r_{i,j}(\nu_i, \nu_j)\delta\nu_i\nu_j. \tag{12}$$

In the interconnection network it can be safely assumed that the service benefits are independent if the services are differentiated in terms of their prices and performance. For those services that can be replaced by other services, we can assume joint functional services. In the following it is assumed that all the different service classes are able to have independent density functions as in Equation 10. The optimal service price for the independently differentiated services can be calculated as follows:

$$\arg\max_{p_i} \pi = \sum_{i=0}^{n} \int_0^T [p_i - f_i] R_i dt \tag{13}$$

$s.t.\ G_i = [\sum_{i=0}^{n} Ri] - C \leq 0,$

where $T$ is the pricing and provisioning interval, $f_i$ is the opportunity cost, and $C$ is the interconnection interface capacity. This is an optimization problem of pricing over time. $G_i$ should be less than zero, since the sum of the served demand, $\sum_{i=0}^{n} R_i$, cannot exceed the interconnection capacity $C$.

In the case where an ISP can exercise control over service pricing (as exists in a monopoly), we apply the result from [16], and the optimal pricing schedule $p_i^*$ must satisfy

$$p_i^* = \frac{-R_i}{\frac{\delta R_i}{\delta p_i}} + f_i - \psi_i, \tag{14}$$

where $\psi_i$ is the costate variable for the Hamiltonian,

$$H_i(p_i^*, \psi_i, t) = (p_i - f_i + \psi_i)R_i. \tag{15}$$

The Lagrangean for the above maximization can be expressed as

$$L_i(p_i^*, \psi_i, \lambda_i, t) = H_i(p_i^*, \psi_i, t) - \lambda_i G_i. \tag{16}$$

At the optimal $p_i^*$, $\frac{dH_i^*}{d\psi} = \frac{dL_i}{d\psi} = R_i^*$, and the following first-order condition should be satisfied:

$$\frac{\delta H_i}{\delta p_i} = R_i + (p_i - f_i + \psi_i)\frac{\delta R_i}{\delta p_i} = 0, \tag{17}$$

which produces optimal price $p_i^*$ shown in Equation 14. From the optimal price, it is apparent that $\frac{\delta R_i}{\delta p_i}$ represents the function of the demand elasticity

$$\varepsilon_i = \frac{\delta R_i}{\delta p_i}\frac{p_i}{R_i}, \tag{18}$$

on the price and that $\psi_i$ satisfies the differential equation ($x_i$ is bandwidth unit used to represent $R_i$):

$$\frac{d\psi_i}{dt} = -\frac{\delta H_i^*}{\delta x_i} \qquad \psi_i(T) = 0. \tag{19}$$

The previous price equation at its optimal point can be rewritten as follows:

$$p_i^* = -\frac{R_i}{\varepsilon_i \frac{R_i}{p_i^*}} + f_i - \psi_i. \tag{20}$$

The equation derives the optimal price for service $i$ with demand elasticity $\varepsilon_i$ at

$$p_i^* = \frac{\varepsilon_i}{1+\varepsilon_i}(f_i - \psi_i). \tag{21}$$

An optimal price schedule that maximizes the networks' profits is dependent on the demand elasticity of the price and the opportunity costs of that service as shown in Equation 21.

In the interconnection market, the networks negotiate a form of settlement with respect to the services to be provided at the network boundary. This settlement is commonly called a Service Level Agreement (SLA), and includes the terms and conditions regarding the interconnection, encompassing service availability, pricing, QoS specifications, and other legal and business issues. SLAs can be offered through various topological points. For example, a network can provide SLAs at an interconnection point to either a specific point (Point-to-Point SLA), multiple egress points (Point-to-Many Funnel SLA), or groups. The subset of the SLA that provides technical specifications such as QoS, aggregate traffic profiles, and service resource allocation is referred to as the Service Level Specification (SLS) [17]. The SLSs can take on various forms such as static or dynamic. Static SLSs are negotiated on a long-term regular basis (monthly or yearly). Dynamic SLS implementations require autonomous systems and signaling to update the interconnection SLSs dynamically (by the minute or hour).

In the QoS Interconnection networks, the absence of settlements may lead to unintended economic distortion among the network providers due to the asymmetry in the volume and values of different cross-service traffic over the QoS interconnection. For example, interconnection pricing is a way to explicitly motivate network providers to provide interconnections for the benefit of indirect network users and consistently maintain the quality of the connections, even for transit connections. Therefore, there should be certain financial interconnection settlements required to make the network of QoS networks work with reasonable economic incentives. The QoS interconnection pricing model is a financial form of the QoS interconnection settlement.

Payoff maximization can be an economically reasonable form for QoS interconnections where individual QoS networks will set the schedule for network service allocation (and accordingly provision the individual service classes differently) and pricing strategy. We consider three settlement types: S-K-A (Sender-Keeps-All), N-C-B (Non-Cooperative-Bilateral) and C-B (Cooperative-Bilateral):

**S-K-A** There is no financial settlement in place.

**N-C-B** The price and service and the selection of the interconnections are done in a selfish way by each domain. This is considered to be the most typical model of the interconnection settlements.

**C-B** Cooperative incentive exists for both of the interconnecting networks.

These different settlement scenarios characterize the optimization strategies of both inter-domain and intra-domain operations differently. The following equations are interconnection models for bilateral interconnection in different settlement types.

For S-K-A, an interconnecting network will try to minimize the loss of opportunity cost to serve the interconnecting demand, because there is no price settlement for each service class in place. In the case of N-C-B with price settlement, each interconnecting network will try to maximize the gross profits generated from all interconnection services. In C-B, two interconnecting networks try to maximize their joint gross profits. The discount factor function for monetary price and cost, $e^{\tau_i t}$, and the volatility function of the demand, $\upsilon_i(t)$, are used to represent the dynamic nature of the considered interconnection economy.

S-K-A: (Sender-Keeps-All)

$$\arg\min_{i} \sum_{i=n}^{n} \int_{t}^{t+T} \left[ f_i R_i e^{\tau_i t} \upsilon_i(t) \right] dt; \tag{22}$$

N-C-B: (Non-Cooperative-Bilateral)

$$\arg\max_{i,p_i} \sum_{i=1}^{n} \int_{t}^{t+T} \left[ [p_i - f_i] R_i e^{\tau_i t} \upsilon_i(t) \right] dt; \tag{23}$$

C-B: (Cooperative-Bilateral)

$$\arg\max_{i,j,p_i,p_j} \sum_{i=1}^{n} \sum_{j=1}^{m} \int_{t}^{t+T} \left[ [p_i - f_i] R_i e^{\tau_i t} \upsilon_i(t) + [p_j - f_j] R_j e^{\tau_j t} \upsilon_j(t) \right] dt; \tag{24}$$

where:
$i$: network service class selection $i$,
$j$: network service class selection $j$,
$p_i$: optimal price selection for service class selection $i$,
$p_j$: optimal price selection for service class selection $j$,
$f_i(t)$: marginal opportunity cost for service $i$,
$f_j(t)$: marginal opportunity cost for service $j$,
$e^{\tau_i t}$: discount functions for service $i$ with discount rate $\tau_i$,
$e^{\tau_j t}$: discount functions for service $j$ with discount rate $\tau_j$,
$\upsilon_i(t)$: volatility function for service $i$ demand,
$\upsilon_j(t)$: volatility function for service $j$ demand,
$R_i$: aggregate marginal data rate (instant demand) for service $i$,
$R_j$: aggregate marginal data rate (instant demand) for service $j$.

From the objective functions of three settlement scenarios, service $i$ denotes the service selection of one network and $j$ denotes the service selection of the other network that is involved in the bilateral interconnection contract. Assume that a network tries to maximize total payoff over the period of interconnection contract period $T$. The price rule for each service in the interconnection will be set to a default market price at the beginning of time $T$ and will be updated or controlled over time $T$ within SLA, depending on market conditions. The SLA conditions for the new pricing rules (not a specific price, but a rule for defining the price) might be applied to the next interconnection period. The pricing rule determines the price $p_i$ of the unit aggregate bandwidth over an interconnection provisioning update interval for service class $i$. An interconnection provisioning update interval should be selected that is stable enough to avoid demand depression on the user side due to uncertainty and provide more scalability of network management. On the other hand, this interval period must be dynamic enough to optimize the interconnection payoffs due to market value fluctuation. Let $R_i$ be the aggregate marginal data rate for service $i$. Then the ISP's economic behavior would be to allocate the preferred services $i$ up to the optimal resource allocation $R_i$, first using $p_i$.

## 3 Model Results

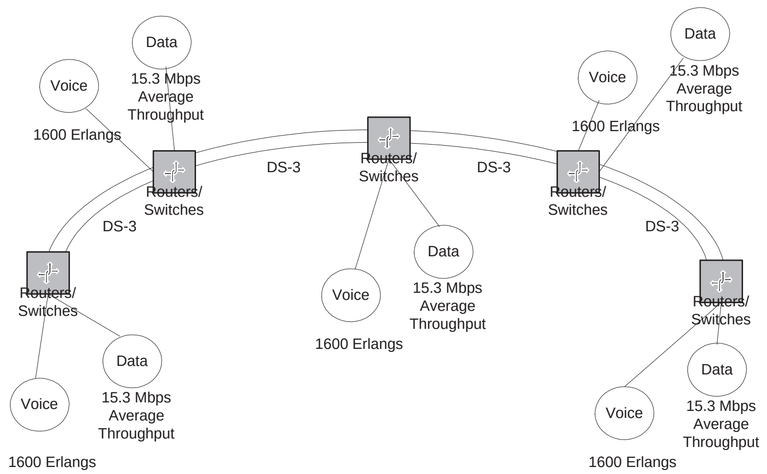
### *3.1 Simulations*

In order to analyze the economic behavior of different QoS networks, we perform computer network simulations to dimension the network resources and numerical simulation for the economic models. As shown in Figure 1, we used the COMNET III to capture the opportunity costs of different service classes of DiffServ QoS interconnection networks. The base network we simulated is a five-node network that is interconnected with DS3 links. We offered an average of 1600 Erlangs[5] of voice traffic and 15.8 Mbps of data traffic throughput as a maximum offered load to each switch. For voice traffic, we assumed that the traffic is compressed voice using G729A.[6] Using the statistics of recent traffic data measured from the Internet backbone trunks, we modeled the integrated service traffic as a cross section of the Internet backbone traffic and computed the intensity of this traffic relative to voice-call demand. We assumed the QoS traffic load to be offered to each of the core switches and to be traversed a various number of hops in the model. By monitoring the QoS measures of voice traffic which traversed various hops, we found the equivalent utilization that provides an acceptable QoS for the specific service. We ran the baseline

[5] Traditionally, one Erlang represents the traffic load that fully utilizes one telephone circuit.
[6] A G729A voice compression coder compresses a telephone-quality voice call at the maximum rate of 8kbps.

workload and, to capture the opportunity costs of different service classes of DiffServ networks, varied the workload for the given network capacity (typically at DS3) and measured the total usable idle bandwidth for different PHB service classes with various levels of QoS requirements and workloads.

Fig. 1. Simulation Model




For the traffic demand function of price, we used the following form of demand model function for modeling, and numerical simulation in the rest of our study:

$$R_i(p_i, t) = R_{i-max}\left[1 - \left(\frac{p_i(t)}{p_{i-max}(t)}\right)^{\alpha_i(t)}\right]. \qquad (25)$$

### *3.2 Single Network*

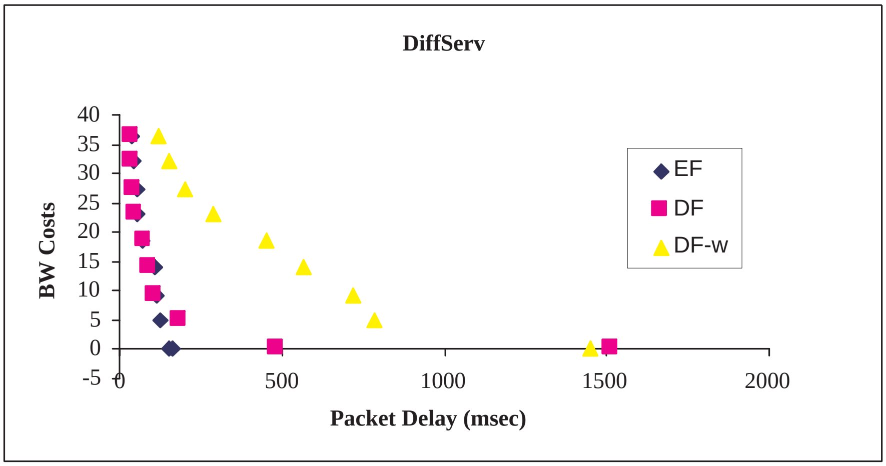
We simulated a simple DiffServ QoS network that treats each packet differently, based on the priority set in traffic packet. We simulated the voice load and data traffic load as EF (Expedite Forwarding) [7] and DF (Default Forwarding) [8] service class of the DiffServ network, respectively. In a similar way, we monitored the usable bandwidth and the delay for the different relative loads of base-line traffic loads.

Figure 2 shows the opportunity cost trends of the QoS requirements for those two DiffServ service classes. The bandwidth opportunity costs are derived from the usable idle bandwidths for different levels of QoS with different QoS requirements.

The usable bandwidth of the EF is measured for various 99th percentile packet delay requirements. Similarly, the usable bandwidth of the DF is measured

[7] A DiffServ service class designed for a real-time network service application.
[8] A DiffServ service class designed for best-effort service traffic.

Fig. 2. Usable Bandwidth Measurement for Different Service Classes




for various mean packet delay requirements, and for DF-w, the worst packet delay. Using our model for market opportunity costs, we calculated the DS-3 interconnection opportunity costs for EF-50 msec and DF-500 as follows:

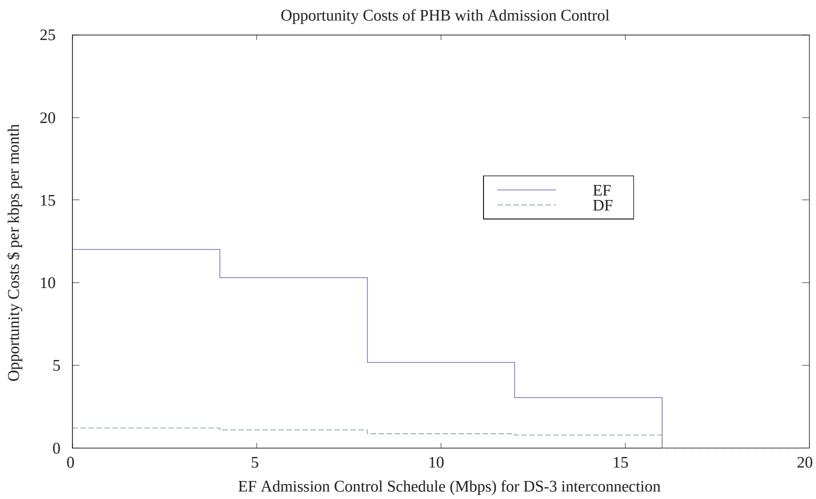
- Find the raw DS-3 interconnection market price. We assumed the found price is \$54,000 in the example.
- Calculate a market opportunity cost for each service class at different service strategy (admission control).
- Optimize the model and find an optimal price and resource allocation for each service schedule.

The optimization problem that must be solved is:

$$\begin{array}{ll} \max & \int (p_{EF} - f_{EF}) R_{EF} e^{-\gamma t} v(t) + (p_{DF} - f_{DF}) R_{DF} e^{-\gamma t} v(t) dt \\ \text{s.t} & f_{EF} R_{EF} \leq IC_0 \\ & f_{DF} R_{DF} \leq IC_0 \\ & f_{EF} R_{EF} + f_{DF} R_{DF} \geq IC_0 \\ & f_{EF} \leq p_{EF} \leq p_{EF-MAX} \\ & 0 \leq p_{DF} \leq p_{DF-MAX} \\ & 0 \leq R_{EF} \leq R_{EF-MAX} \\ & 0 \leq R_{DF} \leq R_{DF-MAX} \\ & R_{EF} + R_{DF} \leq C. \end{array}$$

The stated constraints imply the QoS networks bandwidth management mechanisms and the network characteristics. The first constraint says that the individual opportunity costs $f$ are limited to quantify the costs, which are less than the market value of the interconnection $IC_0$, which is \$54,000 in our example. The second constraint represents the gain of the interconnection due to the differentiation of the service traffic, since the $IC_0$ is the cost for best-effort interconnection data from the current market. The third and fourth constraints state that the selected price schedule $p$ from the DiffServ interconnection is bounded by individual opportunity costs and demand price. The fifth and sixth constraints state that the allowable bandwidth for each service $R$ is constrained by the demand characteristics for the interconnection shown in the previous demand model. The last constraint shows the capacity limit $C$ of the interconnection, which is 45 Mbps in the example case of the considered model.

Fig. 3. Opportunity Costs for Service Schedules with Admission Control




Using the opportunity costs for the different admission control schedules specified in Equation 3, we simulated the following base case model, which specifies parameters of demand for different service classes. We used the parameter set and solved the optimization model numerically. The following tables summarize the base model parameters and the resulting service schedules for the base DiffServ network. The demand function parameters in Table 1 are used for the optimization model. We assumed that the demand of EF is much less elastic than DF. No discount rate is assumed for each optimization. The volatility of the network average demand is assumed to be the gamma distributed with the mean of 1 and variance of 0.1. The gamma distribution is a non-integer generalization of Erlang distribution, which is often used in a queueing model. The mean value of 1 means that the average demand of a service for a given interval is gamma-distributed over a specified average estimated demand value of $R_i$.

For the above example, the parameters are chosen to represent a probable DiffServ network. Unlike the general best-effort interconnection, this QoS interconnection requires a fair amount of voice traffic (which is EF) to be transported over the DiffServ network. Assuming the \$1800 per month for a T1 PSTN connection, we assumed that \$20 per kbps per month for IP telephony

Table 1
Demand and Network Base Parameters for DiffServ Classes

| parameter | EF Values | DF Values |
|---|---|---|
| Maximum Rate, $R_{i-max}$ | 22 Mbps | 45 Mbps |
| Maximum Price, $p_{i-max}$ | 20 | 1.2 |
| Demand factor, $\alpha_i$ | 0.9 | 0.3 |
| Market Interconnection Value, $IC_0$ | 54,000 | 54000 |
| Interconnection Capacity, $C$ | 45 Mbps | 45 Mbps |
| Discount rate, $\tau$ | 0 | 0 |
| Volatility function, $\upsilon_i(t)$ | $\gamma(0.1, 10)$ | $\gamma(0.1, 10)$ |

service would be the maximum price the network is willing to pay for voice service. However, the maximum price for DF is considered to be bounded with the average price of the general best-effort interconnection, because the differentiated DF might have a more degraded network service quality than the general best-effort network interconnection. We assume that the best effort DF is more price-elastic than the EF in the QoS-interconnection. With the above parameter values and demand, the optimization of the DiffServ bandwidth management can be solved. Table 2 summarizes the numerical solution for the base case.

Table 2
Numerical Solution of the Base Case

| Variables | EF Values | DF Values |
|---|---|---|
| Profit Maximizing Price, $p_i$ | 10.05 | 0.6 |
| Network Profit, $F$ | 48104.92 | 675.2 |
| Bandwidth Allocation, $R_i$ | 10.16 Mbps | 8.44 Mbps |
| Opportunity Cost Price, $p_i = f_i$ | 3.08 | 0.52 |
| Network Revenue, $E$ | 52424.68 | 5418.39 |
| Bandwidth Allocation, $R_0$ | 17.02 Mbps | 10.39 Mbps |

The key observations from the example results are as follows. Note first that the profit-maximizing price of the EF deviated more from the opportunity cost than the price of DF. We observed that the profit of the DF service over the opportunity costs was very stable for any prices in the range of \$0.52 and \$1.2. These results are consistent with the economic theory of profit maximization pricing. The price schedule of inelastic EF service is more deviated from the opportunity cost than that of elastic DF service. With a profit maximizing price schedule, the ratio of the EF price to the DF price was \$16.75. The comparative ratio for the opportunity-cost pricing was only \$5.90. At the different prices (profit maximization and opportunity cost pricing), we computed demand. With opportunity cost pricing, the service utilization is maximized for each service, but the profit is not maximized. With the profit maximizing pricing, the network service provider could generate an additional 90% of profit in addition to the network opportunity costs of a DS-3 interconnection. [9]

### *3.3 Two Networks*

Using the same simulation model, we interconnected two network domains with different interface capacities and service demands. We also considered three service class [10] models for interconnection services. We considered several different demand assumptions for different DiffServ service classes, and different opportunity costs for different capacities. Here, we used a linear demand function, which is a special form of the normalized power demand function with $\alpha = 1$ that was used before.

We focused on the consumer surplus effects due to the number of networks in the QoS interconnection. For different demand scenarios, we numerically simulated the price schedule and bandwidth allocation of each network. Using the profit-maximizing price schedule and bandwidth allocation of a direct interconnecting network, we computed consumer surplus. Two cases (Single Network, and N-C-B Two Networks) were compared and tested. We simply varied $p_{i,max}$ to represent different demand patterns. We considered eight demand scenarios by varying the elasticity of demand (Elastic, Inelastic) of each DiffServ service class for three types of interconnection capacity. Table 3 summarizes the values used for different demand scenarios.

As shown above, two maximum price values were assigned to each service class to reflect the different elasticity measures of each class. Since three service classes per interconnection were considered, the total number of demand scenarios considered was eight. For example, demand patterns 1, 2, 3, are 4 when the EF service demand is price-inelastic. So demand patterns 5, 6, 7, and 8 are when the EF service demand is price-elastic. Similarly, demand patterns 1, 2, 5, 6 have inelastic AF service and the rest have elastic AF service. Demand patterns 2, 4, 6, and 8 have inelastic DF service, and the rest have elastic DF service.

[9] Notice that opportunity cost pricing generates the integrated revenue that recovers its opportunity costs of \$54,000 with slight additional deviation.

[10] AF (Assured Forwarding) service class is additionally considered. This service class is designed to support loss-sensitive QoS data services in the DiffServ service architecture.

Table 3
Demand Parameters for Different Service Class

| $R_{max}$(kbps) | Service Class | Elastic $P_{max}$ | In elastic $P_{max}$ |
|---|---|---|---|
| 1544 | Expedite Forwarding | 18.404 | 23.074 |
| 1544 | Assured Forwarding | 6.106 | 8.366 |
| 1544 | Default Forwarding | 3.241 | 4.412 |
| 44736 | Expedite Forwarding | 7.702 | 11.537 |
| 44736 | Assured Forwarding | 2.753 | 4.183 |
| 44736 | Default Forwarding | 0.621 | 2.206 |
| 155000 | Expedite Forwarding | 4.210 | 7.702 |
| 155000 | Assured Forwarding | 1.225 | 2.753 |
| 155000 | Default Forwarding | 0.587 | 0.621 |

The optimization input variables shown in Table 4 were used for each of the above demand scenarios. The opportunity cost set was selected to represent the interconnection service requirements of EF(100msec)-AF($2*10^{-2}$)-DF, and EF(50msec)-AF($10^{-2}$)-DF for the interconnections of one network and two networks, respectively. Using these parameters, we gathered the optimization pricing schedule and service allocation data to calculate the amount of consumer surplus created. We compared the consumer surplus values when the QoS interconnection involves one or two networks in order to satisfy the same end-to-end service quality across the interconnecting network.

Table 4
Baseline Parameters Inputs From Measurement

| Network | Opportunity Costs EF-AF-DF | Interconnection Costs | Capacity $R_{max}$ |
|---|---|---|---|
| one-network | 8.333-3.020-1.530 | 4400 | 1544 |
| | 1.755-0.621-0.306 | 27800 | 44736 |
| | 1.567-0.569-0.289 | 91062 | 155000 |
| two-network | 8.636-3.105-1.585 | 4400 | 1544 |
| | 1.817-0.656-0.331 | 27800 | 1544 |
| | 1.674-0.610-0.326 | 91062 | 155000 |

In Figures 4, 5, and 6, mean consumer surpluses of different network interconnection scenarios (one-network, two-network, two-network-adjusted) were assessed for different demand patterns. For each interconnection with different demand conditions, the study showed differences in the means of the consumer surpluses that different interconnection conditions generated.

Fig. 4. Mean Consumer Surplus Values of Network Interconnections: DS1

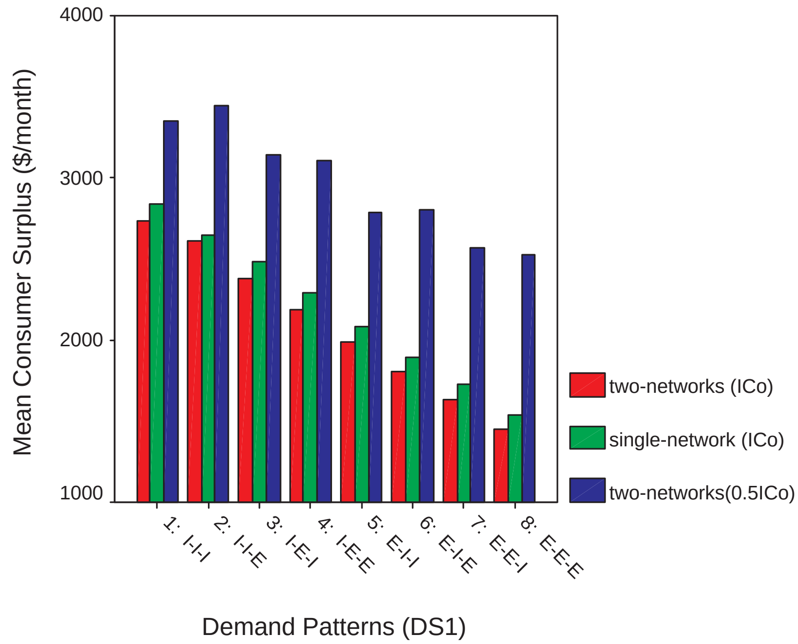
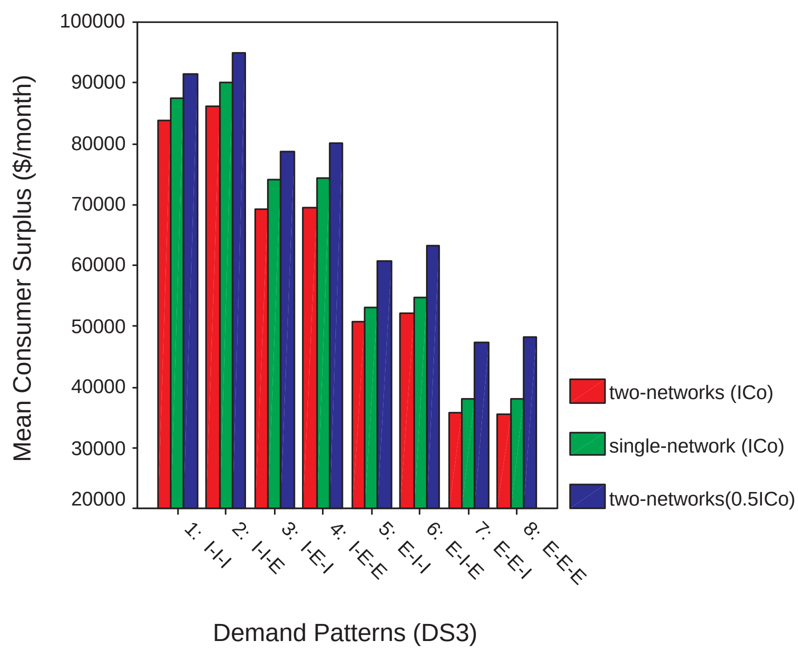
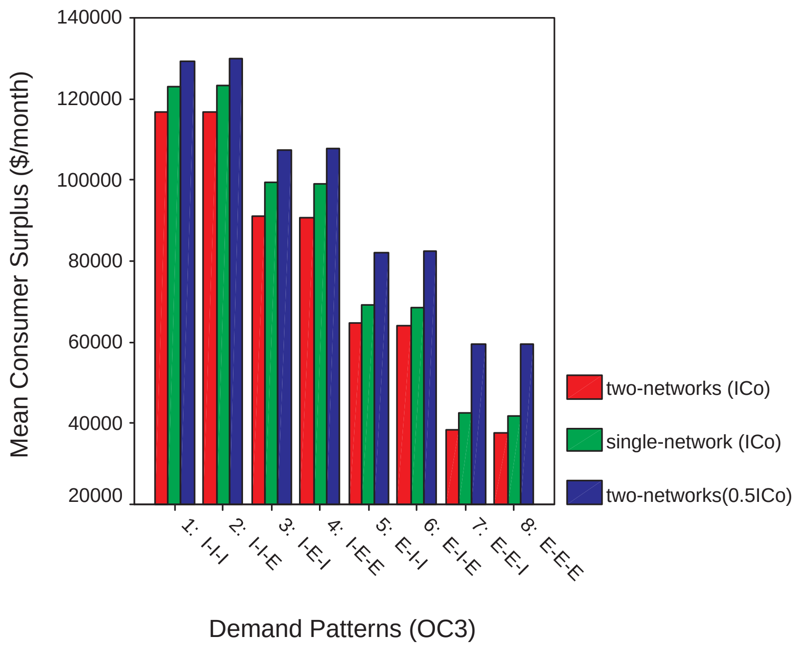



Fig. 5. Mean Consumer Surplus Values of Network Interconnections: DS3




Assuming the same capacity options, demand patterns, and market conditions, we performed data analysis on the consumer surplus effect of interconnection

Fig. 6. Mean Consumer Surplus Values of Network Interconnections: OC3




network configurations. The study compared three network interconnection configurations such as one-network interconnection with standard market price (ICo), two-network interconnection with standard market price (ICo), and two-network interconnection with adjusted market price (0.5*ICo).[11] End-to-end network performance was assumed to be 99 percentile 100 msec and average 500 msec-$2 * 10^{-2}$ for EF and AF services, respectively. For a two-network interconnection configuration, these requirements were evenly shared between two networks, since we assumed the same QoS technology and network topology for each network, and each network is likely to be motivated to provide QoS services. The measured opportunity costs for these network configurations are given in Table 4. Using optimal pricing for these interconnection opportunity costs, we calculated the consumer surplus for each network service, and the sum of the individual service consumer surpluses were compared among different network interconnection configurations. Tables 5, 6, and 7 represent descriptive mean statistics, a summary of the analysis of variance, and comparison tables of the consumer surplus of different interconnection network configurations, considering their interactions with interconnection capacity choices and demand patterns.

[11] The rationale for an adjusted market price scenario is because many of the market interconnection settlements are distance-adjusted. And, with a two-network interconnection, the market interconnection price can be adjusted to 50% in the market if we assume each network has half of the distance of the end-to-end interconnection distance.

Table 5
Descriptive Mean Statistics of Consumer Surplus of Interconnection Network Configuration

| Variable | Capacity | Network | Mean | Std. Dev. |
|---|---|---|---|---|
| consumer surplus | DS-1 | two-network-ICo | 2098.475 | 445.478 |
| | | one-network-ICo | 2189.375 | 442.114 |
| | | two-network-ICo | 2964.506 | 405.866 |
| | | Average | 2417.452 | 576.641 |
| | DS-3 | two-network-ICo | 43155.750 | 9036.954 |
| | | one-network-ICo | 43975.375 | 9017.138 |
| | | two-network-ICo | 46436.438 | 10543.215 |
| | | Average | 44522.521 | 9459.104 |
| | OC-3 | two-network-ICo | 77568.313 | 31034.648 |
| | | one-network-ICo | 83400.750 | 31806.609 |
| | | two-network-ICo | 94744.313 | 30153.632 |
| | | Average | 85237.792 | 31183.016 |

Data results indicate that the mean consumer surpluses of a two-network, market-based QoS interconnection can be significantly different from those of a single-network interconnection. We compared three different network configurations. Among the comparisons of those configurations, the mean consumer surplus of one network was not significantly different when the bandwidth cost was not adjusted from the bandwidth market. However, when the bandwidth cost was adjusted (half-cost in this case), significant differences were observed for all the demand patterns considered.

These results have some implications for the market-based interconnection economy. With the existence of more market signal (price) information and market products (interconnection) in the bandwidth economy, a dynamic pricing and provisioning interconnection mechanism will help the networks take the best advantage of the opportunity and availability of the market. This dynamic interconnection mechanism will be strongly supportable by ISPs that do not own their own carrier network facilities. Further study of the market mechanics and business implications of the market-based dynamic interconnection between backbone carriers and ISPs is clearly necessary.

Table 6
Summary Table for ANOVA on Consumer Surplus of Interconnection Network Configuration

| Source | Type III Sum of Square | d.f. | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Corrected | 2.10E+11 | 71 | 2.96E+09 | 50.203 | .000 |
| Intercept | 2.80E+11 | 1 | 2.80E+11 | 4737.619 | .000 |
| NETWORK | 1.27E+09 | 2 | 6.34E+08 | 10.737 | .000 |
| CAPACITY | 1.65E+11 | 2 | 8.23E+10 | 1395.145 | .000 |
| DEMAND | 2.21E+10 | 7 | 3.16E+09 | 53.511 | .000 |
| CAP*DEM | 2.09E+10 | 14 | 1.49E+09 | 25.245 | .000 |
| CAP*NET | 1.27E+09 | 4 | 3.19E+08 | 5.401 | .001 |
| NET*DEM | 63302101 | 14 | 4521579 | .077 | 1.000 |
| CAP*NET*DEM | 1.15E+08 | 28 | 4090930 | .069 | 1.000 |
| Error | 4.25E+09 | 72 | 59003349 | | |
| Total | 4.94E+11 | 144 | | | |
| Corrected Total | 2.15E+11 | 143 | | | |

Table 7
Multiple Comparison Table on Consumer Surplus of Interconnection Network Configuration

| (I) Network | (J) Network | Mean Diff. (I-J) | Sig. | 95 % C.I. | |
|---|---|---|---|---|---|
| | | | | L.B. | U.B. |
| two-ICo | one-ICo | -2247.654 | .329 | -5999.971 | 1504.663 |
| | two-0.5*ICo | -7107.573 | .000 | -10859.9 | -3355.256 |
| one-ICo | two-ICo | 2247.654 | .329 | -1504.663 | 5999.971 |
| | two-0.5*ICo | -2859.919 | .008 | -8612.236 | -1107.601 |
| two-0.5*ICo | two-ICo | 7107.573 | .000 | 3355.256 | 10859.890 |
| | onw-ICo | 4859.919 | .008 | 1107.601 | 8612.236 |

## 4 Discussion and Future Work

In this paper we have developed a framework for a market-based QoS interconnection. The motivations for these interconnection and related problems were presented at the beginning of the paper. We used the market-based network economic model as the implementation mechanism for optimizing interconnection policies through the bandwidth management of DiffServ networks. The economic model encompasses cost, demand, pricing, and settlement. Theoretical network economic models were developed to characterize different economic variables and properties of QoS network services.

First, we propose to capture some of the important economic cost variables of the network services for problem formulation. For example, the bandwidth management agent in the DiffServ network domain may assess the costs of the individual network services based on the estimates or measurement of network usage and the market value of the telecommunications transport and interconnection capacity. An estimate of network usage can be obtained by measuring different service traffic loads at the edge of the interconnected networks. The market value of the transport capacity can be assessed from price information of the online bandwidth commodity market or the bandwidth management agent's collocation or interconnection price exchange. Therefore, the economic cost assessment is the market opportunity cost for the fixed capacity network with fixed interconnection link capacity. We used computer simulation to capture the idle usable capacity of the fixed network interconnection link capacity for different QoS requirements, for example, different delays for EF services of DiffServ. We present a potential difference in the bandwidth opportunity costs of QoS interconnection among different QoS networks. The usable bandwidth measurement from simulation results for different types of QoS networks provides a proxy for the opportunity costs for QoS interconnection. Even though the bandwidth opportunity cost is a function of the quality of integrated service traffic, it is mainly determined by the most QoS-aware services, since such service is most sensitive to changes of quality. The optimal interconnection strategy for the QoS networks will be to minimize this cost by allocating the quality optimally.

For extended study, we consider the multiple networks' bandwidth management trading situation where individually assessed opportunity costs of different types of service groups for different bandwidth management network domains can be represented as indifference characteristic cost functions when the interconnections are between networks. The indifference curves for different QoS requirements can be used to define the exchange rates among interconnecting networks of bandwidth by capturing the slopes of indifference curves.

## References


[1] M. Kende, The digital handshake: Connecting internet backbones, in: OPP Working Paper No. 32, Office of Plans and Policy, Federal Communications Commission, 2000.

[2] J. K. MacKie-Mason, A smart market for resource reservation in multiple quality of service information network, in: Technical Report Draft of University Michigan and NBER, 1997.

[3] K. Kumaran, M. Mandjes, D. Mitra, I. Saniee, Resource usage and charging in a multi-service multi-qos packet network, in: Proc. of MIT/Tufts Workshop on Internet Service Quality Economics, Cambridge, Massachusetts, 1999.

[4] J. Hwang, M. B. Weiss, On the economics of interconnection among hybrid qos networks in the next generation internet, in: Proc. of International Telecommunications Society 2000, Buenos Aires, Argentina, 2000.

[5] A. Chiang, Fundamental Methods of Mathematical Economics: Third Edition, McGraw-Hill, 1984.

[6] R. Cocchi, S. Shenker, D. Estrin, L. Zhang, Pricing in computer networks: Motivation, formulation, and example, IEEE/ACM Transactions on Networking 1 (6) (1993) 614–627.

[7] A. Gupta, D. Stahl, A. Whinston, Priority pricing of integrated services networks, Internet economics (1996) 323–352.

[8] J. Mackie-Mason, H. Varian, Pricing the internet, Public Access to the Internet .

[9] D. S. A. Gupta, B. Jukic, A. B. Whinston, Impact of congestion pricing on network infrastructure investment, in: Internet Service Quality Economics, Cambridge, Massachusetts, 1999.

[10] Q. Wang, J. Peha, M. Sirbu, Dynamic pricing of integrated services networks, presented at the Telecommunications Policy Resarch Conference (October 1995).

[11] M. Mandjes, N. van Forest, Aspect of pricing in an integrated services network, in: Proc. of the 16th International Teletraffic Congress, 1999.

[12] S. Shenker, D. Clark, D. Estrin, S. Herzog, Pricing in computer networks: Reshaping the research agenda, presented at the Telecommunications Policy Resarch Conference. (October 1995).

[13] J. Bailey, J. Nagel, S. Raghavan, Ex-post internet pricing, in: Internet Service Quality Economics, Cambridge, Massachusetts, 1999.

[14] V. Siris, D. Songhurst, G. Stamoulis, M. Stoer, Usage-based charging using effective bandwidth: studies and reality., in: Proc. of the 16th International Teletraffic Congress, Edinburgh, UK, 1999.

[15] L. W. McKnight, J. Boroumand, Pricing internet services: After flat rate, in: Proc. of MIT/Tufts Workshop on Internet Service Quality Economics, Cambridge, Massachusetts, 1999.

[16] S. Kalish, Monopolist pricing with dynamic demand and production cost, Market Science 2 (1983) 135–159.

[17] Y. Bernet, J. Binder, S. Blake, M. Carlson, B. Carpenter, S. Keshav, E. Davies, B. Ohlman, D. Verma, Z. Wang, W. Weiss, A framework for differentiated services, Internet Draft, Internet Engineering Task Force, work in progress (March 1999).