Differentiated Services and Pricing of The Internet

by

Atanu Mukherjee

Submitted to the System Design and Management Program
in partial fulfillment of the requirements for the degree of

Master of Science in Engineering and Management

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1998
[June 1998]

Signature of the Author..
System Design and Management Program
May 20th, 1998

Certified by..
Dr. Michael S. Scott Morton
Jay W. Forrester Professor of Management, Sloan School of Management
Thesis Supervisor

Accepted by..
Dr. Thomas A. Kochan
George M. Bunker Professor of Management, Sloan School of Management
Co-Director, System Design and Management Program

Differentiated Services and Pricing of the Internet

by

Atanu Mukherjee

Submitted to the System Design and Management Program
on May 20th , 1998, in partial fulfillment of the
requirements for the degree of
Master of Science in Engineering and Management

Abstract

The convergence of communication mediums and the diffusion of the Internet into the society is blurring the distinction between voice, video and data. The ability to use a single communication infrastructure is very desirable as it brings down the cost of communication while increasing the flexibility of use. However, voice, video and data have different latency and bandwidth requirements. The varying characteristics leads to inconsistent delivery of the traffic streams in the presence of congestion in the network. In a best effort network, the quality of service is unpredictable if the network is overloaded.

This thesis analyzes the market and bandwidth characteristics of the Internet and develops a service differentiation model, which provides preferential treatment to traffic streams using statistical drop preferences at the edges. The drop preference scheme provides assured service levels by classifying traffic streams into classes of bandwidth demanded by the users. It then analyzes the effect of pricing schemes in efficient distribution of differentiated services and outlines a model for pricing the differentiated services based on auctions.

Thesis Supervisor : Dr. Michael S. Scott Morton
Title : Jay W. Forrester Professor of Management, Sloan School of Management

1 Introduction

There are few technological and commercial success stories as dramatic as that of the Internet. In 1997 alone the number of users connected to the Internet was estimated to be 50 million and that number is slated to reach 150 million by the year 2000. Originally, designed to link a small community of researchers, the Internet has grown into a social and commercial institution having a significant impact on business and society.

The Internet is fundamentally different from traditional telephone and cable networks, in that it does not provide a guaranteed capacity or pre-allocated circuit. The current Internet offers a single class best effort service. The Internet offers no assurance about when, or even if, the messages will be delivered. The applications above the Internet are responsible for reliability of the transmission of messages. This mode of packet switched transmission, offers remarkable efficiencies in transmission over traditional networks. The ability to transmit messages at any time without establishing a connection, the high degree of "statistical gain" and the availability of standards based applications has led to the explosion in the adoption of the Internet as the standard medium of communication.

In the meantime the convergence of communication mediums is blurring the distinction between voice, video and data. It is conceivable that in the very near future, the Internet will provide the capability to transmit voice, video, data and other enhanced multimedia applications. The rise of IP telephony, real-time multicasting and integrated applications are all pointers in that direction. This is a very desirable characteristic as it brings down the cost of transmission[1] for the end user.

---

[1] Fundamental Design Issues of the Future Internet: Appendix A, Shenker, Scott; ARPA research report under contract DABT63-94-C-0073

This report examines the factors related to pricing and differentiated service, which will allow the transition of separate communication networks to a single all digital Internet. Specifically, it analyzes the following elements necessary for the successful transition of communications to the new Internet:

- The Quality of Service and service guarantee enhancements for providing integrated communication services on the Internet
- The emerging structure and the competitive dynamics in the Internet Service Provider (ISP) marketplace.
- Mechanisms for pricing the Internet connectivity by ISPs based on classes of service and statistical guarantees on the quality of service.

This report validates the hypothesis that the Internet needs to have mechanisms for service differentiation. However, it argues that absolute guarantee levels are not only unnecessary and inefficient but may well be impossible to implement given the current status of the Internet. It further analyzes the ISP market and validates the hypothesis that there will be a consolidation of ISPs if the ISPs are unable to price differentiate their service based on service quality levels. Finally, the report suggests a model of pricing which allows the ISPs to provision their networks most efficiently, while offering differentiated service to its users.

2 The Internet and Connectivity

The Internet is a network of networks that interconnect by using a common addressing scheme and a standard suite of protocols. These protocol suites are layered and the layered abstraction of the Internet creates the scalability and the robustness of the Internet infrastructure. This chapter describes the key constituents of the Internet technology and how a user gets an end to end service on the Internet. It begins with an overview of the technology and then analyzes the constraints in providing a guaranteed end to end service based on the current, best effort delivery model.

2.1 Overview

The architecture of the Internet is based on the concept of successive layering of functionality. The layering of functionality in the protocol suites allows the change and/or addition of additional modules within a layer without affecting the adjacent layers. This powerful abstraction has created the proliferation of thousands of applications and has made it possible for the Internet suite of protocols to run over physical mediums ranging from the Plain Old Telephone System (POTS) to high speed fiber at multi-terabits per second.

The suite of protocols which make this possible is the *Transmission Control Protocol (TCP)* and the *Internet Protocol (IP)*[2]

Broadly speaking the layered architecture of the Internet has three generic layers. The bottom layer is the underlying physical layer, also known as the *network substrate.* The network substrate can exist over a variety of physical network mediums ranging from a few kilobits per second to terabits per second. These mediums could be xDSL over copper, Asynchronous Transfer Mode

(ATM) running between 155 Mbps to a few Gbps, Synchronous Optical Network (SONET) on fiber at a few Gbps or a high speed satellite connection.

The middle layer is the IP layer responsible for connectionless delivery and routing of packets. This is analogous to the *bearer service* of the post offices. This layer is responsible for unreliable delivery of message through the network by routing.

The layer above the bearer service layer is the TCP layer which is responsible for *reliable delivery* of data. Simultaneously, a framing layer known as the User Datagram Protocol (UDP) coexists with the TCP layer to support high speed streaming applications.

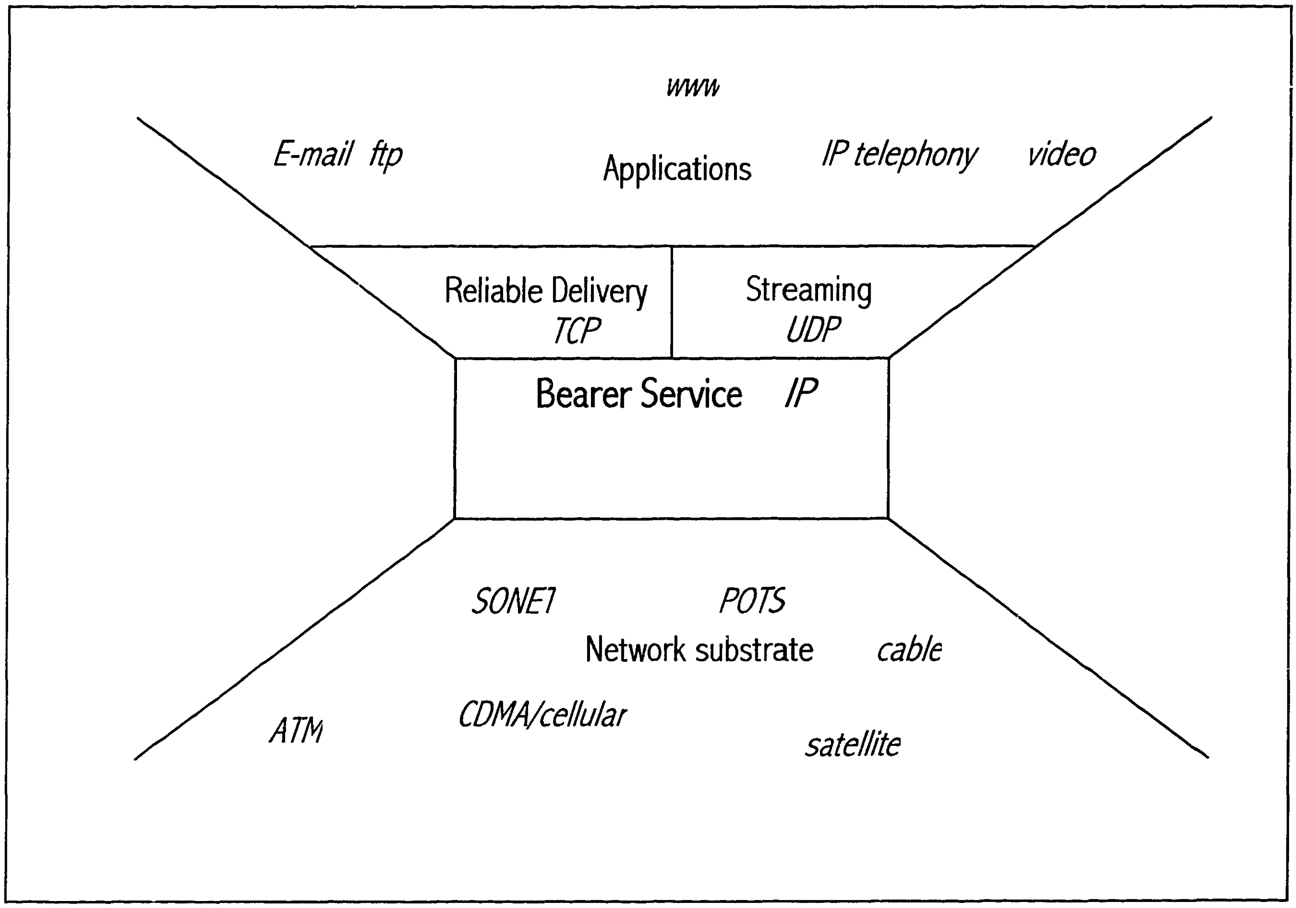
The applications use the TCP/IP suite of protocols to communicate between end nodes. A wide variety of applications exist today which harness the TCP/IP suite of protocols. These applications range from the relatively elastic applications like e-mail and file transfer protocols to the emerging applications like IP telephony and streaming video over the Internet. The layered architecture and the standard interface between layers makes it possible for the Internet to provide a variety of services over an equally large variety of underlying physical transport systems.

As discussed earlier, the layered architecture of the Internet Protocols lays the framework for providing a diverse set of services ranging from real-time video and audio to more traditional data oriented services like the web traffic and e-mail. Figure 2.1 illustrates the layered architecture of the Internet.

---

[2] See RFC 791, RFC 793 Internet Engineering Task Force

Figure 2.1 Layered Internet Architecture



2.2 The Current Internet End-to-End Architecture

The end-to-end service provided by the Internet today is based on best effort delivery. The packets traverse through various routes before they are delivered to the end receiver. This is achieved through a set of coordinating *access points*, *switches* and *routers* over a set of interconnected networks. Typically, the traversal of an end-to-end path is based on hierarchical routing. Typically, the networks are organized in three in the hierarchy.

2.2.1 The Customer Premise Network/ Local Access Point

The end user is located is either a consumer or a business user is located in the local customer premise network or the local ISP network. Each end user's computer has a unique IP address. This unique IP address may be permanently assigned or may be dynamically managed[3]. ISPs typically manage a block of floating addresses and dynamically assign these addresses to end users as and when they request an access service. This optimizes the number of IP addresses an ISP needs. The corporate networks on the other hand usually run on high speed Ethernet and FDDI (Fiber Distributed Data Interface) based local backbone networks operating between 10 Megabits to 100 Megabits per second. Current, enhancements in Gigabit Switching has potential of extending the LAN capacities to about a gigabit per second[4] by the year 2000. There are other options available as well. ATM can be used as the local backbone switching technology, providing speeds of upto 622 Mbps in the customer premises[5]Thus, a business customers local network has a sufficiently high bandwidth connection for it's end users within the customer premise.

---

[3] See RFC for DHCP, RFC 2131, Internet Engineering Task Force

[4] Cisco Systems, Gigabit Switching; Also conversations with Andreas Bechtolsheim, Vice President of Gigabit Switching, Cisco Systems.

[5] Fore Systems, Forerunner ATM switches http://www.fore.com

The end user's of the ISPs however are connected to the ISP access point at various speeds depending upon the available "last mile" bandwidth. The last mile bandwidth varies depending upon the physical medium. In POTS connections the best one could achieve is no more than 40 Kilobits per second using an analog modem[6]. The digital ISDN connections over the telephone lines can achieve a maximum of 144 Kbps using digital signaling, while the Asymmetric Digital Subscriber Lines (ADSL) can achieve upto 1.5 Mbps with current xDSL based technologies[7]. Using cable modems and existing cable technology is another option for providing higher speed connections by leveraging the existing cable infrastructure. By using two 6Mhz channels, one for upstream and one for downstream, one could achieve a last mile bandwidth of anywhere between 4 Mbps to 10 Mbps[8] using Ethernet over cable. However, this is the theoretical maximum limit. The asymmetric nature of the cable Ethernet infrastructure creates congestion problems as subscribers increase over 500. A solution to this problem is to segment the cable infrastructure and connect the head-ends of the segments by fiber over ATM or SONET. This has deployment issues in terms of sunk costs and time to implement the backend cable infrastructure with fiber using the Hybrid Fiber Coaxial (HFC) systems.

Alternatively, the end-user can connect to the ISP through a dedicated T-1 connection at 1.5 Mbps, but the cost of such a connection is prohibitive for a single user. Direct T-1 connections can thus be best used by small offices/ businesses.

---

[6] The telephone channel has 4Khz bandwidth. Maximum bit rate is given by Shannon's theorem as being equal to Bandwidth * $\log_2$(Signal to Noise Ratio). In POTS lines typical SNR is about 1000, thus the maximum achievable bit rate is limited to about 40 Kilobits per second.

[7] TR-001 ADSL Forum System Reference Model, http://www.adsl.com/tech_info.html

[8] Cable Data Modems, http://www.cablelabs.com/Cablemodem.pdf

The following table lists the characteristics of various Customer Premise Network/Access Points

Table 2.2.1

Technology	Access Speed	User Type	Cost/User/Month	Issues
Ethernet/FDDI	10Mbps-200Mbps	Corporate		Preferred Mode
Gigabit Ethernet	1 Gbps	Corporate		Unproven Technology
POTS/Analog	40 Kbps	Consumer	$20	Last Mile Bottleneck
ISDN	144 Kbps	Consumer	$100	Low Price Performance
T-1	1.544 Mbps	Small Business/ISP	$1,500 - $2,500	Expensive
Ethernet on Cable	4Mbps – 10 Mbps	Consumer/Small Business	Starting $40	Congestion
xDSL/ADSL	500Kbps – 1.5 Mbps	Consumer/Small Business	$750	Price/Performance, cannibalizes T-1

The analysis of the Customer Premise/Access Point network shows that the possible bottleneck for the end users is the last mile. Last mile bandwidth availability is unlikely to be an issue for corporations with well provisioned corporate networks, but consumers and businesses connecting to the Internet through an ISP is likely to have significant last mile bottlenecks for providing integrated services. Although, Cable based Internet connections has the capability to provide sufficient last mile bandwidth, it is still unclear as to how the effect of congestion and hence degraded service quality can be mitigated in a population of large number of subscribers.

The customer premise network or an ISP is connected to the Internet through either a regional network or directly to the Internet backbone. The traffic is handed off through sets of high speed routers from the Customer Premise Network/ISP to the regional network using T-3/ATM/SONET circuits.

2.2.2 Regional Networks

Most of the Internet's networks are Customer Premise or ISP networks connecting users within a company or a community. The ISP networks and Customer Premise networks are in turn connected to "regional networks which serve a region. NYSERnet (New York State Education and Research Network) connects campuses and industrial customers in the New York State. Similarly, BARRnet connects the customer and ISP networks in northern California. Regional networks provide the connection between campus networks and the national backbone carriers that carries traffic to other regions.

The regional networks, provide a variety of hand-off speeds from the local ISP networks or the customer premise networks. Figure 2.2.2 illustrates the typical handoff characteristics of a regional network provider[9].

Table 2.2.2

Technology	Bandwidth	Price/month	Comments
Tiered T-3 Dedicated	4Mbps – 45 Mbps	$7,500 - $41,500	High Cost/ High Volume Required
Metered T-3 Dedicated	4 Mbps- 45 Mbps in 1 Mb increments	$13,500 - $41,500	Available in 2/3 Mbps increments
Frame Relay Dedicated	56 Kbps- 128 Kbps	$ 600 - $ 1200	Low Bandwidth, but dedicated information rate
T-1 Dedicated	1.544 Mbps	$ 2,100	Better Price/Performance

Depending on the traffic characteristics the ISPs and the Customer Premise networks can provision for the appropriate capacity to connect to the regional networks. This is a very difficult provisioning problem. The connectionless nature of the Internet makes the traffic on the network bursty. This can result in a very wide range of bandwidth demands ranging from a few kilobits per sec to a few megabits per second. In the event of multiple high bandwidth transfers through an ISP, the ISP will be overwhelmed with the peak traffic. The unpredictable nature of the bandwidth demand makes it extremely difficult to manage the bandwidth requirement. This means that the ISPs would need to provision for peak capacity, which may be several times the average demand. This would result in a high operation costs. The need for service differentiation through pricing mechanisms is the only way that an ISP can provide differentiated service levels to different consumer classes while provisioning the network adequately.

The regional networks are currently provisioning their networks with high bandwidth fiber connections. Typically, the regional provider's backbone

[9] http://www.cerfnet.net

operates at OC-3/OC-12 (155Mbps/622 Mbps) speeds. Further, the regional network providers aggregate traffic from several ISPs and Customer Premise networks. This results in lower burstiness and variability of the traffic, helping the regional providers to achieve more predictable provisioning. The likelihood of congestion in the regional network provider is thus lesser than the congestion possibility within an ISP or from an ISP to a regional provider hand-off link. The regional providers hand-off inter-regional traffic through the national network providers.

2.2.3 The National Backbone Providers

The foundation of the Internet in the United States is made up of commercial Internet backbone providers -- MCI, UUNET, Sprint, PSINet and others -- connecting regional, ISP and national backbone networks. Backbone providers interconnect with each other in two ways: 1) by direct interconnect; and 2) through Internet Exchange points (IXs). IXs are data interchanges similar to the interchanges on the U.S. highway system. These points facilitate the exchange of Internet traffic by operating the switching platforms needed to handle the different speeds and connection types coming in from the backbone providers and various Internet Service Providers (ISPs), regional providers and other national backbone providers.

The backbone networks operate at anywhere between OC-12 (622 Mbps) to OC-48 (2.4 Gbps) range. Recent technologies like Dense Wavelength Division Multiplexing (DWDM) allow us to increase the backbone bandwidth further in the range of 1700 Ghz per second over a single strand of fiber[10]. It is only likely that backbone bandwidths will keep increasing over time.

[10] http://www.ciena.com . Using DWDM, Ciena is currently able to offer switching in the range of OC-48 speeds.

The even higher degree of aggregation of traffic in the backbones, creates the need for such high bandwidths. On the other hand traffic in the backbone is rarely bursty[11] .This characteristic makes it easier to make longer term traffic growth forecasts and long term provisioning. This would also mean that the backbone networks will be heavily over-provisioned to amortize the sunk costs of laying the fiber, over a longer time horizon. The longer-term, aggregate nature of traffic growth makes over-provisioning possible. Accordingly, the backbone providers will be more interested in entering into long term contracts for aggregate bandwidth with the regional providers and the ISPs. The pricing strategy in the backbone will hence be designed to reduce churn and transaction costs while providing a near guaranteed service quality through over-provisioning. This seems to be the case. MCI, Sprint, AT&T and others have laid thousands of miles of dark fiber to over-provision the backbone [12]. Similarly, the backbone carriers are entering into long term tiered contracts for the backbone bandwidth provided to ISPs[13].

2.2.4 Network Access Points (NAP)

The Network Access Point (NAP) is an exchange point for Internet traffic. Internet Service Providers (ISPs) connect their networks to the NAP for the purpose of exchanging traffic with other ISPs. This allows regional ISPs to connect to large backbone providers (aka: National Backbone Providers); the Routing Arbiter[14] for address resolution and assistance, and to other regional ISPs for the exchange of Internet traffic. National Backbone Providers connecting at the NAP obtain the benefit of the regional ISPs' customer bases, access to content providers, and redundancy of connection to other backbone providers.

[11] Traffic Characteristics in the MCI Backbone, July 11 1997, by Rick Wilder, End to End Conference, MIT Laboratory for Computer Science.

[12] Telecosm , George Gilder

[13] Tiered Contracts with Backbone Providers, Srinagesh, Draft Proceedings of Workshop on Internet Economics , March 1995

[14] Route Arbiter Project at http://www.merit.net

The exchanging of Internet traffic is generally referred to as "peering".

2.2.4.1 Interconnection

In 1994, the NSF awarded contracts to replace the NSFNet (Internet backbone). These contractswere for backbone transport, the routing arbiter and traffic exchange points (NAPs).

There are four original NSF sponsored NAPs, located in the Bay Area Pacific Bell, Chicago (operated by Ameritech), New York (operated by Sprint), and Washington DC (operated by MFS). Due to the explosive growth of the Internet, additional unsponsored NAPs, commercial Internet exchanges (generally operating at DS1 and lower access rates), and private exchange points have appeared.

2.2.4.2 NAP Peering Agreements

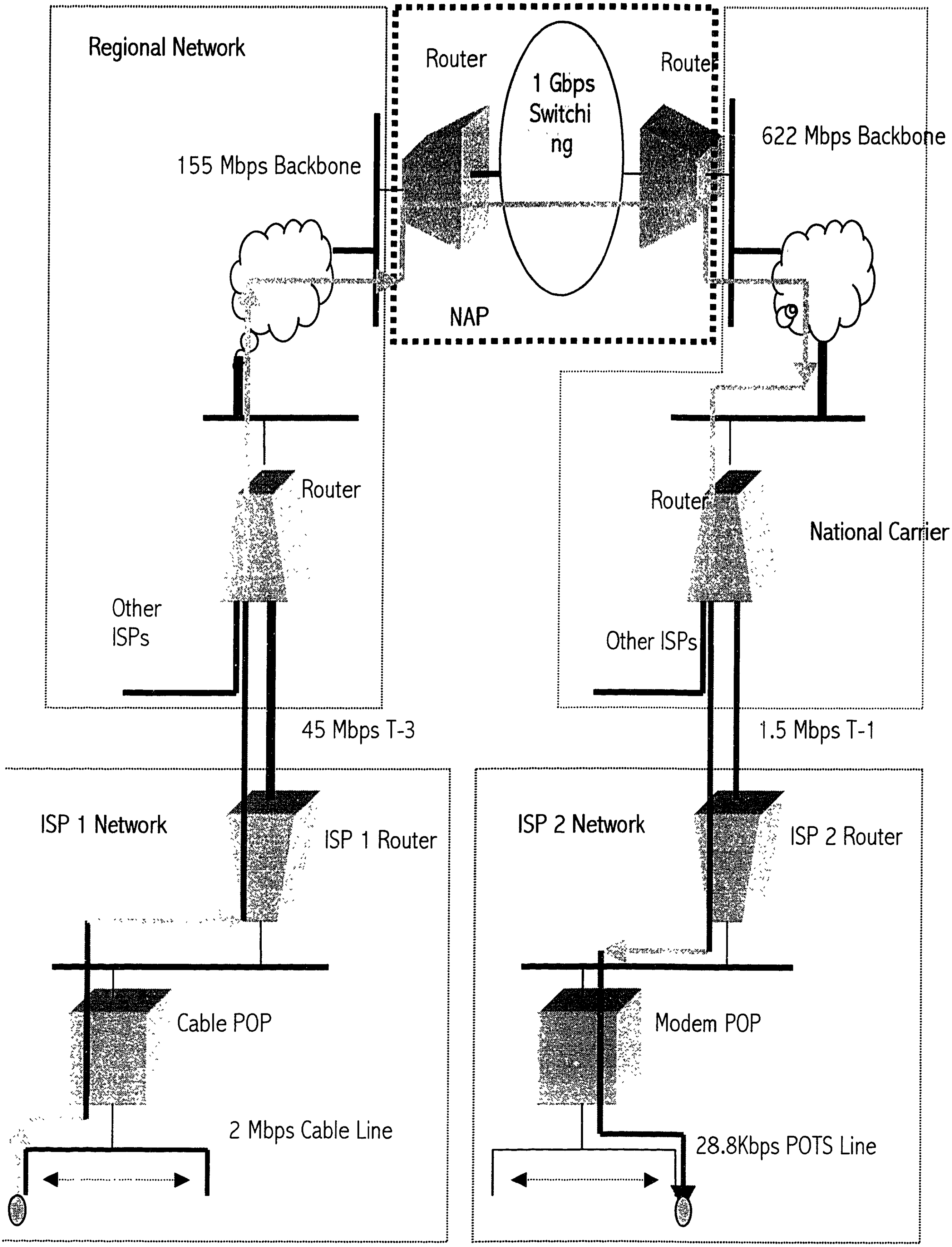
NAP clients may negotiate their own bilateral agreements with other NAP clients for the exchange of TCP/IP routing information and traffic. Alternatively, NAP clients may sign the "MultiLateral Peering Agreement" (MLPA) which established mutually acceptable rules for TCP/IP information and traffic exchange.

In addition, the NSF sponsors a "Routing Arbiter" service at each NAP. Clients should make arrangements with the Routing Arbiter for dissemination of routing information among participants in the service. Upon the mutual consent of customers or the signing of the MultiLateral Peering Agreement, the NAP will establish a circuit between customer sites reflecting the peering agreement, as well as a PVC from the customer site to the Routing Arbiter.

The variation in the nature of the traffic at different points in the Internet, create the need for service quality and pricing differentiation. Given the nature of traffic aggregation and variability, different degrees of congestion occurs at various points. This would mean that the pricing and service differentiation strategies at the edges (ISPs) of the Internet may need to be based on quality of service and variable price, whereas the center (backbone providers) have long term contracts and no service differentiation.

Figure 2.2.4 illustrates the end-to-end transfer between a sender and a receiver. The sender is connected to the ISP 1's point of presence (POP) on a cable network at about 4 Mbps. The colored lines indicate the route taken by the messages as it reaches the sender connected to ISP 2's POP by a lower 28.8 Kbps POTS connection. The red lines indicate the likely bottleneck points and the green lines indicate that congestion is unlikely to occur in any of these areas. However, any redline along the path will hamper the end-to-end quality of service. The first bottleneck could be at the T-3 connection between ISP-1 and the Regional network provider. As multiple bursty 4 Mbps streams are coming into ISP-3 , the 45 Mbps T-3 connection can be overwhelmed at times. The regional network provider's backbone is provisioned at 155 Mbps and the aggregate traffic from multiple ISPs is much more predictable and smooth. So the peak provisioning would be closer to the long term average provisioning of the network for the regional provider, resulting in lesser congestion. The NAP is the peering point which is switching the traffic at 1 Gbps to the National provider and congestion is likely to be lesser here compared to the edges. The national provider is provisioned at 622 Mbps on it's backbone cloud . The likely congestion point is going to be in the 1.5 Mbps T-1 link to ISP-2 and then the 28.8 Kbps POTS lines. This would mean that the pecking order of congestion and quality of service degradation in an end-to-end service will depend on the weakest member in the link. Although, congestion can occur anywhere in the net resulting in service degradation, the likelihood of it occurring at the edges of the Internet is higher compared to center of the Internet. This is also corroborated by the fact that, the edge ISPs frequently under-provision the network in a quest to provide access at a low price. Thus a mechanism of traffic prioritization at the edges along with a variable pricing scheme can relieve the end-to-end congestion.

Figure 2.2.4



Regional Network
Router
1 Gbps Switchi ng
Router
622 Mbps Backbone
155 Mbps Backbone
NAP
Router
Router
National Carrier
Other ISPs
Other ISPs
45 Mbps T-3
1.5 Mbps T-1
ISP 1 Network
ISP 1 Router
ISP 2 Network
ISP 2 Router
Cable POP
Modem POP
2 Mbps Cable Line
28.8Kbps POTS Line

3 Service Quality and Congestion

In the previous chapter we have seen that variations in the end-to-end bandwith due to the structure of the Internet, will create congestion problems resulting in varying qualities of service to different traffic streams. Some traffic streams can tolerate the delays due to the congestion like fax or file transfer, while others like voice traffic have bounds on the extent of delays that they can tolerate. Or in other words, the traditional Internet applications like e-mail etc. are rather elastic in nature, in that they tolerate delays and losses in packets rather gracefully, and are hence well served by the current Internet's best effort service. Moreover because of their elasticity, they can decrease their transmission rate in the presence of overload or congestion. On the other hand the World Wide Web, IP telephony and video conferencing are less elastic and hence less tolerant of delay variations and hence do not perform adequately when running over the current Internet. In addition, if these transmissions are contending for bandwidth with traditional data applications, the data applications end up with very little bandwidth. The Internet thus has various classes of users who have preferences and values attached to different classes of services ranging from video to e-mail. However, the current Internet provides no mechanisms for the end-user to realize these preferences due to the classless nature of the Internet.

One can always increase the efficiency of the current Internet to support different classes of application by increasing the bandwidth. Alternatively, one could increase the total end-to-end utility of the Internet by keeping the bandwidth fixed and delivering a wider variety of services than just a single class best effort service. This enhancement has been proposed in many of the IETF drafts[15] and will allow different applications to get different qualities of service. Thus for

[15] Integrated Services in the Internet Architecture, RFC 1633, Internet Engineering Task Force

example if we have IP telephony service in one class and e-mail in the other class which have very different sensitivities to delay, then offering two different priority classes will likely increase the overall efficacy of the network.

Given the emerging nature of the applications, the increase in fiber capacities in the backbone and regional providers the question comes down to the degree of tradeoff between a) The cost of adding bandwidth and b) the cost of adding extra mechanism needed to create classes of service. The amount of bandwidth needed to offset the benefits of enhancing the service model depends on the nature of the utility functions of the applications.

3.1 Bandwidth Enhancement

Enhancement of bandwidth is the option of heavily over-provisioning the network so that the bandwidth and delay requirements of all applications can be met regardless of their utility functions. If the network traffic flow is aggregate then the average aggregate bandwidth will have reasonable utilization of the network while delivering to all traffic service suitable for all class of applications. This will mean that in these areas of the Internet, adding bandwidth is the primary method of providing the service to different types of applications having varying delay bounds. As a second order enhancement, services can be classified at an aggregate level to provide closer conformance to the service requirements of the applications.

The bandwidth enhancement scheme, is thus suitable for backbone providers like MCI, Sprint, Uunet etc who handle large aggregrate traffic flows. Large aggregate flows, having less bursty traffic is well served by over-provisioning. As a second order enhancement, the backbone networks can implement service classes at aggregate levels of flow between providers.

3.2 Service Class Extensions

On the other hand, a network with bursty traffic flows with minimal aggregation will be better off by implementing service class extensions which classify traffic based on the delay requirements. It is not possible to operate and provision for an "edge" ISP network at reasonable utilizations while delivering to all traffic a service suitable for a real time application like video-on-demand. By taking advantage of the extreme elasticities of the comingled applications like e-mail or bulk file transfer, they will be able to utilize the amount of leftover bandwidth as long as they preferentially allow the real-time application to pass through. In this scenario, the payoffs in terms of bandwidth saved by offering multiple classes of service will more than outweigh the cost of implementing the extensions.

The service class extension scheme is very well aligned in providing services based on classes and incentives at the edge of the Internet with minimal aggregation ie. ISPs. The cost of implementation of class extensions on the edges is easier due to the fewer number of network elements involved.

3.3 Service Classification Mechanism

The service classification can range from a two level preferential drop scheme to a complicated mechanism of multilevel priorities. The service can be either implicitly classified by the network or the application can explicitly choose the service. In the former the network can classify the service based on the application characteristic into voice, video or e-mail. The application characteristic may be derived from some implicit attribute of the application traffic eg. Web traffic transmits on a specific port number. This approach does not require any change in the end application and hence is seamless in it's implementation.

However, this approach has some severe shortcomings. The network based classification scheme assumes a fixed set of service classes given it's knowledge of applications at any point of time. It will be unable to classify other applications which it does not know of. As new applications emerge in the integrated networks, the network cannot service the new applications into service classes. Hence embedding the application characteristic in the network is detrimental to servicing new applications as and when they emerge.

Hence the alternative of applications explicitly requesting the service classes is more appropriate in the context of end-to-end services integrated services on the Internet. In this case the network offers a set of service classes and the applications request the service class requirement. In our case the edge ISP can broadcast the classes of service it can support and an application can pick the one best suited to it's need. Depending upon the ISPs infrastructure, the competitive positioning and the application segment (voice, video etc.) in which the ISP intends to compete, it will provide service classes best aligned to the corresponding segments.

3.4 Pricing and Incentives

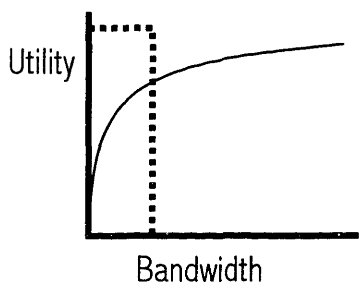
If the applications request service classes from the network or the ISPs , the users who control the should be motivated to ask for lower quality of service . If there are no incentives for asking a lower class of service, the users will always ask for the highest class of service no matter what the application requirements are. The network or the ISP should hence provide some other system of incentives to encourage users to ask for the appropriate class of service. Differential pricing of the service classes is the mechanism which will enforce such a behavior. Users will request for higher quality of service for delay sensitive applications only if this class of service is charged more. Similarly, a lower quality of service will be compensated by a lower price. Pricing also allows the ISPs to differentiate their service offerings and position their service offerings into different segments of the market. Thus depending upon the pricing structures, it is entirely possible that the ISP market reorganizes focussing on different market segments.

3.5 Defining Service Classes Through Utility Functions

3.5.1 Elastic Application Class

As discussed in the previous sections, traditional applications like file transfer, electronic mail, fax are tolerant of delays and are called elastic. This means that adding an additional second of delay to the application hurts much more when the delays are small than when the delays are large. The utility functions of elastic applications look like the following

Figure 3.5.1

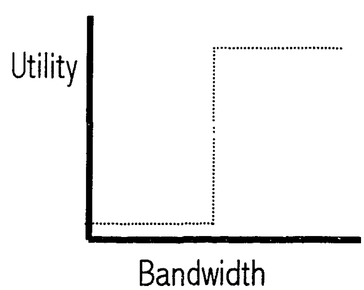


The elastic application class there is diminishing marginal rate of perceived performance as the bandwidth is increased. As the utility is the overall utility is always maximized all users get access to the network . This is the classic application for best effort service Internet.

3.5.2 Inelastic Application Class

On the other end are the delay sensitive applications which have very hard bounds on the delay and they expect their data to be within a delay bound. If the packets arrive earlier, the application does not care as it is below the minimum perceivable resolution of the application. However, if the packets arrive later than this bound it performs badly. Voice telephony, cable video are some such applications with hard bounds . The utility function for this class of application looks something like this:

Figure 3.5.2



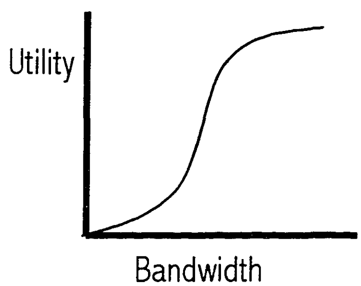
As long as the delay bounds are met the application performance is constant, but as soon as the bandwidth share drops below a certain bandwidth level, the performance falls to zero. Traditional voice and video applications over circuit switched networks have this characteristic . However, most of the real-time applications can sustain a certain degree of elasticity without perceptible degradation in performance. By playing the application elasticities and by relaxing the hard bounds on real-time applications a wide variety of applications classes can be supported on the Internet.

Service class definitions by ISPs will then be defined by the relative elasticity of the application and it's ability to provision for a set of application elasticity limits. This means that provisioning for applications need not be hard real-time or totally elastic. By providing service classes which are rate adaptive or delay adaptive, a wider class of services with varying performance requirements can be supported by the ISP.

3.5.3 Bounded Elasticity

The utility functions for delay adaptive applications do not have extreme elasticities. The utility curve is smoother.

Figure 3.5.3



The performance degradation in this case is not as sharp as that of hard in-elastic real time applications. This curve is typical of current IP telephony or IP video transmissions. The applications degrade gracefully, depending upon bandwidth availability. However, given the nature of this utility function, there is a tendency of the network to overload the network with services due the S shaped nature of the curve[16].

The analysis of the traffic generated by different application types using the utility functions explains the challenges that the Internet and the ISPs need to overcome to offer integrated services over the Internet. The original Internet was designed to handle totally elastic class of applications and hence the best effort service model was appropriate. The telephone and cable companies built their networks to support totally inelastic class of applications with hard real time bounds.

This had a fundamental effect on the mechanism to achieve these ends. The cable and telephone companies implemented hard real time through *admission control.* This meant that the telephone networks had to have an initial call setup and if an incoming call could not be given the quality of service guarantee, the call would be blocked. The current Internet does not have admission control and it would accept any traffic, without guaranteeing any quality of service, given it's original design premise.

This suggests that with the advent of integrated services and given the economic benefits of offering integrated services, the Internet has to offer some form of admission control to accommodate real time applications.

This means, that the premise that the ISPs on the edges should offer service classes and preferentially drop traffic in the case of overloading of the network, is correct. The other option is for the ISPs to overprovision such that overloading of the network never occurs and not to implement service classes.

[16] This utility function is convex but not concave in a neighborhood around zero bandwidth. The exact shape of the overloading point will depend upon the shape of the utility function.

3.5.4 Overprovisioning of the Edge ISPs

One can always increase the efficacy of the Internet and the ISP infrastructure by supplying more bandwidth ; faster speeds mean lower delays and therefore higher utility values. However, there are two problems with this approach of over-provisioning.

If the ISPs could cost-effectively over-provision the network so that under natural conditions the offered load rarely exceeded the overloading point for all classes of services then the ISPs can create a healthy competition while offering required quality of service for all classes of applications. This is however unlikely although bandwidth soothe-sawyers like George Gilder predict of bandwidth storms and free bandwidth availability [17]. While it is true that fiber and fiber based electronics cost less and less and technologies like WDM provide terabits of bandwidth, what is not talked about is the huge sunk costs of cabling that the edge ISPs would have to incur. These sunk costs could probably be recovered at the backbone regional and national providers, where the aggregation of traffic from the edges can spread the sunk costs over a large population of users. It is highly unlikely that such sunk costs can be recovered from individual consumers by providing WDM based access to the home on fiber by an edge ISP. The takeoff price per fiber point to the home is about $1500 at penetration levels of 30% [18]. Unless some dramatically altered high bandwidth local loop wireless technology emerges, it is unlikely that there could be significant downward revision in the takeoff price in the foreseeable future. This automatically puts a cap on the bandwidth at the local access point preferring xDSL or Cable Technology, rather than fiber. As the service requirement is end-to-end, the weakest link becomes the edges, limiting the extent of over-provisioning that the ISPs can afford.

Secondly, it is questionable as to whether over-provisioning at intermediate points of aggregation can be afforded by the ISPs to support various traffic types

---

[17] Telecosm, George Gilder

in a best effort network. It can be argued that just like the mature telephone network has been successfully over-provisioned with very low call blocking rates, the Internet can also be successfully over-provisioned with very low packet rejection rates. However, there are some important differences between the phone network over-provisioning and the over-provisioning of the Internet. The bandwidth usage of a phone call is fixed and the invocation of the call requires human action. Both these factors limit the variability in the aggregate phone usage on the CXO trunk lines. In contrast a packet switched mixture of video, audio and data will be much more variable ranging from a few kilobits per second to a few gigabits per second. Thus, the variability in each individual user means that the aggregate usage will be more variable [19].The burstiness of the different types of traffic and it's aggregation creates a highly variable Internet traffic load which renders the over-provisioning model ineffective. In fact a small population of high bandwidth bursty user with bandwidth requirements a few orders of magnitude higher than the normal users like e-mail and voice. The capacity required to over-provision for one video-on-demand user running at 100 megabits per second would be around a gigabit per second [20] even when the average use of the high bandwidth system is much lower than the normal Internet traffic ! The question then becomes as to who pays for the over-provisioning. As the high bandwidth users cause the tremendous variance in demand, if the ISPs spread the costs of capacity expansion equally among all users then the normal users are subsidizing the high bandwidth users.

In a competitive ISP market other ISPs will offer alternative pricing schemes which concentrate the cost of expansion of capacity to the high bandwidth users thus drawing the normal users away from the ISPs which subsidize the high end user at the expense of the normal user. Thus, the alternative for the high bandwidth users is to either accept a significant call blockage/packet drop rate or bear the cost of over-provisioning by the ISP. Given that the bandwidth needed

---

[18] The Prospects of Fiber in the Local Loop, FCC Report September 1993, David Reed
[19] Central limit theorem : Variance of the aggregate is proportional to the variance of the individual distributions
[20] This can be shown using Queueing Theory and assumed arrival, departure rates

to over-provision is orders of magnitude larger than the average demand , this cost will be quite large. Hence a class/priority based transmission mechanism by the ISPs will be far more effective than brute force over-provisioning. Depending upon the load on the ISP network, preferences and budgets of the consumers, and competing classes of traffic a market based pricing will determine the admission of a users traffic stream.

In light of these two factors, the adoption of a service class based priority pricing scheme by the edge ISPs is the most effective way to maximize the overall utility of the Internet and promote competition between the edge ISPs rather than using over-provisioning. Such a priority class based model will allow different users to get different qualities of service which are better aligned to their application needs while increasing the efficacy of the network. This will also enable the ISPs to maximize their profits by designing various quality of service based price profiles addressing the needs of the customer base.

3.6 Service Classes in the Internet

One possible way to support the consumer need is to define class of service profiles to identify the needs of applications for bandwidth across the network, internal and external. Broken into six classes, each class covers a wide range of services, defined in Table 1. With edge network speed spanning from 1 kilobits per second to 1 gigabit per second in the class 6 category, the applications covered are everything reasonably imaginable today. These class definitions will aid ISPs and users in matching application requirements to network capabilities, and allow the ISPs to implement technologies that will guarantee these classes of service.

Implementing these service classes can be done in various ways. One possible way to achieve this goal is by over-provisioning the network. The other way is to extend the Internet service model to include service classes which can attract differential pricing by exploiting the preferences of the customer. While the former sounds easy it may be practically infeasible and the latter mechanism may be the preferred solution. The latter mechanism for service differentiation and related pricing will help in providing a continuum of services ranging from commodity basic access to near guaranteed quality of service.

The increase in bandwidth which enables new applications typically occurs in an order of magnitude[21]. Thus between 1 and 10 Kbps , basic applications like e-mail, file transfer became available. Similarly, between 10 Kbps and 100 kbps applications like IP telephony and the World Wide Web becomes common over the Internet. IP video takes 1 Mb and so on. Research has shown, that there is little change in the user's perception of the value of a stream, if the change does not occur in an order of magnitude.

Table 3.6

Class	Bandwidth	Type of Application
1	1 Kbps - 10 Kbps	Electronic Transactions, Mail, WWW
2	10Kbps – 100 Kbps	IP Telephony, WWW Graphics
3	100 Kbps – 1 Mbps	IP Video, Games, WWW Multimedia
4	1 Mbps – 10Mbps	Video on Demand, IP Multicast
5	10 Mbps- 100 Mbps	High Quality Imaging, Uncompresses Video
6	100 Mbps – 1 Gbps	HDTV, Video Broadcast

[21] Branco Gerovac, David Carver – Research Notes , Research Program in Communications Policy, MIT

3.7 Integrated Services Model

The above analysis suggests the use of a combination of fine grained service classes, incentives through pricing and short term contracts as the key to determine the business and technology infrastructure of the ISPs at the edges. On the other hand, coarse grained service classes, bulk pricing , over-provisioning and long term contracts will determine the business and technology infrastructure of the regional and national backbone providers. The combination of short term and long term contracts at the edges and the core are the derivatives of the nature of the traffic, the nature of bandwidth concentration and the user perception of traffic elasticity. The short term contracts is through differentiated services at the edges and the long term contracts is through network provisioning by the backbone ISPs.

4 Current Market Structure and Competition in the ISPs

The Internet Service Provider (ISP) market has evolved from an assortment of small regional networks dependent on government funding and serving the research and education communities to a collection of global backbones run by hard-driving, well-financed start-ups and monolithic telephone companies serving Fortune 1000 companies. There is an enormous range across the ISPs. Some provide global networks and an enormous array of services. At the other end are the small operators in a single area code with a single POP (Point of Presence) and a few dial-up modems. The growth of ISPs to a large degree driven by the relatively low entry costs, at least initially. A small ISP needs to spend no more than low six figures to start off but the long term economies of scale lie in the ability to scale the infrastructure and more importantly the operational infrastructure of order processing, billing and customer service. Given the low entry costs and high growth rates, the ISPs frequently oversubscribe while the infrastructure cannot keep up with the growth. This causes congestion, outages and results in a high churn rate effectively increasing the cost of servicing the subscribers.

A lack of service differentiation has led the prices of the ISPs to be driven down to commodity levels. As we have seen in the last chapter, service differentiation and differentiated pricing is the key to the sustained growth of the ISPs. This chapter analyzes the technology infrastructure requirements, the market structure and competition and the key determinants of competitive advantage of the ISPs.

4.1 The ISP Technology Infrastructure

The Internet Service Providers provide primary connectivity to users on the Internet. The current connectivity is based on the IP protocols best effort service. The basic building block of an ISP is the Point of Presence (POP). The POP consists of network hardware and software for seamless Internet connectivity, namely

- Modem Banks/Head End Switches
- Routers and Switches
- Domain Name Servers (DNS)
- News and Mail Servers

4.1.1 Points of Presence (POPs)

A POP serves as a gateway to the rest of the Internet for any subscriber who can connect to it. Connection to the POP is currently achieved through the local telephone lines through dial-in or through cable companies. In the former case the POP has a set of modem banks which allow the subscribers to connect at speeds upto 56 Kbps with the current technology. This may be upgraded to ADSL technology as and when it becomes available. Usually, the POTs based POPs have a modem to subscriber ratio of 10:1[22]. Additionally, the POP may also have direct connections for T-1, E-1 and ISDN connectivity from the subscriber premises. The modem banks are connected to switching hardware which switch the connections to the ISP's backbone infrastructure. Depending on subscriber density and traffic loads, some POPs can have a minimal infrastructure of modem banks and a switch, with the routers and other servers

[22] Interview with Mahesh Bhave, Director of Strategic Planning, Citizen's Telecom, Connecticut

located in a centralized places which afford better scale economies and manageability. These POPs are also know as virtual POPs or vPOPs. The POPs for Internet access via cable on the other hand are also known as head-ends. The head-ends connect to the regional cable networks, which in turn connects to the NAPs.

4.1.2 Routers

The Internet traffic generated by a subscriber is directed along a set of links over the Internet until it reaches it's destination address. The hop-by-hop path determination is done by routers. Routers connect one network to the other and the path that is undertaken by a message is embedded in it's routing tables. The routing path information is dynamic and changes with time , and routers update their routing tables periodically over the Internet[23]. The route determination is usually based on an end-to-end metric like the shortest path to the destination. Each ISP has one or more routers responsible for routing the messages generated by a subscriber. The routers are present at each POP unless it is a vPOP. The vPOP switches the traffic directly to a POP which then routes the traffic.

4.1.3 Domain Name Servers

The Domain Name Servers (DNS) are responsible for resolution of the destination name in a message to an Internet address. The names are resolved using a distributed name resolution mechanism through Domain Name Servers. The DNS is a hierarchical name resolution system which has top level hierarchies as .com, .org, .mil, .edu etc. These top level hierarchies map into lower level hierarchies till the actual name is resolved[24]. The responsibility of the

---

[23] There are different mechanisms to update routing tables, but the best known one is the Bellman Ford algorithm based on a Shortest Path Algorithm.

[24] RFC 1033, Domain Names Concepts and Facilities, Internet Engineering Task Force

resolution is distributed across the Internet, and ISPs can have DNS servers which resolve parts of the domian name hierachy. Thus ISPs usually have DNS servers which help resolve some of the names which is a part of the hierarchy hosted by the ISP. This helps in speedier resolution of the domain names.

4.1.4 Mail and News Servers

Apart from the access to the Internet, the ISPs typically provide the capability to send and receive e-mail as well as subscribe to the Internet news groups. The mail service provided the ISPs is based on the Post Office Protocol (POP) for the end user. The ISPs host mail servers which receive and store the SMTP (Simple Mail Transfer Protocol) based e-mail on SMTP servers at their POPs. The mail is retrieved by the user using Post Office Protocol at his or her end machine. The news servers is a repository of news from different newsgroups which, the subscribers download depending upon their interest. The news servers can grow to be very large. A few gigabytes of news feeds onto the news servers of an ISP is not uncommon.

4.2 The ISP Business Process Infrastructure

Apart from the basic hardware infrastructure requirements, the ISPs require an integrated business process infrastructure. This can often be a key service differentiator for ISPs. The following are the basic elements of the systems that are required for a succesful ISP operation.

4.2.1 Billing Systems

ISP's charge their customers based on access and services provided on the Internet. The dial-up access services are charged on a combination of flat and variable fee based on connect times. The billing systems are an integral part of realizing the revenues maintaining the customer service. Currently, the billing systems are mostly homegrown, and with a few exceptions, and not robust. This is particularly the case with smaller ISPs who do not have the capital infrastructure to create such billing systems. There are a few shrink wrapped billing systems[25] which provide the scale and required robustness, but they are way overpriced at more than $200,000.

The connectionless nature of the IP protocol is another bottleneck to variable rate billing. The ISPs providing service based on dial-up lines can charge variable price based on connect times, but a cable based ISP cannot. This would mean that in the longer term, the billing systems would have to be based on quality of service, otherwise given a competitive market everybody will bill based on a flat rate charge. The ability of the billing systems to adapt to a quality of service based model is crucial for future competition in the ISP marketplace. Other mechanisms of charging including, per packet traffic, has very high implementation overheads[26]. Further, packet based billing systems also do not relate to consumer value. The end user's value is in the value of the traffic stream and not on the individual packet and hence this mode of billing is not sustainable in the long run in a competitive environment.

4.2.2 Call Center and Customer Care Systems

The call center and customer care systems is the other building block in delivering service to the subscriber. The call center acts as the central point of coordination for all end user problems. The call center is responsible for trouble ticketing of a user problem, problem escalation and problem resolution. The call center is thus key to customer satisfaction and minimum customer churn rate.

---

[25] Specification for Arbor from Kenan Systems, http://www.kenan.com

4.2.3 Network Management and Operations

A cost effective and reliable 24X7 hour operation of the ISPs infrastructure is essential for the long term success of the ISP. Network management and operations provide that support by proactively monitoring the network for faults. The effectiveness of a an ISPs network management and operations center is based on it's ability to manage a multitude of network elements remotely and centrally with possible lights out operation. This aspect is a key element of reliable operations at the minimal cost.

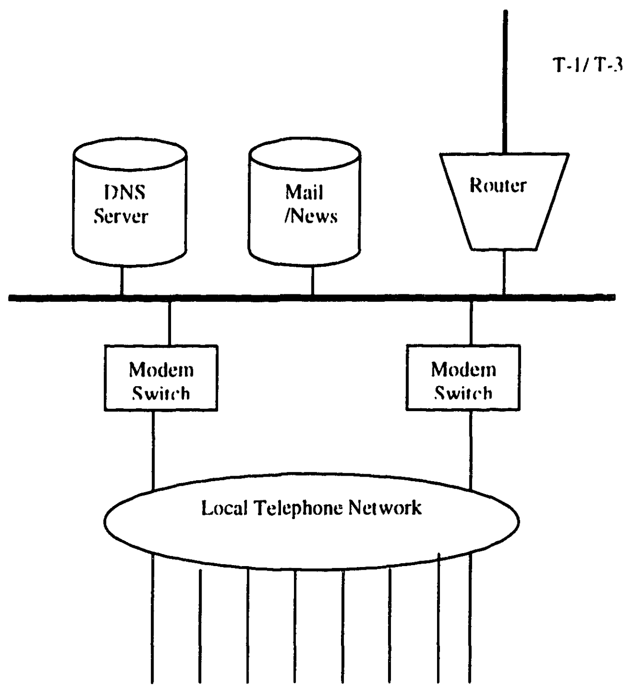
Finally, the integration of the three systems is vital for the seamless operation of the ISP, while providing end-to-end service to the subscriber.

Figure 4.2.3 illustrates the business and technology infrastructure required for a minimal ISP operation

---

[26] Billing Users and Pricing for TCP/IP, Richard Edell, Nick McKeown et. Al., IEEE JSAC Special Issue on the Fundamentals of Advances in Networking, September 1995

Figure 4.2.3 ISP Technology Infrastructure



5 A Service Model for Differentiated Services

This section outlines a model for providing what has been called differentiated service on the Internet. The goal of the mechanism is to allocate the bandwidth of the Internet to different users in a controlled way during periods of congestion such that the utility of each of the users is maximized. Using this model an ISP can differentiate services and price differentiated services to maximize the profit by moving away from marginal cost pricing.

This framework up builds on the the Committed Information Rate mechanism of Frame Relay based systems and on similar models suggested by David Clark for the Internet[27]. However, the model is simplified by taking into account that the congestion problems mostly occur on the edge of the network and the core of the network is usually over-provisioned. This allows easier implementation of the service differentiation strategy by providing preferential drop of packets mostly on the edges.

The mechanism also provides useful information to providers about provisioning requirements. With this service differentiation model service providers can more easily allocate specific levels of assured capacity to customers, and can easily monitor their networks to detect when the actual service needs of their customers are not being met.

---

[27] An Internet Model for Cost Allocation and Pricing , Internet Economics , Joseph Bailey and Lee McKnight.

5.1 The Base Framework

The general approach of framework will be to define a service profile for each user, and to design a mechanism in each router on the end-to-end path to favor traffic that is within those service profiles. The core of the idea is very simple -- monitor the traffic of each user as it enters the network, and mark packets as being "green" or "red" depending upon whether they are within or outside their there service profiles. Then at each router, if congestion occurs, preferentially drop traffic that is marked as being "red".

However, in this scheme while the marking of the packets is done at the edges and inter ISP transfer points, the dropping of the packets can be done only at the edge routers. This results in much simpler implementation by taking advantage of the provisioning characteristics of the network. As the edge ISPs need to provision for bursty traffic, drop mechanisms need to be designed at the edge ISP networks. The 2 nd. Tier ISPs who receive aggregate traffic, plan for an overprovisioned infrastructure on longer time horizons and thus the probability of a congestion inside the core of the network is relatively low. This implies that, it is not necessary to implement packet dropping schemes on the routers inside the core of the network. Inside the edge ISP network, at the routers, there is no separation of traffic from different users into different flows or queues. The packets of all the users are aggregated into one queue, just as they are today. Different users can have very different profiles, which will result in different users having different quantities of "green" packets in the service queue. A router can treat these packets as a single commingled pool. This attribute of the scheme makes it very easy to implement, in contrast to a scheme like RSVP reservations, in which the packets must be explicitly classified at each node. The figure 5.1.1 illustrates the model.

To implement this framework on the edge networks, the edge routers need to be enhanced to implement the dropping scheme, and a new function must be implemented to mark the traffic according to its service profile. This algorithm can be implemented as part of an existing network component (host, access device or router) or in a new component created for the purpose. This will be referred to as a distinct device called an "edge meter". The service profiles will be downloaded to the edge meters.

The concept of a service profile can be applied at any interconnect point in the network where a subscriber-provider relationship exists. A profile may describe the needs of a user connected to an ISP or the traffic handling agreement between a edge ISP and a backbone ISP. In figure 5.1.1, the aggregate inter-ISP edge meter stores the aggregate service profile in terms of aggregate traffic passed between ISP 1 and the Backbone ISP. In all likelihood the inter ISP edge meter on the backbone ISP will have an service profile for edge ISP which is the sum of all the traffic that the edge ISP is passing to the backbone ISP.

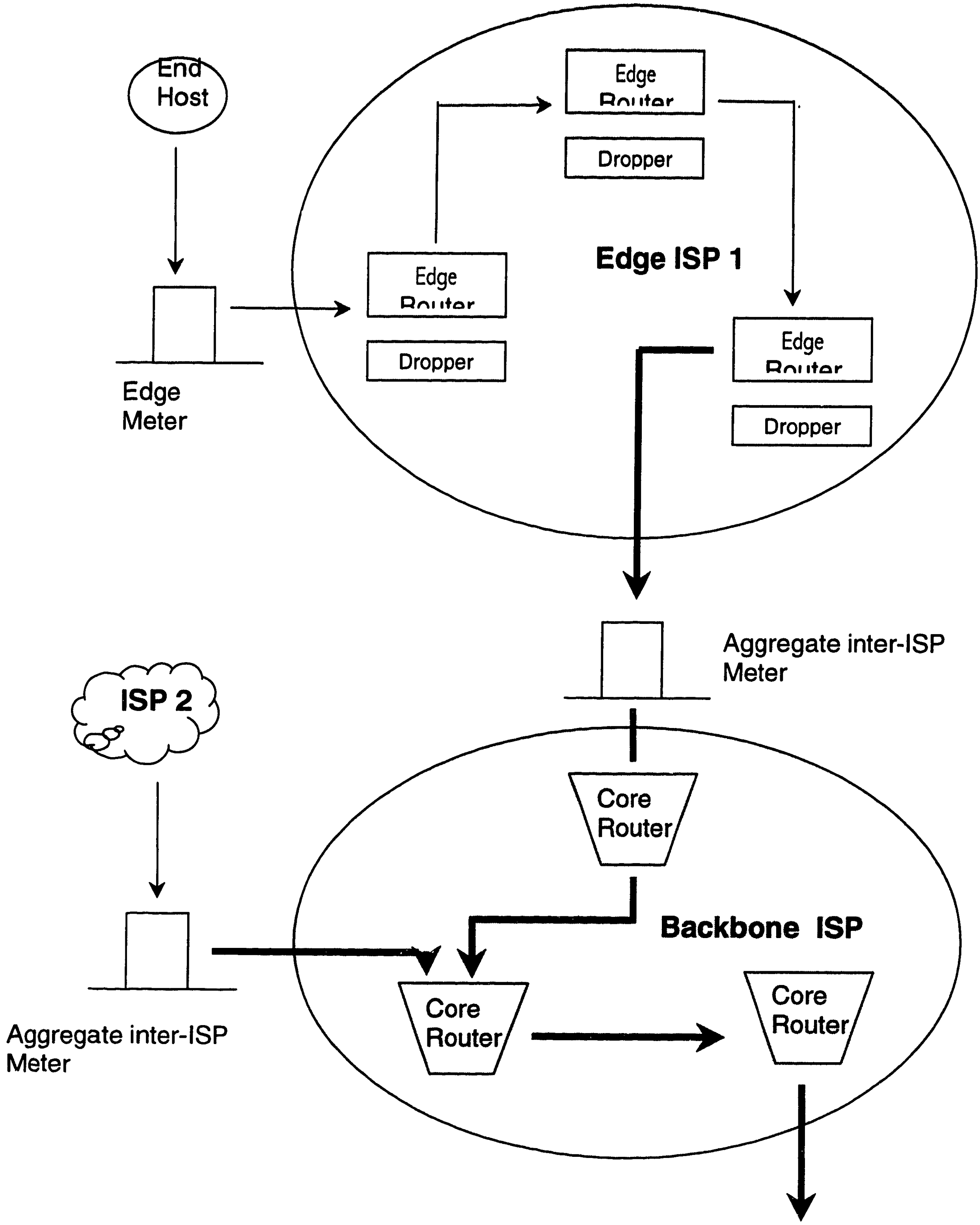

End Host
Edge Meter
Edge Router
Dropper
Edge Router
Dropper
Edge ISP 1
Edge Router
Dropper
Aggregate inter-ISP Meter
ISP 2
Core Router
Backbone ISP
Aggregate inter-ISP Meter
Core Router
Core Router

Figure 5.1.1

5.2 **Service Portfolios**

As discussed above, there are two general issues concerning service models. It will be easier to start by implementing a few simple set of services, which are useful and easy to understand. As discussed earlier, embedding these services into the framework, reduces the flexibility to change the service definitions instead a general mechanism that allows the change of services as new services are defined and added.

The definition of "service profiles" allows that flexibility by pushing the definition of the service to the edges. Thus a service is defined by the edge meter, which implements the user's service profile. To change the service, it is only necessary to change the service profile definition on the edge meter. The edge routers in the interior of the network implement a common mechanism which is used by the different edge meters to provide different services.

The following are the key variables which define can define a service

- Average expected throughput (an example might be "two megabits per second of average bandwidth)
- Peak burst rate and time (an example would be "5 mbps peak rate for no more than 5 ms at a maximum of one burst per hour")
- destination of the service (examples might be a specific destination, a group of destinations, all nodes on the local provider, or "everywhere")
- with what level of assurance is the service provided (or alternately, what level of performance uncertainty can the user tolerate)

5.3 A Service Model

As a starting point the edge ISP could offer a simple service which essentially is an equivalent of a dedicated link service. This means that the end user connected to the ISP has fixed bandwidth all the time from his point of attachment to the Internet . This often known as the establishment of a virtual connection with a fixed bandwidth.

This model has been implemented in a number of network architectures, with different "enhancements". The CBR service of ATM is an example, as is (to some extent) the CIR mechanism of Frame Relay. However, there are some issues and limitations to this very simple model.

A limitation of the virtual link model is that the user may not wish to purchase this virtual link full time. He may need it only some of the time, and in exchange would hope to obtain a lower cost. This would mean that the edge ISP offer a service profile definition which is dynamic. This could be a service profile which is effective for a limited time horizon. The end user could buy the service profile on demand depending on the availability of the required profile. The edge ISP would provision the network based on the number and types of profiles he would like to sell at any point of time At any instant the sum of the profiles should not exceed the overall network capacity.

A second issue is whether the user can exceed the capacity of the virtual link when the network is unloaded. Today, the Internet allows its users to go faster under that circumstance. Continuing to capture that benefit may be important in user acceptance. The CIR of Frame Relay works this way, and it is an important aspect of its market appeal.

An equally important issue is that the user may not wish to set up different distinguished committed bandwidth flows to different destinations, but may prefer to have a more aggregated commitment. There are several drawbacks to making distinct bandwidth commitments between each source and destination. First, this may result in a large number of flow specifications. If the user is interested in 1000 network access points, he must specify one million source-destination pairs. Frame Relay has this specification problem.

Second, the sum of the distinct commitments for any source (or destination) cannot exceed the physical capacity of the access link at that point, which may force each individual assurance to be rather small. Finally, the source-destination model implies that the user can determine his destinations in advance, and in some cases that he understands the network topology; two situations which are not universally true.

In fact, several variations of service commitment might make sense to different users; from one source to a specific destination, from a source to a pool of specified destinations (one might configure a Virtual Private Network in this way) and finally from a source to "anywhere", which could mean either all points on the ISP, on a collection of ISPs, or any reachable node.

The latter sorts of commitments are by their nature more difficult to offer with high assurance. There is no way to know for sure what the service will be to any specific destination, because that depends on what other traffic is leaving the source, and what other traffic is arriving at the destination. Offering commitments to "anywhere within the ISP" implies that the ISP has provisioned its resources adequately to support all in-profile users simultaneously to the same destination. Offering commitments to "anywhere at all" implies that all ISPs in any reachable path from the user have provisioned sufficiently, which is most unlikely.

5.4 Bursty Traffic

Not all Internet traffic is continuous in its requirement for bandwidth. Indeed, based on measurements on the Internet, much of the traffic is very bursty. It may thus be that a service model based on a fixed capacity "virtual link" does not actually meet user's needs very well. Some other more complex service profile that permits bursty traffic may be more suitable.

It is possible to support bursty traffic by changing the edge meter to implement this new sort of service. The key issue is to insure, in the center of the network, that there is enough capacity to carry this bursty traffic, and thus actually meet the commitments implied by the outstanding profiles. This requires a more sophisticated provisioning strategy than the simple adding up tactic needed for constant bit-rate virtual links. This sort of analysis can be employed as a way to predict the capacity that must be provided to support profiles describing bursty traffic. As a practical matter, in the center of the existing Internet, at the backbone routers of the major ISPs, there is such a high degree of traffic aggregation that the bursty nature of individual traffic flows is essentially invisible. So providing bursty service profiles to individual users will not create a substantial provisioning issue in the center of the network. It is the edge of the network which needs to be provisioned to accommodate traffic bursts .

5.5 Levels of Service Assurance

The next aspect of sorting out services is to consider the degree of assurance that the user will receive that the contracted capacity will actually be there when he attempts to use it. Statistical bandwidth allocation allows the Internet to support an increased number of users, use bandwidth vastly more efficiently, and deal flexibly with new applications and services. However, it does lead to some uncertainty as to the bandwidth that will be available at any instant. The approach to allocating traffic is to follow this philosophy to the degree that the user can tolerate the uncertainty. This means that the capacity allocation scheme should provide a range of service assurance. At one extreme, the user may demand an absolute service assurance, even in the face of some number of network failures. Less demanding users may wish to purchase a service profile that is "usually" available, but may still fail with low probability. The presumption is that a higher assurance service will cost substantially more to implement.

These statistically provisioned service profiles can be called "expected capacity" profiles. This suggests that the profiles do not describe a strict guarantee, but rather an expectation that the user can have about the service he will receive during times of congestion. This sort of service will somewhat resemble the Internet of today in that users today have some expectation of what performance they will receive.

Statistical assurance is a matter of provisioning. In this scenario, an edge ISP can track the amount of traffic tagged as “green” crossing various links over time, and provide enough capacity to carry this subset of the traffic, even at times of congestion. This is how the Internet is managed today, but the addition of tags gives the ISP a better handle on how much of the traffic at any instant is "valued" traffic, and how much is discretionary or opportunistic traffic for which a more relaxed attitude can be tolerated.

For traffic that requires a higher level of commitment, more explicit actions must be taken. The specific sources and destinations must be determined, and then the paths between these points must be inspected to determine if there is sufficient capacity. There are two approaches. The static approach involves making a long term commitment to the user, and reserving the network resources to match this commitment. This involves some computation based on the topology map of the network to allocate the needed bandwidth along the primary (and perhaps secondary) routes. The dynamic approach involves using a setup or reservation protocol such as RSVP to request the service. This would explore the network path at the time of the request, and reserve the bandwidth from a pool available for dynamic services. Information concerning this pool would have to be stored in the routers themselves, to support the operation of RSVP. Within one ISP, the reservation could be submitted to a central location for acceptance, depending on the design adopted for bandwidth management.

It is important to note that traffic requiring this higher level of assurance can still be aggregated with other similar traffic. It is not necessary to separate out each individual flow to insure that it receives it promised service. It is necessary only to insure that sufficient capacity is available between the specific sources and destinations desiring the service, and that the high-assurance packets can draw on that capacity. This implies that there would be two queues in the router, one for traffic that has received a statistical assurance, and one for this higher or "guaranteed" assurance. Within each queue, “green” and “red” tags would be used to distinguish the subset of the traffic that is to receive the preferred treatment.

An ISP could avoid the complexity of multiple queues and still provide the high-assurance service by over-provisioning the network to the point where all “green” traffic is essentially never dropped, no matter what usage patterns the users generate. It is an engineering decision of the edge ISP whether this approach is feasible.

5.6 Service Profile for End-to-End Access Path

In some cases, what the user is concerned with is not the end-to-end behavior he achieves, but the profile for utilizing his access path to the network. For example, users today buy a high-speed access path for two different reasons. One is to transfer a continuous flow of traffic, the other to transfer bursts at high speed. The user who has bursty traffic might want on the one hand an assurance that the bursts can go through at some known speed, but on the other hand a lower price than the user who delivers a continuous flow into the Internet. Giving these two sorts of users different service profiles that describe the aggregated traffic across the access link will help discriminate between them, and provide a basis for differential charging.

A service profile of the sort discussed here is a reasonable way to capture this sort of requirement. By marking the traffic that crosses the access path according to some service profile, the ISP commits to forward that subset of the traffic within its region, and only delivers the rest if the network is underloaded. It is instructive to compare this approach to pricing an access path to the more traditional "usage-based" scheme. In the traditional scheme, the actual usage is metered, and the user is charged a fee that depends on the usage. If the user sends more, he pays more. In contrast, a service profile allows two users with different needs to be distinguished (and charged differently) but each user could be charged a known price based on the profile. If the traffic exceeds the profile, the consequence is not increased fees, but congestion pushback if the network is congested.

5.7 Location of Service Profiles in the Network

In the simple sender-based scheme described so far, the function that checks whether traffic fits within a profile is implemented by marking packets as in or out of profile at the edge of the network. A profile describes an expectation of service obtained by a customer from a provider. These relationships exist at many points in the network, ranging from individual users and their campus LANs to the peering relationships between global ISP's. Any such boundary may be an appropriate place for a edge meter.

Further, the packet marking associated with this service profile will, in the general case, be performed by devices at either side of the boundary. One sort, located on the traffic sourcing side of a network boundary, is a "policy meter". This sort implements some policy by choosing the packets that leave the network (or user's machine) with their in-profile bit set, and thus receive the assured service. Another sort, a "checking meter", sits on the arriving- traffic side of a network boundary, checks the incoming traffic, and marks packets as out of profile (or turns off excess in-profile bits) if the arriving traffic exceeds the assigned profile.

A general model is that the first meter that the traffic encounters should provide the highest degree of discrimination among the flows. A edge meter could be integrated into a host implementation , where it could serve to regulate the relative use of the network by individual flows. The subsequent meters, looking only at larger aggregates, serve to verify that there is a large enough overall service contract in place at that point to carry all of the traffic tagged as "green" (the valuable traffic) at the interior points. When a traffic meter is placed at the point where a campus or corporate network connects to an ISP, or one ISP connects to another, the traffic being passed across the link is highly aggregated. The ISP, on the arriving- traffic side of the link, will check only to verify the total behavior. On the traffic sourcing side of the link, an additional edge meter can be installed to verify that tags have been applied according to policy of the user.

Additional edge meters installed at intermediate points can provide good feedback on network demand. Consider a specific situation, where traffic is marked at individual hosts according to policies specific to these hosts, and then passes through a second meter at the point of attachment from the private network to the public Internet. If the number of "green" packets arriving at that point exceeds the aggregate service profile purchased at that point, this means that the user has not purchased enough aggregate capacity to match the needs of his individual policy setting. In the long run, this provides a basis to negotiate a higher service level with the ISP. So traffic meters actually provide a basis for monitoring user needs, and moving users to a higher service profile as needed.

5.8 Scope of Profiles

Even in the case where the user wants to obtain a service profile that is not specific to one destination, but rather applies to "all" possible destinations, it is clear that the "all" cannot be literally true. Any service profile that involves an unspecified set of destinations will have to bound the scope of the agreement. For example, a single ISP or a set of co-operating ISPs may agree to provide an assured service profile among all of their end points, but if the traffic passes beyond that point, the profile will cease to apply.

The user might be given further options in the design of his profile. For example, if there are regions of restricted bandwidth within the Internet, some users may wish to pay more in order to have their “green” tags define their service across this part of the net, while others may be willing to have their “green” tags reset if the traffic reaches this point.

This could be implemented by installing a edge meter at that point in the network, with explicit lists of source-destination pairs that are and are not allowed to send “green” traffic beyond this point.

5.9 Details of the Mechanism

The following sections describe the details of the differentiation and drop mechanisms at the routers. It is a variant of the Random Early Detection algorithm implemented in routers today, coupled with the differentiation of traffic at the edges using profiles.

5.9.1 Design of the dropper

One of the key parts of this scheme is the algorithm in the router that drops "red" packets preferentially during times of congestion. The behavior of this algorithm must be well understood and agreed on, because the marker at the edge of the network must take this behavior into account in their design. There can be many markers, with different goals as to the service profile, the degree of aggregation etc. There is only one dropper, and all the routers have to perform an agreed behavior.

The essence of the dropper is an algorithm which processes all packets in order as received, in a single queue, but preferentially drops "red" packets.

The primary design goals of the dropper are the following:

- It must allow the markers to implement a range of service profiles in a useful and understandable way.
- If the router is flooded with "red" packets, it must be able to discard them all without harming the "green" packets. In other words, it must deal well with non-conforming flows that do not adjust their sending rate when they observe packet loss.
- If the router is receiving a number of "well-behaved" traffic flows, which will (as TCP always does) speed up until they encounter a lost packet, it must have enough real buffering available that once it starts to get overloaded with packets, it can discard "red" packets and still receive traffic bursts for a round trip until the affected TCP slows down.

5.9.2 Weighted RED

The specific dropping scheme is an extension of the Random Early Detection scheme, or RED, that is now being deployed in the Internet. The general behavior of RED is that, as the queue begins to build up, it drops packets with a low but increasing probability, instead of waiting until the queue is full and then dropping all arriving packets. This results in better overall behavior, shorter queues, and lower drop rates as the mechanism is anticipatory in nature rather than reactive[28].

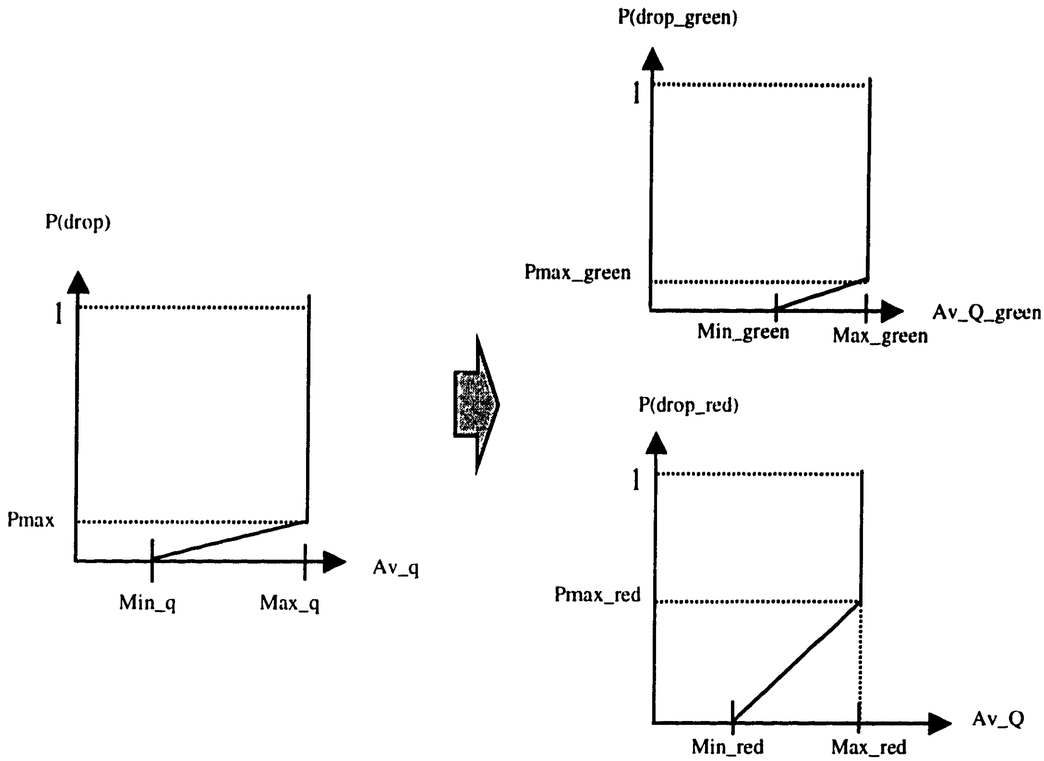
The approach here is to run two RED algorithms at the same time, one for "green" packets, and one for "red" packets. Figure 5.2 illustrates this situation. The "red" RED algorithm starts dropping at a much shorter average queue (Min_red) length, and drops much more aggressively than the "green" algorithm (i.e. it's probability of drop P_{drop_red} is much higher than the green packet's probability P_{drop_green}). With proper setting of the parameters, the "red" traffic can be controlled before the queue grows to the point that any "green" traffic is dropped.

The "green" dropper must look at the number of "green" packets in the queue. The "red" dropper must look at the total queue length, taking into account both "green" and "red". This is because the link can be subjected to a range of overloads, from a mix of "green" and "red" traffic to just "red". In both cases, the router must start dropping "outs" before the "green" traffic is affected, and must continue to achieve the basic function of RED; preserving enough free buffer space to absorb transient loads with a duration too short to be affected by feedback congestion control.

---

[28] Random Early Detection Gateways for Congestion Avoidance, Floyd S, Jacobson V., IEEE/ACM Trans. on

Figure 5.9.2



5.10 Non-Responsive Flows

A well-behaved traffic source that responds similarly to congestion signaled by packet loss, will respond well to the dropper. As more of its packets are marked as "red", one will eventually be dropped. At this point, the source will back off. As a result, most of the time a network of well-behaved controlled flows will contain just enough "red" packets to use up any excess capacity not claimed by the "green" packets being sent.

However there could be scenarios where the source does not adjust, such as a video data application, that does not or cannot adjust.

Networking, August 1993

In such situations, if the unresponsive flow's packets are marked as out of profile, the flood of "red" packets will cause the dropper to operate in a different way, but well behaved controlled flows must continue to receive good service.

5.10.1 Robustness against non-responsive flows

In this scheme, once the level of "red" packets exceeds a certain average level, all the incoming "red" packets will be discarded . This behavior has the consequence of increasing the router's queue length. The average queue length will increase by the number of "red" packets that are allowed to sit in the queue before the dropper switches over to the phase where it drops every "red". There must be enough physical room in the buffer so that even when there are this many "red" packets present, there is enough room for the normal instantaneous bursts of "green" packets which would be seen in any event.

5.10.2 Filtering out non-responsive flows

Although the dropper can be reasonably robust against overload from non-responsive flows, it may be useful to consider the alternative strategy of singling out non-conforming flows and selectively dropping them in the congested router. There has been work towards enhancing the traditional RED scheme with a mechanism to detect and discriminate against non-conforming flows[29]. Discriminating against these flows requires the installation of a packet classifier or filter that can select these packet flows, so that they can be discarded. This adds complexity and introduces scaling concerns to the scheme. These concerns can however be mitigated because only the nonconforming flows, not the majority of flows that behave, need be recognized.

5.11 Other Mechanisms

Schemes for differential service or capacity allocation differ in a number of respects. Some standardize on the service profiles, and embed them directly in the routers. As discussed above, this scheme has the advantage that the actual service profile is not a part of what is standardized, but is instead realized locally in the traffic meter, which gives this scheme much greater flexibility in changing the profile.

5.12 Drop preference vs. priority

One possible difference is what the router does when it is presented with an overload. This scheme is based on a specific algorithm for drop preference for packets marked as “red”. An alternative would be to put packets marked as “red” in a lower priority queue. Under overload that lower priority queue would be subjected to service starvation, queue overflow and eventually packet drops. Thus a priority scheme might be seen as similar to a drop preference scheme.

However, the priority scheme has the consequence that packets in the two queues are reordered by the scheduling discipline that implements the priority behavior. If packets from a single flow are metered such that some are marked as “green” and some as “red”, they will in general arrive at the receiver out of order, which will cause performance problems with the flow. In contrast, the diffrentiated service scheme always keeps the packets in order, and just explicitly drops some of the “red” packets if necessary.

The priority scheme is often proposed for the case of a restricted class of service profiles in which all the packets of a single flow are either “green” or “red”. These schemes include the concept of a "premium" customer (all its packets are “green”), or a rate-limited flow (packets that exceed the service

---

[29] Router Mechanisms to Support End-to-End Congestion Control, Fall K, Floyd S, available at http://www-nrg.ee.lbl.gov/nrg-papers.html

profile are dropped at the meter, rather than being passed on.) These proposals are valid experiments in what a service profile should be, but they are not the only possibilities. The drop preference scheme has the advantage that it seems to support a wider range of potential service profiles (including the above two), and thus offers an important level of flexibility.

5.13 Deployment Issues

5.13.1 Incremental deployment plan.

No scheme like this can be deployed at once in all parts of the Internet. It must be possible to install it incrementally, if it is to succeed at all.

The obvious path is to provide these capabilities first within a single ISP. This implies installing routers within the ISP, and tagging the traffic at the access points to that ISP. This requires a edge meter at each access link into that ISP. The meter could maintain a large amount of user-specific information about desired usage patterns between specific sources and destinations (and this might represent a business opportunity), but more likely would provide only rough categories of traffic classification. A user of this ISP could then install a edge meter on his end of the access link, which he controls and configures, to provide a finer- grained set of controls over which traffic is to be marked as “green” and “red”. Eventually, meters might appear as part of host implementations, which would permit the construction of profiles that took into account the behavior of specific applications, and which would also control the use of network resources within the campus or corporate area.

At the boundary to the region of routers implementing the dropping scheme, all traffic must be checked, to make sure that no un-metered traffic sneaks into the network tagged as “green”. So the implementation of this scheme requires a consistent engineering of the network configuration within an administrative region (such as an ISP) to make sure that all sources of traffic have been identified, and either metered or "turned out".

If some routers implement the dropper, and some do not, but just implement simple RED, the user may fail to receive the committed service profile. But no other major failures will occur. That is, the worst that the user will see is what he sees today. One can achieve substantial incremental improvements by identifying points of actual congestion, and putting statistical drop preference routers there first.

In the proposed scheme the incremental deployment of the edge meters and weighted RED mechanisms at the edge ISPs, is likely to provide a significant improvement in the end-to-end traffic assurances. As discussed earlier, the structure of the provisioning in the Internet is such that congestion in the edges is more likely than in the interior of the network. As the center of the network is heavily over-provisioned, implementing drop-preference scheme on the edges provides a high probability of end-to-end assurance even if the interior routers do not immediately implement the drop-preference scheme. Eventually, as the edge ISPs with the drop-preference scheme implementations increase, they will enter into aggregate traffic contracts with the backbone providers. The incremental deployment thus gracefully scales to include all the network providers in the end to end path.

5.14 Pricing Issues

The scheme described above here has been conceived in the context of the public commercial Internet, where services are offered for a price. It also works in the context of private or corporate networks where other more administrative allocations of high-quality service may be used. But it must work in the context of commercial service. It is therefore crucial that it take into consideration the varying business models of Internet service customers and providers, and that it be consistent with pricing principles

This is discussed briefly below as a precursor to the pricing model developed later.

5.15 Congestion pricing

The first economic principle is that there is only a marginal cost to carrying a packet when the network is congested. When the network is congested, the cost of carrying a packet from user A is the increased delay seen by user B. The traffic of user B, of course, caused delay for A. But if A somehow were given higher priority, so that B saw most of the delay, A would be receiving better service, and B paying a higher price, in terms of increased delay and (presumably) dissatisfaction. According to economic principles, A should receive better service only if he is willing to pay enough to exceed the "cost" to B of his increased delay. This can be achieved in the marketplace by suitable setting of prices. In principle, on can determine the pricing for access dynamically by allowing A and B to bid for service, although this has many practical problems.

When the network is underloaded, however, the packets from A and from B do not interfere with each other. The marginal or incremental cost to the service provider of carrying the packets is zero. In a circumstance where prices follow intrinsic costs, the usage-based component of the charge to the user should be zero. This approach is called "congestion pricing".

The scheme described here is consistent with the framework of congestion pricing. What the user subscribes to, in this scheme, is an expectation of what service he will receive during times of congestion, when the congestion price is non-zero. When the net is underloaded, this scheme permits the user to go faster, since both "green" and "red" packets are forwarded without discrimination in that case.

This line of reasoning has some practical implications for the design of service profiles. If a provider sells a profile that meters usage over some very long period (so many "green" packets per month, for example) then there will be a powerful incentive for the user not to expend these packets unless congestion is actually encountered. This consequence imposes an extra burden on the user (it is not trivial to detect congestion) and will yield no benefit to either the user or the provider. If there is no cost to sending traffic when the network is underloaded, then there is no cost to having some of those packets carry "green" tags. In fact, there is a secondary benefit, in that it allows providers to track demand for such traffic during all periods, not just during overload. But profiles could be defined that would motivate the user to conserve "green" tags for times of congestion, and these seem misguided.

5.16 Getting incentives right

The second economic principle is that pricing can be used as an incentive to shape user behavior toward goals that benefit the overall system, as well as the user. The "incentive compatibility" problem is to structure the service and pricing in such a way that beneficial aggregate behavior results.

Service profiles represent an obvious example of these issues. If a profile can be shaped that closely matches the user's intrinsic need, then he will purchase that profile and use it for those needs. But if the only profile he can get provides him unused capacity, he will be tempted to consume that capacity in some constructive way, since he has been required to purchase it to get what he wants. He may be tempted to resell this capacity, or use it to carry lower value traffic, and so on. The capability to buy a profile and resell unused capacity is fundamental to achieve greater efficiencies and utilization of the ISPs network provisioning. This aspect will be explored further in the following chapters.

5.17 Inter-provider payments

One of the places where a traffic meter can be installed is at the boundary between two ISPs. In this circumstance, the purpose is to meter how much traffic of value, i.e. "green" packets, are flowing in each direction. This sort of information can provide the basis for differential compensation between the two providers. In a pure sender-based scheme, where the revenues are being collected from the sender, the sender of a packet marked as "green" should presumably pay the first ISP, who should in turn pay the second ISP, and so on until the packet reaches its final destination. In the middle of the network, the ISPs would presumably negotiate some long term contract to carry the "green" packets of each other, but if asymmetric flows result, or there is a higher cost to carry the packets onward in one or the other direction, this could constitute a valid basis for differential payment.

The most general model requires both sender and receiver based payments, so that payments can be extracted from all participants in a transfer in proportion to the value that each brings to the transfer. In this case, the direction of packet flow does not determine the direction of value, and thus the direction of compensating payment.

The basic sender-based scheme considered makes sense in many business contexts. For example, a user with multiple sites, who wants to connect those sites with known service, can equally well express all of these requirements in terms of behavior at the sender, since the senders are all known in advance.

In contrast to this "closed" system, consider the "open" system of a node attached to the public Internet, who wants to purchase some known service profile for interaction with other sites on the Internet. If the primary traffic to that site is incoming (for example, browsing the Web), then it is the receiver of the traffic, not the sender, who associates the value with the transfer. In this case the receiver-based scheme, or a zone scheme, may best meet the needs of the concerned parties.

6 Pricing Service Profiles

Differentiating services by the ISP or the inter-ISP connection points allows us to provide various qualities of service to the end users with statistical guarantees. Differentiating services statistically at the edges using profiles is a practical alternative to the hard guarantees and the best effort service. Using service profiles it is also possible to charge consumers differently. This chapter discusses the issues related to pricing and proposes a mechanism based on service profiles.

6.1 Pricing for congestion

The role of service differentiation for traffic on the Internet is to statistically prioritize a service such that , in the event of congestion the traffic with the higher priority receives the service as requested. However, to establish differentiated services a system of incentives is required such that it drives the right behavior within the end user community. Various approaches have been proposed earlier which range from pricing for guaranteed soft circuits to individual packet pricing based on spot markets. These methods are impractical to implement in the current Internet. A pricing mechanism which leads to the most efficient use of the resources while maximizing the overall utility of the network is obviously the most preferred mechanism.

Currently, the Internet uses a mix of non-price resource allocation mechanisms including a First Come First Served (FCFS) scheme. With FIFO all packets are queued as they arrive; if the network is congested every packet's delay is based on it's arrival time in the queue. This scheme is not very efficient as delay sensitive applications, like real-time video, suffer. A pricing mechanism on a per packet basis tries to levy an entry price per packet based on the level of network load or congestion. When the backbone is congested, the cost of the service will be high and if prices reflect cost then only those packets with high value will be sent until congestion diminishes. The revenues collected from congestion pricing can be used by the ISP to provision the network for higher capacity.

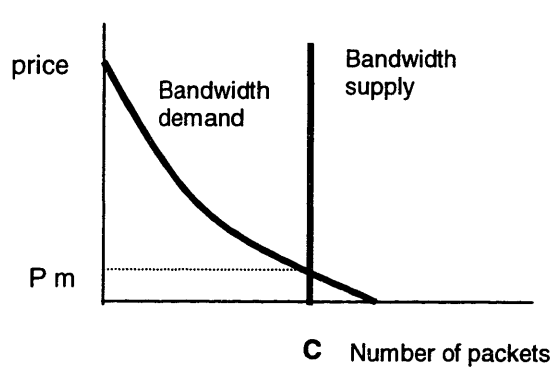
Today, the ISPs have the same type of pricing i.e. a flat access charge for unlimited access. These pricing mechanisms provide no incentives to control traffic bursts or any mechanism for allocating network bandwidth during periods of congestion. If the traffic were monitored by an ISP and each packet counted then, it would be possible to account and bill the customer based on usage. But this mechanism is flawed as the marginal cost of sending a packet over the network is nearly zero. Thus an efficient pricing mechanism should reflect the availability of bandwidth.

6.1.1 Price Components in a Packet Based Analysis

The various components of price on the Internet is a function of the cost and the following costs are relevant

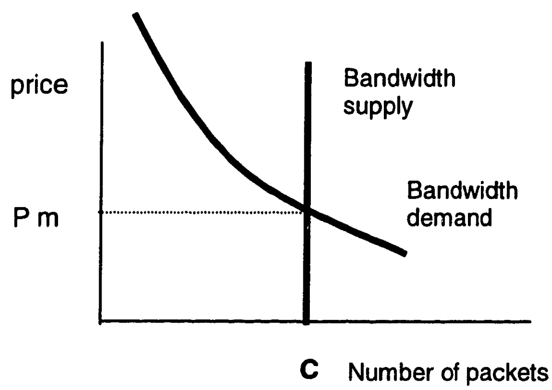
1. The marginal cost of sending traffic : The incremental cost of sending traffic on the Internet is only relevant, if the network is congested. In our case if the profiles marks a packet as "red" and if the weighted RED mechanism drops the marked packet we can assume that the network is congested. This can be shown by using the demand-supply curves of the demand for bandwidth and the capacity. In this diagram when the demand is low, the price,Pm, of the packet is low and the marginal cost of sending a packet approaches zero.

Figure 6.1.1.a



However, a surge in demand moves the entire demand curve shifts to the right and the price of sending a packet increases

Figure 6.1.1.b



This analysis however assumes fixed Supply curves with no elasticities. If an increase in packets from some users imposes a delay on other users, then the delay will vary with the number of packets. The marginal cost of delay is hence the cost added by the next packet. The efficient price then is where the user's willingness to pay for an additional packet equals the marginal increase in delay costs generated by that packet. The dynamics of this situtation can be captured by the underlying figures

Figure 6.1.1.c

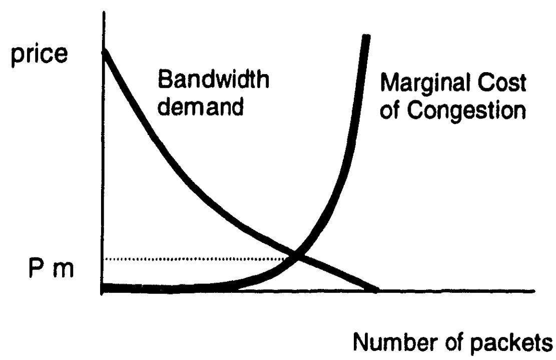
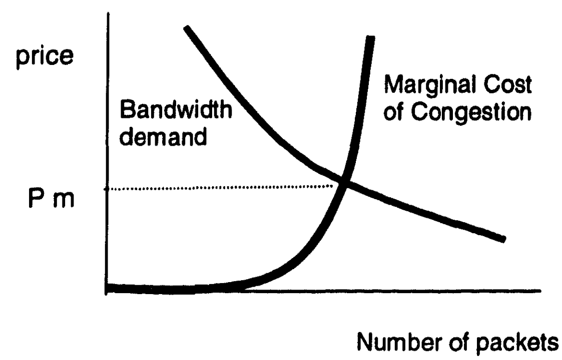


Figure 6.1.1.d



In the first figure, the network is unloaded and hence the marginal cost of congestion is very low resulting in a low price of sending the packet. On the other hand an increase in the network load moves the demand curve to the right and as the marginal cost of congestion increases exponentially as a function of the number of packets in the queue, the entry price for sending an additional packet increases dramatically.

2. The fixed cost of provisioning : The ISP incurs a fixed cost of providing the various network elements, the circuits, the gateways and the administrative and support costs. It is intuitive to think that paying for the network with a flat access fee would cover for the fixed costs. However, if the marginal congestion costs are not included then, the ISPs will be driven down to flat pricing and most of them will go out of business if they are not able to differentiate the services.

This analysis suggests that, if the marginal costs of sending a packet are available and known then the price of sending traffic a traffic should be equal to the sum of attachment costs and marginal cost of congestion.

Although this model explains the dynamics of pricing on a per packet basis on the Internet, it has the following shortcomings:

- The marginal cost of congestion is dynamic and is extremely difficult to compute due to rapidly changing congestion states in the network.
- Implementation of a packet pricing model using marginal cost pricing will need to have spot pricing and auction mechanisms. This means that in an end to end network, a user will have to participate in a per packet auction at each router. This is like coordinating a series of sequential auctions across the network for each packet in a users traffic stream. The feasibility of solving such an auction is practically remote due to the excessive coordination and computation costs.

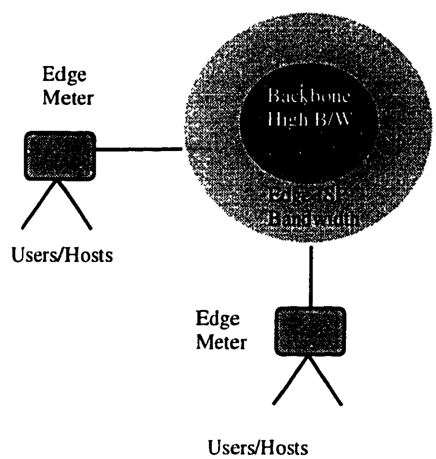
If the hard optimality criteria is relaxed and we take advantage of the structure of bandwidth concentrations in the Internet and assume that the edges are important from a pricing perspective, then a much simpler pricing model can be derived. Secondly, computing the true congestion costs requires that one can compute other user's loss in utility due to one user's use. This is practically unknowable. Further knowing the detailed knowledge of congestion across the entire path is extremely difficult due to the changing dynamics of congestion. As a second practical and significant simplification it is assumed that *expected congestions* are fair reflections of *approximate congestion* conditions in the network. The second assumption is like time-of-the-day pricing of telephone networks where, congestion is approximated based on expectation. Thirdly, instead of taking a link by link pricing model for the entire path in the network, it is assumed that the user only considers an expected path which is lumped . Given, the high concentration of the bandwidth in the backbone of the network and the high level of aggregation, it is a fair assumption. Thus the pricing is reduced to source destination pricing only at the edge of the networks. Based on these assumptions an “Edge Pricing” model can be derived.

An “Edge Pricing” model coupled with drop-preference profiles only attached to the edges is a more pragmatic and feasible alternative to the spot market and marginal cost pricing models. Although , the spot pricing model may provide optimal efficiency for a packet based pricing scheme it is probably not possible to implement such a scheme. The next few sections describe the proposed scheme.

6.2 Edge Pricing Model

Combining the assumptions and approximations made earlier, the price that a user pays is then based on expected congestion along the expected path along the traffic's source and destination. Hence the pricing for a differentiated quality end-to-end service can be determined at the periphery of the network, locally rather than computing it using a distributed coordination mechanism like sequential auctions in spot markets for packets.

Figure 6.2



The above figure captures this abstraction. In this figure, the high bandwidth backbone is in the center and is a part of the expected path that a differentiated traffic will take across the network. However, the high bandwidth backbone has low congestion partly due to high bandwidth and partly because of the high level of aggregation from the edge ISPs. The edge ISPs have bursty traffic and lower bandwidth and this results in higher probability of congestion and hence degraded quality of service. The edge ISPs and backbone providers have long term contracts as the rate of change in provisioning of the backbone and burstiness is lower. The differentiation thus needs to occur at the edges. The drop-preference based service model discussed earlier achieves this. The edge meters implement the drop preference scheme and the routers in the edge ISP implement the weighted RED for drops. The edge meters have the service profiles that the end host buys. The pricing of differentiated services is reduced to pricing of profiles instead of pricing of packets. As users relate to traffic streams, rather than packets each profile expresses the utility or preference of the end user for a basket of traffic types that the users wants to send or receive. This has the attractive property that the pricing can be done locally. Secondly, the pricing has converted the problem of congestion to a problem of provisioning. At any point of time, the total capacity of the edge ISP is equal to the total bandwidth capacity offered by all the profiles for a given duty cycle.

As the pricing decisions are now local, an edge-ISP provisions the network for a fixed capacity and issues a combination of profiles depending upon the traffic mix that the ISP wants to support. This allows the ISP to focus on different market segments. Local price setting allows the service provider or the ISP to focus on specific market segments with different pricing policies and thus pricing policy becomes an important service differentiation tool.

6.2.1 Pricing and Distribution of Profiles

The profiles are basically capacity filters which when summed at any point of matches the capacity of the edge-ISP network. The capacity filter, or profiles, as discussed earlier classifies flows as confirming or non-confirming. The filters can measure flow characteristics over a different time horizons like controlling long-term average rates or short term peak rates. Using bandwidth as the primary driver for defining capacity filters, a profile can be expressed in terms of two variables:

1. Average sustained rate
2. Peak Rate and period

An edge ISP can then create various combinations of profiles by combining these two variables, upto the capacity of the edge network. Depending upon the demand of these capacity filters, different profiles may be offered at various times while at all times the sum of the offered capacity remaining equal to the actual edge capacity of the ISP. Additionally, ISPs could create bundles of profile based on the users budget and preferences.

6.2.2 Edge Spot Markets

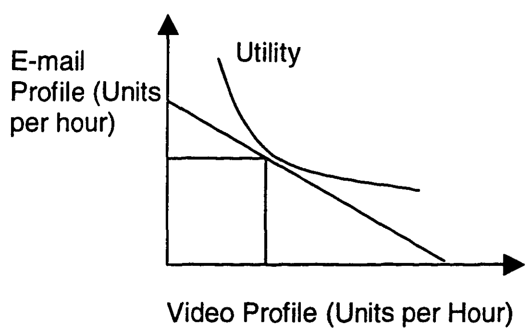
The optimal distribution of the profiles can be best achieved through bidding of the profiles at the edges. This is best achieved through auction of profiles at the edges instead of sequential auctions for packets across all the gateways. The distribution of resources through packet auction has been converted to distribution of resources through profile auction at the edges only. Most of the classic auctions examined in introductory surveys of auction theory[30] are one-sided, in that a single seller (or buyer) accepts bids from multiple buyers (or sellers). Two-sided, or double auctions in contrast,

[30] Auctions and Bidding, R Preston McAfee and John McMillan, Journal of Economic Literature 25:699-738, 1987

permit multiple buyers and sellers to bid to exchange a designated commodity. The continuous double auction (CDA), matches buyers and sellers immediately on detection of compatible bids. A periodic version of the double auction (sometimes termed a call market [31]or clearinghouse) instead collects bids over a specified interval of time, then clears the market at the expiration of the bidding interval.

In the profile based pricing framework, an edge auction agent of the user can make buy and sell offers on the profiles, based on the users budgets and preferences. The preferences are determined by utilities of a basket of profiles that a user wishes to use. The basket of profiles shaping the overall utility function for each user is dependent on the application characteristics and the individual application utility functions as discussed in Chapter 4. The following Figure indicates the structure of the profile preferences.

Figure 6.2.2



The auction of the profiles based on the budgets and the overall utility is a periodic version of the double auction, which is mediated by the ISP. The auction agents submit bids for profiles at periodic intervals for both buying and selling profiles and the ISP acts as a mediator of the auctions.

Thus, intermediation process through double auctions can be modeled as a set of user agents who have the preferences, budgets and bandwidth profiles at the edges interacting with an "auction manager" at the edge ISP. The auction manager's task is more infrastructural, serving individual functions as generally as possible, in the context of supporting an overall profile exchange process. In particular, generic infrastructure for exchanging profiles can not assume a market model where vendors announce a fixed price for consumers to take or leave. Rather, there might be many modes of negotiation, which market-matching services should take into account in identifying potential matches for buyers and sellers of profiles.

The auction manager thus has to operate within a dynamic environment, by matching descriptions of profiles to existing markets and, when appropriate, creating new markets. Implicit in the auction manager's support for market-matching operations are questions of market policy. For example, when and for what kinds of profiles should new markets be started? How does the system account for market creation costs? How are community rules, norms, and objectives (if any) expressed--- through regulations or incentives? These are issues in auction design that needs to be addressed for efficient creation and distribution of profiles between ISPs and users. The exact mechanisms for the spot markets for profiles is beyond the scope of this report and is a future area of research.

---

[31] Auction Institutional Design: Theory and Behavior of Simultaneous Multiunit Generalizations in English and Dutch Auctions, McCabe, Rassenti and Smith; American Economic Review : 80(5) 1276-1283, 1990

7 Summary and Future Work

This section describes the summary of the research and modeling for creating differentiated services and pricing of the Internet.

7.1 Summary

This thesis developed a model for differentiating services based on the concept of drop preferences on the edges using a weighted drop mechanism at the edges by taking advantage of the underlying bandwidth concentrations, varying user preferences and the market structure of the Internet. This model provides assured service for bandwidth requirements by end users and hosts with a statistical duty cycle rather than providing a guaranteed service level. By relaxing the guaranteed service level requirement, the model simplified the end-to-end guarantee requirement with low marginal impact to the service level requirements. The thesis also analyzed the pricing schemes necessary for setting the right incentives to provide the assured levels of service. In the end an edge spot market model is proposed for the efficient distribution of bandwidth profiles.

7.2 Future Work

The differentiated service model at the edges for providing end-to-end assurance requires extensive simulation before implementation. This is an open area of research. Developing and validation of the models for edge spot markets for bandwidth profiles is a fertile area of research which may have significant potential. Finally integrating the edge spot market models with the service differentiation models for end-to-end assure services will require extensive simulations and test-bed implementations. It is hoped that this research fuels further thinking and research in the emerging areas quality of service and pricing of access to the Internet.

THESIS PROCESSING SLIP

FIXED FIELD: ill. ______ name ______

index ______ biblio ______

► COPIES: Archives Aero Dewey Eng Hum

Lindgren Music Rotch Science

TITLE VARIES: ►☐ ______

NAME VARIES: ►☐ ______

IMPRINT: (COPYRIGHT) ______

► COLLATION: ~~lvr~~. 77 l

► ADD: DEGREE: ______ ► DEPT.: ______

SUPERVISORS: ______

NOTES:

cat'r: date:

page:

► DEPT: SD+M ► J159

► YEAR: 1998 ► DEGREE: S.M.

► NAME: ~~Mukherjee, Atanu~~

MUKHERJEE, Atanu