ISSUES IN PRICING OF AND ACCOUNTING FOR INTERNET PACKET DELIVERY SERVICE

Alexander Kogan, Ephraim F. Sudit and Miklos A. Vasarhelyi

Faculty of Management, Rutgers University

180 University Ave., Newark, NJ 07102

{kogan, sudit, miklosv}@andromeda.rutgers.edu

June 1998

Introduction

The Internet is a global network interconnecting networks of computers. The users of the Internet have access to a variety of services ranging from traditional electronic mail, remote login, and file transfer to the World Wide Web (WWW). All these services are based on the use of a common technology and language over telecommunication links. These services are rendered and priced by private providers who own and/or operate Internet Protocol (IP) networks, which comprise the Internet.

Internet service provisioning is becoming a sizable business. According to *Data Communications* (see [8]), Internet service providers (ISP) in the United States generated more than $3 billion in 1997 revenues for providing Internet connections, and these revenues may reach $13 billion by 2001. According to the fall 1997 survey of the Internet

service providers by the Boardwatch Magazine (see [17]), the number of ISPs in North America reached 4,354.

Internet service providers and equipment vendors have been concerned that the growth in the demand for Internet bandwidth is outpacing the expansion of the capacity of Internet backbones. Alan Taffel, vice president of marketing and business development at UUNet Technologies reports (see 21) that Internet traffic used to double every year, but now it is doubling every three to six months: "We have to radically alter our backbone very, very regularly. We and everybody else are going to have a difficult time keeping up with bandwidth demand." The problems faced by the ISPs are largely attributable to new real time multimedia applications straining the current Internet technology and infrastructure, and the difficulty in predicting the future demands for particular Internet applications. Most of the participants in the Spring Internet World'98 conference agreed that usage-based pricing would eventually prevail.

At present, the end users who generate the demand for Internet services are charged for access but not for usage. With the Internet experiencing increasingly frequent episodes of traffic congestion, the economic efficiency of the pricing system which rations demand on a first come first serve basis becomes questionable.

Sophisticated economic pricing mechanisms have been suggested for the Internet (see [8,9,10]). However, the accounting systems required to support billing based on these mechanisms have not been developed (see [1,2,3,6,7,11]). Moreover, existing Internet service providers by the Boardwatch Magazine [illegible], the number of ISPs in North America reached [illegible].

Internet service providers and equipment vendors are concerned that the growth in the demand for Internet bandwidth is outpacing the growth of the capacity of Internet backbones. Alan Taffel, vice president of marketing and business development at [illegible] Technologies, reported [illegible] that Internet traffic [illegible] every [illegible] doubling every three to six months. "We have to find out where our bandwidth [illegible] regularly. We [illegible] that are going to [illegible] keeping up with bandwidth demand." The problems faced by the [illegible] time multimedia applications, [illegible] the current Internet technology and infrastructure, and the difficulty in [illegible] the [illegible] for [illegible] Internet [illegible]. Most of the participants in the Spring Internet World [illegible] agreed that [illegible] based pricing would eventually prevail.

As [illegible] the end users [illegible] generate the demand for [illegible] the [illegible] with the [illegible] Internet [illegible] traffic congestion [illegible] the economic efficiency of the pricing system [illegible] first-come first-served [illegible].

Such [illegible] mechanism [illegible] [illegible]. However, the [illegible] mechanism [illegible] [illegible].

instrumentation may limit the capability of accounting systems to support complex pricing mechanisms. Yet, without an efficient pricing system, investment in capacity expansion may suffer due to lower expected returns. Absence of adequate activity and cost accounting systems precludes economically efficient allocation of costs among providers of Internet services.

All major users of the Internet are likely to be profoundly affected by the efficiency of the Internet pricing system. The more they rely on the Internet in the conduct of their business, the higher their stake is in the economic consequences of the Internet pricing system. In this article, we assess the economic efficiency and technical and accounting feasibility of alternative systems for pricing Internet services.

How the Internet Works

Internet services are rendered by the transfer of information among its users. The transferred bites of data are clustered in packets of various lengths that typically average about several hundred bytes, and generally do not exceed 1.5 KB. Depending on the amount of data it contains, a message delivered over the Internet is comprised of few or many packets. Each packet consists of a header and a body of data. The header includes addresses of source and destination(s) and information about data length, and type of services (precedence, reliability, delay, and throughput).

Packets are transmitted over telecommunication lines and directed (switched) to their destination by a sequence of routers. Routers are usually special purpose computers designed for packet switching. They route the packets by reading the address information stored in the packets headers. Routing algorithms use the so-called "next hop routing" (i.e. they route a packet to the next available router on its path to the final destination). Different packets belonging to the same message may be routed along different routes to be reassembled at their final destination into the original message. Some networks on the Internet may impose stricter limitations on the maximum packet size than the network from which a packet originates. As a result, a packet can be occasionally fragmented en route into a number of smaller packets. The latter will travel independently over the Internet to be reassembled at the final destination.

Packets are routed on a first come first serve basis. Providers of Internet services are not obligated, either collectively or individually, to transfer all packets to their destinations safely, reliability, and in a timely manner. They are bound only by the so-called "best effort delivery" principle. They simply have to try their best. Thus, when errors, breakdowns or congestion occur, packets may be lost or corrupted.

The network of wires and routers comprising the Internet is "dumb" in the sense that it will not detect loss or corruption of packets. The detection task is left to the smarter end point computers, known as the Internet "hosts", who can perform quality control. The transmission control protocol (TCP) software implements this quality control by creating and maintaining virtual connections over the Internet. TCP uses acknowledgments of

Packets are transmitted over telecommunication lines and directed (switched) to their destination by a sequence of routers. Routers are usually special purpose computers designed for packet switching. They route the packets by reading the address information stored in the packets headers. Routing algorithms use the so-called "next hop routing" i.e. they route a packet to the next available router on its path to the final destination. Different packets belonging to the same message may be routed along different routes to be reassembled at their final destination into the original message. Some networks on the Internet may impose stricter limitations on the maximum packet size than the network from which a packet originates. As a result, a packet may be occasionally fragmented en route into a number of smaller packets. The latter will travel independently over the Internet to be reassembled at the final destination.

Packets are routed on a first come first serve basis. Providers of Internet services are not obligated, either collectively or individually, to transfer all packets to their destinations safely, reliably, and in a timely manner. They are bound only by the so-called "best effort delivery" principle: They simply have to try their best. Thus, when errors in transmission or congestion occur, packets may be lost or corrupted.

The network of wires and routers comprising the Internet is "dumb" in the sense that it will not detect loss or corruption of packets. The detection task is left to the smarter end point computers, known as the Internet "hosts", who can perform quality control. The transmission control protocol (TCP) software implements the quality control by creating and maintaining virtual connections over the Internet. TCP uses acknowledgments of delivered packets, and retransmission of those packets whose acknowledgments are not received. This ascertains end-to-end reliability of Internet services. Internet services built on TCP include WWW, e-mail, file transfer and telnet. A number of Internet services (in particular those used for real time multimedia communications over the Internet) cannot afford the delays and overhead associated with acknowledgments and retransmissions. Such services do not use TCP, and are implemented on top of the user datagram protocol (UDP), which does not guarantee reliability.

How Individual Users Access the Internet

Many individuals access the Internet through their employers or organizations they belong to. Others connect through the Internet Service Providers (ISPs) who typically supply their users with basic Internet software package programs. Online services (e.g., America Online, Microsoft Network) provide another popular access option. They offer to subscribers their own set of services, while providing gateways and access to the Internet.

The online services that currently provide access to their proprietary materials as well as to the Internet are evolving in the direction of becoming pure content providers, i.e. Internet media companies similar in nature to TV networks. The provisioning of access will be separated from the provisioning of content, and the former will increasingly become the domain of telecommunication companies.

For users who own an appropriate computer and have access to a telephone line, the cost of connecting to the Internet is reduced to monthly ISP access charges which can be as low as $20 a month. Employees of organizations connected to the Internet may have "free" access depending on the Internet usage policies of their employers.

Providers of Internet Services

Naturally, Internet providers are the ones responsible for the pricing of their services. The Internet is not owned or operated by a single company. Rather, many parts of the Internet are owned and/or operated by different providers. It is useful to think of a hierarchy of Internet providers. This hierarchy reflects the architecture of the Internet, which has at its base four all-purpose Network Access Points (NAPs) which allow various computer networks within the Internet to interconnect and exchange traffic.

The Internet service providers are a diverse group. Currently they can be viewed as a hierarchy. Their hierarchical relationship is established by their peering agreements and access payments. At the top of the ISPs hierarchy are the National Service Providers (NSPs). The NSPs (e.g., MCI, Sprint, ANS, UUNet, PSI) maintain their own large-scale wide area networks and use high speed lines (i.e., T3 lines with the speed of about 45Mbs, OC3 lines with the speed of about 155Mbs, or even OC12 lines with the speed of about 622Mbs). Next in the hierarchy are Regional Network Providers (RNPs) which usually connect to the Internet through NSPs, but may also have direct connections to NAPs. Large organizations usually connect to the Internet either directly through NSPs or

through RNPs. Individual users or small organizations connect to the Internet through intermediaries, known as Internet Access Providers (IAPs). IAPs range from very small enterprises to very large telecommunication firms. IAPs connect to RNPs and resell Internet access over telephone dial-up connections. Most of the dial-up connections are relatively slow channels that use special versions of the Internet Protocol (IP) called SLIP and PPP.

The ISPs are either dedicated Internet companies or vertically integrated telecommunication companies. In its early days, the Internet was dominated by dedicated Internet companies like Advanced Networks and Services (ANS) or Uunet Technologies. Later the trend shifted toward merging the provision of the Internet service with other telecommunication services, which resulted in the current domination of vertically integrated companies on the Internet. For the companies that own the network infrastructure (e.g. long distance lines) substantial portion of the cost is sunk. Therefore, they can provide value-added services at prices lower than those of companies without their own communication lines who must incur the incremental costs of leasing long haul lines. At the same time, it is often attractive for the telecommunication providers to acquire the dedicated Internet providers for the sake of their expertise, their Internet infrastructure (e.g. routers), and their Internet customer base. These acquisitions of dedicated Internet companies by major telecommunication companies are designed to appropriate the benefits of economy of scope in the provisioning of various telecommunication services.

through RNPs. Individual users or small organizations connect to the Internet through intermediaries known as Internet Access Providers (IAPs). IAPs range from very small enterprises to very large telecommunication firms. IAPs connect to RNPs and resell Internet access over telephone dial-up connections. Most of the dial-up connections are relatively slow channels that use special versions of the Internet Protocol (IP) called SLIP and PPP.

The IAPs are either dedicated Internet companies or vertically integrated telecommunication companies. In its early days, the Internet was dominated by dedicated Internet companies like Advanced Networks and Services (ANS) or Uunet Technologies. Later the trend shifted toward merging the provision of the Internet service with other telecommunication services, which resulted in the current domination of vertically integrated companies on the Internet. For the companies that own the network infrastructure, e.g. long distance lines, substantial portion of the cost is sunk. Therefore, they can provide value-added services at prices lower than those of companies without dedicated communication lines who must incur the incremental costs of leasing long haul lines. At the same time, it is often attractive for the telecommunication providers to acquire the dedicated Internet providers for the sake of their expertise, their Internet software, technologies, and their Internet customer base. These acquisitions of dedicated Internet companies by giant telecommunication companies are designed to appropriate the benefits of economy of scope in the provisioning of various telecommunication services.

The August 1996 merger of UUNet with MFS Communications, and the December 1996 merger of the combined company with WorldCom (the nation's fourth largest long distance carrier) are some of the best examples of this process. Another example is the May 1997 merger of BBN (the oldest Internet company) with GTE (the largest independent local telephone company). Additionally, major telecommunication companies have recognized the importance of the Internet, and started actively developing their own Internet offerings. An impressive example of such developments is MCI which had hardly any presence on the ISP market until late 1994, and is now one of the largest ISPs in the world. The proposed merger of WorldCom and MCI will consolidate the Internet service provisioning market even further. This latest merger demonstrates the effect of economies of scale on the major telecommunication providers.

Other acquisitions of ISPs by telecommunications companies which have been reported in the press (see 23) "... include ICG Communications, a small Colorado-based carrier, buying Netcom On-Line Communications Services, and IXC Communications in Austin taking a 20% stake in PSINet. The few remaining national, publicly held ISPs are Concentric Network Corp, MindSpring Enterprises Inc. and EarthLink Network Inc. "To stay independent, you need to be huge," says a Forrester Research analyst."

Some ISPs opt for establishing various forms of partnerships with telecommunication companies without being formally acquired. According to WEBWEEK (21), PSINet Inc. has entered an agreement with IXC Communications Inc., giving the former the rights to use the latter's high bandwidth long haul lines – 10,000 miles of OC-48 fiber (2.4 Gbps).

The August 1996 merger of UUNet with MFS Communications, and the December 1996 merger of the combined company with WorldCom (the nation's fourth largest long distance carrier) are some of the best examples of this process. Another example is the May 1997 merger of BBN (the oldest Internet company) with GTE (the largest independent local telephone company). Additionally, major telecommunication companies have recognized the importance of the Internet and started actively developing their own Internet services. An interesting example of such developments is MCI which had hardly any presence in the ISP market until late 1994, and is now one of the largest ISPs in the world. The proposed merger of WorldCom and MCI will consolidate the Internet service market even further. This latest merger demonstrates the effect of economies of scale on the major telecommunication providers.

Other acquisitions of ISPs by telecommunications companies which have been reported in the press over 1997 include RCN Communications, a small Colorado-based carrier, [illegible] Communications Services, and IXC Communications in Austin taking a 20% stake in [illegible]. The few remaining national, publicly held ISPs are Concentric Network Corp, MindSpring Enterprises Inc. and EarthLink Network Inc. "To stay independent, you need to be big," says a Forrester Research analyst.

Some ISPs opt for establishing various forms of partnerships with telecommunication companies without being formally acquired. According to WEBWEEK [11], PSINet Inc. has entered an agreement with IXC Communications Inc., giving the former the rights to use the latter's high-bandwidth long haul lines — 10,000 miles of OC-48 fiber (2.4 Gbps).

As of the time of this writing, PSINet remained the only major ISP not acquired yet by a major telecommunication company.

WEBWEEK (21) quoted from a recent Forrester Research Inc. report, "Internet Carrier Gear," that predicted heavy investment in the Internet infrastructure. In particular, they estimated that " … because of rapid increases in the number of Internet users, ISPs will spend almost $8 billion over the next five years on high-density access servers and gigabit routers. Forrester expects that ISP spending per year on high-bandwidth equipment will rise from about $360 million in 1997 to $2.6 billion in 2000."

Interconnection Agreements and Settlements among Providers

The Internet Service Providers (ISPs) interconnect their networks at NAPs and exchange Internet traffic on the basis of the so-called peering agreements discussed below (see 20).

Multilateral and bilateral interconnection agreements include reciprocal arrangements, access payments and membership fees. Settlements based on the balance of traffic accounting, if any, are rare. This situation is partly due to historical reasons. Until recently, the Internet was significantly subsidized by the Federal Government. As a result, no elaborate technological infrastructure was developed to support commercial provision of Internet packet delivery services. This made it impractical for early ISPs to implement any "balance of trade" type settlements, which would have required packet and/or byte

counting. Hence, a basis for the ubiquitous acceptance of peering – i.e. agreement to freely exchange traffic without any reciprocal payments, has been created.

Most peering agreements are bilateral: companies willing to exchange traffic sign an agreement and set up interconnections. These interconnections can be either dedicated lines connecting one network to the other, or co-allocation of equipment and connection to a LAN in some common facilities (like a NAP, where all the participants have to enter into bilateral peering agreements). Another existing model is a multilateral peering agreement. A good example is the Commercial Internet Exchange (CIX), where ISPs join a trade association by paying a membership fee for the opportunity to freely exchange network traffic through the CIX routers.

The absence of settlements may create incentives for free riders. Small providers may route packets through other providers instead of expanding their own capacity. There is a tendency among larger ISPs to limit peering with smaller networks. This was manifested in the May 1997 decision by UUNet to limit its peering to its "true peers", which caused uproar among smaller ISPs, some of whom were forced to pay connection charges for their links to UUNet. On the other hand, there is a counter-trend by some major ISPs toward universal peering. For example, following UUNet's decision, PSI announced its willingness to peer with any ISP, regardless of their size or reach.

Congestion on the Internet

Driven by technological developments, communication needs, and media attention, Internet usage is exploding exponentially (see [4]). Currently, most users of the Internet access it by using telephone dial-up connections. These connections are low bandwidth – either 33.6 Kb per second, or up to 56 Kb per second, if the newest modems are used. This low bandwidth naturally limits the amount of Internet traffic a user can generate. However, much higher bandwidth alternatives for Internet access are already being deployed. The so-called cable modems offered by cable companies in association with startups like @HOME provide one of the most compelling options. They provide very high bandwidth of up to 10 Mb per second while charging only twice as much as a dialup connection. Most telephone companies contemplate another high bandwidth alternative – the digital subscriber line (DSL) for the deployment in the near future.

At the same time, the nature and mix of Internet applications are rapidly changing. There is a formidable growth in demand for real-time video and audio over the Internet (desktop video teleconferencing, video on-demand, virtual marketing). These new real-time multimedia Internet applications consume substantially larger bandwidth, and are capable of generating very large amount of sustained Internet traffic over long time intervals. The expansion of Internet infrastructure is unlikely to keep up with the explosive growth in demand. If a majority of Internet users switch over to the high bandwidth alternatives for Internet access, the amount of traffic over the Internet may rise to the point where frequent and major Internet "brownouts" are experienced.

The Internet packet delivery service may become unavailable for two major reasons: the inability to connect to the Internet and the inability to transmit packets over the Internet. The former is due to the shortage of Internet entry points caused by inadequate dial-up connections. The latter is usually attributable to bottlenecks in the Internet infrastructure caused by packet congestion.

At times of congestion, blocked packets may be rerouted, delayed or lost. When Internet traffic congestion occurs, end users experience significant slowdowns in establishing connections with remote computers, in the speed of information flows (the number of bytes transferred per unit time), and in the most extreme cases, the unavailability of certain Internet services. Technically, congestion happens when the connection capacity is not high enough, or the routers are overwhelmed by the number of packets they get (more important). Internet traffic congestion rarely occurs in arteries. Fiber optics telecommunication wires of the major Internet backbones are wide enough to accommodate simultaneous transmission of large number of packets. The routers, when overwhelmed by high volumes of packets, are often the bottlenecks.

Pricing the Internet Delivery Service

Users are presently charged for access to the Internet, but not for usage. Access prices are normally determined by the upper bound on the bandwidth of the line purchased by the user. The higher the purchased capacity is, the higher will be the price. However, quantity

discounts are common. For example, a T-3 line, which is equivalent in bandwidth to 28 T-1 lines, may be priced only 5 to 6 times higher.

Strictly speaking, the present system of Internet access pricing is usage sensitive to a certain degree: priced by the upper bound on the capacity that one can consume (the higher the bound the higher the price), i.e. it is based on the potential rather than actual volume of Internet traffic. There is only one dimension – volume, and it is potential rather than actual. The capacity of the connection also controls to a certain extent the type of usage: one will not use real-time video over a regular telephone line.

It is important to distinguish between the usage of the entry point (e.g., a dial-up modem or an ISP's port) and that of the Internet. There may be time measurement and time related charge for the usage of the telephone line that connects the user to the Internet. However, the Internet packet delivery service itself is not subject at present to usage sensitive pricing. Also, the pricing of the Internet packet delivery service should be clearly distinguished from the pricing of specific services provided by individual Web sites over the Internet. Individual Web sites may opt to charge for services or content by subscription or usage sensitive prices.

Providers of Internet services settle with each other through a system of bilateral and multilateral agreements. Any two providers who estimate the exchange of packets between their networks to be roughly equal in volume are likely to enter "peering

agreements" to reciprocate waivers of charges. Smaller providers may be charged access charges by larger providers.

Advantages and Disadvantages of Current Pricing

From the view point of economic efficiency, the current pricing system for the Internet delivery service works best when the bandwidth of the transmission lines and the capacity of the routers are adequate to meet user demand. Under these circumstances, experimental use of the Internet is encouraged because the fixed charge is sunk upon subscription and the incremental cost of usage (limited to charges, if any, for the time the telephone line is in use) is low. Transaction costs are also low because this pricing system requires neither measurement of flows of Internet traffic (e.g. number and size of packets) nor any complex accounting, billing and administration.

The current pricing system however, is not based directly on the actual usage of Internet resources (i.e. the utilization of the bandwidth of telecommunication lines and the computational resources of the routers). In the absence of usage sensitive pricing, Internet traffic flows are not directly subject to economic rationing. During congestion, queuing of packets at the routers is resolved on a "first-come-first-served" basis. All packets are treated equally despite the fact that they are not of equal value to the end user. The cost of delay to the end customer is significantly higher for some packets relative to others. For example, delays in transmission or loss of packets carrying desktop video teleconferencing session can affect the end user more severely than a comparable problem in file transfer.

Yet, the current pricing system does not differentiate between higher valued and lower valued (to the users) packets of information. In that sense, it is analogous to a mail system where all mail deliveries are handled in exactly the same way, with no provisions for any kind of priority mail. As a result, the current pricing system is likely to be economically inefficient. It does not allow its users to pay different prices for different quality of service, and prevents the providers from realizing higher returns on their investments by matching prices to quality of service.

Additionally, increased utilization of Internet resources may not generate sufficient revenues to provide incentives for adequate and timely investment in the development of the Internet infrastructure. This may lead to wrong incentives and encourage free riding. For example, it has been reported that significant amounts of Internet traffic originating in and destined for Mexico have been travelling over the United States Internet backbones (see [20]). In the 30 minute sample of Internet traffic taken on June 21, 1995 at FIX-West (Federal agency intereXchange point, west coast) about "... 4% of the traffic was both sourced and destined from a foreign country" (see [15]). It may be cost-effective under the current pricing system for smaller ISPs to over-sell Internet access without investing in creating a network infrastructure.

It has to be remarked, however, that in a new market like the Internet it may be economically optimal to create over-capacity of the infrastructure in the anticipation of market growth and new customers. Then this access capacity may justify flat fee access pricing since the abundance of resources reduces the marginal cost of using a resource to

close to zero. This may however change over time as the system starts approaching the steady state solution, although rapid technological advances in the area of computers and telecommunications may result in the steady state solution being an option of a very remote future.

Alternative Pricing Systems

We differentiate between pricing systems that are *adaptive* to changing demand-supply conditions on the Internet, and pricing systems that are not (called *non-adaptive*). Adaptive systems have *built-in* mechanisms to change per packet (or per byte) prices in response to the fluctuating packet queues at the routers.

Adaptive systems follow two major approaches to implementing adaptation: the *smart market* approach (see [11,12]) and the *welfare maximization* approach (see [9]). The former is a decentralized system, based on auctions at the routers where packets bid for passage. The latter is a system that centrally sets and periodically adjusts prices to maximize welfare across all Internet computers and routers.

Under the smart market system, packet headers carry user-assigned bids. The bids contain the maximum dollar amount that the user is willing to pay for that packet to be processed by the router. If upon its arrival to a router, the packet does not encounter a queue, it can be either processed free of charge or be subject to a predetermined minimum charge. In case of a queue, the auction takes place in the following manner. All packets are ranked by

their bids, and the cutoff point is determined by the processing capacity of the router. The packets whose bids are above the cutoff amount are processed and are each charged the same amount equal to the bid of the highest ranked packet below the cutoff amount. Hence, the charges for the processed packets reflect the congestion or queuing cost, and are always below the packets' maximum bids. The packets that are not processed may remain in the buffer and be processed later, or may be discarded if the router is overwhelmed by traffic.

Under the welfare maximization approach, prices are computed by maximizing the total economic benefits to all Internet users. Users are presented with prices for packet delivery services with alternative priorities to choose from. Prices are recomputed periodically, based on estimates of production and cost functions for all Internet computers and routers, and estimates of packet arrival rates and waiting times for all Internet computers and routers.

Quasi-adaptive pricing systems make prices depend on the *expected* congestion of the network. The most common example of a quasi-adaptive pricing system is Peak Load Pricing, which, in the case of the Internet, makes prices vary by the time of day in accordance with the expected congestion of the network.

Non-adaptive systems include Post Office Priority pricing and Paris Metro Pricing. Under Post Office Priority (POP) pricing system, an order of service classes is defined (e.g. regular mail, first class mail, express mail, etc.), with the condition that a higher order

class (e.g. express mail) is always served before (i.e. preempts) a lower order class (e.g. first class mail). As a result a higher order class renders higher value to the customer, and is therefore priced higher. This is the type of priority pricing which is the simplest conceptually. It is based on priority queues and first come first served principle applied to packets with the same priority level. The pricing can be either usage-based – prices are per packet with higher prices for higher priority packets, or flat fee – prices are for the maximum access bandwidth in each priority class (with higher prices for higher priority). Although the pure breed priority systems are non-adaptive (i.e. the prices of priority classes do not adjust to changing Internet conditions), the users of priority systems will exhibit adaptive behavior: higher priority packets will be sent only in the case of congestion or significant queuing delays of lower priority packets.

Under Paris Metro Pricing (PMP) system (see 17), several service classes are defined without ordering. These classes are differentiated only by the price, e.g. the price of the first class metro ticket is higher than the price of the second class ticket, while the cars are exactly the same. Thus, the quality of service is affected only by the price differential (i.e. the congestion in second class is *expected* to be higher than in the first class, because of the lower price of the second class).

The Paris Metro Pricing system produces better economic incentives than the current system, but not as strong as POP incentives. The best effort of the latter is likely to be at a higher level, because under POP a packet with higher priority is always routed before packets of lower priority, while this is not always the case for PMP. By being non-

adaptive, both POP and PMP do not provide direct incentives to reduce low value usage at times of congestion.

In addition to adaptability, another important dimension of a pricing system is the loci of the charges:

- end (i.e. entry/exit) points only;
- edge (i.e. borders between ISPs) points only;
- every router.

Table 1: Loci \ Adaptability Matrix of Pricing Systems

Loci \ Adaptability	(1) Adaptive	(2) Quasi-adaptive	(3) Non-adaptive
(1) All routers	Smart market, welfare max	Peak load	Post office priority, PMP
(2) Border routers	Smart market, welfare max	Peak load	Post office priority, PMP
(3) Entry/exit routers (usage-based)	Smart market	Peak load	Post office priority, PMP
(4) Entry/exit routers (flat fee)	N/A	Peak load	Current, post office priority, PMP

If transaction costs are not taken into account, then the economic efficiency of the pricing cells decreases from the top down and from the left to the right. On the other hand, the transaction costs are clearly increasing from the top down and from the left to the right

(i.e. in the opposite direction to the economic efficiency). This implies the necessity of finding a tradeoff between the economic efficiency of the pricing system and the amount of the transaction costs.

It is probably not very meaningful to rank the pricing systems presented in Table 1 by the complexity of their implementation, since those systems are pure breed, while real life implementations will most probably be hybrid systems. For example, it is difficult to imagine an implementation of peak load pricing, which is not combined with either POP or PMP. As a general guideline, it is arguable that non-adaptive systems are significantly less complex to implement than adaptive ones.

Except for the fourth row of the matrix, all the other pricing schemes will bill the user even for those packets that were lost on their way. This implies that under those schemes the end user will bear the financial uncertainty of the actual IP transport cost.

If the charges are imposed at the entry/exit points only, absence of settlements between different ISPs is implied, since otherwise some sort of border pricing would be required. The absence of settlements may attenuate economic incentives of the pricing mechanism. The pricing systems in (3,3) and (4,3) require routing to be dependent on the class of service. In the absence of settlements, an ISP need not honor the priority level of a transient packet. Therefore, the pricing systems in (3,3) and (4,3) may lack the necessary economic incentives.

Technological Requirements and Problems of Pricing Mechanisms

Economic analysis of Internet pricing usually assumes technological feasibility and insignificant implementation costs. However, in view of the technological complexities characterizing the Internet environment, technological feasibility and implementation costs are likely to be dominant factors in determining the choice and viability of alternative Internet pricing mechanism.

Efficient pricing systems should be based on measuring actual (rather than potential) Internet traffic volume. This can be done by counting the number of packets and/or bytes generated by the user at the entry point and by counting the balance of trade at each router connecting networks of different Internet service providers. The flow balance counting can then be used for settlements between the Internet service providers. Presumably, usage sensitive pricing should generate more revenues per unit capacity. When applied to TCP/IP networks, this approach has a number of inherent problems:

- Too many packets to count.
- It is impossible to track all the routes packets follow (combinatorial explosion).
- Since the network is best effort delivery only, some packets will not get through (will be discarded by routers or expire), and they have to be retransmittted. It is therefore arguable that users should not be billed for lost packets. However, it may not be feasible to separate them out from the other packets and relate them to individual customers. It is not clear who has to pay for those lost packets.

- Due to different constraints on the Maximum Transmission Unit, some packets may have to be split into smaller ones somewhere in the middle of their route, and travel separately afterwards (*fragmentation*). How will such fragmentation affect the balance of trade accounting? For example, what if one packet enters network A and is fragmented into 3 packets? In this case network A may be shortchanged by being compensated for handling one packet only, while it processed 3, and was charged for 3 packets at the point of exit.
- Given the client/server architecture of Internet applications, packets are frequently sent over the Internet by a server in response to a request from a client. Arguably, in most cases, the client should be charged for these packets. This may be viewed as an analogue to a collect call. Who should keep track of these transactions and how? Realistically, this burden cannot be shifted to the end users. Most of them will be unable to administer their own billing system so as to recover the cost of sending packets.
- Any kind of usage sensitive pricing that attributes the cost of delivery to the sender of the packet will make it prohibitively costly for many, if not most, providers of information on the Internet either to absorb the cost of the recipient requested packets or to bill numerous recipients. Alternatively, billing recipients as well as senders is also impractical because the former may be then charged for unwanted packets (e.g. junk e-mail).
- Internet multicasting used by most real-time group communication services is based on replicating packets along the way over the Internet as needed. This introduces an imbalance of trade analogous to the effects of fragmentation.

Smart market pricing has not been fully worked out for successive auctions. This poses a serious problem. To be deployed on the Internet, smart market specification requires a strategy of allocating bids to successive auctions. The latter is complicated by the possibility that packets, which paid to get through earlier auctions in the sequence, may be lost before reaching their destinations. As a result, the economic efficiency of smart market pricing cannot be fully evaluated.

The welfare maximization pricing system may not be feasible computationally, unless it is appropriately decomposed into small subsystems that can collect information and compute welfare maximizing prices independently. However, decomposition may require restrictive assumptions about the properties of the system (e.g. the form of the utility functions).

Accounting and billing for all pricing systems of the first row of Table 1, pose difficulties, which may be insurmountable. In the first place, every router must keep accounting in such ways that end (i.e. paying) customers are identified for each packet. This requirement can be avoided only when packets pay electronic cash, which will require enormous changes in the Internet protocol, may present serious financial risks, and will undoubtedly exacerbate the problem of wasted payments in the case of packet loss. Second, problems resulting from erroneous charges plague all first row pricing schemes, with the end user bearing the financial responsibility for routing mistakes, packet fragmentation and packet loss which are due to failures of the Internet infrastructure.

The accounting and billing problems inherent in the first row pricing systems of Table 1 are alleviated in the second row pricing systems by the decomposition of charges for packet delivery into pairwise balance of traffic settlements between ISPs, with end users being billed only by entry/exit ISPs.

The pricing systems in the third column of Table 1 are the simplest to implement. For a single ISP, it may be feasible to implement and deploy both POP and PMP, as costs seem to be manageable and counting, being required only at the edge of the network, is relatively easy. However, it is not clear how to deal with packet loss and retransmission (without sampling). Sampling may be a feasible approach to this problem. With many providers, additional complications associated with settlements arise: counting becomes much more difficult because of large numbers of packets/bytes, and it is not clear how to handle packet fragmentation. Consequently, even if POP and PMP are feasible, the cost/benefit balance is uncertain.

New Technological Developments

Quality of service (QOS) guarantee over the Internet and Service Level Agreements (SLA) are currently among the most active areas of technological development on the Internet. A new Internet protocol called Resource ReSerVation Protocol (RSVP) has been proposed. It will make it possible for a certain amount of bandwidth to be reserved at each router along the path from the sender to the receiver at a request of an application program. RSVP in effect establishes a virtual temporary private line between the end

points, thus providing the same connection bandwidth guarantee as the telephone network. This is currently considered as the most promising way to guarantee acceptable performance of real-time multimedia applications like desktop teleconferencing over the Internet. If RSVP is supported and deployed by all major ISPs, it will make it possible to buy a virtual circuit over the public Internet instead of buying (or leasing) dedicated private lines.

Only a few major ISPs currently offer RSVP within their networks. No inter-ISP agreements to settle charges for this service have been reported. Clearly, for this protocol to be deployed widely over the Internet, settlement payments will be required. In the near future RSVP will be used only within parts of the Internet belonging to the same ISP, since this avoids settlement problems, and makes it feasible to guarantee that all routers support it. The pricing of RSVP may be complicated by the fact that if bandwidth is allocated but not used (e.g. no packets are transferred when both parties are silent), the router may use this unused bandwidth for routing other packets. On the other hand, if bandwidth is reserved, and the router bandwidth is close to being used up, other bandwidth reservation requests may have to be turned down. This, in combination with the fact that maintaining bandwidth reservation requires some processing on the router, shows that a reservation price should be charged. This price however may depend on how congested a router is, because other reservation requests are very unlikely to be turned down if the router is not congested.

Another important technological development, which can significantly speed-up access to Web materials, is caching of Internet Web sites by ISPs (see [16]). To the extent that the issues of intellectual property ownership are resolved, this technique can also be encouraged by the appropriate pricing system. If settlements among ISPs existed, the revenues generated by Internet traffic would have to be shared among all the ISPs along the path from the browser to the server over the Internet. When a cached version is used, this path is limited to one ISP that will receive all the revenues. Currently, caching does make a contribution to the quality of services provided by an ISP, and may generate additional customers and revenues indirectly. In the future, however, caching may be regarded as inconsistent with basic notions of copyright protection. Even for public Web sites, caching results in some visits to the site that are not recorded in the log of the site. This may result in potentially lower advertisement revenues. It is conceivable that in the future, caching agreements may require log exchanges. In this case, caching could generate incremental advertising revenues even when the caching sites do not charge subscription fees.

References

1. H.W. Braun, K.C. Claffy, and G.C. Polyzos, "A Framework for Flow-Based Accounting on the Internet", *Proc. of the Singapore International Conference on Networks (SICON'93)*, pp. 847-851, September 1993.
2. N. Brownlee, "New Zealand Experience with Network Traffic Charging", *ConneXions*, The Interop Company, Vol. 8, No. 11, November 1994.

3. N. Brownlee, "Accounting Requirements for IPng", *RFC 1672*, Internet Engineering Task Force, Network Working Group, August 1994.
4. K. C. Claffy, H.W. Braun, and G.C. Polyzos, "Tracking Long-Term Growth of the NFSNET", *Communications of the ACM*, Vol. 37, No. 8, pp. 34-45, August 1994.
5. D.E. Comer, *Internetworking with TCP/IP - Vol. I: Principles, Protocols, and Architecture*, 3rd Ed., Prentice-Hall, Englewood Cliffs, NJ, 1995.
6. R.J. Edell, N. McKeown, and P.P. Varaiya, "Billing users and pricing for TCP", *IEEE Journal on Selected Areas in Communications*, **13** (1995), pp. 1162-1175. (Available at http://tiny-tera.stanford.edu/~nickm/papers/IEEE_JSAC_95_Billing.ps)
7. D. Estrin and L. Zhang, "Design Considerations for Usage Accounting and Feedback in Internetworks", *ACM Computer Communication Review*, Vol. 20, No. 5, pp. 56-66, October 1990.
8. R, Gareiss, "Is the Internet in trouble?", *Data Communications*, September 21, 1997. http://www.data.com/roundups/trouble.html
9. A. Gupta, D.O. Stahl, and A.B. Whinston, "Priority pricing of integrated services networks", *Internet Economics* (L.W. McKnight and J.P. Bailey, eds.), MIT Press, 1997, pp. 323-352. (Preliminary version is available at http://www.press.umich.edu/jep/works/GuptaPrior.html)
10. F.P. Kelly, "Charging and accounting for bursty connections", *Internet Economics* (L.W. McKnight and J.P. Bailey, eds.), MIT Press, 1997, pp. 253-278. (Preliminary version is available at http://www.statslab.cam.ac.uk/~frank/charge.ps)
11. J. Mackie-Mason and H. Varian, "Pricing the Internet", In: *Public Access to the Internet* (B. Kahin and J. Keller, eds.), MIT Press, 1995, pp. 269-314.

12. J. Mackie-Mason and H. Varian, "Economic FAQs About the Internet", *Journal of Economic Perspectives*, Vol. 8, No. 3, pp. 75-96, Summer 1994.

13. C. Mills, D. Hirsch, and G. Ruth, "Internet Accounting: Background", *RFC 1272*, Internet Engineering Task Force, Network Working Group, November 1991.

14. T. Monk and K. Claffy, "Internet Data Acquisition & Analysis: Status & Next Steps", http://www.nlanr.net/Papers/data-inet97.html

15. National Laboratory for Applied Network Research (NLANR), "Does the U.S. export anything anymore?", http://www.nlanr.net/NA/Learn/export.html

16. National Laboratory for Applied Network Research (NLANR), "A Distributed Testbed for National Information Provisioning", http://ircache.nlanr.net/Cache/

17. A. M. Odlyzko, "A modest proposal for preventing Internet congestion", http://www.research.att.com/~amo/doc/modest.proposal.ps

18. J. Rickard, "Fall 1997 Directory of Internet Service Providers", *Boardwatch Magazine*. http://www.boardwatch.com/isp/

19. S. Shenker, D. Clark, D. Estrin, and S. Hertzog, "Pricing in computer networks: Reshaping the research agenda", *Telecommunication Policy*, **20** (1) (1996), pp. 183-201. (Available at ftp://parcftp.xerox.com/pub/net-research/picn.ps)

20. P. Srinagesh, "Internet cost structures and interconnection agreements", ", *Internet Economics* (L.W. McKnight and J.P. Bailey, eds.), MIT Press, 1997, pp. 121-154. (Preliminary version is available at http://www.press.umich.edu:80/jep/works/SrinCostSt.html)

21. J. Sweat, "Internet demand is moving faster than technology, panel says", *Information Week*, March 16, 1998. (http://www.techweb.com/wire/story/TWB19980316S0017)

22. WEBWEEK, ISP Spending Boom Predicted, July 28, 1997,

http://www.webweek.com/current/infrastructure/19970728-system-briefs.html

23. Investor's Business Daily, "Phone Companies Gobble Up ISPs", Nov. 26, 1997.

X-Sender: miklosv@andromeda.rutgers.edu
Date: Tue, 16 Jun 1998 16:47:21 -0400
To: k._casey_bennett@smtp.aicpa.org
From: "Miklos A. Vasarhelyi" <miklosv@andromeda.rutgers.edu>
Subject: system reliability project / writeup
Cc: rzeibig@kpmg.com, <lwilliamson@kpmg.com>, miklosv@andromeda.rutgers.edu

Hi Casey;

following up our conversations relative to the write-ups for the systems reliability committee I suggest the following:

that I take responsibility for leading the write-ups of three separate documents:

1. main committee document proposing the product(s) using materials and ideas discussed in the committee. (principles and criteria)

2. practice manual: a one day course based on the above document to serve as the initial training tool (about 100 slides, a couple of short illustrative cases, and guidance for instructors)

3. resource guide - work papers and other documents from actual engagements -- contingent in help from committee members in obtaining and preparing these materials

being a committee member i do not feel that i should charge for my work. i have asked Cathy Dodson (a phd student) and Peter Gillett on our faculty to help with parts of this document and would pay for their effort. I would suggest a budget of $20,000 that includes a 15% overhead charge for the Rutgers Accounting Research Center. we would keep you and Bob appraised of our progress as it is very difficult to estimate costs on projects of this nature.

Miklos A. Vasarhelyi
KPMG Peat Marwick Professor of Accounting Information Systems
315 Ackerson Hall, 180 University Avenue
Rutgers University, FOM
Newark, NJ 07102
973-353 5002 fax 973-3531283