AN OPTIMAL PRICING MECHANISM FOR INTERNET'S END-USERS

by

SANTIAGO J. VILLASIS

A Thesis

Submitted to College of Graduate Studies

in Partial Fulfillment of

the Requirements for the Degree of

Master of Science in Economics

at the

University of Idaho

May 1996

Major Professor: John T. Wenders

AUTHORIZATION TO SUBMIT THESIS

This Thesis of Santiago J. Villasis, submitted for the degree of Master of Science with a major in Economics and titled "An Optimal Pricing Mechanism For Internet's End-Users," has been reviewed in final form, as indicated by the signatures and dates given below. Permission is now granted to submit final copies to the College of Graduate Studies for Approval.

Major Professor ______________________________ Date ____________________

John T. Wenders

Committee

Members ______________________________ Date ____________________

John Knudsen

______________________________ Date ____________________

David Van Over

Department

Administrator ______________________________ Date ____________________

Shaikh M. Ghazanfar

Discipline's

College Dean ______________________________ Date ____________________

Byron Dangerfield

Final Approval and Acceptance by the College of Graduate Studies

____________________________________ Date ______________

Jean'ne M. Shreeve

ABSTRACT

AN OPTIMAL PRICING MECHANISM FOR INTERNET'S END-USERS

by

SANTIAGO J. VILLASIS

Submitted to College of Graduate Studies

in Partial Fulfillment of

the Requirements for the Degree of

Master of Science in Economics

at the

University of Idaho

May 1996

This thesis provides an economic understanding of the Internet. It describes the structure and function of the Internet. It outlines the history of the Internet. It shows the derivation of the end-user's demand for Internet access and usage. It illustrates how the Internet is priced at present. In considering the aforementioned, it investigates the need for traffic-based pricing and current pricing schemes on the Internet. The thesis specifically focuses on the implementation of the rate-of-transfer concept as an optimal pricing mechanism and its limitations.

Thesis Chairman: John T. Wenders, Professor of Economics.

ACKNOWLEDGMENT

I want to express my deepest gratitude to Dr. John Wenders, the chairman of the thesis committee, for his very helpful assistance and critique throughout the investigation. I would like to thank Dr. John Knudsen for his thorough review of this thesis and Dr. David Van Over for his constructive comments. Also, I offer my sincerest thanks to Dr. Raymond Dacey for his guidance throughout my graduate studies, for his important comments and contributions to this thesis, and for providing me with a research assistantship without which none of this would have been possible.

I am very grateful to the Faculty and Staff of the Department of Economics, in particular to Dr. S. M. Ghazanfar for his support and trust, and Dr. Felix Kwan for his guidance in putting the literature review together. I also want to acknowledge my appreciation to the Department for granting me with a fellowship. A special “thanks” to Karen Gillespie, the secretary, for lending me a computer that was used in preparing this thesis.

I am also fortunate to have worked for the Computer Services Department. This experience allowed me to deepen my knowledge of computers and learn about the Internet. I especially appreciate the support of Sandy Dennis, who is a great boss. Further, I am in debt to Sharee Paull, of the McCarthy lab, for her kind help in editing this thesis.

I would also like to thank my friends the Kirklands for offering me a place in their home and my friends the Macraes for sharing some great times with my wife and I. Most importantly, I want to thank my parents Fausto and Nancy for their unmatched support throughout my academic career.

Dedication

This thesis is dedicated to Sylvia, my wife:

- Who is always there with her love, encouragement, understanding, and exceptional sense of humor which made the writing of this thesis substantially easier,

- Who helped me set up the slide projections for the thesis defense, and

- Whose dedication to me fulfilling this project is embodied in every word written.

TABLE OF CONTENTS

AUTHORIZATION TO SUBMIT THESIS......ii

ABSTRACT......iii

ACKNOWLEDGMENT......iv

DEDICATION......v

TABLE OF CONTENTS......vi

TABLE OF FIGURES......ix

LIST OF TABLES......xi

1. INTRODUCTION......1

1.1. PROBLEM STATEMENT......4
1.2. STRUCTURE OF THE STUDY......5
1.3. METHODOLOGY......5
1.3.1. Literature Review......*6*
1.3.2. Collection of Data......*8*
1.3.3. Interviews......*8*

2. UNDERSTANDING THE INTERNET......10

2.1. WHAT IS THE INTERNET?......10
2.2. HOW THE INTERNET WORKS......11
2.3. INTERNET CONNECTIVITY......14
2.4. INTERNET USES......18
2.5. INTERNET SERVICES......18
2.6. COMMON INTERNET TOOLS......18
2.6.1. E-Mail......*19*
2.6.2. File Transfer......*19*
2.6.3. Telnet......*19*
2.6.4. Newsgroups......*20*
2.6.5. Gopher......*20*
2.6.6. World Wide Web......*20*
2.7. COMMON WAYS TO ACCESS THE INTERNET......21
2.7.1. Dial-up Internet Connections......*22*

2.7.1.1. Terminal Access (via vendor-provided host VT-100 shell accounts)....22
2.7.1.2. Workstation SLIP (Serial Line Internet Protocol) or PPP (Point to Point Protocol) Access....23
2.7.1.3. Internet Gateway Access (via commercial on-line provider)....23
2.7.2. Leased Line Internet Connections....24
2.7.2.1. OPAC Gateways....25
2.7.2.2. Local Area Network (LAN) Access....25
2.8. HARDWARE AND SOFTWARE COMPONENTS....26
2.8.1. Minimum Hardware Requirements....28
2.8.2. Additional Hardware Equipment....29
2.8.3. Software Requirements....31
2.9. SUMMARY....31

3. BASIC TOPOLOGY OF THE INTERNET....32

3.1. INTERNET SERVICE PROVIDERS (ISPS)....36
3.1.1. Difference Between ISP And On-Line Service....40
3.2. NAPS-NETWORK ACCESS POINTS....40
3.3. VBNS-VERY HIGH SPEED BACKBONE NETWORK SERVICE....41
3.4. THE INTERNET AS AN ORGANIZATION....42
3.5. SUMMARY....44

4. THE ROAD TO PRESENT: INTERNET HISTORY....45

4.1. PACKET SWITCHING AND THE EXPERIMENTAL NETWORK....45
4.1.1. The Advanced Research Projects Agency Network (ARPANET)....46
4.1.2. Internetting And Standardization....47
4.1.3. Transition From ARPANET To NSFNET: Increased U.S. Government Interest In Networking....48
4.2. NSFNET BACKBONE....49
4.2.1. Expansion Need....49
4.3. NSFNET: COMMERCIALIZATION, PRIVATIZATION, AND DECOMMISSION....50
4.3.1. Commercialization....51
4.3.2. Privatization....53
4.3.3. Decommission....54
4.4. THE INTEGRATED NETWORK AND WHAT'S AHEAD....55

5. END-USER'S INTERNET DEMAND FOR ACCESS AND USAGE....58

5.1. INTERNET EXTERNALITIES....59
5.2. END-USER'S INTERNET DEMAND FOR ACCESS....61
5.3. END-USER'S ACCESS COSTS....62
5.4. SOME STATISTICS ON INTERNET ACCESS....63
5.4.1. Exponential Growth....64

5.4.2. Number Of Users On The Internet...65
5.4.3. Access Locations...67
5.4.4. Computer Type...69
5.5. END USER'S DEMAND FOR INTERNET USAGE...69
5.6. END-USER'S USAGE COSTS...76
5.7. SOME STATISTICS ON USAGE...77
5.7.1. Access Speed...77
5.7.2. Uses Of The Internet...78
5.7.2.1. Internet Tool Usage By % Of Total Internet Users 16+ Using The Internet In The Past 24 Hours...........78
5.7.2.2. Internet Tool Usage By % Of Total Users 16+ Using The Internet Over The Last 24 Hours......................79
5.8. SUMMARY...80

6. HOW INTERNET PROVISION SHOULD BE PRICED...82

6.1. BACKGROUND...82
6.2. ISP'S ACCESS COSTS...84
6.2.1. ISP's Installation Costs...85
6.2.2. Customer-Activation Costs...86
6.3. ISP'S USAGE COSTS...86
6.3.1. Maintenance Costs...87
6.3.2. Network Load Costs...87
6.4. SOME ISP STATISTICS...89
6.5. MOTIVATION FOR TRAFFIC-BASED PRICING...89
6.5.1. Traffic-Sensitive Pricing For End-Users...91
6.5.2. Traffic-Sensitive Pricing Among ISPs (Interconnection)...91
6.5.3. Traffic-Sensitive Pricing For Institutions...92
6.6. AN OPTIMAL PRICING SCHEME...94
6.6.1. An Optimal Pricing Tariff...94
6.7. CURRENT NETWORK LOAD: RANKING THE USE OF CONGESTED RESOURCES...96
6.8. RATE-OF-TRANSFER PRICING...98
6.9. SMART MARKET VS. RATE-OF-TRANSFER PRICING...100
6.10. PROBLEMS OF RATE-OF-TRANSFER PRICING...102
6.11. THE PROBLEMS OF VOLUME DISCOUNTS...103
6.12. WHEN PRICES ARE NOT OPTIMAL...103
6.13. WHEN MARGINAL-COST PRICING REVENUES DO NOT COVER COSTS...104
6.14. SUMMARY...106

7. HOW THE INTERNET IS PRICED AT PRESENT...110

7.1. BACKGROUND...110

7.2. ISP INTERCONNECTION: PROMOTING NON-USAGE SENSITIVE PRICING MECHANISMS ON END-USERS...113
7.3. OTHER CHARACTERISTICS OF THE CURRENT INTERNET 117
7.4. CURRENT PRICING MECHANISMS 117
7.5. ADVANTAGES AND DISADVANTAGES OF NON-TRAFFIC SENSITIVE PRICING SCHEMES 120
7.6. HOW CONGESTION IS PRESENTLY DEALT WITH....120
7.7. SUMMARY 121

8. FROM NON-TRAFFIC-SENSITIVE TO TRAFFIC-SENSITIVE PRICING MECHANISMS: THE TRANSITION....123

8.1. IMPLEMENTATION OF AN OPTIMAL PRICING SCHEME....128
8.2. PROBLEMS OF THIS PRICING SCHEME 131
8.3. AN ALTERNATIVE TO DEALING WITH CONGESTION 131
8.4. SUMMARY 132

9. CONCLUSIONS AND FURTHER AREAS OF RESEARCH....133

APPENDIX A: THE U.S. TELECOMMUNICATIONS SECTOR....136

FACTORS PROMOTING COMPETITION IN TODAY'S TELECOMMUNICATIONS MARKET 137

APPENDIX B: INTERNET SERVICE PROVIDER (ISP) PEERING....139

COMPARING SKA PEERING MODEL WITH THE FINANCIAL SETTLEMENTS STRUCTURE....140
AN EXAMPLE OF AN SKA MODEL: THE CIX INTERCONNECTION MODEL 141
ISP PEERING PROBLEMS: THE CIX-ANS DISPUTE....142

APPENDIX C: CHARACTERISTICS OF THE INTERNET PROTOCOL VERSION 6 (IPV6)......144

THE IMPORTANT FEATURES OF IPV6 INCLUDE:....144

10. GLOSSARY....147

11. BIBLIOGRAPHY....167

TABLE OF FIGURES

FIGURE 1. POINT-TO-POINT VS. PEER-TO-PEER CONNECTION 12
FIGURE 2. BANDWIDTH RATE OF TRANSMISSION 13
FIGURE 3. DATA TRANSPORT SKETCH 17
FIGURE 4. INTERNET SERVICE PROVIDER (ISP) NETWORK WITH A DIAL-UP CONNECTION 27
FIGURE 5. NEW NSFNET BACKBONE (APRIL 1995) 35
FIGURE 6. MCI BACKBONE—THE LARGEST BACKBONE (JANUARY 1996) 36
FIGURE 7. BASIC INTERNET TOPOLOGY 38
FIGURE 8. VERY HIGH-SPEED BACKBONE (VBNS) 42
FIGURE 9. OBTAINING THE DEMAND FOR ACCESS FOR THE USER'S OWN INITIATED USAGE 62
FIGURE 10. EXPONENTIAL GROWTH OF INTERNET HOSTS 65
FIGURE 11. INTERNET ACCESS METHOD 69
FIGURE 12. MAP OF INTERNET NETWORKS (PLACE OF REGISTRATION) IN THE U.S 74
FIGURE 13. OBTAINING THE DEMAND FOR USAGE 76
FIGURE 14. INTERNET TOOL USAGE IN THE PAST 24 HOURS 79
FIGURE 15. INTERNET TOOL USAGE OVER THE LAST 24 HOURS 80
FIGURE 16. TOP FIVE ISPS RANKED BY NUMBER OF SUBSCRIBERS (IN THOUSANDS) 89
FIGURE 17. TYPES OF INTERNET TOOLS 100
FIGURE 18. NONOPTIMAL PRICING OF ACCESS AND USAGE 104
FIGURE 19. AN INDIVIDUAL'S INTERNET ACCESS DEMAND 106
FIGURE 20. LAYOUT OF AN OPTIMAL PRICING MECHANISM 129

LIST OF TABLES

TABLE 1. 1996 INTERNET SOCIETY'S REVENUE SOURCES....44
TABLE 2. ACCESS LOCATIONS....68
TABLE 3. ACCESS SPEED....77

1. Introduction

> **"[E]xponenetial growth is going to exacerbate the need for understanding the economics of these systems and how to deal with charging for use of resources."**
>
> **(Vinton G. Cerf, *The On-Line User's Encyclopedia*, p. 532.)**

As a result of the initiative of the U.S. government, the rapid fall of technological barriers, and the more active role of industry, global access to the Internet is becoming much more widespread. It is speculated that within the next decade a sizable fraction of homes in the industrialized world will have Internet access. In the United States, this widespread connectivity rests on two pillars. One pillar is the government's double-plan of supporting the idea of public access to the Internet by increasing Internet connectivity to schools, libraries, hospitals, etc., and its backbone capacity-expansion projects compensating for the increase in demand of Internet access. The other pillar is the vast increase in commercial provision of Internet access. The deployment of this revolutionary technology will undoubtedly produce major changes in people's lifestyles in the coming years. Further, the most significant benefits will be realized when the Internet and other networks, such as telephony, cable, satellite, and radio finally merge, forming an integrated network such as the one envisioned in the concept of the Information Superhighway. This thesis will focus on the development of one of these networks: The Internet in the United States.

The Internet is an enormous computer network for *the transmission of digitized data*. Given present technology, the Internet is used as: An educational tool, a business tool, a political tool, and an entertainment tool, among others. From this set of uses, more categories are plausible as human creativity expands. Nevertheless, the fundamental purpose will still be the same: a tool for communication.

From an economic perspective, the Internet is a valuable resource that has been gradually made available to the public. It started as a network reserved for scientific research and with time has turned into a very powerful computer network for the public community. Paul Baran, first advocate of this network, envisioned the early Internet as a possible "non-existent public utility." Many experts have followed suit, comparing the Internet to public utilities and developing similar schemes of operation and management. The truth is that the Internet began after a government initiative, and through another government initiative, in 1990, it started a commercialization and privatization stage. Yet, today, the market structure of the Internet is not clearly defined, but this market in many ways resembles that of the telephone network.

Because, at present demand exceeds supply, the major problem the Internet faces is congestion. For a short or long period, congestion will be present in parts of the Internet at one point or another. Different tools designed to fight this problem have been used but with no observable results so far. There is an inability of technological resources to control congestion, an inefficient non-usage based pricing mechanism that makes usage costless for the demanders, and a set of social forces resisting what economic theory would predict to be an appropriate pricing scheme. Appropriate pricing—traffic-based pricing based on the user's rate of transfer requested—may help straighten out three problems (all related to congestion) of today's Internet: (1) The efficient allocation of scarce resources; (2) the extant free-rider problem on the Internet, where low-usage users are subsidizing high-usage users; and (3) the exclusion cost that one user imposes on another during congestion times.

Economic theory suggests that individuals must face actual costs of resources in order to have an appropriate pricing mechanism. A specific type of traffic-based pricing (to be described later) will allow suppliers to efficiently recover costs for capacity expansion. Economic efficiency and network efficiency will be obtained by allocating resources according to the value users place on their requests. Users who place a higher value on the packets requested should pay a higher price. This practice should help control the above-mentioned problems.

Pricing mechanisms are, however, constrained by two issues: Technology and the policy of Universal Service. On the one hand, current technology does not allow the charging of users on a rate-of-transfer basis. In a rate-of-transfer pricing mechanism users would be charged by the value they place on rate of transfer of their packets. Packets from different applications can obtain both the desired or required rates of transfer and the required type of service (ranging from elastic to inelastic) for optimum functioning. And, since a better (faster) service, presumably, will be selected by most users, prices should vary positively with a faster rate of transfer. In this way, scarce resources, such as bandwidth, will be efficiently allocated as this pricing mechanism allows users to select their choice of rate of transfer, according to the valuations they place on their desired service. Additionally, since this type of pricing scheme charges users by their own personal valuations about the speed of the service required, cross-subsidization and the free-rider problem that take place with current pricing schemes can be controlled. Furthermore, since the rate of transfer mechanism works together with an admission control protocol, which admits a certain number of requests, this pricing mechanism will be able to guarantee a rate of transfer satisfying the needs of users.

On the other hand, special interest groups are lobbying in favor of a policy of Universal Service by which every individual should be granted Internet access by means of a subsidy. While the technology issue refers to usage pricing, the Universal Service issue refers to access pricing. Offering Internet access at no cost to authorized individuals both encourages undisciplined use of the Internet and inhibits private investment in that market. If Universal Service is to remain a goal, the least disruptive subsidization scheme should be considered.

The complete picture is, however, more complicated. Pricing may occur at different levels (national, regional, or local) on the Internet market. It would be beyond the scope of this paper to attempt to study the whole market. In this paper I will specifically describe current pricing schemes and proposed pricing solutions and their limitations between the last bridge of the Internet interconnection chain: Local Internet Service Providers (ISPs) and end-users[1]. In so doing, I will attempt to show how the critical pricing problem at this level of the Internet arises. I will briefly refer to the problem of pricing interconnection on the Internet, as interconnecting networks are the key to the Internet. If it were not for interconnection, the Internet would not be nearly as big and useful as it is.

1.1. Problem Statement

The general objective of this study is to describe **an optimal pricing mechanism implemented by local Internet Service Providers (ISPs) on end-users of the Internet**. In order to do so, it is important to first understand the basis on which the Internet works. Only then can one explore the economics of the system and identify its inefficiencies, if any, and hence propose economically sound solutions to the problems.

Specific objectives include:

1. To understand how the Internet came into being, its structure and how it works, why a user may decide to obtain access, and the different types of access to this network;
2. To describe how the demand for Internet access and Internet usage are derived;
3. To indicate the current pricing mechanisms and their characteristics;

---

[1] Local ISPs provide end-users connectivity to the Internet via a dial-up or a leased line and they interconnect with other ISPs for the transport of data. There are two types of end-users, institutions such as, universities and corporations, and individuals at homes and other places with access to the Internet.

4. To present an optimal pricing method to be implemented by local Internet Service Providers (ISPs) on end-users that would allow an efficient allocation of resources; solving the most critical economic problem surrounding today's Internet: Congestion.

1.2. Structure of the Study

Chapters 2 and 3 attempt to set up the basis for an understanding of the Internet's structure. Chapter 4 shows how we have arrived to the current stage. Chapter 5 describes the derivation of the demand for Internet access and usage, which will be the basis for the pricing analysis. Chapter 6 proposes an optimal pricing mechanism of how the Internet should be priced. Chapter 7 presents an overview of how the Internet is priced at present. Chapter 8 describes the transition process to attain the proposed traffic-based pricing scheme. Finally, Chapter 9 presents the conclusions of this study and critical areas of research.

1.3. Methodology

The building of this study followed three basic steps: Review of the literature, collection of data, and interviews. The review of literature allowed me to obtain a general understanding of the Internet, from a technological perspective to an economic perspective. On the technical side, it has shown me that current technical standards would have to change in order for the rate-of-transfer pricing to be instituted. On the economic side, it showed me that the Internet field is in its early theoretical and empirical stages. Several pricing mechanisms have been proposed; none of which has considered the rate of data transfer as proposed in this investigation. I have also learned that there is a great need of empirical studies verifying many of the hypotheses proposed.

The data collected permitted me to observe that the major economic problem of today's Internet is congestion. Also, I can attest to the need for more sources of data, as the vast majority of the sources found were either repetitive or little reliable. I have also found that, as a student, it is very difficult to obtain data to work with. Specifically, I could not find any accessible data on end-user's usage measured on a per-packet, type-of-service, or rate-of-transfer requested basis that would have had allowed me to fit an econometric model for pricing purposes. As a result, this study is at large theoretical. Interviews were conducted to clarify certain aspects of the investigation and obtain more data in certain areas, such as costs; however, much of the data is proprietary and a competitive environment does not favor openness.

1.3.1.Literature Review

The bulk of the literature review for this thesis essentially came from three areas of study—history, economic theory, and telecommunications. Literature on the history of the Internet constituted the first targeted area. Vinton Cerf (1989, 1993, 1995) wrote simple overview papers that provide a general idea of how the Internet has come to its present state. Henry E. Hardy (1993) Master's thesis, *The History of the Net* represents a very informative piece that describes the evolution of the network until 1992. Michael Hauben and Ronda Hauben (1996) present a comprehensive paper (electronic document) on the history of the Internet called *Netizens: On the History and Impact of the Net*, which depicts the Internet up to the beginning of the commercialization stage of the Internet (1990). In addition, Karen Frazer's *NSFNET: Final report 1987-1995*, details the transition period from the government-funded backbone (1987) to its decommission in April 1995, when the Internet is said to have been officially privatized. Supplementary papers and on-line documents were used to corroborate the information used in this study.

Literature on the technical functioning of the Internet was the second area researched. Carl Grzybowski (1994) wrote a paper titled *Internet: Connectivity Handbook*. It presents basic layouts of the different types of access to the Internet. Other documents made available electronically by the National Science Foundation, MERIT, Inc., MCI, Ameritech, Pacific Bell were used to understand the underlying technical infrastructure of the Internet. Personal interviews with Michael Lasher, Assistant Director of Telecommunications Technology, and John Dickinson, Chair of the Department of Computer Science, both from the University of Idaho, helped in this area too. Economic theory was another area targeted for this study. Basic economic theory, for example, *Microeconomic Theory: Basic Principles and Extensions* by Walter Nicholson, 1995 allowed an understanding of the extant and proposed pricing models. Internet pricing models have been written, especially over the last five years, during the commercialization period of the Internet. Bohn, et al. (1993) suggested using a mixture of altruism and quotas to implement priority-based routing. They proposed charging users based on the requested priority. MacKie-Mason and Varian (1992) introduced the idea of the *Smart Market mechanism* that followed the ideas of Bohn, et. al., but accounted for congestion externalities. Cocchi, et al. (1991, 1992) interlaced the idea of efficient pricing policies with network efficiencies concluding that both are compatible under the *type-of-service* (TOS)[2] pricing mechanism. Faulhaber (1992, 1995) analyzes the type of subsidy that should be put in place and he also speculates on the future integrated communications market. Shenker (1993) and MacKie-Mason and Varian (1993, 1994, 1995) have written several papers on pricing issues related to the Internet, including some analyses of advantages and disadvantages of usage-based and flat-fee based models. MacKie-Mason, Murphy, and Murphy (1995) present a very comprehensive paper summarizing previous issues on the role of pricing on the Internet. Additional economic papers contributing to the above ideas were also taken into account. Sandra Schickele (1993) presents a case for the subsidy of the Internet by pointing out its benefits and market failure problems.

Additional important economic literature came from the proceedings of various Internet-related conferences. In particular, the proceedings from the *Public Access to the Internet* conference (held at Harvard University on May 26-27, 1993), and the *Internet Economics Workshop* (held at MIT on March 9-10, 1995) were used. Other important documents were the Request For Comments (RFC) posted on-line on various topics related to the Internet.

---

[2] Type-of-service (TOS) pricing is an interchangeable term with quality-of-service (QOS) pricing.

These economic and technical papers constituted the bulk of the literature used for this study.

A technical understanding of the telecommunication industry was also necessary. For this purpose electronic documents made available by MCI and Sprint as well as books by Wenders (1987) and Taylor (1993) were reviewed. Also, handouts obtained from Professor Wenders contributed to the understanding of the technical aspect. A number of other papers and documents available electronically were reviewed.

1.3.2.Collection of Data

I started collecting data during the Fall of 1993. The gathering of bibliographic references on these topics was completely carried out using Internet services: File Transfer Protocol (FTP), World Wide Web (WWW), Telnet, Gopher, and E-mail. Working papers, documentation on presentations, conference proceedings, RFC (Request for Comments), and library searches became the main source of information. They provided me with an understanding of the current status of the Internet. A great part of the sources was downloaded, saved, and printed. Most of them were found at the computers of other U.S. universities, research centers, and government agencies. Other sources were gathered at the university library or requested from the owners. Assembling the data was largely a personal effort. However, I was always guided by my thesis committee. In addition, I enrolled in a directed-study course on Fall 1995, which refined my understanding of the subject.

1.3.3.Interviews

Interviews were conducted personally and through e-mail. The following people were interviewed:

- John Wenders, Professor of Economics, thesis chairman and telecommunications consultant, University of Idaho.
- Raymond Dacey, Professor of International Business, University of Idaho.
- John Dickinson, Chairman and Professor of Computer Science, University of Idaho.
- David A. Staudt, Associate Program Director, NSFNET, National Science Foundation.
- Michael Lasher, Assistant Director of Telecommunications Technology, University of Idaho.
- Zbigniew J. Tyrlik, Founder and System Administrator, APK Net Ltd., Cleveland, Ohio.

2. Understanding The Internet

Understanding the structure of the Internet is critical for the analysis of pricing. In this chapter, I will describe what the Internet is, how it works, and why a user may want to access the Internet. In so doing, I will show the different ways of connecting to the Internet, the equipment required to connect, and the different services that the Internet can provide. The reader must be advised that many technical terms are introduced here, so he is advised to refer to the Glossary of this study for further details.

2.1. What Is The Internet?

The Internet is often described as:

A government initiative
The network of networks
The most powerful tool ever invented
A spider web
A national obsession
A new democracy
A major cultural event
A cyberspace of living knowledge
A tool with rich information channels.

All the above descriptions have been gradually thrown out, but a more precise description is provided by the Federal Networking Council (FNC). The FNC, an organization that provides Internet coordination for the U.S. government, stated (on October 24, 1995) that

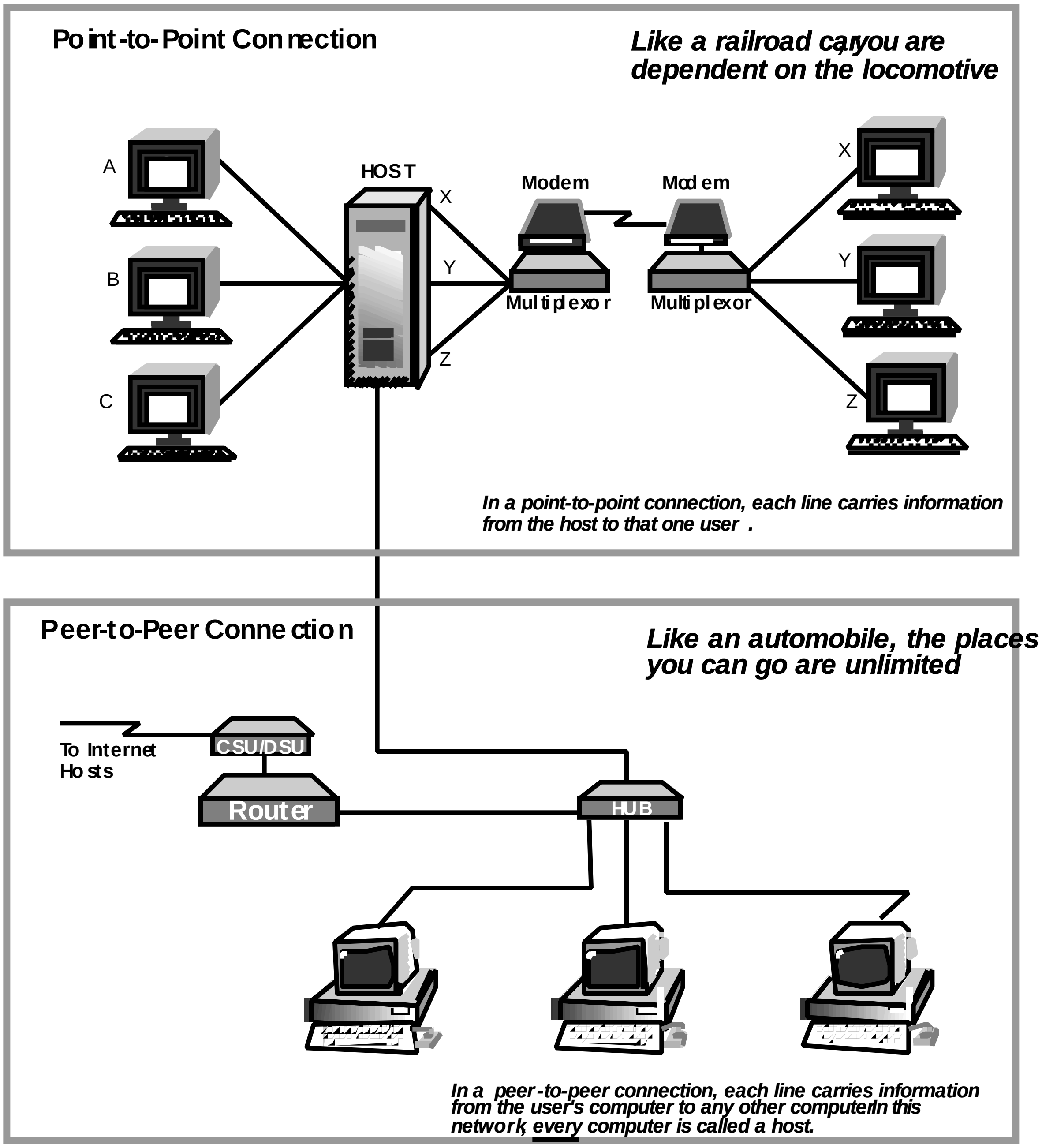
> *Internet* refers to the global information system that: (i) is logically linked together by a globally unique address space based on the Internet Protocol (IP) or its subsequent extensions/follow-ons; (ii) is able to support communications using the Transmission Control Protocol/Internet Protocol (TCP/IP) suite or its subsequent extensions/follow-ons, and/or other IP-compatible protocols; and (iii) provides, uses or makes accessible, either publicly or privately, high level services layered on the communications and related infrastructure described herein.[3]

2.2. How The Internet Works

The Internet is a network of thousands of interconnected networks around the world. Internet Service Providers (ISPs) and phone companies, similar to the post office and transportation systems, work together to deliver data packets around the world. Today, the physical link connecting two networks can be a full time, *point-to-point*, dedicated phone line, or a *peer-to-peer* connection, which are two different types of types of common network arrangements (see Figure 1).

[3] See FNC Resolution at http://www.fnc.gov/Internet_res.html

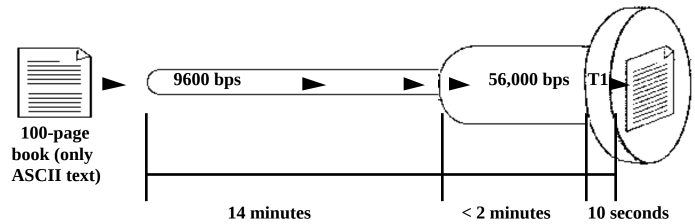
Figure 1. Point-to-Point vs. Peer-to-Peer Connection



Source: Internet: Connectivity Handbook, Carl Grzybowsky, 1994

A digital service with *bandwidths* (speeds) of 56 Kbps or T1 (1.5 Mbps) is the most common type for these dedicated lines (see Figure 2).

Figure 2. Bandwidth Rate of Transmission



Source: NSF: Final report 1987-1995, 1995.

The core technical concept of the Internet is packet switching. A *packet* is a small amount of digitized data ranging between 50 and 1500 characters in length (depending on the capacity of the network), and with a destination address on the header, which allows it to be put in line with packets from other hosts (computers) and sent in turns down a line.[4] Several packets make up a digital file. Packet-switching technology was proposed about 30 years ago by Bolt Beranek and Newman, Inc. (BBN) as an efficient way of using expensive long-distance telecommunications circuits (Hauben and Hauben, 1996). Packet switching is a communications process in which all files are broken up into equal size packets that are transmitted, scattered and then reassembled. In this way, short, medium, and long messages are transferred with minimum delay (Kleinrock, 1976), and the cost of the link is divided among several users.

Packet-switching technology differs from circuit-switching technology—the one used by telephone network. A voice phone line is an analog sound-wave or digital connection that is switched (only in use when you lift the hand set). When you place a call the phone company temporarily connects your circuit to another by setting up a full path from the sender's telephone to the receiver's telephone. Packet-switched technology does not set up the entire

---

[4] A link is the physical copper/fiber optics line connecting to computers or routers.

route at once, rather each packet is independent from each other and is moved in steps from one computer to the next until it gets to the final destination.

The use of this circuit-switched system is the most common method for an individual's computer to *dial up* to an Internet Service Provider (ISP). There are different types of ISPs (local, regional, and national), which will be illustrated in detail in Chapter 3. For now, it is important to note that commercial on-line providers such as, America On-Line, CompuServe, and Prodigy are certainly a type of Internet Provider, but are not Internet Service Providers (ISPs) in the sense I will use the term ISP in this study. For more detail see Section 3.1.

Just like interstate highways, the Internet uses the telephone's high bandwidth service for high speed backbones that interconnect many other major networks. The U.S. government began the Internet project back in the early 60's, funded several networks, and during the 1980's turned it over to the private sector for provision as part of the U.S. National Information Infrastructure Program. The government-funded backbones, like the Advanced Research Projects Network (ARPANET) and the National Science Foundation Network (NSFNET), are restricted to education and research.[5] Commercial use of the Internet is now driving the growth of the Internet with backbone providers that include MCI, Sprint, AT&T, and WilTel.

2.3. Internet Connectivity[6]

Businesses, universities, on-line providers, and ISPs among others install and maintain the equipment and *routers* that direct traffic on the Internet. The number of ISPs has grown significantly over the last five years; at present there are over 3,200 at present.[7] In addition, cable companies, telephone companies and other individuals running operations out of their

---

[5] See Chapter 3 for a more detailed description.
[6] This section is based on Carl Grzybowsky's *Internet: Connectivity Handbook*.
[7] *Boardwatch Magazine*. December 1995.

homes have become Internet providers.[8] For one computer to send data to another it uses an interface that splits the to-be-transferred file (e.g., an e-mail message) in data packets. Like envelopes sent through the post office, each packet *header* contains source and destination information. Each packet goes independently from the source computer (home A) to the local ISP's router (local post office) to the regional ISP's router to the national ISP's router (regional and national post offices), and then finally to the regional ISP's router to the local ISP's router (local post office) closest to the destination computer (home B).

A packet may take different paths to the same destination, depending on traffic. Or, a packet may never need to get to an ISP's router. For example, if the destination user is within the same regional ISP's reach, then packets do not leave to the national ISP's router, unless bad routing occurs. At the destination ISP's router, they are checked for accuracy and reassembled into the file sent. Until now, Internet's network technology design has been very democratic: All packets are created equal. However, attempts are being made to recognize real-time packets and move them along more quickly than others that do not require low latency (intensive-bandwidth use) (Staudt, 1996). Given the type of service required (e.g., e-mail, video, telephone-on-the-Internet) and increasing use of low latency (little delay) applications, packet price discrimination (pricing of packets from different applications) may begin to be more common.

When airborne, all pilots in the world are required to speak English. In the same way, a common standardized communication protocol allows different kinds of computers using different operating systems to communicate with each other over the Internet. This protocol is called TCP/IP. The Transmission Control Protocol (TCP/IP) contains rules to divide large files into smaller packets at the same source and reassemble them later at the destination. The Internet Protocol (IP) contains the rules that manage the addressing scheme to deliver streams

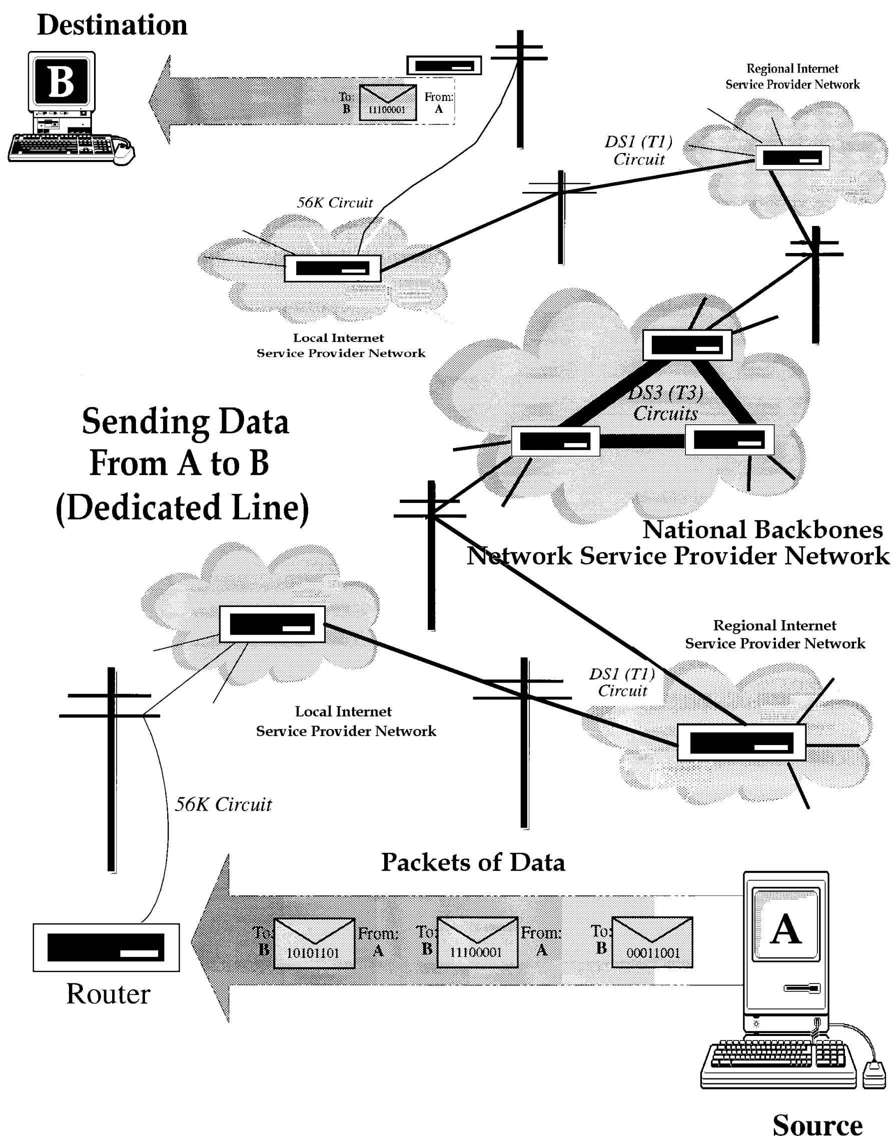
---

[8] Do not confuse the term Internet provider, which refers to all entities providing access to the Internet, with the term Internet Service Provider (ISP), which refers to one type of Internet provider.

of packets from one computer to another.[9] Inside the package are instructions relating to the purpose of the transmission. Like pigeon holes, *ports* identify the type of application, i.e., E-mail, gopher, FTP, HTTP, and Telnet (see Figure 3).

[9] New protocols, such as the ATM/SONET protocol, currently used by big capacity routers, are evolving, but in the end they will all serve the same purpose: connect computers.

Figure 3. Data Transport Sketch



Source: Carl Grzybowsky, 1994.

2.4. Internet Uses

Why would I want to use the Internet? The reason why people use the Internet is because it is an affordable means to other goods such as, a fax, a telephone call, a library, a video game, a magazine, a newspaper, a book, etc.

2.5. Internet Services

What service can I get? An individual can use the Internet to obtain uses that include:

- Mail service
- Advertising
- Contacting people
- News
- Research
- Entertainment.

Therefore, Internet substitutes include the postal service, newspapers, magazines, telephone, libraries, VCRs, movies, and books.

2.6. Common Internet Tools

Despite the Internet's presumed anarchic system, standards or protocols allow the Internet to interconnect different computer brands for various purposes. Internet tools, such as e-mail, FTP, Telnet, Gopher, and WWW use specific protocols that allow computers to be compatible with each other. At present, all these Internet tools use one single protocol—the

TCP/IP—to transit the Internet. Other protocols are used in certain stretches, but TCP/IP is the most common one.

2.6.1.E-Mail

Electronic mail programs may use protocols, such as SMTP (Simple Mail Transfer Protocol), POP (Post Office Protocol), IMAP (Interactive Mail Access Protocol). Electronic mail is a communications tool that allows messages to automatically be passed from one computer user to another, often through computer networks and/or via modems over telephone lines. An electronic message usually begins with several lines of headers giving the name and electronic mail address of the sender and recipient(s), the time and date when it was sent, and the subject. There are many other headers that may get added by different message handling systems during delivery. The message is eventually delivered to the recipient's mailbox; a file which can be read by using a mail reading program. Common e-mail applications include: Pine, Eudora, Pegasus, and Netscape.

2.6.2.File Transfer

File transferring is based on the FTP—File Transfer Protocol—which allows a user on one computer to transfer files to and from another computer over a TCP/IP network. These files can be any type of digitized data: Text, pictures, voice, and video. Common FTP applications include, Win Sock FTP, UNIX FTP, and FSP.

2.6.3.Telnet

Telnet is the Internet standard protocol for remote login (access to a computer), which runs on top of TCP/IP. Some Telnet applications include: Lynx, Unix Telnet, Novell Telnet and Netscape. Telnet accessing allows the user to run commands and perform tasks while not being physically present at the accessed computer.

2.6.4.Newsgroups

Newsgroups option is based on the NNTP (Network News Transfer Protocol), a distributed bulletin board system supported mainly by Unix machines and the people who post and read articles thereon. Originally implemented in 1980 by Steve Bellovin, Jim Ellis, Tom Truscott, and Steve Daniel at Duke University, it has swiftly grown to become international in scope and is now probably the largest decentralized information utility in existence. Some newsgroup applications include: News X, Gopher, Netscape, and Mosaic.

2.6.5.Gopher

Gopher accessing users the Gopher protocol, and it is a document retrieval system that started as a Campus Wide Information System at the University of Minnesota. Despite its graphic limitations, many hosts on the Internet still run Gopher servers which provide a menu of documents. A document may be a plain text file, sound, image, sub-menu or other Gopher object type. It may be stored on another host or may provide the ability to search through certain files for a given concept. Some Gopher applications include: Gopher, Lynx, Netscape, and Mosaic.

2.6.6.World Wide Web

It is uses the HyperText Transfer Protocol (HTTP). World Wide Web (WWW) is an Internet HyperText distributed information retrieval system that originated from the European Laboratory for Particle Physics (CERN) in Geneva, Switzerland. On the WWW everything (documents, menus, indices) is represented to the user as a HyperText object in HyperText Markup Language (HTML) format. HyperText links refer to other documents by their Uniform Resources Locators (URLs) to connect to other places. These can refer to local or remote resources accessible via FTP, Gopher, Telnet or news, as well as those available via the HTTP protocol used to transfer HyperText documents. Some WWW applications include: Lynx, Netscape, and Mosaic.

On of the most interesting features of the World Wide Web is that it supports what is termed *Home Pages*. A home page is a file written on a sub-directory of an Internet host normally dedicated to the support of home pages only. A home page is the entry point for other users to start browsing the information that the owner of that home page has authored by using HTML. From the home page, Internet users are able to follow links to other sources of information that the owner has posted. Typically, home pages contain:

- A formal resume to be accessed by potential employers,
- Pictures of people or interesting objects,
- Links to other favorite sites on the Web to visit,
- Addresses, telephone numbers, or e-mail addresses to access the home page author,
- Other personal information that the author may want to share with home page readers.

2.7. Common Ways To Access The Internet[10]

It is possible to conceptualize Internet access in several ways. One useful approach is to compare dial-up versus leased line connections. Dial-up access is typically associated with relatively low-speed (2400 to 28,800 bits per second) connection available using a modem with an ordinary phone line. However, ISDN (integrated services digital network) high-speed dial access is being implemented in some areas supporting 128,000 bits per second data transfer rates. Leased-line access is most often associated with leased *point-to-point* circuits that connect the user's location (LAN) with the Internet provider's network (ISP). For the following description refer to Figure 4.

---

[10] This section is based on Carl Grzybowsky's Internet: Connectivity Handbook, 1994.

Many new high-speed leased line connections are being introduced that may prove even more cost-effective than simple leased lines. Some of them include, frame relay, SMDS (switched multi-megabyte data service), and ATM (asynchronous transfer mode). These leased services vary greatly in capacity (speed of service), price, and how they are billed. Often there will be multiple communication charges assessed. For example, there may be a fixed monthly fee for the "local loop"—a *point-to-point* connection from the ISP's network to the vendor's service point-of-presence (wholesaler or regional ISP), and an additional monthly charge (fixed or variable) for the switched service (phone company's leased line).

2.7.1.Dial-up Internet Connections

All of the communication alternatives referenced above can be configured to support a variety of Internet service requirements. However, low-speed and relatively low-cost dial-up access is typically associated with supporting either one or a small number of *terminals/workstations*. The equipment necessary for this type of connection includes a computer, a modem, and a phone line. The phone line does not have to be "data grade," a regular voice grade line is usually adequate. Telecommunications software, which is usually bundled with the modem, distributed as freeware or shareware, or available at low cost, is also required. Three types of service often associated with dial-up access are:

2.7.1.1.Terminal Access (via vendor-provided host VT-100 shell accounts).

The computer workstation runs software that allows it to emulate a terminal. The terminal accesses a vendor-provided host computer connected to the Internet. The Internet protocol known as TCP/IP stops at the host. Therefore, it is not very useful. Because TCP/IP does not reach all the way to the computer workstation, the user will not be able to use graphical applications such as, Netscape or Mosaic with a terminal connection. Users also will not be able to use packet video teleconferencing, which requires TCP/IP all the way to the desktop.

However, users will be able to use the WWW through text-based browsers like Lynx or "The Internet Adaptor" (TIA). Furthermore, users will be able to use the entire Internet suite of tools: Electronic mail, Telnet, FTP, Gopher, WWW, and others. Pricing options include both a flat fee with unlimited connect time and a monthly charge plus connect time services.

2.7.1.2.Workstation SLIP (Serial Line Internet Protocol) or PPP (Point to Point Protocol) Access.

These types of connections give a computer workstation its own *Internet address.* This makes the computer an *Internet host* running TCP/IP (for the duration of the connection[11]). Unlike terminal access, SLIP and PPP allow users to use graphical Internet tools such as video teleconferencing and multimedia access to the WWW. SLIP and PPP connections have become more accessible for Windows or Macintosh users.

It is possible for multiple workstations on a local area network (LAN) to share a SLIP or PPP dial-up connection, allowing several users to access the Internet simultaneously. With this connection, more than a handful of concurrent users can slow network speed to the point where it is of dubious value. Pricing of SLIP and PPP connections is similar to that of terminal accounts, using both flat fee with unlimited connect time and monthly charge plus connect time models.

2.7.1.3.Internet Gateway Access (via commercial on-line provider).

Some commercial on-line services (e.g., America On-Line, Prodigy, or CompuServe) offer users a graphical interface (dial-up connection) through a terminal connection to the Internet. Sometimes while using this access type, users have only an e-mail gateway. In other

[11] Unlike a direct connection, where a computer is a permanent Internet host, through SLIP/PPP dial-up a computer turns into a temporary Internet host.

instances, users may be able to use the service provider's Gopher, but not use other, off-site gophers around the world. Or perhaps users will not have access to the WWW or FTP. Most of these types of dial-ups will not provide TCP/IP all the way to the desktop, so packet video and other applications requiring TCP/IP will not work in this configuration. As ISDN services become available, dial-up connectivity will be able to support some of the more robust connection alternatives described below.

2.7.2.Leased Line Internet Connections

Leased line access describes a variety of communication services. This type of connection typically requires a dedicated *point-to-point* communications circuit from the service provider's network (ISP) to the user's LAN.[12] The user pays the ISP a monthly fee for this line plus initial installation charges. Typically, the ISP will get in charge of paying any additional fees to the phone company on behalf of the user, so the user needs to deal only with the ISP. The lines come in various *speeds* or *bandwidth*, such as 56K (56,000 bits per second) or T1 (1.5 million bits per second). The user will generally need some special CPE (or customer premise equipment), called a *CSU/DSU*, to use this service. A special type of communications port may also be required for the library's computer or router.

This type of connectivity will almost always include TCP/IP or similar protocols, which will allow the user to run sophisticated software, use packet video, and operate Gopher or WWW tools, which would reside in his own machine. Generally there are no connect time charges, but this is not always the case. Computers connected to the user's local area network (LAN) can use the Internet at the same time. Two common types of connectivity associated with leased line service are:

---

[12] Recall the set-up of a LAN is similar to that of an ISP, except that it is only the ISP network that connects directly to the Internet (through a wholesaler or regional ISP).

2.7.2.1.OPAC Gateways

A number of on-line public access catalog (OPAC) vendors now include an Internet gateway as part of the OPAC. For example, a user at an OPAC terminal might encounter a menu of Internet services such as Telnet, FTP, Gopher sites, or other applications would appear. The user then selects the desired application from the menu. As with terminal *shell accounts*, users are limited to text-based Internet tools (TCP/IP is usually not passed to the desktop). When graphical user interfaces (GUI) are offered, they are typically proprietary to the OPAC vendor.

Developments in this type of connectivity are occurring rapidly. Many new applications and gateways are appearing as of this writing. To some degree Internet-based OPAC gateways are a “turnkey” solution (providing both software and hardware) for large users that may want connectivity and services without managing the connection directly themselves. Use of this type of connectivity, however, makes the user very dependent on the vendor and the degree to which the vendor will upgrade and support the system for the specific needs of the user.

2.7.2.2.Local Area Network (LAN) Access.

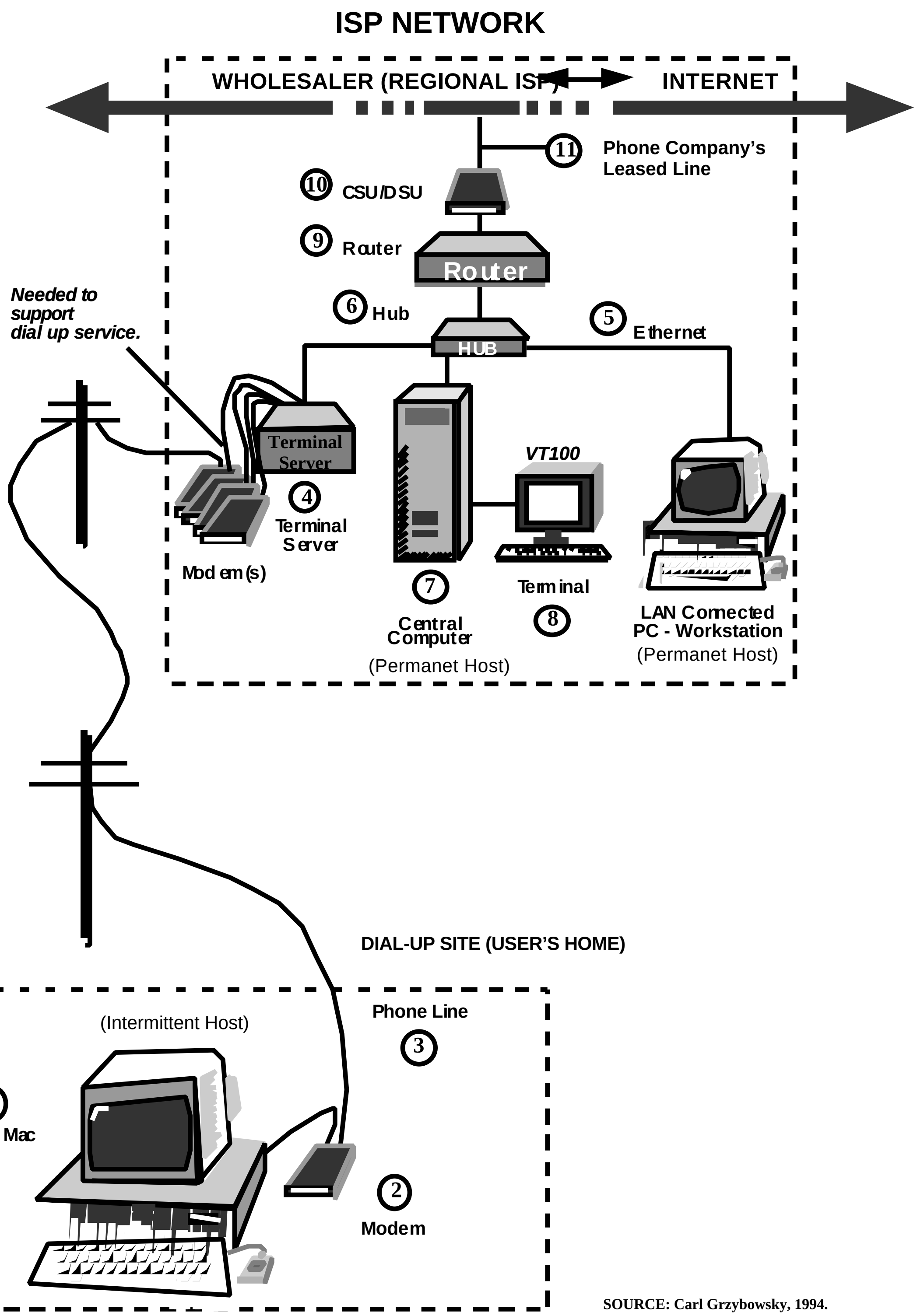
Other large users have installed LANs to provide computer workstations with shared access to CD-ROMs. A user can install an Internet gateway router to connect these LANs to the Internet. Computer workstations can be configured to support TCP/IP and will be able to access the full range of Internet services via the LAN’s gateway router. This configuration would also be capable of supporting Internet Gopher or WWW servers—making their content available not only to directly to the connected patrons but to remotely (dial-up) connected Internet subscribers as well.

2.8. Hardware and Software Components

From here on, I will focus the description only on the types of connections between an individual (end-user) and an Internet provider (ISP). The hardware and software required for a connection depends on the type of access selected by the user. They mainly enjoy two types of Internet access: A dial-up connection and a dedicated line connection. These methods will be detailed in the next section. However, it is worth noting that while dial-up connection links a user (computer-phone line-modem) to the ISP's network; a dedicated or direct connection links a user's Local Area Network (LAN) to the Internet. The latter allows larger speeds of transfer than the former.

If an individual selects a dial-up connection, then minimum hardware is required. If an individual decides to be an ISP or connect his LAN to the Internet then additional hardware is required. Here I will show some basic set-ups for these two types of connections. Notice that the set-up of a user's LAN is similar to that of an ISP. In fact, this is the reason why I will argue that businesses and universities are a type of Internet provider. See Figure 4 for an illustration.

Figure 4. Internet Service Provider (ISP) Network with a Dial-up Connection



SOURCE: Carl Grzybowsky, 1994.

2.8.1.Minimum Hardware Requirements

The following minimum computer hardware is necessary for the dial-up connection.

1. **PERSONAL COMPUTER**

   The computer consists of a monitor, Central Process Unit (CPU), keyboard and often a mouse. Most common personal computers connected to the Internet are IBM-compatible and Macintosh, which can be connected in two ways: (1) Remote (dial-up), via a modem and phone line, or (2) direct, via a network interface card. A computer directly connected to the ISP network is a *permanent host.* A computer remotely connected to the ISP network is an *intermittent host*, as it's on only during dial-up. For example, the computers at a university's computer lab are often permanent hosts, and an individual user connecting through a modem from his home is an intermittent host.

2. **MODEM**

   Modulator/demodulator: An electronic device for converting between serial data from a computer and an audio signal suitable for transmission over telephone lines. In one scheme the audio signal is composed of silence (no data) or one of two frequencies representing 0 and 1. Modems are devices that allow computers to make a call and send data by sound waves over standard voice grade telephone lines. A pair of modems is needed for computers to communicate. Modem speeds vary and the connection speed is determined by the slowest modem of the pair.

3. **PHONE LINE**

   Just like any other call made, long distance charges apply if the ISP is not in within the user's local area.

2.8.2. Additional Hardware Equipment

As mentioned, if a user decides to connect his network (LAN) to the Internet or if he himself is an ISP, additional hardware is necessary. The following is a description of this additional hardware.

4. TERMINAL SERVER

If the central site (ISP) wants to accommodate dial-up connections (PCs at the user's home), a terminal (remote access) server is used. There is a great variety in remote access equipment, from dedicated boxes to others that integrate the function via a PC or router.

5. ETHERNET

It is the protocol that allows computers within a LAN to function. The LAN cabling type defines the network. Older versions of Ethernet use coaxial cable that loops from one computer to another, while other Ethernet connections use twisted-pair cable (similar to phone wire and also known as 10 base T). Newer Ethernet connections use fiber.

6. HUB

A hub is a switching device for networks that run a wire for each workstation and for networks that use a twisted-pair wiring system. Each device can be plugged and unplugged into the network without affecting other computers since they are independently connected to the hub.

7. CENTRAL COMPUTER

A centralized storage and processing machine can run applications that serve other workstations in the ISP network or LAN. Traditional services include on-line catalog, circulation, and backup. Internet services managed by the central computer include: terminal gateway access, Domain Name Service, Post Office, Usenet, Gopher, and WWW servers.

8. TERMINALS

An electronic device for entering data into a computer or a communications system and displaying data received. These simple devices are dependent on the single server to which they are attached. They are not stand-alone computers, like PCs, as they use the central computer to function.

9. Router

The device that connects networks together and passes packets among them is called a router. Internet routers are commercial products, and many thousands of routers currently exist within the Internet. As the name indicates, routers deal with routing, addressing, management of telecommunications links, and other issues of operation. In particular, routers maintain routing tables, which direct packets to reach a specific network (Kleinrock, 1976). As a traffic cop, the router identifies network segments and routes data packets to their destinations.

10. CSU/DSU

Channel service unit/data service unit: The CSU/DSU takes data from the router and translates it into the format needed by the telephone company leased line to connect to the Internet and viceversa. That is, the Channel Service Unit (CSU) is used to terminate a T-1, DS-1 or DS-0 (1.54 Mbps, 64 Kbps, 56 Kbps) digital circuit. The Data Service Unit (DSU) terminates the data circuit and converts the customer's data stream into a bi-polar format for transmission. Instead of using a row of multiplexors and analog modems, the CSU/DSU relays the data streams into a single digital connection.

11. Leased Line

A leased line is provided by the telephone company to connect the ISP network with the wholesale provider (regional ISP). A wholesale provider connects the ISP or LAN to the Internet.

2.8.3. Software Requirements

Depending on the user's needs different applications are necessary. Typically, Internet multimedia-applications, such as Netscape, are enough for most users. By itself, a browser such as Netscape allows: E-mail, FTP, Gopher, Newsgroups, pictures, and simple text usage. So, buying and installing the application are the first steps. The software requirements to run a LAN or an ISP network are different (more expensive) than those to connect through a dial-up connection.

2.9. Summary

The Internet is a technology providing goods and services that are substitutes of other telecommunication technologies, such as fax, telephone, and cable. Users need to determine the capacity of their Internet connection according to their expected needs. In doing so, they will identify the hardware requirements for their connection. Dial-up and leased-line connections are the two more common types of end-user Internet access choice. The latter allows faster speeds of transfer than the former, and is thus more expensive in both dimensions access and usage. Internet Service Providers (ISPs) allow Internet connectivity for users. There are three main types of ISPs: Local, regional, and national. Local ISPs are the ones directly providing Internet connectivity of Internet demanders (end-users).

3. Basic Topology Of The Internet

In this chapter I will describe the basic interconnection levels that make up the Internet. I will introduce the concept of backbone, Internet Service Provider (ISP), Network Access Provider (NAP), very high-speed Backbone Network Service (vBNS), and will briefly refer to the Internet as an organization. Additionally, I will indicate the types of Internet Service Providers (ISPs) and will disclose the difference between ISP and on-line provider.

Defining the topology of the Internet with accuracy is almost impossible. The number of routers and remote (dial-up) users as well as the core structure are continuously changing (Gwertzman, 1995). Moving routers from one place to another is also common. For example, after the National Science Foundation Network (NSFNET[13]) privatization, the Federal Networking Council proposed a plan describing the continued interconnection of government networks, which comprised the moving, creation and/or termination of ISPs (EOWG, 1995). Additionally, when there are no strong attachments between private firms and their Internet providers (ISPs) switching ISPs is easy. This is generally the case, except when long term commitments are signed.

The connection of networks on the Internet is correctly called *Internet cloud*. Despite routing tables and the designation of routes for transmitted packets, when a router is down packets take the second-best route as specified in an algorithm. Therefore, no longer can one tell what lines go where, one just knows that packets go where they are supposed to go. In this *Internet cloud* (see Figure 7), backbones (referred also as national ISPs or Network Service Providers) are the core structure. When the NSFNET (one of the first Internet public backbones) first started in 1987, it consisted of individual 56 Kbps and, around 1990, of T1 leased lines between specific connection points for the NSF-supported regional networks—Seattle, Houston, Pittsburgh, etc.—so that one could see the NSFNET Backbone on a map. With the need for expansion, many new links were added, which made it very difficult to plot the new NSFNET backbone structure in detail. In the United States, the national ISPs—Sprint, MCI, AGIS, AT&T, and WilTel—have developed their own national infrastructures so that they can serve as backbones. Thus, there is not a single backbone anymore.

By 1991, Microwave Communications Inc. (MCI), International Business Machines (IBM), and the Michigan Education and Research Internet Network (MERIT), won the NSF backbone award[14] to expand the Internet. Their upgrade was successful, as MCI had built up the infrastructure to the point that they just hooked the NSF connection points and the first

---

[11] See next chapter for more details.
[14] "[T]he National Science Foundation shall upgrade the National Science Foundation funded network, assist regional networks to upgrade their capabilities, and provide other Federal departments and agencies the

national Internet backbone—the Advanced Research Projects Agency Network (ARPANET) into their infrastructure. As a result, the NSF began calling it the NSFNET Backbone. At this time, the Internet backbone was expanded from T1 to T3 capacity also. Then, in May 1993, NSF outlined the basic elements of the current Internet structure in the US:

> The implementation of the architecture includes four separate projects for which proposals are herein invited: One or more Network Access Point (NAP) Managers; a Routing Arbiter (RA) organization; a provider organization for very high-speed Backbone Network Services (vBNS); and a set of Regional Networks which connect client/member institutions and which provide for inter-regional connectivity by connecting to NAPs and/or to Network Service Providers (NSPs) which are connected to NAPs. [15]

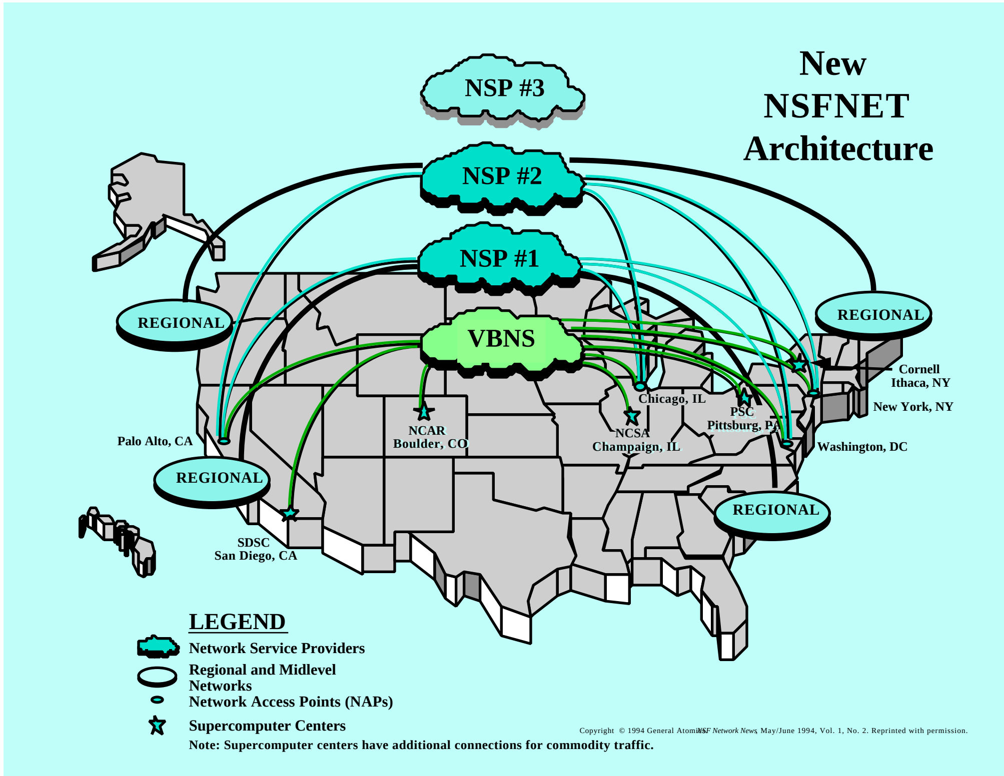
Figure 5 shows a simplified graph of the NSFNET backbone as of April of 1995.

---

opportunity to connect to the National Science Foundation funded network." Public Law 102-194--December 9, 1991. 15 USC 5521 (Section 201).

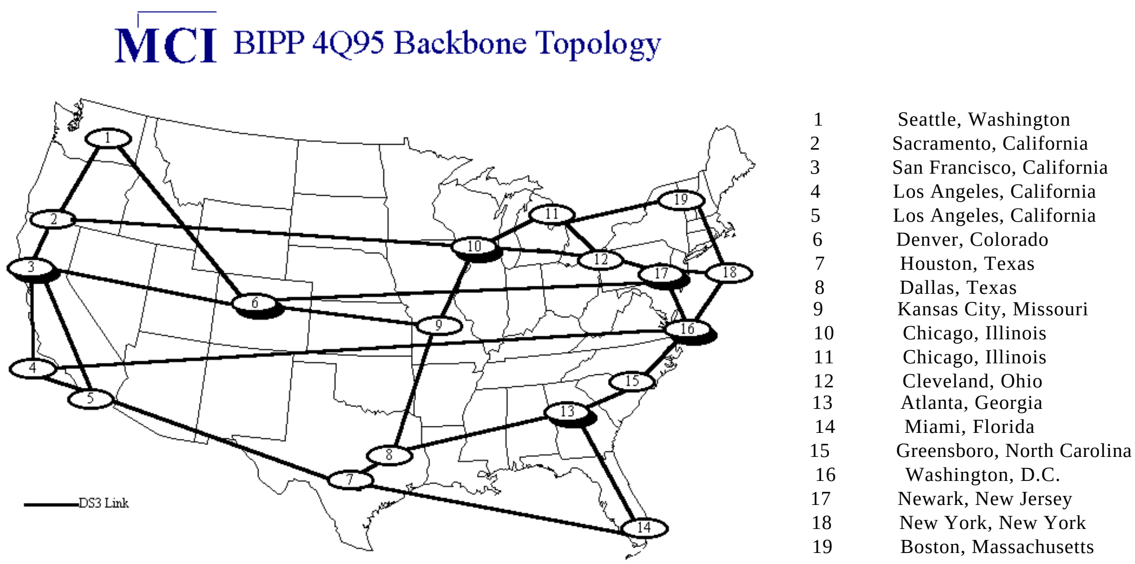
[15] For further details on this issue see: NSF's solicitation # 93-52. Network Access Point Manager, Rounting Arbiter, Regional Network Providers, and very high-speed Backbone Network Service. Program Guideline. National Science Foundation: CISE/NCR. Distributed electronically.

Figure 5. New NSFNET Backbone (April 1995).



Many other backbones have been attached to this infrastructure since. Hence, NFS' as well as backbones from other private and government institutions such as Sprint, WilTel, AT&T, AGIS, ARPANET, and NASA are the core structure of the *Internet cloud*. In early 1996, the largest public backbone was maintained by MCI (see Figure 6).

Figure 6. MCI Backbone—The Largest Backbone (January 1996).

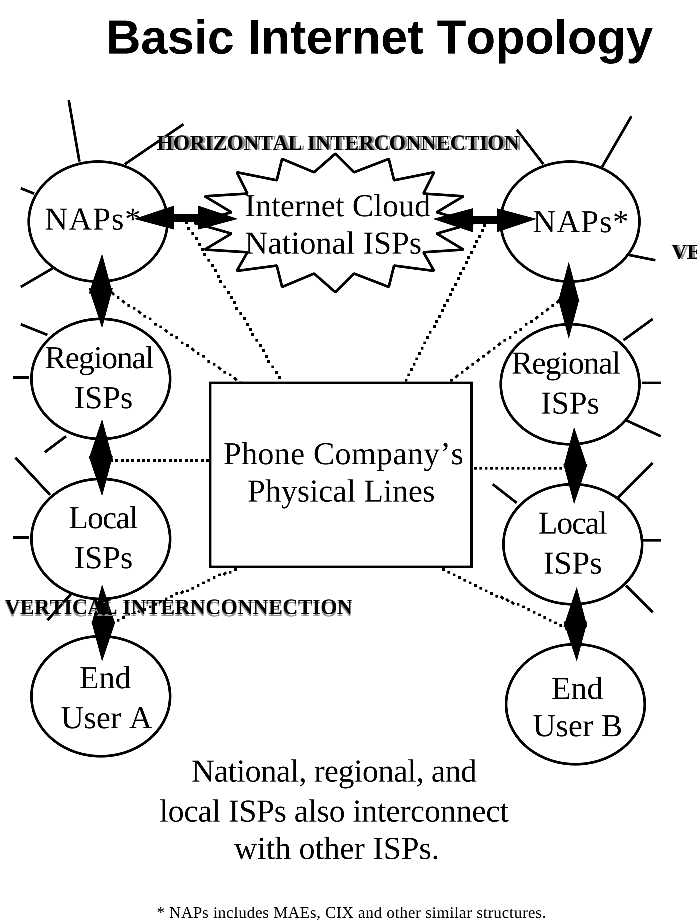


3.1. Internet Service Providers (ISPs)

'Internet Service Provider' (ISP) is a generic term used to denote an entity that provides either Internet access or Internet information services. Recently and with the purpose of taking advantage of the presence of economies of scope and scale, the Internet market has experienced mergers and acquisitions, which has led to the difficulty of classifying ISPs. First, Ramsoy (1995) asserts that the essential criteria for distinguishing among ISPs is the kind of infrastructure they own. Second, Economides and Woroch (1992) distinguish between two types of networks: An integrated network and unintegrated network. I combine their ideas to add that each of these networks has three levels of interconnection—National, regional, and local—which may be distinguished by the kind of infrastructure owned. An integrated network refers to a network that can provide service without need for vertical interconnection with competitors. An unintegrated network in turn refers to a network that needs vertical interconnection with competitors to provide its product (see Figure 7). Horizontal interconnection is generally necessary, though.

While it is typical for national ISPs to concentrate on providing horizontal interconnection, regional and local ISPs (see below) often focus their resources on vertical interconnection. However, with the Telecommunications Act of 1996, long-distance carriers and local phone companies are allowed to enter each other's markets. As a result any ISP (national, regional, or local) could provide end-to-end service without having to lease lines or equipment from competitors—this is what I call an integrated ISP. Most often, though, this is not the case and interconnection agreements with competitors (see Appendix B) play a significant role among ISPs. When an ISP (national, regional or local) does not own the whole end-to-end network and thus has to pay interconnection fees, I call it an unintegrated ISP at its corresponding level, e.g., regional-integrated ISP, or local-integrated ISP. America On-Line (AOL), for example, after many mergers and acquisitions has become an unintegrated-national ISP. MCI, Sprint, are examples of integrated-national ISPs. Also, phone companies play an essential role on the Internet. They allow interconnectivity to happen by providing (leasing to ISPs) the physical lines needed. Interconnection lines come in different speeds or bandwidth sizes, as described in Figure 2. Finally, the end-user is a single individual, or a group of individuals (connecting through a LAN). It is in this layer that the decision of using a dedicated line versus dial-up connection is generally made. End-users pay their fees directly to the ISP and do not deal with the phone company despite it is the latter that provides the lines. Also, end-users typically purchase their own equipment for dial-up access or LAN access and computer peripherals, which needed for connection.

Figure 7. Basic Internet Topology



The first national ISPs were the government-funded ARPANET and the NSFNET, but with the commercialization of the Internet, additional national ISP's were built with funds and technology from telephone companies. Long distance phone companies own national interconnection lines, which are dedicated for Internet purposes. Investments on national ISPs are large.

The first regional ISPs also started with government funds to provide access to research centers, supercomputer labs, and universities. They included NorthWestNet, CERFnet, SURAnet, NYSERnet, BARRnet, NEARnet, among others. Here too the commercialization stage of the Internet (since 1990), allowed private companies to enter this market. Regional

ISPs are the wholesalers of the system, providing interconnection between the national ISPs and local ISPs. Their investment level is medium. Unintegrated-regional ISPs lease lines from national ISPs and sell service to local ISPs and also to end-users. Integrated-regional ISPs would not have to lease any infrastructure to provide end-to-end service, as they would own it. Some other examples of regional ISPs include, Pacific Bell, Ameritech, among others.

The first local ISPs were universities and other research and education institutions that bought service from regional ISPs and provided Internet connectivity for end-users. Today, this market is the largest of the three and there are many local ISPs. These are the retailers of the system. Generally their level of investment is smaller than national and regional ISPs. Unintegrated-local ISPs lease lines from regional ISPs and sell service to end-users. Integrated-local ISPs would not have to lease any infrastructure to provide end-to-end service, as they would own it.[16]

It is interesting to notice that one could ask how did so many people became technically knowledgeable all of a sudden, so as to open these brand new firms called Internet Service Providers. The reality is that it has been the local Bulletin Board System (BBS) operators who have increased the number of ISPs (especially, local ISPs[17]) by changing their name and acquiring new equipment so as to offer Internet access. They also use the Internet as an inexpensive medium to interchange data. BBSs are like On-Line providers, e.g., CompuServe, but they operate at a smaller scale.

Despite my attempt to lay out this topology, there are many exceptions and constant changes. For example, (1) an end-user through the phone company may have a direct line to the national ISP, skipping the local and regional ISPs; (2) a local ISP may also skip a regional

---

[16] From now and to simplify things I will consider all ISPs to be unintegrated, that is, all ISPs are dependant on an interconnecting partner to exchange data: national ISP depend on regional ISPs, who in turn depend on local ISPs to carry data from end-to-end.

[17] Jack Rickard suggests that in December 1995 there were 3, 240 ISPs (local, regional, and national). The Internet by the Numbers. Boardwatch. December 1995. http://www.boardwatch.com/mag/95/dec/bwm1.htm

ISP and connect directly to the national ISP; and (3) a user may connect directly to the regional ISP. That is, if the destination user is beyond the vertical interconnection scheme, then the data needs to go through the national ISPs infrastructure to reach the end-user; otherwise, it may not need to leave the regional ISP reach.

This complex structure is what describes the Internet. An important remark is that on the Internet market there are numerous players that play the role of both suppliers and consumers, e.g., regional and local ISPs, which serve end-users buying from their suppliers. Full suppliers are the national ISPs and full consumers are the individual users. Also, while the national ISP's market type may be that of a monopoly or oligopoly, the regional and local ISP's market structure is more competitive.

3.1.1.Difference Between ISP And On-Line Service

A person can have dial-up access to the Internet in two ways. Some people have *direct, dial-up, or both types of* access through an Internet Service Provider (ISP) like Pipeline, Netcom, or Alternet. These ISPs typically offer no additional information services to their subscribers. Their primary function is to provide a connection to the Internet. Others access the Internet through *indirect dial-up* by using an on-line service like CompuServe, America On-Line or Prodigy. Commercial on-line services make accessing the Internet easiest by offering special features such as parental controls for kids, large databases, and their prices are competitive, but generally they do not offer direct access.

3.2. NAPs-Network Access Points

NAPs are critical interchange points in the Internet architecture. Although they started as a government-funded project, NAP has become a term used to describe other national interconnection points, such as MAEs (Metropolitan Area Ethernets), CIX (Commercial

Interconnection Exchange[18]), and FIX (Federal Interconnection Exchange). A NAP provides a mechanism for national, regional, and local ISPs to interconnect (see Figure 7). It is made up of equipment provided by the interconnecting parties who typically pay a flat fee for interconnection. The only restrictions on traffic flow are those resulting from bilateral agreements between the ISPs or from legal restrictions. ISPs must have at least one bilateral arrangement with another ISP in order to attach to the NAP.

3.3. vBNS-Very High Speed Backbone Network Service

The main idea for a vBNS is to provide for Internet expansion. At the initiative of governmental organizations (mainly the NSF and ARPA) and private groups, the Gigabit Testbed Initiative[19] was started in Spring 1990. It had two goals. The first goal was to develop a wide-area gigabit network to serve the research and education communities. The second goal was to understand the importance of these networks for the research community and the larger society. The gigabit technology uses the ATM/SONET protocol, which refers to transmitting Asynchronous Transfer Mode (ATM) packets over a Synchronous Optical Network (SONET) by transmitting data across fiber optic cable. Several experimental sites already have gigabit speed. These testbed networks were named AURORA, BLANCA, CASA, NECTAR, and VISTAnet. It is the linkage between these networks that comprise the vBNS.

So, in 1995, as part of the National Research and Education Network, the vBNS began stable operation. Its linked supercomputer centers and other sites reserved for meritorious research (the gigabit networks) to the NAPs. The vBNS was funded by NSF and provided by MCI.

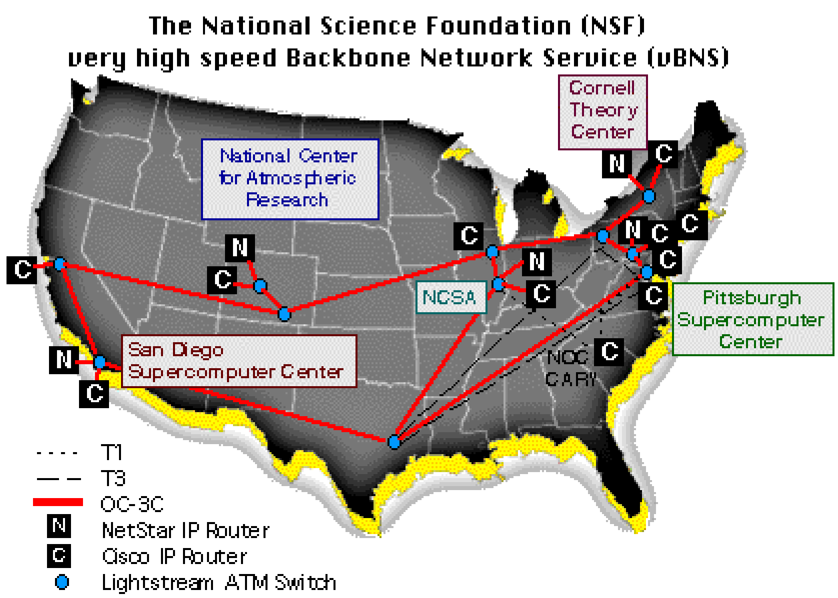
---

[18] See Appendix B.

[19] Most testbeds operate at 2.4 billion bits per second. A typical comprehensive encyclopedia contains about one billion bits (HPCC, 1996).

The vBNS will initially run at 155 Mbps, and in 1996 MCI will increase network speeds to 622 Mbps (MCI, 1995b). (See Figure 8). The High-Performance Computing Act[20] of 1990 states that this research and education network "shall be phased out when commercial networks can meet the networking needs of American researchers." (RFC 1192, 1990). In 1996, NASA will demonstrate 622 Mbps interoperation (HPCC, 1996).

Figure 8. Very High-Speed Backbone (vBNS)



3.4. The Internet As An Organization

The Internet as an organization has been, until recently, mainly guided by the U.S. government—ARPA, DCA and NSF. In 1989, the Federal Research Internet Coordinating Committee (FRICC)—a U.S. Federal government committee composed of the research program managers at the NSF, ARPA, the Department of Energy, and NASA—invested in the Internet for purposes of conducting research activities. Later, FRICC became the Federal

---

[20] For further reading see http://www.hpcc.gov/.

Networking Council (FNC), which had more Federal agencies among its members,[21] and whose mission was:

> To act as a forum for networking collaborations among Federal agencies to meet their research, education, and operational mission goals and to bridge the gap between the advanced networking technologies being developed by research FNC agencies and the ultimate acquisition of mature versions of these technologies from the commercial sector (FNC, 1995).

In the same year, the FNC favored the connection of commercial providers to the so-called academic Internet. The provisions to carry out this plan were developed according to the NSF Acceptable Use Policy for the Internet[22]. During the last five years, under a government plan, the Internet sustained a major change—it was privatized. Vinton Cerf suggests that the privatization project was in part to compensate for the difficulty of NFSNET backbone operators to supply connections to all interested parties (Rams∅y, 1995).

At present, the Internet is coordinated by the Internet Society. The chartering of the Internet Society in 1992 represented an important step in the coordination process of this privatization. It was "recognized that the Society was a critical component necessary to evolve and globalize the Internet and internet technologies and applications, and to enhance their availability and use on the widest possible scale." (ISOC, 1995). The Internet Society is a relatively new, professional, non-profit organization with the general goal of fostering the well-being and continued interest in, and evolution and use of, the Internet. The Internet Society (often abbreviated ISOC) coordinates entities such as, the Internet Architectural Board (IAB), the Internet Engineering Task Force (IETF), and the Internet Research Task Force (IRTF). The funding from the Internet Society comes from several sources as the following figure shows:

---

[21] FNC is currently chaired by Tony Villasenor of NASA. For an update of its current members, see http://www.fnc.gov/FNC_Members.html.

[22] For further detail on the Acceptable Use Policy see ftp://nis.nsf.net/nsfnet/acceptable.use.policy.

Table 1. 1996 Internet Society's Revenue Sources

Organizational Dues	**$825,000**
Individual Dues	**$245,000**
INET96 Conference	**$185,780**
INET96 Workshop	**$31,650**
NDSS 1996	**$59,610**
Publications/Advertising	**$8,000**
Donations	**$5,000**
Other Revenue	**$10,000**
Interest	**$15,000**
Rent	**$36,000**
Total Revenue	**$1,421,040**

Source: Internet Society, 1996.

For a comprehensive study on the organizational structure of the Internet see Tor Ramsϕy's Master's Thesis: *The Governing Structure of the Internet*, MIT, 1995. It is also important to notice that with the Telecommunications Act of 1996, the Federal Communications Commission (FCC) was granted power over ISPs to control Internet provision, which may greatly affect the development of this network.

3.5. Summary

The Internet's main players are national, regional, and local Internet Service Providers (ISP). Interconnection, among these players is what allows for exchange of data on the Internet. NAPs are the most important structure allowing for this interconnection. The level of competition on each of these levels decreases from national to local structures. Phone companies lease lines to regional and local ISPs for transport of data. National ISPs own their own lines as well as networks. The very-high speed Backbone Network Service (vBNS) is the government's project to anticipate for scarcity of bandwidth. The vBNS will be turned into the private sector when necessary. The Internet is coordinated by the Internet Society, a non-profit organization, but the Federal Communications Commission (FCC) has also been granted power to control Internet provision.

4. The Road To Present: Internet History

From a U.S. government project secretly developed more than 30 years ago, today the Internet has become a public world wide used resource that may bring important changes for civilization in years to come. It started with four computers and today there are millions and growing. Its core structure is masterminded at large by private companies. Over five years of private firms' involvement have passed and the Internet, today, is attracting even other industries, such as cable, cellular, telephone, and satellite. The amalgamation of these networks, of which the Internet is just one part, is expected to form the *Information Superhighway,* a network allowing a highly sophisticated method of telecommunications for the future.

4.1. Packet Switching And The Experimental Network

The Internet was envisioned by a group of researchers, led by Paul Baran of the Rand Corporation, at the Advanced Research Projects Agency (ARPA), a division of the U.S.'s Department of Defense.[23] The Internet was built with the idea of a decentralized network (a fishnet, or geodesic network) that could survive nuclear attack. This network was assumed to be unreliable since its conception: Like postcards in the postal system they would have a maximum length and would not necessarily be reliable. All nodes would have an equal status to handle data—there was no hierarchical system. Making use of a technology named packet-switching, files would be divided into packets—with headers showing destination and origin—and each packet would be independent from the others. Packets would be forwarded

---

[23] Paul Baran. (1964). *On Distributed Communications Networks*. RAND Corporation.
Paul Baran is currently involved in two companies (Com21 and Metricom) pursuing what's known as "the last mile problem"--how to get high-speed connections between the Internet and the home. While the first researches on technology using cable hook-ups, the latter focuses on wireless technology..

from one computer to another until they reached their destination. The route a packet would take would be determined by a computer algorithm at each node, as the packet is transferred from one node to another. The specific route of each packet really did not matter as long as it finished in the targeted place. If a node went down, packets would still stay airborne until a route would open to their final destination, or if packet was lost, they would be re-sent by the originator (Cerf, 1995). Baran suggested that the network be viewed as "a new and possibly non-existent public utility . . . designed specifically for the transmission of digital data among a large set of subscribers." (Hauben and Hauben, 1996).

In 1962, J. C. Licklider left Bolt Beranek and Newman Inc. (BBN), a Massachusetts-based firm, to lead the computer science research at ARPA. Rather than concentrating on isolated areas, he guided the research towards a broad framework. Soon, others were attracted by his insights on computer networking. For instance, Larry Roberts another pioneer devoted his time to start a new field: *Computer Networking*[24]. In 1966, based on experiments at the MIT and Dartmouth, Roberts and his colleagues set up a test network between Lincoln Labs in Lexington, Massachusetts and System Development Corp. (SDC) in Santa Monica, California. The main problem he and his team found was that the capacity of the phone lines at the time was insufficient for their goals (IEEE, 1986). By 1968, Bolt Beranek and Newman (BBN) won an award by ARPA to set a small network, which it called Interface Message Processors (IMPs).

4.1.1.The Advanced Research Projects Agency Network (ARPANET)

On September 1, 1969, BBN delivered the first computer for the experimental network involving the Network Measurement Center at the University of California, Los Angeles, the

---

[24] Roberts defined *Computer Networking* as the ability to access one computer from another easily and economically to permit resource sharing.

Network Information Center at the Stanford Research Institute, the Culler-Fried Interactive Mathematics Center at the University of California at Santa Barbara, and the Graphics Program at the University of Utah. ARPANET (ARPA Network) was born. ARPANET would help the researchers work more efficiently and deepen the study on computer networks. Because each site used a different computer brand, some interfacing problems arose. Standards[25] were required to solve these inconveniences. These agreed upon standards were called protocols (Cerf, 1989).

By 1972, both a rapid increase of network traffic and system incompatibility produced problems with the reliability of data transferring. Then, in 1973, in an attempt to achieve uniformity, Vinton G. Cerf[26] outlined a standardized algorithm—the Internet Protocol or IP. ARPANET grew from four nodes in 1969 to roughly one-hundred in 1975. By mid-1975 ARPANET's management was transferred to the Defense Communications Agency (DCA) (Kahn, 1995).

4.1.2.Internetting And Standardization

Internetting refers to the linkage of technologically different packet networks together. Since the early 1970's, having realized that ARPANET was a success, ARPA furthered research on similar networks at sea and ground mobile units. (Satellite Network) SATNET, Packet Radio Network (PRNET), and Ethernet were created. To allow these networks to interconnect to each other, another standard of communications was developed—the Transmission Control Protocol/Internet Protocol (TCP/IP). By 1977 Cerf and his team had successfully connected these three networks together with the ARPANET (Cerf, 1995).

---

[25] Standards are the glue that holds the Internet together. Developing new standards is the responsibility of volunteer members of groups like the Internet Engineering Task Force (IETF).
[26] Graduated at UCLA, Dr. Vinton G. Cerf is currently Senior Vice President, Data Services Division at MCI Telecommunications Corporation.

4.1.3. Transition From ARPANET To NSFNET: Increased U.S. Government Interest In Networking

Both ARPA's and DCA's research funding allowed TCP/IP to become the most robust protocol among others. On January 1, 1983, it became the standard for the ARPANET and other networks replacing older protocols like the Network Control Protocol (NCP) (Kahn, 1995). This transition was obtained by preventing NCP packets to move through the ARPANET, while giving free routes to packets using the TCP/IP system (Cerf, 1993).

At about the same time as the Internet was coming into being, Ethernet local area networks (LANs) were developed. This technology matured quietly, until desktop workstations became available around 1983. Most of these workstations came with Berkeley UNIX (an operating system), which included Internet Protocol (IP) networking software. This created a new demand: Rather than connecting to a single large timesharing computer per site, organizations wanted to connect the ARPANET to their entire local network. This would allow all the computers on that LAN to access ARPANET facilities (Kahn, 1995).

During the early 1980's the National Science Foundation (NSF) started the effort called the Computer Science Network (CSNET) to interconnect the computer science departments around the US. In similar manner, other organizations started building their own networks using the same communications protocols as the ARPANET and perfecting the use of the other protocols. Fiber optics[27] were laid across the United States and elsewhere in unprecedented quantities. In 1984, the Internet Domain Name Services (DNS[28]) was specified. DNS corrected the limitations of the ARPANET host table. Host tables required all hosts addresses to be listed in a single file, kept by a single organization, which limited resources for expansion.

---

[27] As a reference fiber optics were invented in 1955 by it was not until 1976 that they were commercially used.
[28] DNS provides the means by which the name of an Internet host (www.uidaho.edu) is resolved to an IP address (such as 129.101.119.228).

By mid-1980s, several U.S. government agencies already had their own national networks—the Department of Defense, Department of Energy, and NASA, among others. Federal agencies competed among each other regarding networking projects. As a result, network protocols among these were different and machines had trouble communicating with each other. Nevertheless, while the National Science Foundation (NSF) had its interests in Science and Engineering, it had somewhat experience in networking—the CENSER and the 1986 National Science Foundation Network (NSFNET). So, it was NSF staff that envisioned the overall designed of the future Internet at the time (Frizzier, 1995). Between 1988 and 1990 the transition from ARPANET to the new and faster NSFNET backbone took place. ARPANET was decommissioned during 1990.

4.2. NSFNET Backbone

4.2.1.Expansion Need

The first NSFNET backbone service went on-line late in 1986, connecting six sites the five supercomputer centers (Cornell Theory Center, Pittsburgh Supercomputing Center, John Von Neumann Supercomputer Center, National Center for Supercomputer Applications, and the San Diego Supercomputer Center) and the National Center for Atmospheric Research. It included the facilities of the ARPANET as a major part of the structure and had a 56 Kbps speed.

Due to overwhelming demand on the backbone, NSF saw the need for an improved version, one that would link educational and research institutions together to be a part of the National Research and Education Network (NREN[29]) proposal. So, on June 15, 1987, NSF issued a solicitation (NSF 87-37) for management and operation of a new version the NSFNET

[29] To find out the principal objectives of the NREN see: http://www.hpcc.gov/imp95/section.4.3.html. March 1996.

Backbone Network. There were three main components in this proposal: The packet-switching nodes, the circuit switches interconnecting these nodes, and a network management system. This NSFNET backbone had a three-layer structure (Backbone, Regional Networks, and Campus Networks) and ran at a speed of approximately[30] 1.54 Mbps, or T1 (Frizzier, 1995). On November 24, 1987, NSF awarded a contract to Merit Network, Inc. (the Michigan Education and Research Infrastructure Triad), working in partnership with the State of Michigan, MCI and IBM, to manage and operate the NSFNET backbone (Gilster, 1994).

The 56 Kbps NSFNET backbone was decommissioned on July 24, 1988, and the T1 NSFNET backbone began operation in the same month. It had over 170 networks connected and linked 13 sites: Merit, the NSF supercomputer centers, the National Center for Atmospheric Research, and the following regional and mid-level networks: BARRNet, MIDNET, Westnet, NorthWestNet, SESQUINET, and SURAnet. NYSERNet and JVNCnet were also served, as the were part of a supercomputer center (Frizzier, 1995).

By 1989, and once again due to the exploding growth in usage of the NSFNET backbone, the capacity of the network had to be expanded. In just one year a 500% increase (of 500 million packets per month) had taken place. NSF had encouraged use of the network by funding many more universities and colleges through its Connections program.[31]

4.3. NSFNET: Commercialization, Privatization, And Decommission

While commercialization of the Internet refers to the action of permitting commercial users and providers to access and use Internet facilities and services, privatization refers to the

---

[30] It was not until 1989, when the NSFNET backbone ws re-engineered, that full T1 speed was possible. Routers were the bottleneck.

elimination of the federal role in providing or subsidizing network services. Commercialization was achieved by the creation of a market for private vendors. Privatization was attained by shifting the federal subsidy from network providers to smaller users, and was eventually terminated. This last action spurred private sector investment in network services (RFC 1192, 1990).

4.3.1.Commercialization

In the late 1980's, with many devastating congestion problems, the NSF realized that more investment funds were needed to cope with increasing demand of backbone access. So, the regional networks, which until then were government-funded and by agreement were supposed to become self-supporting (Kahn 1995), took on non-academic customers as an additional revenue source. In addition, non-academic organizations willing to pay commercial prices increasingly desired Internet connectivity, but the NSF Acceptable Use Policy restricted backbone usage for academic purposes only. This led to the formation of two privately sponsored Internet carriers: Unix-to-Unix Copy Program Network Technologies (UUNET) and Performance Systems International, Inc., or PSI[32] (Kahn, 1995). The commercialization stage of the Internet hence began.

Around 1989 also, a U.S. Federal government committee, the Federal Networking Committee (FNC), decided that commercial Internet suppliers would officially be allowed to connect to the Internet (the NSFNET backbone). The FNC's provisions were published in the NSF acceptable use policy document[33] (NSF 1990). As a result, an experimental MCI e-mail

---

[31] In 1995, NSF continues its Connections program to support the use of innovative technologies for schools and libraries, and institutions of higher education. It also contributes funding high bandwidth Internet connectivity for qualifying applications.

[32] A for-profit subsidiary of New York Regional Network (NYSERNET).

[33] The NSF agreed to restrict the transportation of packets from the following countries: Bosnia-Hercegovina, Iran, Iraq, Libya, Macedonia, North Korea, Sudan, Syria Arab Republic, Vietnam, and Yugoslavia (i.e. Serbia and Montenegro). However, some NSFNET-sponsored networks may have agreements with other network providers (e.g. CIX or ANS) to exchange traffic with networks in these countries. NSF: Distributed electronically.

system became the first attempt of a commercial firm connecting to the Internet. Shortly, thereafter CompuServe, ATTmail and Sprintmail (Telemail) followed suit. The growing use of the Internet for purposes other than research and education together with the challenges presented by an upgrade of such magnitude, stimulated a transition from public funding to private funding—the beginning of the Internet industry.

As a result, in 1990, Merit, IBM and MCI spun off a new organization known as Advanced Network and Services (ANS), which had the charter to commercialize and upgrade to 45 Mbps the backbone network, providing the NSFNET service along with other customer traffic on the same system. Merit Network's Internet Engineering group provided a Policy Routing Database, routing consultation and management services for the NSFNET, while ANS operated the backbone routers and a Network Operations Center (Ameritech, 1995). So, by 1991 the T3 NSFNET backbone network became operational with 16 nodes all running at 45 Mbps.

Then, in May 1991, ANS extended a for-profit subsidiary called ANS CO+RE (Commercial and Research[34]) Systems. This entity focused its efforts on upgrading the network further with the help of the rest of the private industry, so that the research and education community would benefit from their provision of technology. ANS then joined UUNET Technologies and PSI for commercial service (HPCC, 1995). NSFNET traffic had grown drastically from 195 million packets in August 1988, to almost 24 billion by November 1992 (Gilster, 1994). To help transportation, Sprint, Inc., Pacific Bell, Ameritech, and Metropolitan Fiber Systems[35] augmented the networks around the nation (Kanaley, 1994).

---

[34] ANS CO+RE Systems, Inc., was established by ANS in June 1991 to target the special networking and security needs of the business community. In February 1995, America Online (AOL) acquired the assets of ANS CO+RE Systems, Inc. See *Interactive Age*, March 13, 1995. Issue 210, page 42.

[35] AT&T management was skeptic about the future of the Internet and consequently did not take much initiative (Cerf, 1993). Today, AT&T wants to be a major player on the Net, but it has been the last big communications company to introduce a strategy and product line. For further detail see *Information Week*, Issue 542, page 14. August 28, 1995.

4.3.2.Privatization

Right behind the commercialization came the privatization. For years, the networking community wanted the telephone companies and other for-profit ventures to provide an Internet connection just like they provided a telephone jack. Except for BBN, there were no takers. But with the sudden jump in interest of businesses and the increased number of pro-competitive policies for the market, phone companies and other for-profit organizations changed their attitude.

By 1994, the U.S. national information system was characterized by a rapidly growing and diversifying user population. The information world also experienced pressures for deregulation of telecommunication providers (NRC, 1994). On September 15, 1994, as part of the National Information Infrastructure (NII),[36] the U.S. government announced a plan to privatize the NSFNET backbone and put network development into private hands subject to so-called *market forces*, thereby subordinating an advanced sector of the U.S. economy to a more backward sector (Hauben and Hauben, 1996). With this transition, the government focused its resources towards other highly advanced projects, such as the Gigabit Testbed Network. During the same year, other gigabit networks were added: MAGIC, BAGnet, ATDnet, DREN, ACTS AAI, NYNET, CalREN, NCIH, NLANR, and TBONE (CNRI, 1995).

In an attempt to cope with both the termination of the cooperative agreement between NSF and Merit and the much debated question about commercializing the backbone, NSF decided to allow a greater involvement to the community by releasing its proposal for the new architecture of the backbone to the general public in the summer of 1992. This proposal was based on the ideas of Bob Aiken, Hans-Werner Braun, and Peter Ford, who at the time

---

[36] Vice-President's Al Gore project to put vast amounts of information at people's fingertips. It essentially consist of a seamless web of communications networks, computers, databases, and consumer electronics so that all Americans may benefit from the communications revolution.

worked at the National Science Foundation, the San Diego Supercomputer Center, and the Los Alamos National Laboratory, respectively.

In May of 1993, after a draft solicitation, a request for comments period, and the incorporation of comments, a final solicitation (NSF PR 93-52) was issued. This document requested proposals for four different activities—a very high-speed backbone (vBNS), Network Access Points (NAPs), a Routing Arbiter, and Regional Network Providers (RNPs), and stated that there would be a transition period during which the current NSFNET backbone service would be phased out.

The winners for the award were revealed throughout 1994. Merit and the Information Sciences Institute (ISI) at the University of Southern California would become the Routing Arbiter team. The vBNS award went to MCI. Network Access Point Manager awards were given to Sprint in New York; Metropolitan Fiber Systems Datanet in Washington D.C.; and Bellcore for Chicago (jointly with Ameritech) and California (jointly with Pacific Bell). Awards for regional network providers were given to seventeen regional and mid-level networks.

4.3.3.Decommission

From 1994 through the first half of 1995, Merit and NSF's Networking and Communications Research and Infrastructure division (NCRI) worked closely with the regional and mid-level networks to ensure a smooth transition, helping them to disconnect from the NSFNET. The T1 NSFNET backbone was finally decommissioned on April 30, 1995, and the NSF major funding for the Internet backbone ended. Regional networks terminated their original connections from the *cloud* and established new connections to the network service provider of their choice, and were using their provider's infrastructure. Of the seventeen regionals NSF is supporting on a declining basis (two more years to go), nine chose MCI, six chose Sprint, one chose ANS, and one had two connections, one each to Sprint and MCI. They are free to change providers as they wish; the one choosing ANS switched to MCI last October (Staudt, 1996).

NSF continues to fund some regional networks, but this funding will steadily decrease to zero over four years (NSF PR 95-37, 1995). With federal programs, such as the Telecommunications and Information Infrastructure Assistance Program (TIIAP) and new industry efforts, such as the Pacific Bell plan to wire up schools in its service areas, the so-created distribution of resources will help sustain the dissemination of information around the U.S. However, President Clinton has claimed that funding to extend connectivity to schools, libraries, hospitals and others will have to come from the state and local level (NRC, 1994).

4.4. The Integrated Network And What's Ahead

The NSFNET as a public Internet provider is now gone—it devotes its resources to connecting research and educational institutions only. The NSFNET was initiated to act as a market developer and catalyst for the growth of the Internet, and as the Internet grew the NSFNET became a smaller and smaller part of the Internet. By 1995, the NSFNET had served its purpose and, as planned, NSF phased it out (Staudt, 1996). There is not a single backbone in the post-NSFNET stage.[37] Private companies have built their own backbones, thereby creating a competitive market.

The Internet has become fully commercialized, and a mixture of state and industry ownership. It is evolving in a competitive environment that is growing exponentially (see Chapter 5) causing harmful congestion problems. The Internet is, however, just one part of a larger telecommunications system. One that includes phones, databases, cable, TV, Internet, and other information transferring systems running on a network thousands of times faster

---

[37] It is important to notice that the backbone infrastructure also extends to other parts of the world, with a big network in Europe.

than at present, perhaps on a gigabit network. When operational, this network will provide services to academic and research institutions, government agencies, and businesses. This integrated network is what has been dubbed the *Information Superhighway.*

The Internet application called World Wide Web (WWW) may be the first step towards this integrated network. The WWW was created by Tim Berners-Lee and his team from the European Particle Physics Laboratory (CERN), in 1989. The World Wide Web has become the universe of network-accessible information. It assembles other Internet services, such as Telnet, FTP, Gopher, News, video, voice, and others into one single application (Berners-Lee and Cailliau, 1990). Thus, it may soon integrate other networks in a stable manner. Some companies are already engaged in providing electronic funds transfers on the Internet in support of electronic commerce. Others are exploring the provision of packetized video, video-conferencing, packetized voice, and increasingly secure tools for corporate Internet operations.

It is hard to speculate or say how long it will take to have an integrated network, such as the one depicted by the term *Information Superhighway.* Both the problems of dealing with a continuously increasing Internet community and the lack of the necessary technology may be in the way for quite some time. First, today the Internet allows access to between 20 to 30 million people around the world, of which 18 million are World Wide Web users.[38] Second, despite the great cooperation by the research community to improve Internet services, which will attempt to harmonize telephone, video, cable, etc. in an integrated network, there is a general consensus at present that the necessary technology is still not in place to make this integrated network operational.

---

[38] See GVU's 4th WWW User Survey Home Page at http://www.cc.gatech.edu/gvu/user_surveys/survey-10-1995/. This survey was carried out from October 10, 1995 through November 10, 1995.

Finally, because both Internet backbone's capacity (supply side) and user's Internet access[39] (demand side) are increasing exponentially,[40] one may expect that the Internet and an integrated network will produce unprecedented changes in people's lifestyles worldwide. Thus, the potential *Information Superhighway* may be then considered as one of the three major U.S. government's achievements of this century, coming behind the setting up of the first linear accelerator and the landing of man on the moon!

[39] Usage has also been increasing exponentially. It is important to notice though that usage is barely related to access as the incumbent users themselves (being more knowledgeable than newcomers) are the ones who chiefly make use of the popular, but nonetheless, bandwidth-intensive applications that have caused Internet traffic to proliferate so fast.

[40] See the next chapter.

5. End-User's Internet Demand For Access And Usage

Based on the belief that demand creates its own supply, this chapter will concentrate on end-users Internet demand. End-users are the ultimate consumer of Internet services. They can obtain access through several media, as shown in previous chapters, but two methods are most common: End-users can have a permanent connection (direct Internet connection—leased line) or a temporary connection (indirect Internet connection—dial-up). Whether connecting through a direct or a dial-up connection, an individual setting up an Internet connection is an Internet end-user. For some users, though, demand for Internet access is obtained through an institution (university, business, etc.). In these cases, it is the institution that establishes the contract with the ISP for Internet access and thus the institution is considered the end-user. This chapter will concentrate on the derivation of an end-user's demand for Internet services.

Demand for Internet services has two main components: Demand for access and demand for usage. An individual can demand Internet access to, for example, receive e-mail only; he can also demand Internet usage to send e-mail messages. That is, Internet demand for access is derived from the demand for usage actual or potential, incoming or outgoing. In this chapter I will start exploring how an individual's demand for Internet access and usage is affected by externalities. I will then describe how to obtain the end-user's demand for Internet access and usage. I will also present an end-user's Internet costs and some related statistics on access and usage.

5.1. Internet Externalities

Individuals on the Internet face three types of externalities: Congestion externalities, linkage externalities,[41] and network externalities. Congestion is a negative externality—also known as the *Tragedy of the Commons*. *Congestion externalities* arise from the fact that, during congested times, one user imposes a cost in the form of degradation of service or delays on other users when the first uses the network and viceversa. There are many network resources whose performance suffers when there is overuse: Bandwidth, routers, computer capacity, etc. Users typically use these resources by only paying attention to their own benefits and costs from usage (when priced), but ignoring the congestion, delay, or exclusion costs that they impose on others. With changes in technology, the availability of these resources will vary over time, with periods of scarcity or abundance (MacKie-Mason and Varian, 1994b).

The increased availability of user friendly interfaces, such as Netscape and other multi-media applications, together with the increased awareness about the potential of Internet services (for example, by-passing telephone tolls) has created serious congestion problems on the Internet, which are ultimately borne by the subscribers in terms of delays and degradation of service.

Like the call externality in telecommunications, *linkage externalities* are an important determinant of demand. A communication between two parties is initiated by one of them. The initiator's decision to demand service is guided by his private benefit. This, however, inflicts benefits (or disadvantages) on the other party. Benefits will fall on the second party if the latter sees the connection (e.g., a *talk* connection, an e-mail message, etc.) as useful and desirable. Costs will fall on the second party if he sees the connection as undesirable. So, the net effect could either be positive or negative, depending on whether the benefit or the cost of the connection were dominant. In other words, an individual's decision about making a

[41] Similar to call externalities in the telephony.

connection on the Internet (demand for usage) may not be consistent with maximum social welfare because the individual's benefit does not necessarily coincide with social benefits. However, this effect would be mitigated if the receiver has a device, such as a "caller ID" by which he would be able to filter end-users according to his wishes.

Linkage externalities are presumed to be positive. It can be expected that most people use the Internet for useful things rather than to impute negative externalities on others. For example, there is a potential benefit of setting up a *talk* connection (similar to place a phone call, except that instead of speaking one writes) as both users may mutually benefit from the link. However, in the case of a potential subscriber, he may receive an overwhelming amount of undesired e-mail messages. Then, if he is aware of this, he may not be willing to pay as much or even nothing for accessing the Internet. So far, no studies on these two issues have been published. Most of the time, though, an individual is willing to pay something to have this access available to him. So, similarly to the telephone industry, this option can be termed *option demand.*

Network externalities arise because a new end-user joining the network increases the range of choice open to other members. An individual's decision to join the network is typically guided only by the benefit to that individual of communicating with others on the network. Again, this private benefit differs from the social benefit, as the former includes the benefits of having all other individuals able to communicate with the first. Hence, network externalities can affect demand for access: The number of other subscribers on the network will increase an individual's desire (demand) for access

Let us concentrate on the latter two types of externalities. Because of the benefits that linkage and network externalities bring to all users, it is suggested that those marginal users should be subsidized. For example, in a world of two users, if the recipient's benefit is of the same magnitude as the initiator's, then the marginal social benefit is twice the marginal cost. The optimal policy would be to set price of access equal to half of the marginal cost. Setting this optimal policy has difficulties: (1) measuring the magnitude of those externalities; (2) the possibility of actually identifying them; (3) the likely losses from subsidizing inefficient connections or end-users; (4) the possibilities for internalizing these externalities (Littlechild, 1977); (5) need the existence of a benevolent mechanism with the power to do the right thing but insulated from rent-seeking influences that could do damage (Wenders, 1996).

5.2. End-User's Internet Demand for Access

Internet demand for access is the maximum price an end-user (from home, business, educational or research institutions) would be willing to pay for the right to be connected to the Internet and use it at whatever prices are charged for the various kinds of usage. Knowing that his consumer surplus is the excess of what he would be willing to pay for a certain number of units of consumption over what he actually pays, his demand depends on: (1) the sum of his consumer surplus of the various Internet services he can obtain, which may range from e-mail, to advertising through a web page; (3) the number of users on the network (Internet) who increase the range of choice for each other (network externalities); (2) the value he places on the option to be able to transmit and receive digitized data (option demand); and (3) the value of usage initiated by others (linkage externalities). The important characteristic of Internet demand for access is that it is *binary*: An end-user either has access or he does not.

Figure 9. Obtaining The Demand For Access For The User's Own Initiated Usage

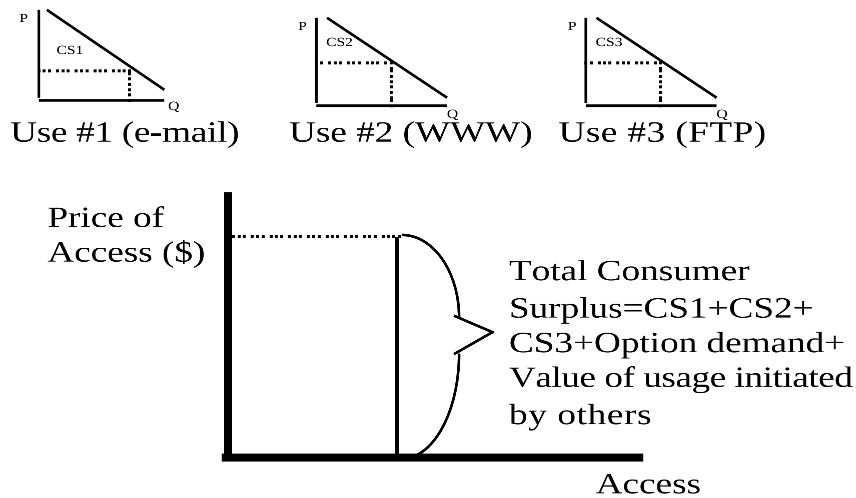


Figure 9 shows the price of access for each end-user. Each end-user has a different consumer surplus for the usage he can get from different Internet applications. The sum of those consumer surpluses, the option demand, and the value of usage initiated by others will give a particular end-user's total demand for access. Notice that the quantities of each application can be measured in time or packets desired per period, but these numbers become important when analyzing usage demand.

5.3. End-User's Access Costs

End-user's access costs refer to those incurred by an individual when obtaining an Internet connection. For this, an end-user incurs in costs of obtaining the computer equipment and the costs of paying for access to the ISP's network (or on-line providers). Both of these are fixed, non-recurring costs, which will chiefly depend on the capacity of the equipment and bandwidth chosen. In comparison, the short-run costs of obtaining the computer equipment (typically over $1,000) are much larger than those of paying to be connected to an ISP's network (around $20 a month).

5.4. Some Statistics On Internet Access

Mark Lottor, who at one time worked at the Defense Data Network (DDN) Network Information Center (now InterNIC[42]), provides the basis for most of the statistics published about the Internet. Specifically, he has published statistics with the number of hosts on the Internet (world-wide) per year since 1981. Lottor shows that there were 6,642,000 hosts on the Internet as of July 1995. Rickard (1995), states that 64% (4,228,207) of those hosts were located within the US.

One important remark has to be made. Hosts are not necessarily end-users. In general, a host is a computer with an IP address, permanently connected to the Internet. On the one hand, when someone uses a dial-up connection, he obtains a temporary IP address and turns into a host for the duration of the connection. On the other hand, a university server (host) is permanently connected to the Internet and has a permanent IP address (but these hosts are continuously shared by several users). Lottor's data was collected by computing the number of hosts that would be connected to the Internet at any particular point in time (both temporary and permanent addresses). As is apparent, this computation requires some adjustment factor to account for multi-users: Those sharing hosts. This is particularly important when observing that the *.edu* domain (i.e. colleges and universities), for example, is the second largest users group on the Internet. These institutions typically have a ratio of around 10 students per Internet host. However, as it is shown below this adjustment factor seems to be much lower for the U.S. as a whole.

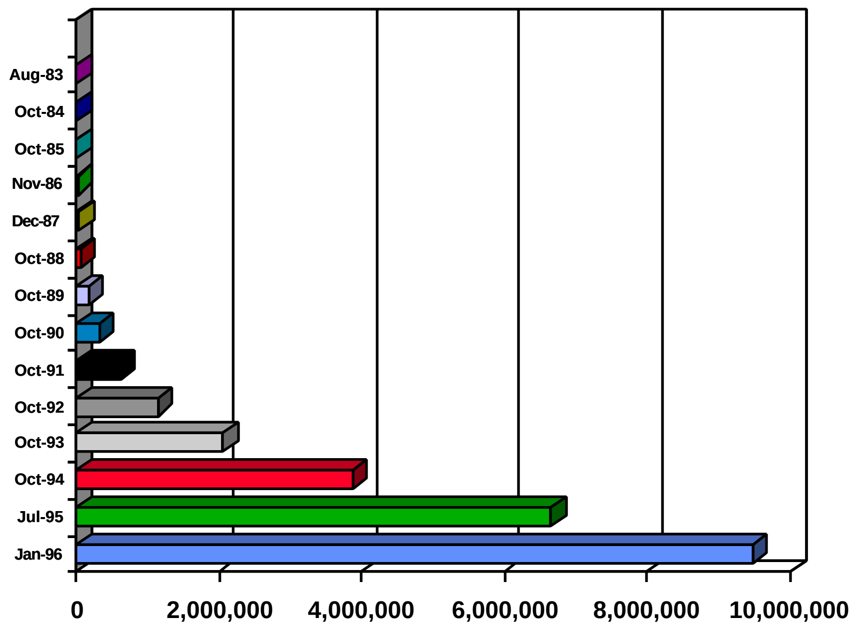
---

[42] InterNIC--Internet Network Information Center.

5.4.1.Exponential Growth

Lottor's data, when plotted, allows a display of the exponential growth that everyone talks about (see Figure 10) and which is mainly the result of: The industry—the wide adoption of the TCP protocol, the reduction in price of microprocessors, the wide availability of fiber optics, government promotion, and technology innovations (such as Domain Name Service).

Figure 10. Exponential Growth Of Internet Hosts.



Source: Mark Lottor, Network Wizards 1996.

5.4.2.Number Of Users On The Internet

O'Reilly's (1995) is one of the most accurate ones as far determining the number of users on the Internet. The results were published in September 1995. The survey's methodology is quite accurate because:

- It is based on a random sample
- It is a large scale sample
- It is independent of other Internet derived information

This survey concludes that there are 5.8 million adults in the US accessing the Internet without using a commercial on-line service. Therefore, with the estimate of the number of users on the Internet in the US (O'Reilly's survey), and the estimate of the number of hosts (or computers) on the Internet in the US (Lottor's data), and by noticing that these two estimates were computed at similar periods in time, one can determine the *ratio* of users per host on the Internet in the US. The result gives 1.37 users per computer in the US. To expand this estimate worldwide, we would multiply this factor by the number of hosts on the Internet world-wide (6,642,000) to give 9,111,096 users on the Internet world-wide as of July 1995. There are some flaws to this calculation, but it turns out that is one of the most accurate ones. For example, it is expected that the number of users per computer outside there is much greater than inside the US. Also, the number of users connected through a commercial on-line service is not included here (Rickard, 1995).

An additional speculative prediction that one can make is about the number of users on the Internet as of January of 1996. By using the same users/computer ratio (1.37) and multiplying it by the number of host from Lottor's data (9,472,000), the conclusion is that there were 12,976,640 users on the Internet world-wide as of January 1996. However, because of the unknown and potentially unlimited numbers of multi-user computers and network or application gateways, as well as the existence of innumerable temporarily connected, non-advertised, or *firewall*[43] protected machines, it is not possible to establish a complete total or correlate any of this information with the number of end users with great accuracy.

5.4.3.Access Locations

Another survey, the Nielsen's study[44] was published on December 1995. This study found that while home Internet connections are important, locations other than home were significant sources of access. Reportedly, 62% of the users said that they had access at

---

[43] See glossary for explanation.
[44] See http://www.nielsenmedia.com.

home. Interestingly, the research showed that 54% of the users had an Internet connection at work and 30% had access at school (as the sample consisted of persons 16 years and older, at school is synonymous with college in most instances). On average, each Internet user had 1.4 different types of access locations.

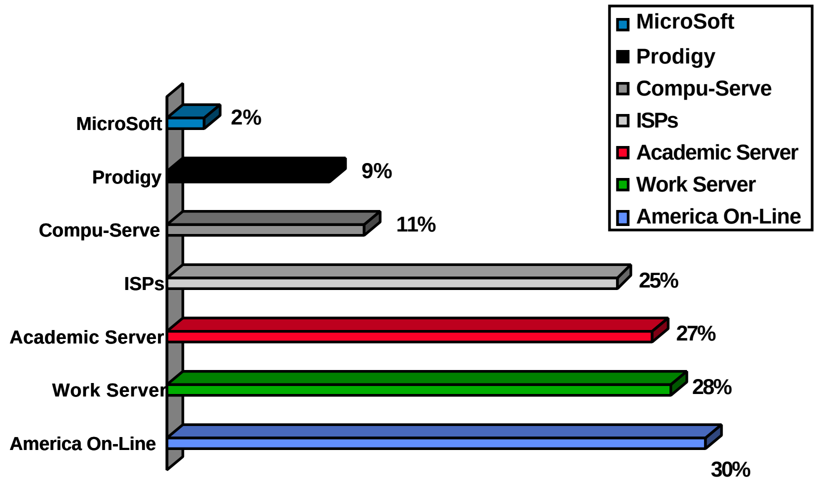
Table 2. Access Locations

	Nielsen	**FIND/SVP**
Home	62%	69%
Work	54%	47%
School	30%	21%

Yet another study conducted by FIND/SVP,[45] a large information services provider, reported that nearly half of all Internet users (46%) use a commercial on-line service to access the Internet. Figure 11 below shows that America On-line was the preferred Internet access method, although nearly a quarter of all users say they have more than one way to get on the Net:

[45] See http://etrg.findsvp.com.

Figure 11. Internet Access Method[46]



Source: FIND/SVP, 1995.

5.4.4.Computer Type

GVU4[47] reports that the most common Internet access platform is Windows (62%), followed by Macintosh (21%) and Unix (7%). By comparison, the NPD Group, Inc.[48] reports that Windows-based PCs now account for 59% of household PCs, while a Macintosh is found in 14% of all U.S. homes.

5.5. End User's Demand For Internet Usage

Bandwidth refers to the maximum speed of transfer of the physical line between two connection points, e.g., a user and an ISP. Since Internet usage refers to the amount of data

---

[46] The sum of percentages is over 100 because some surveyed users had more than one method of access.
[47] See http://www.cc.gatech.edu/gvu/user_surveys/User_Survey_Home.html.
[48] See http://www.npd.com

transferred through a line, which is mainly constrained by its bandwidth capacity,[49] these two concepts are very close related. There are two dimensions to bandwidth: Maximum bandwidth and actual bandwidth or rate of transfer. While the maximum amount of data transferred (per unit of time) at a time through a line depends on bandwidth capacity, the actual amount of data transferred (per unit of time) is the rate of transfer. This difference arises because of the presence of bottlenecks at several points along a user's connection and the fact that a line is shared by several users. This difference is going to be emphasized in the next chapter. For illustration purposes, let us consider usage as simply the amount of data transferred per unit-of-time, i.e., rate of transfer.

Internet demand for usage is an end-user's demand for the different applications that the Internet can support. Recall that Internet services included: E-mailing, advertising (through a home page for instance), researching, and entertainment, among others. It follows that analyzing the demand for usage is complicated, as there are many applications that a user may demand. Because of the similarities with the telephone network, one can use analogous procedures to those used in the telephone network to measure Internet usage. Wenders (1987) points out that an individual's demand for usage for the telephone depends on:

- The price of usage.
- The prices of substitutes for usage.
- The prices of complements to usage.
- The subscriber's income.
- The number of subscribers on the network.
- The tastes and preferences of the individual.

[49] Rate of transfer will also vary with connection bottlenecks. These bottlenecks are difficult to identify, but nevertheless important for an optimal pricing policy, as it will explained later.

Besides these categories one must also consider that demand for usage on the Internet depends on the type of application being used. That is, a user who uses a graphic browser, such as Netscape, will impose heavier usage demands than one only using a text browser, such as Lynx. Internet usage should be measured in terms of packets transferred (sent and received) by each user, and not by the number of minutes connected to the network, as the latter overlooks completely the issue of usage in terms of packets transferred.

First, the price of usage is the price paid for the various forms of usage. At present, for most providers, those who use a non-usage sensitive pricing scheme, this price is zero. The price of usage should, however, vary depending on the type of application demanded. For instance, unlike the telephone network, transmission of video, voice, and other special applications may be charged at different rates. While on the Internet, different Internet services require different speeds of transmission, quality, etc., on the telephone network a phone call is basically a phone call anywhere.

Second, the price (or cost) of substitutes (i.e., postal service, phone calls, faxing, etc.) for Internet usage will also influence the individual's demand for usage. But, unlike the telephone network, Internet services substitutes are diverse and vary with each service. For example, a substitute for an e-mail message could be a phone call, a fax, or a *talk* connection. A substitute for using the application CUSeeMe (similar to teleconferencing), may be teleconferencing itself, and so on for other Internet services. These all have a cost to the user and a quality dimension, both of which must be considered in choosing among the various alternatives. Thus, changes in the price or quality of Internet substitutes will affect the individual demand for Internet usage. It is important to notice that none of the substitutes puts the relevant materials in a computer ready for computer-based analysis. As such, there are poor substitutes (Dacey, 1996).

Third, since goods or services that are used together are complementary goods, demand for usage will additionally depend on the costs of the equipment needed when obtaining a dedicated line or a dial-up connection. Note that not all costs are dollar costs. Internet service's quality dimension serves to illustrate another aspect of demand for Internet usage—a user may also face the cost of congestion. Congestion costs appear in the form of delays and degradation of service that other users impose on the consumer when the network is congested. This is the congestion externality discussed in the previous section.

Applications with time-saving characteristics are complementary to Internet services. For example, commercial on-line providers sell user-friendly interfaces that allow faster FTP methods (no need to type so many commands), easier mechanisms to download files (no need to decode pictures), more attractive methods to obtain current news (screen icons), etc. than other applications (say, gopher). Mosaic and Netscape are examples of these types of applications. They are widely available to educational institutions, but priced to everyone else.

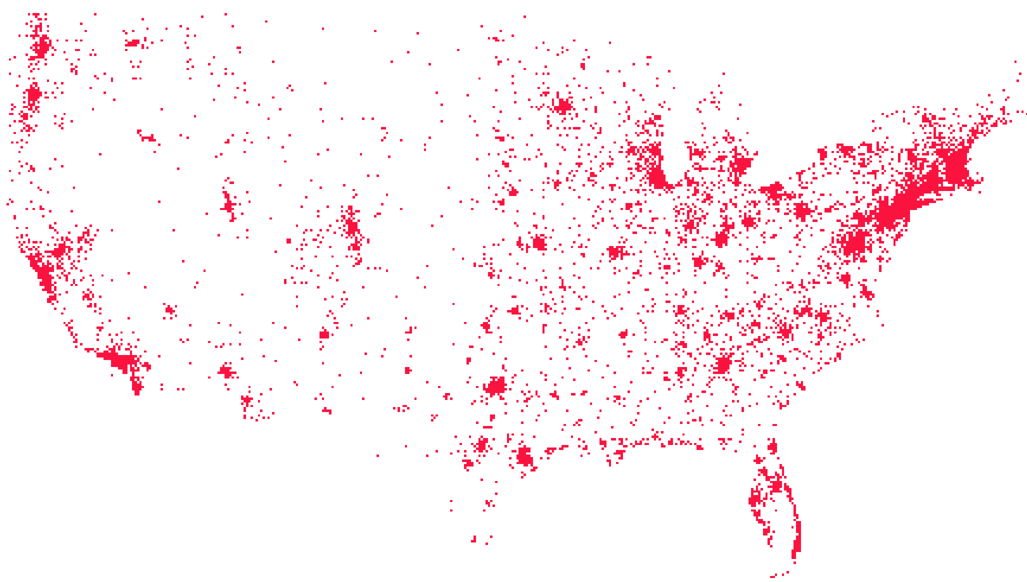
Time is directly complementary to both Internet usage and indirectly to many of its substitutes. This, as Wenders (1987) points out for telephone usage, makes it difficult to predict how a rise in the user's value of time may affect Internet usage, since it will raise the price of its substitutes too. In contrast to the telephone system, empirical endeavor suggests that a rise in the value of a user's time will decrease actual Internet usage, because of congestion, but demand for usage may increase. For example, given the current congestion problems surrounding the Internet, the time spent in doing research through, say, a database service (Lexis), may be less than the time spent in carrying it out with Internet applications, such as Netscape; thus, reducing actual Internet usage. In other words, to the individual user, an increase in congestion is like a price increase (or a decline in quality).

Fourth, income affects the demand of Internet usage by broadening the user's opportunities for the consumption of all goods (normal goods), including Internet usage. But, attempting to classify Internet usage as a normal, inferior or neutral good (service) may be difficult: No studies have been carried out with this specific purpose and the widely used flat-fee pricing scheme does not help for this purpose. Given some usage-based Internet pricing schemes and previous experience with similar networks (telephone), one can state that Internet usage is likely to rise with income. Nevertheless, it may very well be that actual Internet usage decreases with income, mainly due to congestion. For example, in some cases, a user may not care to pay the over-priced cost of a long distance phone call (substitute) to send, say, a voice message, instead of waiting for the Internet service obtained through access, such as a *talk* connection or the CUSeeMe application to function usefully—at acceptable rates. The substitution effect may be stronger once the prices of Internet service substitutes move down.

Moreover, one has to be aware of the possible correlation between income and the user's valuation of time (Wenders, 1987). Most often income comes from the wage rate of work. The latter in turn affects the user's valuation placed on time. As a result, it could be tricky trying to sort out the cause and the effect of a rise in income and a rise in a value of time through the greater time content of the substitutes to Internet usage. Overall, classifying Internet access as an inferior or normal good depends on the type of the Internet service under analysis and whether the substitution or the income effects is stronger.

Fifth, network externalities—both the effect of the number of other subscribers and the information added by them on an individual's demand for usage—causes, other things equal, the more users on the network, the more use the average user will make of the network, and the greater will be the benefits derived from that usage. As pointed out by Wenders (1987), one must also notice that the distribution of the users is of fundamental importance for determining the possibility of internalizing these externalities. The distribution of users refers to whether or not other users lie within the network (community) of interest of the individual in question. Being the network of networks, the Internet connects millions of users around the world, a large portion in the US (around 40-50%, see Figure 12). Much of the traffic, among users crosses the different US backbones and is interchanged at the NAPs. Thus, even users in distant communities (states, countries) may affect an individual user's rate of transfer and hence demand for access and usage.[50] Distance is not an obstacle on the Internet.

Figure 12. Map Of Internet Networks (Place Of Registration) In The U.S.



Source: Internet Business Center, 1996.

Finally, users' preferences account for all other variables not included in the above. Preferences are difficult to estimate and thus use in a demand-estimation model. Some of these categories include, demographics and fashion. However, once the type of preference is identified, say, demographics, [51] then it may be explicitly utilized.

---

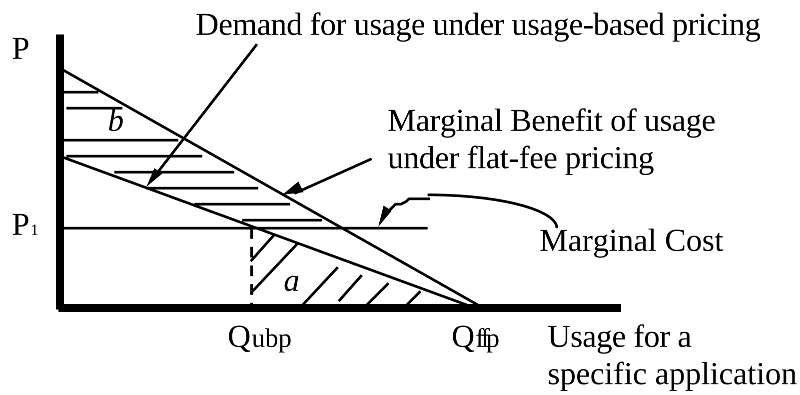
[50] Notice that at times packets from users abroad the US have been transported by US backbones: because using US backbones allowed faster rates of transfer than using an own country's backbone. There have been several attempts to stop this, as outside countries were also congesting the US backbone.

[51] Below I present some statistics about the demographics of the users of the Internet.

With the above aspects in mind, one can now display the graph of the individual demand of Internet usage. Under the common non-traffic sensitive Internet pricing schemes (see Chapter 7), a demand curve reflecting Internet usage is an illusion. Instead, there is a marginal valuation curve, which shows how a user values each unit of usage under flat-fee pricing, but this curve will not show how much usage will be taken when in fact usage is priced. The reason is that demand-side transaction costs under usage-based pricing will reduce the valuation of usage to the subscriber (Wenders, 1987). It must be emphasized, though, that demand for Internet usage, will differ from one Internet service to another. For example, the Internet usage demand for e-mail differs from that of FTP. Just as with the demand for access, it is also necessary to identify demand for usage by specifying the type of Internet service to be analyzed.

Thus, for a specific Internet service, such as e-mail, the demand curve for Internet usage—under a usage-sensitive pricing scheme—will lie below the marginal valuation curve of usage under flat-fee pricing. So, when comparing, i.e., making welfare judgments, between a usage-based pricing scheme and flat-fee pricing, the gains or losses in the demand-side transaction costs should be considered.

Figure 13. Obtaining The Demand For Usage.



Qubp--Quantity Demanded under a usage-based pricing scheme
Qffp-- Quantity Demanded under flat-fee pricing scheme.

In reference to Figure 13, under a non-usage sensitive pricing, the price of usage is zero, so Q_{ffp} is demanded. When a usage-based pricing scheme is utilized and price P_1 is charged, then quantity Q_{ubp} is demanded. The total loss is the sum of the loss in value of repressed number of packets transmitted for a specific application (area a) and the reduced value of transmitted packets due to demand-side transactions costs (area b), i.e., the total loss is the sum of a + b.

5.6. End-User's Usage Costs

End-user's usage costs depend on the will of the user. He decides when to demand Internet services, what kind, and at what quality. These costs continuously vary and depend on his decision of usage as well as on the prices imposed by the ISP network (or on-line provider).

These costs should be usage sensitive, type-of-service sensitive, and current network load sensitive, so as to precisely bill each user for the usage imposed on the network.[52]

5.7. Some Statistics On Usage

Nielsen's study shows that on average, all persons 16+ in US and Canada who have used the Internet in the past three months used it for 5 hours and 28 minutes per week. (FIND/SVP reports that users spend an average of 6.6 hours per week on the Internet). Nielsen reports that 66% of users of the Internet are male. Males tend to use the Internet with both greater frequency and duration than females, accounting for approximately 77% of the total usage.

5.7.1.Access Speed

International Data Corp. predicts that 19 million modems will be sold this year. Since most of these will be of the 28.8-Kbps variety, the make-up of the modem installed base will change significantly this year. Today, about 64% of all modem users surf at 14.4-Kbps according to IDC/Link Resources.[53]

Table 3. Access Speed

Source and Speed	% of Users
ISDN (128 Kbps)	1.4%
LAN Modem/10Mbps	12.5%
28800 Modem	14.8%
14400 Modem	63.5%
9600 or less Modem	3.5%
Other/Don't know/N.A.	4.3%
Total	100%

Source: International Data Corp.

---

[52] See next chapter for further detail.
[53] See http://www.idcresearch.com.

5.7.2.Uses Of The Internet

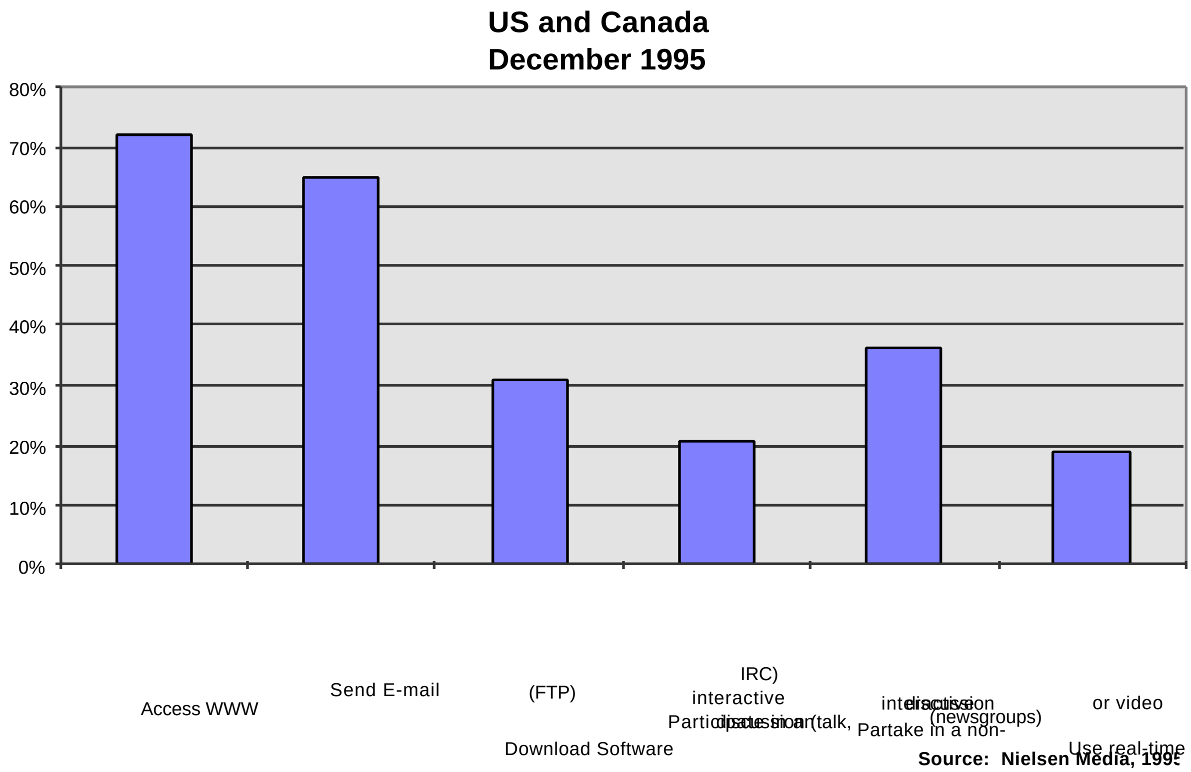
For those persons using the Internet in the past 24 hours, accessing the World Wide Web exceeded E-mail in use. The percentage of Internet users in the past three months who indicated frequent use of Internet applications other than E-mail was also considerable.

5.7.2.1.Internet Tool Usage By % Of Total Internet Users 16+ Using The Internet In The Past 24 Hours

Figure 14 shows the most common used Internet tools over a period of 24 hours. This graph helps get an idea of such usage by recurrent as opposed to one-time users. It shows that while the World Wide Web is the most used Internet tool, interactive talk (IRC[54]) and real-time video are the least used. This may reflect also the popularity and market penetration of Internet tools. The World Wide Web tool utilizes a multi-purpose protocol (HTTP) that allows options such as, transfer of e-mail messages, downloading of files (pictures, video clips, sound clips, text, etc.). IRC and real-time video tools use a specific type of protocol and not many people know they exist. Furthermore, in the case of real-time video, more expensive computer equipment is generally required for their deployment.

[54] IRC stands for Internet Relay Chat.

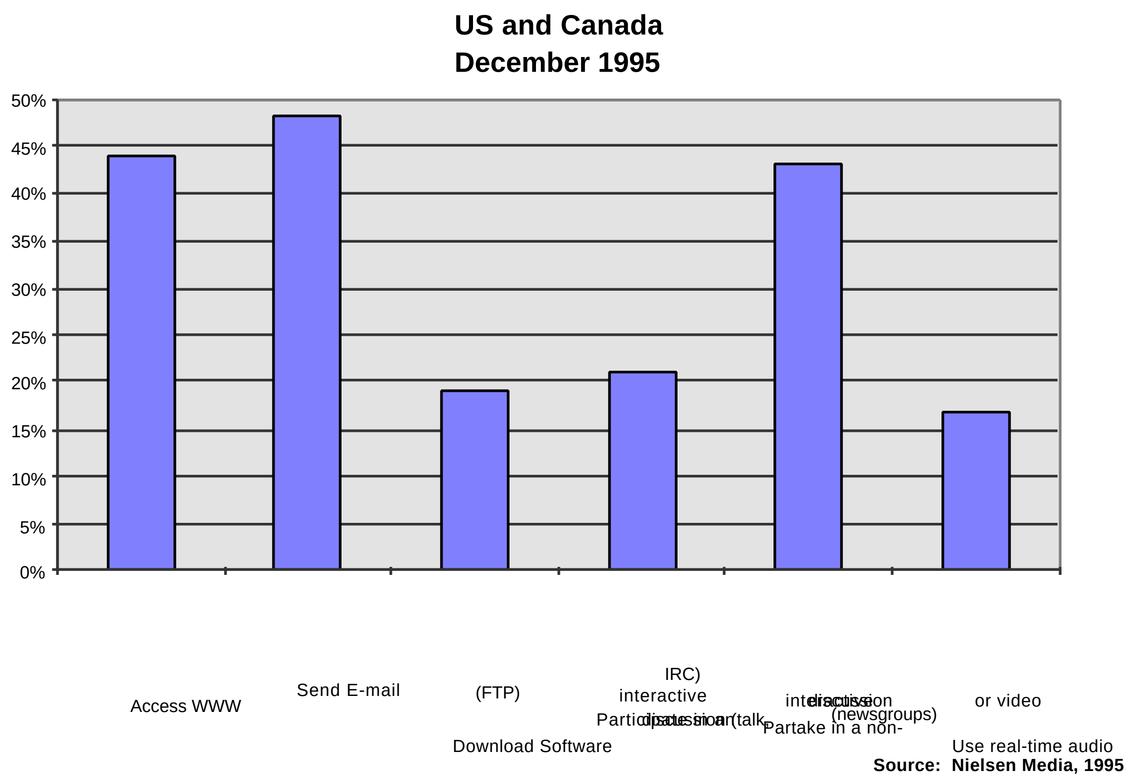
Figure 14. Internet Tool Usage In The Past 24 Hours.



5.7.2.2.Internet Tool Usage By % Of Total Users 16+ Using The Internet Over The Last 24 Hours

Figure 15 shows the most common used Internet tools during a time between the past 24 hours and before. That is, the use of Internet tools by one-time users. The most eye-catching features is that e-mail is the most common tool used by this type of users. This may be explained by the fact that some users get Internet access only to receive e-mail and do not spend much time using other Internet tools. It is also interesting to see that non-interactive discussion (newsgroups) attracts many people. By comparing the two graphs however, one can notice that non-interactive discussion is not used on a continuous basis. Users may be highly attracted by the idea of discussing issues, but after observing the type of discussion (not very intellectual) may decide to drop off.

Figure 15. Internet Tool Usage Over The Last 24 Hours.



5.8. Summary

Three type of externalities (linkage, congestion, and network externalities) have been identified. I have shown how to derive the end-user's Internet access demand curve and indicated that it depends on (1) the various consumer surpluses he can potentially obtained from the various possible usage; (2) linkage externalities, and (3) option demand. I have also illustrated how to obtain the end-user's Internet usage demand and, most importantly, I have emphasized that Internet demand for usage varies with the type of Internet tool a user decides to employ. I have suggested too that an end-user's access (non-recurring) costs are higher than his usage (recurring) costs. Additionally, I have presented some statistics on Internet access and usage, but it is apparent that a set of more accurate and more trustful statistics is needed. In particular, a set showing Internet usage by application measured on a per-packet basis may allow a more clear understanding of an end-user's Internet access and usage. I can also add that this data should be collected from a private Internet provider (e.g., local ISP) as opposed to an academic or work server, as only in the first type of provider are users charged directly, which provides less distorted statistics.

6. How Internet Provision Should Be Priced

Internet access can be obtained through several methods including ISPs, on-line providers, academic servers, work servers, and government servers. This chapter will concentrate on the end-user's Internet demand for access and usage obtained from ISPs. I will first describe the background of the problem; next, I will show the costs incurred by ISPs when offering Internet connection and how these costs vary. Then, I will present the motivation for a traffic-based pricing scheme. In this context, I will propose an optimal pricing mechanism: Because of the two types of individual (end-user) demand revealed in the previous chapter (access and usage), two types of end-user pricing schemes are needed: *Access pricing and usage (traffic) pricing*. Hence, the final pricing toll should consist of a multi-part tariff: A fixed portion to charge for access and a variable portion to charge for usage (traffic). The fixed fee should vary with the maximum bandwidth of a connection and the usage fee should vary with: (1) the user's rate of transfer requested; (2) the current load of the network; and (3) with the volume of packets transferred. I will also show the inefficiencies created with non-optimal pricing schemes and explore a solution for when optimal prices under-recover costs of provision.

6.1. Background

As shown in previous chapters, besides direct connection (from ISPs) an end-user depends on Internet connectivity through several different technologies, including fax, telephone, and cable, among important ones. So, the supply of Internet service is related to several quite different communication-oriented markets. This is because technological advances rapidly blur traditional industry boundaries, and enable competition between firms that did not previously compete with one another. For example, while cable TV providers have been allowed to offer telephone service to their subscribers, numerous telephone companies are delivering video services to households over their networks. In summation, new entrants are using new technologies to compete with incumbents, and incumbents in previously separated industries are beginning to compete with one another (Gong and Srinagesh, 1995).

Because the Internet uses the telephone infrastructure as well as the end-user's computer equipment to transport data, it is considered:

1. A Value-Added Network (VAN) and
2. A Value-Added Service (VAS).

On the one hand, a VAN is generally a commercial entity that uses some pre-existing facilities to function, such as a long-distance telephone company, a computer on-line service, or an ISP. The Internet is considered a VAN, because it provides communications services, electronic store-and-forward mail-boxing, e-mail, WWW service, and other related services for Electronic Data Interchange (EDI) transactions, through telephone networks. VANs are necessary because it would be too expensive and impractical to establish direct connections with all partners involved. VANs are also useful because they are accessible to you regardless of physical location. They support reliable connectivity to all partners via varying communications speeds and protocols. On the other hand, the Internet is considered a VAS, services are accessed by Internet applications such as, Eudora, Netscape, WS FTP. Internet applications are a value-added service, as they reside in the end-user's computer equipment (often purchased in advance), which is also used for functions other than Internet services.

Understanding the environment of the telephone market structure, which has shown consistent consumption growth rates, is the first step to understand the Internet market. As a result of technological advances and deregulation policies, telecommunication markets have left a monopoly type of market and are now immerse in a more competitive environment. The telecommunications industry has experienced economies of scope and scale, which has led to a vertical and horizontal integration of the industry.[55] These factors undoubtedly have an important effect on the Internet market, because of its characteristics of VAN.

The second step is the understanding of how money flows within the Internet market. A typical money-flow system is that in which end-users pay their local ISP, who in turn pays for interconnection to a regional ISP, who finally pays to a national ISPs (NSP[56]). The national ISP has to pay costs of connecting their system to a place to exchange data with other ISPs, such as a NAP, MAE, CIX, etc. There is a recursive relationship in which the cost structure of services provided in any layer is determined by prices charged by providers one layer below (Gong and Srinagesh, 1995) and to providers one layer above.

The third step is the understanding of the ISP cost structure, which is incurred by ISPs (local, regional, or national) when providing Internet connectivity to their customers. Because of the different technologies available and existence of different needs, cost structures among ISPs vary widely. However, in general, ISP's face two main types of costs: Access and Usage Costs. Access costs include two categories: Installation costs and customer activation costs. Usage costs refer to customer support and maintenance costs, and network load costs. In the short-run both access and usage costs are fixed. In the long-run, though, both of these vary with the number of users and with usage demanded. Below, I explain ISP's costs in detail.

6.2. ISP's Access Costs

One type of access cost incurred by ISPs are installation costs. These costs refer to the setting up of the ISP network, including investments in network infrastructure (hardware and software). The other type is customer activation costs, which refer to those costs incurred

---

[55] See Appendix A for more detail.
[56] When a user is connecting through an institution, e.g. a university or business, then it is the institution that generally pays for the individual's access, and usage is currently priced at zero.

by the ISP when connecting the customer to the network via modems, wiring, ISP's server disk space, IP address fee, among others.

6.2.1.ISP's Installation Costs

Installation costs are generally financed by loan investments made by the ISP. They are also non-recurring costs. They refer to the costs incurred by the ISP to achieve operational status. There are two types of installation costs—network infrastructure costs and interconnection costs. In the short-run both of these costs are non-usage sensitive, but in the long-run they will tend to vary positively with the number of users and usage (maximum bandwidth) demanded. Network infrastructure expenses are necessary for building the facilities to enroll in business. ISPs require the installation of fiber optics, routers, software licenses, and other necessary equipment (network capacity) to provide Internet service. Network infrastructure requires large fixed investments. To finance these large investments, some ISPs offer term commitments by which a customer receives discounts if he signs a long term contract (3 to 5 years) of service. With similar purposes, ISPs may also offer volume discounts. Volume discounts mean that the more one transmits the less one pays per packet basis. See section 6.9.

Interconnection costs deal with the agreements between interconnecting ISPs for the transport of data. Although, an ISP has to pay interconnection costs during the installation stage, these costs are most important during the business stage with traffic-sensitive changes in demand. More bandwidth capacity is necessary to avoid aggravating congestion problems, when the number of customers connected to the network increases, and new investments in equipment and lines are compelled. For example, ISP servers are designed to simultaneously support a maximum number of modem connections. Internet interconnection cost schemes can be horizontal or vertical (depending on whether a particular ISP connects directly to a wholesaler, e.g. a NAP, paying a flat fee for interconnection; or that ISP is itself a wholesaler, in which case he could pay a flat-fee, too, or nothing[57]).

6.2.2.Customer-Activation Costs

These costs are a one-time fee charged to the end-user. These are the costs incurred by an ISP in connecting an end-user to its network. Customer activation costs, include hardware, software, fees, wiring, among others. In particular, these costs vary depending on whether the customer chooses a direct connection (leased line) or dial-up connection. The former type is highly expensive, depending again on the speed selected. Srinagesh (1995) points out, for example, that BARRNet (a wholesaler or regional ISP) in 1995, was charging a non-recurring fee of $13,750 for equipment and activation for a T1 customer, and a $1,300 fee for a 14.4 Kbps dial-up connection. The author also mentions that these prices vary widely across ISPs but, in general, are a very significant portion of a customer's costs. Notice that the customer also incurs in costs of acquiring the necessary computer equipment and software to be able to connect to the ISP network. These costs, of course, are not part of the ISP, but nevertheless are considerable for the customer: Computer equipment including a modem is very likely to run over $1,000 (Mitchell and Srinagesh, 1995).

6.3. ISP's Usage Costs

These costs vary with customer usage (maximum bandwidth) of the network in the short- and long-run. There are two types of usage costs: Maintenance costs and network load (the amount of traffic that an ISP network can support at a particular point in time from all users' connections) costs.

---

[57] Notice though that in any case each ISP has to pay the costs of wiring its network to its supplier or interexchange carrier. Refer to next chapter and Appendix B for further details on Internet interconnection schemes.

6.3.1.Maintenance Costs

Maintenance costs refer to those incurred by the ISP because of the customer's usage of the network. When in business, an ISP mainly incurs in customer support and maintenance costs in the short-run; in the long-run capacity expansion may be needed. Customer support costs and maintenance costs should also be charged on a case-by-case basis, and not necessarily on flat-fee basis. Some users will require more help for equipment installation and maintenance than others.

6.3.2.Network Load Costs

Network load costs refer chiefly to server capacity and bandwidth costs (IP transport costs). Network load costs are very peculiar. At present, to the ISP, because of flat interconnection rates, these costs are zero when the ISP's network is either uncongested or congested. To the customer, these costs are considerable when the network is congested. Network load costs are passed on to the customer as delays. One should notice that, when there is congestion, the costs other subscribers, those who value their opportunity time higher are also higher. With excess capacity present, these costs are also zero for the customers, as they would not experience any delay. Excess capacity, in this case, would imply getting rid of non-bandwidth bottlenecks too.

These costs should be charged depending on the load a customer imposes on the network *and* on other subscribers.[58] When a customer requests the use of a real time application such as, live broadcasts, phone transmissions, screen talks, etc., he requires much more network resources and bandwidth than when a user requests the use of a non-real time application, such as e-mail. Also, assuming equal rate of transfer, a customer who requests a larger file transfer than other will be congesting the network for a longer period of time than one who

[58] This practice should be applied to end-users as much as ISPs themselves, as the latter are also customers of larger ISPs.

requests a smaller file transfer. In these extreme cases, the end-user is probably loading the network (routers and bandwidth) to its limits and forcing the network to allocate resources for his own private needs. This instance should be considered and charged as such. That is, as opposed to a flat-fee charge, as it is at present, a user should be charged depending on the current load that he imposes on the network.

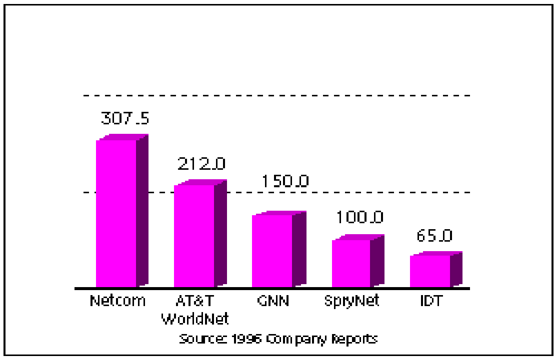
It is important to notice how congestion arises: After a certain point, when the network does not have room for more requests, all users, regardless of their usage impose costs on each other by prompting delays to happen on everyone's packet transmissions. This should also be considered so that everyone pays according to his demand and there are no free-rides. Under a flat-fee based scheme, cross-subsidization would happen when low-usage customers pay the same fee as heavy-usage customers.

Because of delays prompted by congestion customers eventually will force ISPs to expand capacity. This expansion of capacity should be based on prices charged to those connecting to the network during congested times. Economic theory suggests that if the cost of accommodating one more packet (marginal cost) is less than the charged price (marginal benefit) of the packet, then it is appropriate to expand capacity, and not otherwise. One then would expand capacity when the revenues from congestion fees exceed the cost of providing the capacity. If an optimally priced network generates sufficient revenue to pay the cost of new capacity, then it is appropriate to add more. Revenues could come from two sources: (1) higher prices and/or (2) more users. In any case, though, and other things equal, revenues should increase to expand capacity.

Traffic-based pricing permits that only the users who want to use the network when it is at capacity pay for expansion. Those willing to tolerate delay will not necessarily pay for expansion, as they will not have to pay the extra charge to transfer data when the network is congested.

6.4. Some ISP Statistics

Figure 16. Top Five ISPs Ranked By Number Of Subscribers (In Thousands)



6.5. Motivation For Traffic-Based Pricing

Having described the cost structure of Internet provision, I will now discuss how pricing should take place on the Internet. Pricing should reflect costs, so that users can make informed decisions about Internet demand for access and usage. On the present Internet, this does not happen: Prices do not reflect costs. This is because Internet usage is treated only indirectly (i.e., it is targeted, for example, by imposing time restrictions on users). However, the main element affecting Internet usage (network load) is volume (number of packets transferred), not the time of the connection. Time may also be related, but it is not as critical as volume. Besides, very few traffic-sensitive schemes have been implemented on a formal basis; at present, most pricing schemes are non-traffic sensitive. As a result of this disparity, users feel free to demand as much usage as they desire, which in turn gives rise to congestion problems (see Chapter 7 for more detail). Therefore, pricing mechanisms that are not traffic-sensitive cannot control Internet usage efficiently.

It is important to distinguish between usage-sensitive and traffic-sensitive mechanisms. Usage-sensitive mechanisms include traffic-sensitive schemes as well as, for example, those based on duration of the connection (a form of usage), but these do not control congestion efficiently, as congestion is not a function of the duration of the connection (i.e., the amount of time a user's connection lasts), but of the volume of a transfer session. A user can establish an Internet connection but nevertheless remain idle: Maybe just reading a file transferred that, once transferred, resides in the user's computer.

When traffic-sensitive pricing schemes have been introduced, they have been targeted to end-users only. For a successful implementation of a traffic-based pricing scheme on the global Internet, it has to target the *roots of the problem.* That is, it should first be implemented among ISPs (interconnection) and then on end-users. Internet services suppliers have to also face traffic-sensitive pricing schemes, so that on basis of their traffic costs, they are driven to charge their customers on this basis as well—traffic-sensitive.

For prices to accurately reflect costs, a user's Internet usage has to be considered in any Internet pricing scheme. Since Internet usage chiefly depends on the bandwidth available at any particular time, bandwidth is the critical limited resource on the Internet. Despite constant efforts to compensate in the provision of Internet capacity[59], Internet demand for usage is outweighing supply of bandwidth, because of: An increasing number of software applications that make use of all the bandwidth available, e.g., video, voice, interactive games, etc.; an exponential growth in the number of users; congestion problems, among others. So, in order to allocate Internet resources efficiently (at their best use) and in an attempt to solve congestion problems on the Internet, traffic-sensitive pricing has been proposed as one method to simultaneously obtain both economic efficiency[60] and network efficiency,[61] as well as help control the current congestion problem.

For this to take place, suppliers of Internet access (all types of ISPs, institutions, etc.) ought to have the right incentives to charge their customers (end-users, other ISPs, etc.) on a traffic-

---

[59] Through technology: expanding backbone, faster routers, more bandwidth, etc.
[60] If no user currently being denied service values it more than any of the users currently receiving a particular ToS (Type of Service), then that operation is economic efficient.
[61] If a network can maintain a target level of service while minimizing the resources needed to provide this service, we say that its operation is network efficient.

sensitive basis. Local ISPs themselves should face a traffic-sensitive charging scheme. Local ISPs should pay regional and/or national ISPs on this basis. Similarly, transit carried by national ISPs for regional ISPs and other national ISPs should also be charged on traffic-sensitive basis. If this does not happen, there is very little chance that any ISP be willing to implement a traffic-sensitive charging scheme for their customers (the end-users) and that any optimal pricing mechanism take place. One instance when it may happen is upon a regulatory act. However, experience shows that regulatory practices are not necessary the path to take due to public opposition, special interests, inefficiencies, etc.

6.5.1.Traffic-Sensitive Pricing For End-Users

Prices should send signals to originators of Internet usage (end-users) so that they, based on the costs they impose on the network (network load) and on other end-users (social or congestion costs), can decide how much usage to demand. In this way, prices would allow to control between high versus low valuation of usage—ideally allocating resources efficiently.

6.5.2.Traffic-Sensitive Pricing Among ISPs (Interconnection)

Recall, there are three types of Internet suppliers: (1) National ISPs who are the owners of the main transport lines connecting the NAPs, where other ISPs (national, regional, and local) exchange data; (2) Regional ISPs who are the wholesalers of the market and who serve as intermediaries (in general) between national ISPs and local ISPs; and (3) Local ISPs who are the retailers and providers of Internet service to end-users. As it will be explained I detail in Chapter 7, at present, pricing among ISPs (interconnection pricing) is mainly non-traffic sensitive. So, in order for compensation to be efficient, interconnection pricing should also be based on traffic-sensitive schemes. Most important, for the successful implementation of a traffic-based pricing scheme, these traffic-sensitive schemes should be first implemented at this level.

6.5.3.Traffic-Sensitive Pricing For Institutions

Institutions (business, educational, governmental ,etc.) providing Internet access to end-users are also an important source of Internet connectivity. In a sense, institutions are a type of local ISPs too, as they offer the same services as for-profit local ISPs do: direct and dial-up connections for end-users. These institutions, specially universities, allow multiple individuals (students) to share a same host (IP address). Empiricism suggests that the rate of end-users per computer at institutions is much higher (around 15 per computer)[62] than at an individual subscriber's connection (around 1.37 per computer)[63] and that businesses and universities are two of the three largest user's groups (see Chapter 5). However, despite that an institution's local network (LAN) resembles that of an ISP and that there is a significant number of end-users connecting through institutions, institutions are better identified as end-users. They should then be treated as such and should face traffic-based pricing schemes too. This practice would drive institutions themselves to deal with the recovering of traffic-based costs in their own way.

On today's Internet, an institution's member (an Internet end-user) is very likely to pay nothing for connectivity. He, thus, does not face the right costs of usage, which in turn produces inefficiency, such as congestion, free-rides, etc. As mentioned, when a user connects to the Internet through an institution, typically, he does not pay, either for access or usage of the network, at least directly. There may be some institutional fees directed to this effect, but if this is the case, the end-user may be paying a flat-fee for whatever amount of usage he exercises. This scheme is inefficient, as it is not traffic sensitive. Nevertheless, flat-fee charging schemes are preferred by institutions, specially because they allow them to direct a fixed amount of their budget to pay for Internet connectivity. It must be pointed out, though, that just because institutions prefer flat fees it does not mean that they should be

[62] Ratio of computers available to students at the University of Idaho's computer labs, 1996.
[63] As computed by Rickard (1995) and shown in Chapter 5.

provided. An institution's counterpart (the ISP) will also have a say at this issue and so non-usage based pricing may not be preferable.

Binding charges to end-users usage may not be preferred by institutions anyway. For example, electricity fees are assessed on flat fees to each department, telephone fees are assessed depending on usage to each individual. So, institutions may use the option of Internet connectivity to promote research and learning for its members, traffic-based end-user charging is very likely to be too expensive and burdensome to implement. On this basis, institutions may try to get special pricing schemes from Internet suppliers. However, how can private usage be distinguish from educational usage? Even if it is, why should end-users (individual or institutional) be allowed to use the Internet without being charged for usage? After all, when an individual wants to place a personal long-distance phone call, he pays for its costs, personal faxing is individually charged, personal mail is paid individually, etc. No clear answer is apparent for distinguishing usage, but, ideally, either type of usage should be counted and charged with traffic-sensitive schemes.

All Internet users (individual or institutional) should be charged for usage. For institutional subscribers, the pricing mechanism may be different from that applied to individual subscribers. If an ISP charges the university on usage basis, then the institution will be forced to recover those costs from the originators—the end-users; hence, obtaining both economic and network efficiency, and helping to control congestion.

Many mechanisms may be used to this effect. For example, the institution may assign a quota limit to each user. If the user exceeds his quota, then those extra packets transmitted, which should also be traffic-sensitive, should be paid directly by the user. As each individual has his account on a institution's sever, there could be a separate computer dedicated to the charging and billing of users—that is, to track usage.

6.6. An Optimal Pricing Scheme

So far I have described the imminent need for introducing a traffic-sensitive pricing scheme at all levels of Internet connectivity. What needs now to be specified is the type of traffic-sensitive scheme that should be implemented. This section attempts to clarify this aspect. An optimal pricing mechanism should maximize the number of voluntary market exchanges between suppliers and demanders, because such exchanges allow individuals to maximize their individual preferences. An optimum situation (Pareto optimum) is one where individuals have maximized those voluntary exchanges by pricing goods and services at their marginal cost (Wenders, 1987).

Hence, an optimal pricing scheme should target both demand and supply of the Internet market. On the supply side, prices should compensate suppliers (ISPs) of scarce Internet resources. On the demand side, prices should allocate resources efficiently by considering the value each user places on Internet demand: Sending signals to originators of Internet demand (end-users and institutions) and to intermediaries of IP transport (ISPs), so they can make informed decisions about Internet demand. I will focus mainly on the effect of pricing on the demand side. A multi-part tariff pricing scheme, can help attain both of these goals. Economic theory suggests that demand for access should be priced both with fixed recurring and non-recurring fees to recover costs and that demand for usage be charged with an appropriate traffic-sensitive scheme.

6.6.1.An Optimal Pricing Tariff

The optimal pricing tariff should be a multi-part tariff: A flat fee and a varying fee. The flat fee should be directed to cover costs such as monthly customer support, equipment maintenance, billing and accounting. The charging price for the flat-fee portion of a tariff should vary depending on factors such as, the maximum bandwidth (capacity) of the connection and the air mileage from user’s home to the ISP’s network. The variable portion should be targeted to recover usage costs of the network and should vary depending on: ***(1) current network load; (2) rate of transfer requested,***[64] ***and (3) the volume of the transfer session (number of packets transferred—sent/received).*** The expensive investments in network infrastructure comprise the majority of the costs of the Internet service provision. When the network is not congested the marginal cost of transporting an extra packet is essentially zero. When congestion occurs, traffic-based pricing should be used and applied depending on the three factors just mentioned, and if it is heavy-bandwidth traffic that is requested then a higher price per unit should be charged than when the request is for light-bandwidth traffic. In this way, only the users that request service during congestion times will pay to expand capacity. Those who are willing to wait will be rewarded with lower prices or no charge at all for usage.

A pricing mechanism with these features would also solve the problem of worrying about time-of-day pricing, as the proposed tariff would depend on the current load of the network. Notice that this tariff could be applied to end-users and institutions as well as to the different types of ISPs. The latter are just intermediaries in the transfer of data from one end-user to another. In other words, just as in other industries, corporation's customers may face higher rates than residential users.

It must also be observed that while the goal of any Internet pricing mechanism should be to optimize the number of economically efficient connections, it is important to distinguish it from the goal of Universal Service. Though highly unlikely, it may be possible to have full Internet connectivity for all users. However, this goal would not be economically efficient, as there are clear substitutes for Internet communications that may be more efficient for some users. For example, it may be more efficient for an individual to make a telephone call, send a fax, or go to the library than to use the Internet for equal purposes.

---

[64] Type of service requested refers to selecting transfers for different kinds of applications or services, e.g., e-mail, WWW, FTP, etc. Rate of transfer requested refers to the actual speed that the user wishes to obtain (pay).

The issues of economic efficiency versus equity have been endlessly debated for other industries too, such as, telephony, cable, electricity, and water. As it is the case in these industries, no clear answer has been identified to solve the efficiency versus equity trade-off on the Internet either. What it is clear, however, is that Universal Service for the Internet is related to the way income is distributed in society rather than to how the Internet is priced.

6.7. Current Network Load: Ranking The Use Of Congested Resources

Although congestion can be reduced by over-provisioning capacity (Pool, 1983), this is not a cure neither in the short-term nor long term. Congestion can be mitigated by re-configuring the network, but this is only possible in the long-run (Bailey, 1995). However, it is not efficient to have idle capacity, because resources could be put under other uses. So, the costs of over-provisioning capacity versus the benefits of reducing congestion have to be weighed and balanced. Economic theory suggests that a cost-based pricing scheme can help fight congestion efficiently.

The issues of congestion and network efficiency are related. *Network efficiency* refers to the utilization of network resources, such as bandwidth and buffer space. If a network can maintain a targeted level of service (by means of an admission control protocol, which should account for variations in usage demanded) while minimizing the resources needed to provide this service, then that operation is network efficient. Network efficiency is important as scarcity of network resources has been revealed by congestion problems. When a network is congested, one user's packet crowds out another's (social cost), resulting in delayed or dropped transfers.

Besides, two particular problems arise. First, it is the problem of defining how congestion should be measured? The controversy arises because: 1) individual preferences (type of applications, time of day, etc.) vary considerably across users (implying that different users value the same and different applications as well as their usage of time differently); 2) responsibility of detecting congestion may be spread over several network operators; and 3) in a Type of Service network[65] these issues will become more complex. Second, it is the problem of how to allocated limited resources under congestion.

MacKie-Mason, Murphy, and Murphy (1995) suggest that by including users in the pricing model, i.e., allowing them to make decisions of whether to transmit packets or not each time they request service, these problems can be solved. They assume that users are *adaptive* by nature—they response according to their valuations of different situations. During a particular connection adaptive users can adjust their requests to respond to feedback signals from the network.

User's contribution on the pricing process may be attained by introducing an interface (a computer application) in the user's computer so that he can be informed as to the network load and price at the time of his transfers. To be accurate, this interface should consider incoming and outgoing packets to the user's account.

These adaptive users are the ones who can help to increase network efficiency if they are given appropriate feedback signals. When the network load is high, the feedback from the network should discourage adaptive users from increasing requests. This network feedback could be as a price signal based on the level of network congestion: When the load is high, the price is high, and vice-versa. Users who will value the network more will choose to transmit at higher prices (during congestion periods). When the network is lightly loaded the price will be close to zero. This will both have the potential to increase network and economic efficiency—depending on the success of the implementation of a particular pricing mechanism.

---

[65] See next section.

6.8. Rate-Of-Transfer Pricing

Bandwidth generally refers to the maximum bandwidth (or capacity of transfer), which in turn refers to the maximum amount of data that can simultaneously be transferred through a pipe. However, because of the many bottlenecks present in Internet connectivity and the sharing of this resource, what is really important is the actual bandwidth employed. The latter refers to the amount of data the is actually transferred per second during a certain connection.

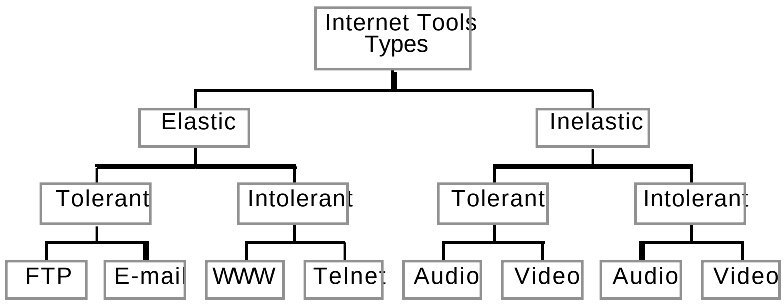
Consider a user connecting to an ISP network through a 14.4 Kbps dial-up connection and one connecting through a T1 leased line. Both users are capable of receiving packets at rates over 10 Kbps, but due to the larger number of possible bottlenecks that a dial-up transmission supports compared to that of a leased line, dial-up transfer rates are lower than T1. Not even a T1 user (say, an employee of a firm) could, at a particular time, get a rate of transmission close to the whole bandwidth capacity of the firm's connection. This is because a bandwidth connection is shared with other users on a two-way traffic basis and also because of the presence of bottlenecks along the connection (e.g., routers). It should also be noticed that the lowest capacity link does not necessarily determine the speed of transfer in a connection. Again, this is because of the many bottlenecks that may exist in the firm's LAN: Switches, bridges, routers, repeaters, wires, etc. Rate-of-transfer pricing would only consider the real speed at which data is moved—solving this disparity.

Wenders (1996) suggests that users will be willing to pay for a guaranteed actual bandwidth for the transfer of their packets. So, it is actual bandwidth of transfer that should be usage priced and maximum bandwidth price should be in the fixed portion of the tariff. I will refer to rate-of-transfer pricing to include two more characteristics of the pricing scheme that I am proposing: Type-of-service requests and traffic-based pricing.

The rate-of-transfer pricing should price users individually on a per-packet transferred (traffic-based). That is, the more packets a user transfers, the more he pays in total. The rate-of-transfer pricing scheme will automatically account for the type of service (ToS) a transmission requires: Some type of packets need faster transmission than others, some need a specific rate of transfer. For example, a real-time video needs a specific rate of transfer in order to not appear distorted. Also, almost all Internet tools (WWW, FTP, Gopher) may require faster speeds of transfer than e-mail applications (unless it is otherwise changed by the user). The inclusion of ToS in the rate-of-transfer pricing scheme is suspected to be automatic, as users, perhaps by experience, will learn that packets from inelastic applications need faster service than those from elastic ones (do not confuse these terms—inelastic or elastic—with elasticity of demand, for a more accurate description of them see below).

In simple terms, type-of-service (ToS) sensitive means that the network considers the different needs of different Internet tools. Now, one can classify Internet tools in two major categories (see Figure 17): Elastic (or delay insensitive) and inelastic (real-time or delay sensitive). Elastic (tolerant) tools are those, such as e-mail and file transferring. They can tolerate delay and their packets can be queued. Elastic (intolerant) tools are those, such as Telnet and WWW. They can still work in the presence of delay, but they would work better with no delays whatsoever. Inelastic applications are those, such as real-time video and a phone or screen conversation. They are also subdidived in tolerant and intolerant. Inelastic (tolerant) are those tools needing immediate transport but that can still work in the presence of delays. For example, an audio or video clip. Inelastic (intolerant) applications are those needing immediate transport and cannot work properly if they experience delays. For example, teleconferencing or live broadcasting. So, a ToS sensitive network should have services priced accordingly: Faster and more reliable service should be charged higher than slower and less reliable service. Importantly, though, the rate-of-transfer pricing scheme considers these needs. As explained above, through a specific rate of transfer requested by an end-user, the type of service (ToS) needed by packets will automatically be considered when the appropriate rate of transfer speed is selected by users.

Figure 17. Types of Internet Tools



6.9. Smart Market Vs. Rate-of-Transfer Pricing

MacKie-Mason and Varian (1993) suggest that economic efficiency is enhanced if the allocating mechanism gives higher priority to uses that are more valuable. That is, if no user currently being denied service values it more than any of the users currently receiving service, then that operation is economic efficient. They imply that users should then show their true preferences as to the type of service required. However, making the users show their preferences through biddings is one thing, but instantaneously comparing such preferences and acting on them is another. A dynamic mechanism allocating service has been described by these authors themselves as hard to implement. Besides, at one point the auction is going to have to close biddings and process information. During, this time, there may be other customers that value service more than the ones already being serviced, thus we fall back into the same problem of not being able to service all those customers that value network service more.

Unlike MacKie-Mason and Varian’s Smart Market, I believe user’s valuation refers more to their transfer’s quality than to their valuation of being admitted to the network or not, because users can always retry later (a network in the rate-of-transfer system should be able to compute the time in which less congestion is expected). When a network is at full capacity, then no guaranteed service should be offered. Hence, I suggest that economic efficiency will be enhanced not so much by just being serviced by the network, but by being serviced at a guaranteed rate of transfer. So, I believe that (as it will be described in detail on Chapter 8) the current network load, through and admissions control mechanism, should be the method by which the network satisfies users' valuations of a transfer connection and is able to guarantee service. The type of service (rate of transfer) requested will be selected by the user. Just like in a membership restaurant, a user that values more being serviced first than other does not necessarily get his way. Rather, if he makes an appointment to obtain service at certain time, his being served there may be more securely guaranteed. Let us assume that, in addition to membership (access) fees, this restaurant charges service based on the number of tables left. Then, if a customer cannot or does not want to pay the price for service (usage) at a particular time, he can always come back later or go to another restaurant.

In similar way, then, a network based on a rate-of-transfer pricing scheme should admit transfer requests based on the network load and on the user's arrival time: First-come-first-serve basis. This does NOT mean that, a user's packets are going to get a first-in-first-come service. Upon a user's request, a user will wait for network acceptance or denial to select from a variety of transfer speeds (each with its own price). Since, each user's request is processed one at a time, the network should be able to guarantee service to all attended users.

In addition, how can network operators control so that not everyone selects the highest priority? Through pricing! Ideally, because better service is hence priced higher, only those users with higher valuation of network usage will be willing to pay more. That is, faster rate of transfer should be charged higher. By focusing on a responsive mechanism that varies according the current network congestion status, MacKie-Mason and Varian (1992) and MacKie-Mason, Murphy, and Murphy (1995) propose charging only when the network is congested, with the fee based on the degree of congestion and on the ToS requested[66], and not on the number of bytes regardless of network conditions. The rate-of-transfer pricing mechanism would also consider these points.

I propose that the usage fee should vary positively with (1) the rate of transfer selected by the user, (2) the size of the transfer request, and (3) the current degree of network load. Therefore, the rate of transfer pricing is optimal because it fights congestion and satisfies users' valuation of their transfers by guaranteeing a rate of transfer, when users' requests are admitted. Finally, when the network is not congested, rate-of-transfer's prices can be low or non-existent at all.

Thus, just as in the Smart Market model, economic and network efficiency are improved by inducing users with low-valued uses to delay requests until bursts of congestion cease. Such time-smoothing will not upset users who can tolerate delay, while it would improve the transfer's value to users that value delay more. Note also, as in the Smart Market model, that a pricing system does not necessarily need to be in monetary terms. In a private network, for instance, the pricing system may work as control mechanisms displaying the current status of the network and using the assignment of quotas. In sum, prices will serve as a mechanism to rank transfer needs. If prices reflect the costs of providing the service, they will require users to compare the value of their packets to the costs they are imposing on the system. Users will then make a decision as to whether transfer packets at a certain price or not. If not, they will wait until congestion decreases.

6.10.Problems Of Rate-of-Transfer Pricing

Although, rate-of-transfer pricing is an optimal mechanism for fighting congestion and allocating resources efficiently, it may still leave some users unserviced when the network is

---

[66] This idea was introduced with the Smart Market mechanism proposed by MacKie-Mason and Varian (1992).

at full capacity. This however, is going to happen with most pricing schemes in the short-run (until capacity is expanded). The biggest drawback of the rate-of-pricing model is that for it to work it needs two fundamental pieces of technology that are still not widely implemented. First, it needs the implementation of the IPv6 (or some other protocol) that would allow to distinguish packets with different needs of service. Second, it needs the introduction of a computer interface that can interact with the ISP network to allow the involvement of the user at the decision-making process.

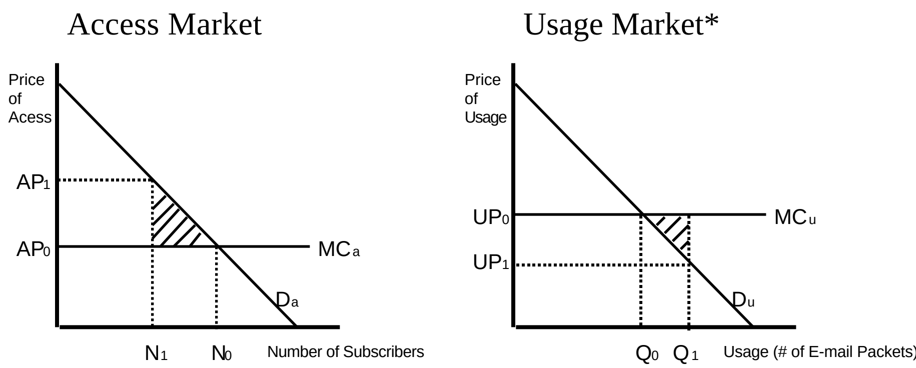
6.11.The Problems Of Volume Discounts

Volume refers to the total number of packets transferred, regardless of the rate of transfer. Volume discounts refers to the inverse relation that some ISPs allow between pricing and the number of packets transferred: More packets, less charge. Because of economies of scale on bandwidth provision and the flat-fee pricing that the ISPs face, volume discounts are often offered by ISPs. However, these discounts are not practical, because the more packets are transferred the more congestion they cause. So volume discounts should not be part of an optimal pricing scheme attempting to control congestion. Additionally, it is obvious that the more packets are transferred the more an individual should pay: Price per unit should vary positively with volume of traffic. Recall, however, that volume (packets transferred) should also be distinguished by the type of service requested by the user: Not merely by quantity, but by quality too.

6.12.When Prices Are Not Optimal

There are certain instances where prices set equal to marginal cost are optimal. In reference to Figure 18, a single individual's optimal prices would be the access price AP_0 and the usage price UP_0. Given a certain network load and a given set of access and usage prices, N_0 subscribers would join the network at a particular point in time. If prices are not marginal-cost based, then welfare losses are created in each market. For instance, if access price were set too high (AP_1) and usage price were set too low (UP_1), then welfare losses are created in each market equal to the representative cross-hatched triangles.

Figure 18. Nonoptimal Pricing Of Access And Usage



* The Usage Market graph varies for each different tool, depending on the rate of transfer requested. So, the righ-hand side graph is one example of usage among many tools.

The above reasoning presumes that it is of net benefit (total benefits > total costs) to measure and price all usage separately for each user and for each different type of service requested. If it were not of net benefit second-best alternatives need to be found.

6.13. When Marginal-Cost Pricing Revenues Do Not Cover Costs

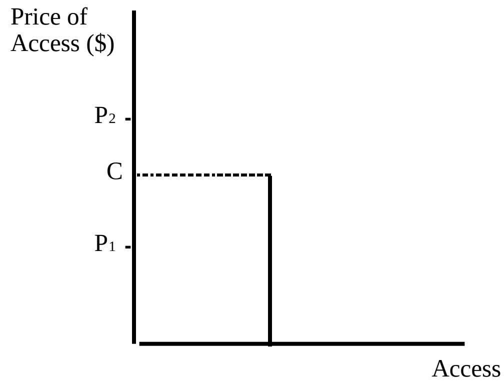
As Wenders (1987) points out that there are instances when marginal-cost pricing may not represent economic efficiency: Demand and supply do not internalize all of the market preferences, not allowing revenues to cover costs of provision. In such situations, the market must be corrected either through voluntary agreements that internalize the externalities or through some type of regulation.

It is likely that even strict marginal-cost pricing, such as the rate-of-transfer pricing, would underrecover revenues for the ISP. For example, the presence of network externalities means that the access-demand curve seen in the market reveals only the individual's private benefit of access, which underestimates the total benefits of access. This is because, an individual does not consider the benefits he provides to other subscribers: When a subscriber joins the network, he makes the whole system slightly more valuable for each of the other subscribers.

In these events, prices would somehow have to depart from marginal cost until the revenue constraint were satisfied. If economic efficiency is to be maintained, only optimal departures from marginal cost pricing are acceptable, so that the resulting welfare losses are minimized. Ramsey pricing is one way to attain these optimal departures from marginal-cost pricing. Recall, from Chapter 5, that an individual's demand for access was partially determined by the consumer surplus obtained from the various kinds of usage. So, if the price of access is less than the user's valuation of access, this user would indeed join the network.

Again, as Wenders (1987) points out for the telephone network, it is unlikely that the price charged for Internet access would be exactly equal to the user's valuation of such access. Usually, there is an access consumer surplus. That is, the user's valuation is greater than the access price. For instance, in terms of Figure 18, if access price P_1 were charged, the total consumer surplus for one individual resulting from access to the network would be measured by the difference between C and P_1. Then, if some additional revenue was necessary from this user, and we wanted to extract it in a way that did not change behavior, raising the access price above P_1 would be the ideal way to get it. Unless the access price was raised above C, say P_2, such a move would not cause this user to change Internet usage, nor would it cause him to drop off the network.

Figure 19. An Individual's Internet Access Demand



Despite that such a rise of access price would make the user worse off in absolute terms, leading him to complain about it, it would not prevent any voluntary exchanges from being lost. In other words, his access and usage will be the same in number, but such a rise would yield less net benefit. A similar reasoning would apply in the case where too much revenue was collected by marginal-cost pricing. This is a second-best solution to the problem. There are several obstacles, as to measuring the appropriate access price, but, for example, this price can be set on the basis of experience or surveys. It must also be noticed that, in order to cause the least distortion and cover the revenue constraint, the rise in price could not only be applied to access but other usage components as well (or to usage components only), as long as distortions are minimum or none.

6.14.Summary

I have suggested that the Internet is a value-added network and service. As it resembles the telephony network in some way, the latter may be used as a guide to understand the Internet's network. I have also identify the type of costs that an ISP (Internet provider) incurs in such a provision. The different types of costs to be considered for Internet pricing are:

- **ACCESS COSTS:**
  - Access costs are the fixed costs of providing the network infrastructure (installation costs) and the fixed costs of connecting a customer to the ISP network (activation costs). The former include ISP's payment for the leased line, cost of the routers. The latter are costs incurred by the customer when connecting to the network. Apart from his computer equipment, a customer will need to pay for obtaining Internet connectivity to the ISP network.
- **USAGE COSTS:**
  - There are two types of usage costs: Network load costs and maintenance costs. While maintenance costs are period fixed charged, network load costs should depend on the current status of the network—congestion.
  - The marginal costs of sending an extra packet. This cost is zero when the network is not highly congested.
  - The social costs of delaying other users' packets when the network is congested. This cost is translated through delay and dropped packets.
  - The marginal costs of connecting users to the network. Each new subscriber implies costs for access lines. The costs are fixed only when excess network capacity (not necessarily bandwidth) is present.
  - The costs to expand network capacity: Adding new routers, more bandwidth, new staff. These costs vary in a lumpy manner—with the increase in number of users connected a larger network is needed.

ISPs run a network service that connects users to the Internet. In the short-run, costs in an ISP network are fixed except for congestion cost. That is, when a user on the Internet sends out a file, he is using up bandwidth that could otherwise be used by other users in the network. All the costs associated with number of customers, electronics, leased lines, etc. have been paid for already and are seen as fixed costs. In the long-run, however, all costs vary. The ISP will need to invest in a larger network as the number of customer increases and as more usage (from the highly appealing intensive-bandwidth applications) is demanded.

There are two basic types of costs that ISPs need to recover: Access costs (recurring and non-recurring) and usage costs. Internet access costs should account for two factors: The setting up of the ISP network and the costs of connecting the customer to the network. Non-recurring-access costs refer to those of recovering the investments of building the network. They should be charged to the customer when he connects to the ISP. Recurring access costs (or customer-fixed costs) do not vary with usage and refer to those of equipment maintenance, customer support, leases, monthly billing and accounting, etc. Usage costs should vary with the customer's usage and the current state of network congestion. They should reflect the cost each end-user imposes on the network *and* on other users when he demands service.

I have also stated that the main reason for some sort of a traffic-based pricing is to control congestion, as current pricing mechanisms are not efficient in such a practice. Controlling congestion is crucial, because it is the most important economic problem of today's Internet. Based on these grounds, I have proposed an optimal pricing mechanism. The first and most important characteristic of this pricing scheme is that is traffic based. It should have a fixed portion varying positively with the maximum bandwidth of a connection and a variable portion varying positively with: (1) the user's rate of transfer; (2) the current load on the ISP network; and (3) with the size of the transfer request. I have also shown the two fundamental elements needed for the functioning of the rate-of-transfer pricing: A ToS sensitive protocol and a computer interface allowing the involvement of the user in the decision making process. Furthermore, I have not only illustrated the inefficiencies created when prices are not optimal but also explored a solution for when optimal prices cannot recover costs of provision.

7. How The Internet Is Priced At Present

Today's Internet faces different types of pricing schemes, most of which are non-traffic based. Traffic-sensitive pricing has been around since the involvement of the private market, but it has not had a good acceptance by the public and has not been largely used either. Yet, none of these pricing schemes are optimal, for as described in the previous chapter, neither accounts for packets from different applications nor the value each user places on them.

Pricing models on the Internet can be implemented on end-users, including institutions (businesses, universities, organizations, etc.), and on ISP's interconnection. In this chapter I will first depict the environment of today's Internet. Then, I will show that what promotes non-traffic sensitive pricing schemes is ISP's interconnection pricing, which is based on non-traffic-sensitive pricing schemes and hence does not provide the incentives for ISPs to charge traffic originators (end-users) on traffic basis. Next, I will illustrate current Internet pricing schemes. I will also identify the advantages and disadvantages of non-traffic-based pricing mechanisms. Finally, I will describe how congestion is treated at present.

7.1. Background

First, special interests are involved in maintaining non-traffic sensitive pricing schemes. For example, one lobbying request that stands out is the one by James Love, of the Taxpayer Assets Project and member of com-priv,[67] who was representing many other supporters of flat-rate pricing. It consisted of a letter sent (in May 1994) to Steve Wolff, the acting Director of the Networking and Communications Division at the National Science Foundation at the time. This letter chiefly presented the view that traffic-based pricing should be

---

[67] Com-Priv (com-priv@lists.psi.com) is a distribution list that discusses issues related to the Internet by using e-mail.

avoided, and that the government should preserve and enhance the free flow of information. The following points were particularly stressed:

- Since, Internet Listservs[68] allow broad democratic discourse, pricing them would very likely endanger or restrict their use.
- Pricing e-mail discussion groups and distribution lists would reduce many sources of data for those who cannot or are not willing to pay.
- With pricing schemes, anti-competitive behavior may be encouraged. Companies operating Network Access Points (NAPs) may practice preferential treatment with selected business partners.

Second, since 1990 the Internet entered a commercialization stage that allowed private entities to provide Internet access. So, because of this commercialization and the anticipated reduction in government subsidy to formerly subsidized ISPs since April 1995, the private market has become the main player in providing Internet service. Hence private providers (ISPs) have implemented a variety of pricing mechanisms.

Third, the Internet has far moved from the dilemma of how to make up for the lost subsidy. Five years of commercial provision have already passed and the major economic problem facing the Internet at present is congestion, which is caused by the failure to price usage.

Fourth, the role of government has been relegated to: Coordinating Internet standards, subsidizing appropriate users (not those that already have access, but those who otherwise

---

[68] An automatic mailing list server, initially written to run under IBM's VM operating system by Eric Thomas. Listserv is a user name on some computers on BITNET/EARN which processes electronic mail requests for addition to or deletion from mailing lists. Examples are listserv@ucsd.edu, listserver@nysernet.org. Some listservs provide other facilities such as retrieving files from archives and database search. Full details of available services can usually be obtained by sending a message with the word HELP in the subject and body to the listserv address. Taken from *The Free On-line Dictionary of Computing*: http://wombat.doc.ic.ac.uk/cgi-bin/foldoc?Listserv

would not connect),[69] promoting a competitive environment with a full realization of benefits for users, and the fostering of an environment that attracts investment for a self-sustaining market. A more active role may be expected with the passing of the Telecommunications Act of 1996, where the FCC is entitled to impose restrictions on ISPs. This issue is still debated.

Fifth, until recently Internet interconnection pricing schemes have not provided incentives to implement traffic-based charging schemes. Instead, they have given rise to a chain of flat-fee charging schemes across the Internet's layers. National ISPs have been charging regional ISPs flat rates for interconnection. Regional ISPs have done the same with local ISPs, who in turn have ended up charging flat rates to end-users. However, as it was reported on *Business Week* on April 22, 1996, Internet access providers such as BBN, MCI, and UUNET have implemented usage-sensitive pricing scheme for certain high-capacity connections. More of this is bound to come as more people realize the need to control congestion.

Finally, being a government initiative and as it is deducible from the above-mentioned points, the Internet was not designed to be a commercial system. So, the problem of pricing the Internet on a usage-sensitive basis arises because of:

- the difficulty of trying to hold users accountable for the resources they use, specifically during congestion periods;
- The trouble of finding a technology to sustain a cost-effective, traffic-sensitive pricing mechanism that allow ISPs for provision of an optimal service; and
- The efforts attempting to find a simultaneous model that allow every citizen to access this communication service (*Universal Service*).

---

[69] Jeffrey MacKie-Mason and Hal Varian. Some FAQs about Usage-Based Pricing. University of Michigan. Also, Scott Shenker, "Service Models and Pricing Policies for an Integrated Services Internet." PARC.

As a result, a large part of the Internet still faces non-traffic sensitive mechanisms, which are largely responsible of the current problems surrounding the Internet: Congestion, cross-subsidization, mainly. Current pricing mechanisms will be described later.

7.2. ISP Interconnection: Promoting Non-Usage Sensitive Pricing Mechanisms On End-Users

To provide for Internet interconnection, which facilitates transit and exchange of data, ISPs have established interconnection arrangements between them. Internet Interconnection basically takes place at two levels: vertically and horizontal. The latter is based on what is called Internet Service Provider (ISP) Peering. The former depends on the ownership structure of the vertically integrated system. ISPs are affected by the forms interconnection agreements take on upper layers of interconnection: Settlements, flat-fees, or no exchange of money. At present, most national ISPs face a non-traffic-based pricing scheme,[70] which starts a flat-rate pricing chain reaching the end-user—to be described later in this chapter.

Early Internet interconnections[71] resulted from bilateral agreements. One of the first was the link between MCIMail and CompuServe in February 1986 (Mitchell and Srinagesh, 1995). With time, a proliferation of interconnection agreements occurred in this highly dynamic market. Some agreements took the form of settlement fees, others required membership fees only, and yet others did not require money exchanges at all. Generally, when the data exchange closely resembled a zero-sum game (i.e., every player benefiting more or less equally) then settlements were not necessary. When it was clearly a non-zero-sum game then incentives led to the implementation of traffic-sensitive pricing schemes for the ISPs.

---

[70] See below for more details.
[71] See Appendix B for further details.

However, it is very difficult to find a cost-effective and implementable accounting and billing scheme on the Internet at present that would allow the implementation of a traffic-sensitive scheme as described in the previous chapter—different router technology (IPv6) would have to be introduced first.

Huston (1994) provides a general framework of the form different Internet interconnection agreements (ISP peering models) can take. These ISP peering models, not surprisingly, draw heavily on models used in the telephony. Three models deserve special attention.

1. In the bilateral settlements model, each provider invoices the originating end user, and then financial settlements are used in accordance with originating traffic imbalance. This arrangement is commonly used between international phone service providers.
2. In the Sender Keep All (SKA) model, each provider invoices their originating client's user for the end-to-end services, but no financial settlement is made across the bilateral peering structure.
3. In the transit fees model, one party invoices the other party for services provided, based in traffic loads. This is the arrangement commonly used as the basis between long distance provider and local access provider in the telephone industry.

Trying to attach a peering model to a specific Internet interconnection layer (local, regional, or national) is futile. All three models are used in the different Internet interconnection levels. Huston (1994) suggests that from experience the SKA model seems to be only sustainable where the parties involved perceive equal benefit from the interconnection. The SKA model has been typical of national ISPs (NSPs).

Then, he states that despite that at present, the dominant peering model on the Internet is one of flat-fees, the bilateral settlements model is increasingly being used at the higher layers of Internet interconnection. Also, he indicates that an increasing number of local and regional ISPs are using some type of financial settlements with national ISPs. But, traffic-based settlements are difficult to implement because they are not popular among users.

Internet interconnection is based on peering agreements, which in turn use interconnection exchanges or network access points (NAPs) to exchange data and cost-effectively share resources. In addressing cost-effective transit, it has been noted that there is a diversity of transit needs: Some networks (public) may request affordable transit, for example, while others (research) may request higher-performance transit even if costs are not minimized. Consequently, NAPs provide a variety of transit options to serve the variety of transit needs. At the same time, NAPs assure that transit does not limit interconnectivity, defeat routing goals, or adversely effect existing networks (Almes, et al., 1992).

Peer networks are those that do not exchange traffic based on fees. Client networks are those that pay a flat fee to access the network through a larger provider. Thus, determining if a party is a peer or a client becomes significant in peering agreements. Its result will define the stability of the SKA model. A party's changing status from client to peer, represents fundamentally a loss of revenue to its counterpart party, as it will no longer pay interconnection fees. The telephony resolution in such matters has been an administrative process of service provider licensing, or fiat (Huston, 1994). Given the highly unregulated environment in which ISPs operate and the value-added network (VAN) status that ISPs get[72] from the regulatory authority,[73] fiats may not be the best option. As a result, other ways to solve the peer status have emerged. Some of those ways are: (1) the compliance of certain

[72] A Value-Added Network (VAN) is generally a commercial entity, such as a computer on-line service, or the Internet that uses some pre-existing facilities to function. It provides communications services, electronic store and forward mailboxing, e-mail, WWW service, and other related services for Electronic Data Interchange (EDI).

requirements imposed by the peering group, such as equipment, personnel knowledge, market reach, etc.; and (2) fee-based acknowledgments of change in status.

ISPs have formed several interconnection exchanges (IXs), such as the Commercial Interconnection Exchange (CIX), Network Access Points (NAPs), Federal Internet Exchanges (FIX), the Global Interconnection Exchange (GIX), the Metropolitan Area Ethernets (MAEs), etc. Little explicit information is available about these arrangements, as they have been subject to a variety of changes since the NSF decommission in April 1995.[74] Thus far, the majority of these IXs appear to be using a flat-fee charging scheme.

Huston concludes that (1) the flat-fee rule reflects the attributes of Internet technology (packet-switching networks[75]), where traditional accounting methods of the telephony do not work; (2) soon, perhaps, the transit model may turn into the dominant model of Internet interconnection, because of increasing transmission demands. However, at present, because of the flat-fee charging that ISPs (providers) encounter, they do not face incentives to implement traffic-based charging schemes to interconnecting parties: ISPs and end-users.

Thus, interconnection pricing is largely responsible for the non-traffic sensitive pricing schemes that dominate today's Internet. A large group of ISPs still face non-traffic-sensitive pricing schemes. ISPs, then, only face very loose incentives to implement traffic-sensitive schemes on their customers (other ISPs and end-users). Besides, when traffic-sensitive pricing schemes have been introduced they have not been of the type proposed in the previous chapter, because routers that support more accurate tracking systems (IPv6) are not yet the standard on today's Internet. In sum, non-traffic sensitive pricing schemes among ISPs starts a chain of this type of pricing, which reaches the end-users.

---

[73] The Federal Communications Commission also oversights the Internet.
[74] Srinagesh (1995) presents a brief overview of some of them.
[75] An accounting and billing mechanism that would be able to cost-effectively track all users' Internet transfers (packet counting) is difficult to implement, especially given the large public opposition and the competitiveness of the market.

7.3. Other Characteristics Of The Current Internet

Today's Internet has four characteristics that are relevant to this study. First, bandwidth is limited. This limited bandwidth prevents the widespread usage of bandwidth-intensive applications on the Internet, without causing congestion. Second, the Internet offers a single type of service, in which all packets are serviced on a best-effort, first-in-first-out (FIFO) basis. This single type of service severely limits the nature of applications that can be adequately supported. Third, there are no traffic fees; users are not charged on the basis of how many packets or the type of packets they send (although they depend on the capacity of the connection to the ISP). Fourth, usage generated by end-users connecting through an institution (educational, commercial, or governmental) is not passed back to them, but paid on a non-traffic sensitive basis by the institution.

7.4. Current Pricing Mechanisms

As mentioned above, the dominant pricing scheme for end-users on today's Internet is a non-traffic sensitive scheme. This is a direct result from the structure of Internet interconnection agreements (ISP peering) and the difficulty of introducing the more sophisticated routers as the Internet standard, which would allow a system to implement more efficient pricing schemes.

MacKie-Mason and Varian (1995) suggest that the most common non-traffic based charging mechanism used by ISPs on is probably a fixed monthly fee based on the capacity of the connection[76] (capacity-based pricing). This is the typical scheme paid by universities, government agencies, and large corporations, where end-users themselves pay nothing.

---

[76] Internet access has always been priced typically with an up-front lump sum to recover fixed costs.

Another alternative is the committed information rate (CIR) pricing, where a user faces a two-part fee for traffic: One part based on the maximum feasible bandwidth of the connection and the other part based on the maximum guaranteed actual bandwidth. Thus, for most traffic, the marginal packet placed on the Internet is priced at zero.

Yet another non-traffic sensitive pricing scheme is one based on the duration of the connection (i.e., per-unit-of-time pricing). In this scheme, users are allowed a certain number of hours per month (e.g., 20) at a flat rate. Extra hours are priced at a higher price. There are some ISPs that have even implemented a time-of day pricing schemes, where pricing varies depending on the hour of the day. However, these mechanisms do not confront congestion directly. That is, they merely attempt to reduce the number of users at peak times, but do not consider the fact that any user, at any time, may overload the network—making the pricing scheme inefficient.[77] This scheme is largely applied to end-users using a dial-up connection.

Srinagesh (1995) carried out a study based on posted prices and assuming that the ISP's market is competitive and thus prices should reflect marginal costs. He notes that a private line tariff often consists of a non-recurring charge, and a monthly charge based on the airline mileage between the two locations. The standard charges vary with the bandwidth (capacity) of the private line, but there are no traffic-sensitive charges. A company may offer a relatively high discount (around 60%) to customers whose monthly bill for a specified bundle of services exceeds $1 million. Those customers must also commit to maintaining that level of expenditure for a certain number of years (typically three to five). Also, ISPs typically charge their T1 customers twice the rate they charge their 56 Kbps customers, even though

---

[77] Even if it is during off-peak times, any user can theoretically download or transmit enormous amounts of data that could use up the whole bandwidth of a connection; thus preventing other users from obtaining access, at least for a while. Examples of overloading a network have continuously happened around the globe. The policy used to deal with this situations is one of warnings coming from the system administrator or even cancellation of accounts. But, unless usage is priced, there are really no enforcement mechanisms to efficiently deal with situations like this.

the T1 customers have 24 times the bandwidth. Term commitments can be seen in BARRNet's price structure. Fifty-six kbps customers are offered a 17% discount over monthly rates if they take a two-year prepaid contract. It is worth it to notice that these discounts arise from the presence of economies of scale and scope, present on the Internet, as a value-added network to the telephony.

There are few exceptions to this dominant non-traffic sensitive pricing mechanism. For example, more experienced digitized-data providers, such as CompuServe (a commercial on-line provider, now offering Internet access), charges per message above a fixed quota. Countries such as Chile and New Zealand have charged their users by the packet for international transmissions. These schemes have worked relatively well in those countries as users did not have other choice: No other pricing mechanism were available to them.

Although better than no charge per packet at all, per-packet charging is not efficient either, as the mechanism does not account for network load or the different valuations users place in their packets (from the same and from different applications). As MacKie-Mason, Murphy and Murphy (1995) point out: Users' valuations are heterogeneous among users and across time. That is, different users value different packets from the same application and/or from different applications, and these valuation vary also depending on a time-of-day basis.

7.5. Advantages And Disadvantages Of Non-Traffic Sensitive Pricing Schemes

The advantages are:

- Desirability by user as they do not feel constraint on strict usage quotas.
- Technical easiness to implement it, as there is no need for complicated measuring and billing mechanisms.
- Provides predictable costs for users, as they know exactly how much they need to pay.

The disadvantages include:

- By ignoring congestion, these schemes do not provide any incentives to flatten peak demands. So, they prevent some users from accessing the network during congestion times.
- They do not rank the value of different service requests (e.g., e-mail versus video), so it does not provide a mechanism for efficiently allocating resources.
- Small (below average users) are subsidizing large users—as all costs are recovered through a flat connection fee (the average connection fee), which keeps some users off the Internet.

7.6. How Congestion Is Presently Dealt With

Congestion is a negative externality that arises due to limited resources and the very core structure of the Internet: Packet-switching network. A packet-switched network is a shared-media technology in which all users share the medium of transportation (bandwidth). This makes bandwidth a very scarce resource. So, unless the technology of the Internet were to change, turning a user's bandwidth into a non-shared resource, other mechanisms have to be found to cope with the current technology. Modifications of flat-fee schemes, such as time-of-day pricing and duration-of-connection pricing are being used in an attempt to deal with congestion. However, these schemes do not help to internalize congestion, as incremental usage of bandwidth is priced at zero.

Also, current Internet technology uses two resource allocation mechanisms, none of these based on a price system: Randomization and first-in, first-out (FIFO). Randomization gives all packets an equal chance of getting through or being dropped. FIFO makes all packets to queue as they arrive. They both are inefficient. Delay is more costly for some packets than it is for others. Randomization and the First-in/First-out system do not account for different types of service request and thus cause inefficiencies.

7.7. Summary

Despite constant efforts to compensate in the provision of Internet capacity, demand is increasingly outweighing supply mainly because of: (1) limited resources, such as bandwidth, routers, IP addresses, among others; (2) an increasing use of intensive-bandwidth software applications (video, voice, interactive games, etc.) that make use of all the bandwidth available; and (3) the exponential growth Internet users. This has caused in acute congestion problem that current pricing schemes have not been able to efficiently deal with. In order to allocate resources efficiently (at their best use) and in attempt to solve congestion problems, traffic-based pricing has been proposed as a method of obtaining both economic and network efficiency.

Current pricing mechanisms are not optimal, they do not maximize the number of voluntary connections. This is specially true during congestion times. At present, technology does not allow to sort packets by the application the belong to or the rate of speed they require. All packets have equal priority in today's routers. As a result, with congestion, when a user who values a transfer more than other, attempts to make a transfer, he may be outcrowded by others who do not value the network as much (which inefficient). Current pricing mechanisms (flat-rate pricing, time-of-day pricing, duration-of-connection pricing, and others) do not allow individuals show their preferences. Despite their advantages, these mechanisms give rise to inefficiencies. This pattern starts within ISP's interconnection, which is at large currently priced on non-usage sensitive basis.

8. From Non-Traffic-Sensitive To Traffic-Sensitive Pricing Mechanisms: The Transition

In this chapter, I will first present the steps towards the transition to the rate-of-transfer pricing scheme. I will then mention some of the problems of this transitional stage. Then, I will show how the implementation of the rate-of-transfer pricing mechanism should be and mention some of the problems of doing it. I will suggest that this optimal pricing scheme is not currently implementable, as technological standards have to be introduced first: IPv6 routers. Despite that current pricing schemes do not directly control the major economic problem of the Internet today (congestion), this technical limitation results in the current non-optimal pricing schemes such as, time-of-day and duration-of-connection pricing schemes. Finally, I will briefly suggest an alternative way to dealing with congestion.

Future traffic control mechanisms are bound to become much more sophisticated. Specifically, the critical change of the future Internet is the adoption of the Internet Protocol version 6 (IPv6),[78] which will fundamentally allow routers to distinguish packets from different applications. This will in turn allow ISPs to discriminate and charge traffic accordingly on a per-user basis. This is the first step towards the implementation of the rate-of-transfer pricing here proposed. The IPv6 routers were reported to have been operational and used by some ISPs, specially in Europe (Com-priv@LISTS.PSI.COM, 1996). However, there is still along way until they become the Internet standard, if at all. The IPv6 transition plan is aimed at meeting four basic requirements:

---

[78] Also called Internet Protocol Next Generation (IPng).

- Incremental upgrade. Existing installed IPv4 hosts and routers may be upgraded to IPv6 at any time without being dependent on any other hosts or routers being upgraded.
- Incremental deployment. New IPv6 hosts and routers can be installed at any time without any prerequisites.
- Easy Addressing. When existing installed IPv4 hosts or routers are upgraded to IPv6, they may continue to use their existing address. They do not need to be assigned new addresses.
- Low start-up costs. Little or no preparation work is needed in order to upgrade existing IPv4 systems to IPv6, or to deploy new IPv6 systems.[79]

An inevitable result of the adoption of the new IPv6 standard is some form of usage-based pricing. A type-of-service sensitive network introduces the issue of performance incentives. Since users will choose good performance from their ISP, and if all else remains constant, there is nothing to motivate an end-user to indicate that his application is less performance sensitive (which would thereby degrade the performance he receives, specially during congested times). Informal enforcement may work in private networks, but in a public network with a large user community, it is highly unlikely that any informal enforcement mechanism will be sufficient to provide with the appropriate incentives to end-users.

Another important issue is that disregarding the level of competition of the Internet, economic efficiency is just as important of a goal as network efficiency is to Internet providers—ISPs. The Internet supplier market (at the national, regional, and local level) is just starting to take shape, it is not defined yet. It is ranging from a competitive structure to that of a regulated monopolist. Because of the social costs of congestion, it is expected that an ISP facing any type of market structure will improve its service—simply because if

---

[79] See Appendix C for further details.

ignored congestion will bring complaints from users, either in the form of regulation or in the form of loss of business.

For these reasons, it is expected that usage-sensitive pricing schemes will become an important part of the future Internet. An optimal pricing mechanism such as the rate-of-transfer pricing will help ISPs attain their goals. Despite that there is a tremendous pressure to offer non-usage based pricing schemes, I believe that they may last for quite some time, but in the long run, as the social costs of congestion become unbearable (too expensive for the end-user) and with the widespread availability of intensive-bandwidth applications, ISPs will no longer be able to support flat-rate prices For any user, the network can be congested at any particular point in time, even with time-of-day pricing, as the latter does not mitigate congestion at all, it only tries to reduce the number of users at a particular time of the day. Despite that usage (maximum and actual bandwidth) and the rate of transfer of a connection could be based on a time-of-day basis, it would not be optimal. Internet pricing should be based in the potential network load of a transfer connection.

It must also be observed that the development of an end-user interface is crucial for the functioning of the rate of transfer pricing. This interface should first allow the participation of the end-user in the decision-making process, which will produce, if successful, an optimal pricing mechanism. Second, it should allow the end-user to reserve and pay a certain rate of transfer. Third, it should allow the end-user to reveal which requests (incoming or outgoing) have been made be him only. That is it should avoid requests made by others, but mistakenly charged to the first. Fourth, this interface should display the different rate of transfer that can be supported by the network (based on current congestion) at that particular time, so that the user can choose. Finally, this interface should also let the user know the predicted time at which he can retry to send transfer requests, if the network is at full load. If the network is not congested, the price of the rate of transfer requested may be set low or could even be zero.

However, congestion is not primarily dependent on the number of users, but on the amount of usage imposed by each user. It is this expected, driven by users' interests, that usage-based pricing schemes will be demanded. Users will end up realizing that only when they are able to select a certain rate of transfer and a certain type of service, they are really satisfied with the service. It will take time, but this seems to be the most probable solution to deal with congestion, unless strong regulatory rules, intervene attempting to obtain similar results.

An additional change that will have to take place is a innovative technology allowing a cost-effective mechanism to account and bill users for their usage. This implies that the benefit from measuring usage (packets) of the marginal user has to be always greater than its marginal cost. In other words, transaction costs have to small so as to allow billing and accounting to be implementable. Until this happens other forms of pricing may be more desirable and everyone will have to live with the associated problems of congestion and inefficiencies.

Yet another change that is necessary is some form of standardization policy to deal with the issue of interoperability among computers. This issue would be solve if the standard Internet protocol becomes the IPv6, or a similar one. A single standard protocol is difficult to enforce in a global Internet, so interfaces may be necessary. Interfaces, though, reduce the welfare of users as they force them to select options that are not the most efficient. It is expected that the protocol that makes the greater market penetration in the next few years will become the *de facto* standard. An important remark has to be made, there is no reason to believe that the market will necessarily choose in time the standard that presumably provides the best results in the long run. So, there may be a case in which government intervention may (assuming good intentions) help solve eventual market failures—after having selected the wrong technology.

Finally, there is some debate as to whether the government should subsidize the network directly and make its own technical choices, or subsidize the users and let them choose the technology. While the subsidization of the network may help to solve the standardization issue, the subsidization of the end-user may help solve the Universal Service controversy. But as Wenders (1996) suggests, “Universal Service may only be an issue of pure politics.” So, the only room for government subsidization may be to concentrate on standardization issues and promoting a competitive market environment.

8.1. Implementation Of An Optimal Pricing Scheme

The fundamental problem is to allocate the most[80] scarce resource (bandwidth) in an efficient way by concentrating on the pricing of Internet usage. Also, given that government subsidies are long gone, ISPs need to implement a self-finance provision mechanism to recover provision costs efficiently. A long-run solution is imminent. Solutions group in two main categories: Non-pricing mechanisms and pricing mechanisms. Let us take the first group. First, administratively assigning different priorities to different types of traffic is appealing, but impractical due to the usual inefficiencies of rationing (MacKie-Mason and Varian, 1993):

- Knowing that a user's valuation of an application varies over time, who is going to decide what is fair? Only the user knows, that is why he must see cost-based prices.
- Special Interests, etc.

Assignment of priority is also difficult to enforce. An opportunistic user may change the administratively attached priority to obtain the highest service. So, it is not a long-run solution.

Another non-price mechanism is a voluntary effort to control congestion, but is not definite. Peer pressure and user ethics are said to be enough to control congestion costs. However, many users may not even understand the costs their are imposing on others when they use the Internet. Peer pressure may be ineffective over any malevolent user who want to intentionally cause damage.

Yet another non-price scheme is an over-provision of resources to maintain sufficient network capacity to support peak demands without degrading service. This has been the

[80] Other important scarce resources are IP addresses, which will be increased with the standardization of next Internet protocol IPv6, and computer infrastructure, which can be expected to increase in the long-run.

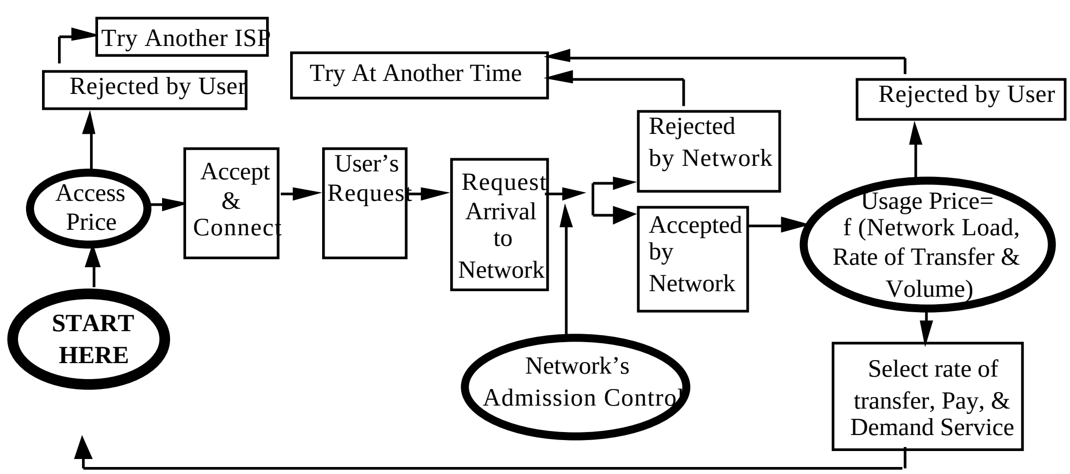
most widely used mechanism used to date in the Internet (MacKie-Mason, Varian, 1994). It is expensive to have idle capacity, but it may be more expensive to supply more capacity—a cost-benefit analysis is necessary to make a decision. It is also not a long-run solution as the increased use of intensive-bandwidth applications and the exponential growth of Internet users will exacerbate the problem.

The second group of solutions are price mechanisms. Not every price mechanism will work, it ought to have the appropriate elements and it has to be implemented successfully. That is, the optimal pricing mechanism should:

- Be traffic sensitive
- Have rate-of-transfer adaptability
- Have an admission control option.

Figure 19 will show how an optimal pricing mechanism should work.

Figure 20. Layout Of An Optimal Pricing Mechanism.



The user would first face an *access price*. This access price should be recovered with non-usage sensitive fees, i.e., charging users a fixed fee depending on the capacity of their connection. If price is rejected then the user may try using a different ISP. If price is accepted, then the user gets connected to the network. Then if he wants to use the network at a particular time, the user would send a signal requesting service (Telnet, FTP, e-mail, WWW). This request then arrives to the network and it is processed based on time of arrival (in relation to other requests) and on the current (including the request) network load. If it is rejected, then user may try at another time. If it is accepted, the network computes rates of transfer available (based on network load), their price (based on network load, rate of transfer, and size of transfer request), and their time of completion. Time of completion can be calculated when the user makes his selection. Next, the network displays the available options to the user, including price and rate of transfer. Then, the user tests different possible times of completion and makes his selection, pays and obtains service. Since the size of the transfer request is generally known and the user has chosen a certain rate of transfer, time can be computed as: time=length/velocity. If usage price is rejected by user, he can try at another time.

The important aspect to achieve an optimal pricing scheme is that users should be part of the decision process, allowing them to decide whether or not they want to pay the usage price at any given time, which in turn depends on the network load. The end-user is assumed to be able to select these pricing options from his own computer, as a computer interface with this purpose should be available. In addition, this pricing scheme would be incentive-compatible if for higher rate of transfer requested, a higher price is charged. Also, to guarantee a certain rate of transfer an admission control system should be implemented in the ISP network. This is the implementation of the proposed rate-of-transfer pricing, which would achieve economic efficiency, network efficiency, and would allow an efficient expansion of capacity.

8.2. Problems Of This Pricing Scheme

The biggest drawback of the proposed mechanism is that it is not implementable yet given the current state of technology. That is, routers capable of distinguishing between different types of packets are not the standard on the Internet yet. They are just beginning to be tested. Nevertheless, these routers, using the IP protocol version 6 (IPv6), are undoubtedly the start for the implementation of more efficient pricing mechanisms, such as the rate-of-transfer pricing.

Also, if the optimal pricing mechanism described in this section is not implemented at a parallel level, i.e., all ISPs use it, then it may not work as Bailey (1995) showed for Italy. In this case, the government may attempt to set some standards, but it is very likely to receive an overwhelming opposition of the people (because users do not like usage-based pricing schemes) and thus may not be able to achieve anything promising. Something to keep in mind is that the more time it takes to make the transition from current inefficient pricing schemes to a traffic-based pricing system, the more difficult it will be and there are less chances of succeeding: The more people in the network, the more people will resist, as the non-usage-based system is very well liked by users, despite congestion delays.

8.3. An Alternative To Dealing With Congestion

An alternative way of dealing with current Internet problems (congestion) may be by combining technologies. That is, circuit switched connections may be used for the most demanding Internet applications (bandwidth-intensive), and packet-switched connections may be used for those less demanding applications (non-bandwidth-intensive). I speculate that the costs of implementing a mixed mechanism of this kind may be more cost-effective at the time of measuring traffic and billing users than one purely based on packet-switched technology. This is because, it is the transfers of bandwidth-intensive applications the ones that are most devastating on network loads. So, if these transfer receive a dedicated line for transfer, like a telephone line, many problems may be solved, but detailed studies will have to test this hypothesis first. Notice, however, that this mixed technology will not be able to omit packet-based transfers, so the rate-of-transfer pricing may still be applicable.

8.4. Summary

I have suggested that the first step towards the implementation of the rate-of-transfer pricing is a change of *status quo*: The introduction of the IPv6. I describe the main step towards this transition and its problems. On this basis, I illustrate the implementation of the rate-of-transfer pricing. Then, I affirm that because of this inability to implement an optimal pricing scheme (rate-of-transfer pricing), second-best pricing schemes appear such as, time-of-day pricing, duration-of-connection pricing, and other non-usage-based pricing schemes. Finally, I suggest an alternative method to dealing with congestion: A network using mixed technology (packet- and circuit-switching).

9. Conclusions And Further Areas Of Research

The Internet is bound to have a large impact on people's lifestyles, by being the coordinating network towards the Information Superhighway, or by simply reducing the overpriced long-distance telephone rates. At any rate, the Internet will find its place in today's competitive world and hence deserves attention. This thesis provides both an economic understanding of the Internet and an optimal pricing mechanism to fight the major problem of today's Internet: Congestion.

The study is divided in two parts: The first part describes the structure of the Internet and the second part discusses the economics of this network. Having started as a project from the U.S. government, the Internet today is a widely used public resource. However, despite the private industry's role as the largest provider of Internet resources, and because of the exponential growth experienced in this newly created and loosely regulated industry, congestion has become its major problem.

The Internet is an amalgamation of computer networks around the world. It is composed of three types of Internet Service Providers (ISPs), national, regional, and local. These entities interconnect and exchange data at the Network Access Points (NAPs) and similar structures by leasing transport lines from the telephone companies. End-users obtain Internet connectivity from local ISPs primarily by dial-up and direct connection.

With the more active role of industry, during the last five years, pricing mechanisms have been sought to fight congestion and efficiently allocate resources; hence, many proposals have been written. However, the inability and reluctance to implement the proposed mechanisms has encouraged Internet providers to utilize second-best pricing schemes. So, flat-fee and time based schemes are the dominant pricing schemes in today's Internet. These are typically based on two elements: The size of the maximum bandwidth connection and the duration of the connection. These alternative mechanisms have not done much towards controlling congestion, because congestion is a function of both the amount and the type of packets transferred, but not a function of the time that a user remains connected to the network, or the time of day users connect to the network. They also have substantial support from users, who have become accustomed to this kind of pricing schemes in similar networks, telephone. While this thesis has considered these aspects, it proposes an optimal pricing mechanism—Rate-of-Transfer Pricing—which is designed to be implemented specifically by local ISPs on Internet's end-user.

The Rate-of-Transfer pricing mechanism is a multi-part tariff. Its fixed fee is based on the maximum bandwidth of the connection and its variable fee is based on: (1) the degree of current network load, (2) the rate of transfer selected by the user, and (3) the size of the transfer request. The Rate-of-Transfer (RoT) pricing is an optimal mechanism because it is a type of marginal pricing (depends on the degree of network load), and it fights congestion efficiently by allocating resources according to the user's and the transfer's needs. So, it achieves economic efficiency.

Nevertheless, despite proposals, such as the RoT, other factors have precluded the implementation of optimal pricing schemes and led to the deployment of second-best pricing mechanisms. The most important factor not allowing the implementation of an optimal pricing scheme is the ignoring by and unavailability to the vast majority of ISPs world-wide of IPv6-routers (routers able to distinguish packets from different Internet tools). There have been some reports of IPv6-routers to have been used in Europe, but their use is not yet widespread. Another factor for the use-of these second-best pricing schemes is the lack of incentives for ISPs to become more efficient. Congestion problems have been approached by simply overprovisioning capacity. It is expected, however, that with the increasing popularity of the Internet, demand of service will outweigh supply, especially because demand prices are not traffic based.

It is important to notice that congestion could be substantially alleviated if a mixture of technology becomes possible. That is, if the circuit-switch technology of the telephone network and packet-switch technology of the current Internet are joined in such a way that the more intensive-bandwidth applications are not dependent on the congestion level of routers but are served instantaneously, end-to-end, by dedicated bandwidth. This may need the introduction of different types of routers than those available with the IPv6 protocol.

Further, as pointed out throughout this investigation, there are numerous areas that need empirical studies. Perhaps, one of the most important empirical studies is the analysis of a particular technology that would allow the implementation of a cost-effective measuring and billing packet system. This would contribute to solve the dilemma of being able to efficiently implement a traffic-based pricing system for the Internet.

Another important area of research is the identification of the most appropriate role of government in contributing to a better interaction of this network with actual or potential users. Today, government efforts related to the Internet are mostly directed to the wide-spread connectivity of public institutions. But, perhaps, some government action is also required in the standardization of protocols and interfaces that would allow the highly-sought introduction of the IPv6 by the Internet community. Yet another area of research is that which focuses on the probabilistic study of network congestion. This type of study would help identify patterns of use and therefore allocate scarce resources such as, bandwidth, in a more efficient manner.

Appendix A: The U.S. Telecommunications Sector

Mitchell and Srinagesh (1995) present a comprehensive survey of the development of the telephone network. In the United States, the public telephone network has developed in four distinguishable phases. First, after the acquisition of the patent in 1876, telephone service was begun by competing private companies linking a community of business subscribers within small central-city areas. The competition by phone networks separated customers from different networks, so businesses had to maintain separate telephones for each network to reach some of their partners. Second, competition proved unsustainable leading to the disappearance of local networks as they were merged or acquired by AT&T. Thus, local telephone networks began to interconnect with distant cities, extending their reach by long-distance lines. Third, the majority of the costs of telephone service in the integrated Bell system were in the network infrastructure, i.e., the distribution of cables and local switching facilities. With these large fixed costs and other large network costs common to supplying businesses and residential service, the industry had many elements of a natural monopoly: The costs of production by a single firm fall below the total costs of two or more firms. Then, a monopoly structure dominated by AT&T and regulated by government commissions overtook the industry. This structure essentially realized the advantages of economies of scale and the balance of political power. With nationwide connectivity achieved, federal policy (in the Communications Act of 1934) proclaimed the goal of *Universal Service* and telephone companies emphasized expanding the number of subscribers and extending service to high-cost and remote areas. Customers placed widely different values on obtaining service, this allowed phone companies to price discriminate their service. Universal Service was then feasible. It was fundamentally financed by excess revenues collected from businesses, and long-distance service and allocated to subsidize the local exchange infrastructure, keeping basic residential rates low. Fourth, in an attempt to restructure economic efficiency, competition policy led to the breaking up of AT&T into separate long-distance and local business in 1984. Price structures that favored large customer classes were replaced by programs to assist marginal subscribers, and barriers to entry of competitors and new technologies were dismantled.

Factors Promoting Competition In Today's Telecommunications Market

Today, once again, the U.S. telecommunications sector is soundly on the path to open competition. First, technological advances and regulatory changes have driven the transformation of the Telecommunications industry from monopoly to competition. The cost structure of most of the pre-1970s telephone network had strong elements of natural monopoly. However, economies of scale for producing telephone terminal equipment did not play an important role, so when the market was deregulated, the successful entry of a variety of suppliers market the first step toward disintegration. Since the 1950s, microwave technology proved a superior and lower-cost technology for long distance transmission and greatly reduced the incumbent's scale and right-of way advantages in inter-city markets. Microwave Communications, Inc. (MCI) then became a powerful national competitor. Recently, technological advances, such as fiber optics and digital switching have enable more competitors to enter formally monopoly markets. Also, technological advances are rapidly changing traditional industry boundaries, which in turn enables competition between firms that did not previously compete with one another. The development of telephony and Internet services over cable television networks are likely to result in additional competition.

Second, pro-competitive public policies have increased in number as technology has changed. The economic benefits of competition with alternatives communication technologies have increasingly been recognized. When political power shifted from states to the federal government, the regulatory barriers facing new entrants were lowered. Regulators have explicitly adopted competitive instruments in important situations. For instance, the NSF issued solicitations for the construction of the current Internet infrastructure and then awarded the winners based on their bids.

Third, there has been a consistent growth in the demand for communication services. Consumption patterns for phone, fax, data services have growth consistently. More and more people have realized the advantages of using innovative technologies for their needs. Business companies have changed the use of their information systems from a local to a inter-company context. An increasing number of households, including those who attempt to run their businesses from their own homes, those who try to keep up with the technological era of the 90s, among others, have opened new markets for telecommunication equipment.

Finally, because of the presence economies of scale and scope in the telecommunications industry, firms wisely have used their facilities to supply different services—mainly fax, phone, IP transport, and cable—simultaneously. The presence of this type of economies has led firms to *integrate both vertically and horizontally*. A vertical integration has occurred as Telecommunications firms have integrated their local, long distance, and international service provision. A horizontal integration has occurred across geographic areas as well as across non-substitutable services, such as cable TV, phone, and IP transport. Local and long-distance companies are racing to enter each other's markets, cable TV operators are experimenting supply of service with switched networks and Internet access, and wireless suppliers are expanding into mobile data and fax messaging services.

Appendix B: Internet Service Provider (ISP) Peering

ISP peering refers to the arrangements that ISPs (national, regional, and local) make between each other to provide for Internet interconnection. Of these arrangements, two types are the most used: The Sender Keep All (SKA) Method and the Financial Settlements Method.

SKA peering is a structure where traffic is exchanged between ISPs without mutual charge. Within a national structure it is typically the case that the marginal cost of international traffic transferred to and from the rest of the Internet is significantly higher than domestic traffic transferred. For this reason, SKA peering is mostly related to domestic traffic—international traffic is either provided by a separate agreement, or provided independently by each party. The essential criteria for a stable SKA peering structure is to establish equality in the peering relationship. There are a number of ways in which this can be achieved, including the use of entry pricing into the peering environment, or the use of peering criteria (e.g., specification of ISP network infrastructure, or network level of service and coverage areas as eligibility for peering).

An alternative structure is to use bilateral financial settlements as the basis of peering. In this model, one party provides common infrastructure services to the other party, using a financial settlement to offer the transfer costs of single service provision. There are a number of ways in which financial settlement can be determined, including direct negotiation, traffic exchange levels, network routing exchanges, and similar. Traffic volume is perhaps the most obvious means of determining a metric to be used as a basis of financial settlement, and such a metric can accommodate a number of variations in peering models. However it must be observed that traffic volumes on the Internet are not as rigidly structured as those within the telephone network environment, and within this environment it is not possible to readily determine which party's client generated a bi-directional traffic flow across the peering structure (e.g., FTP transfers).

Comparing SKA Peering Model With The Financial Settlements Structure.

The major factors in comparing SKA and financial settlement peering models is the determination whether there is an agreeable methodology of determining financial settlements which is sustainable. Within the telephone network domain there is the concept of the originating party being the party that pays for the end-to-end call. A settlement is a balancing of the number of originating calls exchanged between two parties, and a financial reconciliation undertaken at a settlement transfer price.

Within the Internet the concept of an originating party does not readily translate into measurement technologies, or to actual end client billing models. In an FTP transfer, for instance, the originating party requests a file from a remote machine. Under an inefficient settlements model, it would be the remote machine the one charged for the transfer. Accordingly, ISP peering settlement models based on settlements of number, volume, and duration of originating TCP/IP sessions are not readily reconcilable within the Internet domain. A solution is to use a received-traffic statistic. The use of received-traffic volumes is a coarse approximation to volume of originating sessions, based on the observation that the most predominant interaction at the user level is the importation of data.

In sum, the SKA model imposes self-sufficiency on each ISP, which implies that each peering entity uses the peering structure solely for the exchange of user data. This is not generally the case, however, as peering structures typically involve both peer interaction and the provision of transit or infrastructure services. The major challenge with an SKA peering structure is whether it is possible to appropriately define what is encompassed within common peer interactions and what falls within a client/service provider domain. If the relation is more of a client/provider, rather than peer/peer, SKA agreements would not apply, and, perhaps, a financial settlements model would be more appropriate (Huston 1994).

An Example Of AN SKA Model: The CIX Interconnection Model

The Commercial Internet eXchange (CIX) allows a multilateral and bilateral agreements, which include the following actions:

- No restrictions on use to other participants
- No settlement fees
- No liability for traffic
- Participants may involve in any separate contract or agreement with any other participant or third party
- Participants shall bear their own expenses of interconnection and shall agree on the type and speed of such interconnect on a bilateral basis
- A Network Operations Center (NOC) should support in cooperation with other participants the smooth operation of the internetwork services
- Each participant may choose to connect at more than one location
- No participant shall monitor or capture the contents of any data or traffic that is exchanged under interconnections covered by this agreement unless an appropriate legal court order is in force. No statistical information by participant, by company, interconnect point shall be supplied to any third party under any circumstances
- Each participant shall indemnify and hold the non-liable participants harmless from and against any loss, damage, liability, claim or expense which results from a claim or

claims asserted by unrelated third parties concerning an action or omission of such participant; and

- Each participant is responsible for assessing its own need for property, casualty, and liability insurance and each shall obtain such insurance as each sees fit. Each participant shall bear the risk of loss to its own equipment and agrees to not make any claims against other participants for any property loss, nor assign that right to any third party

The duration of the agreement is for two years and will automatically renew for additional two year terms. A participant may terminate its responsibilities under this agreement by giving all other participants 180 days written notice, or sooner by mutual consent of all participants.[81] Some CIX members include, Alternet, ANS CO+RE Systems, Inc., Apex Global Info Systems(AGIS), CompuServe, NorthWestNet, Pacific Bell Internet, and PSINet.[82]

ISP peering problems: The CIX-ANS dispute.

In 1991, a dispute arose between a consortium of commercial networks represented by the Commercial Internet Exchange (CIX) and the government-contractor provider of the network backbone services, Advanced Network Services (ANS). The dispute involved whether ANS had abused its position to gain an unfair competitive advantage.

The dispute arose when ANS formed a commercial subsidiary, called ANS CO+RE (Commercial and Research) Systems, Inc., which competed with the other private network vendors. As a result of its government contract, ANS provided the NSFNET backbone service. ANS implemented the federal government's Acceptable Use Policy (AUP), which

---

[81] Extracted from the 1995-1996 CIX Strategic Plan. Appendix D: Multi-Lateral Peering Agreement V. 1.0. March 30, 1995. Distributed electronically. http://www.cix.org/CIXdocs/sp-toc.html.
[82] For a complete list of members see http://www.cix.org/CIXInfo/members.html.

restricted the NSFNET backbone service from carrying commercial traffic. CIX members charged that this requirement provided ANS with an unfair commercial advantage. ANS claimed that its commercial operations served the public interest by subsidizing research and education users from the fees generated by commercial use of the network.

The dispute erupted publicly when the parties aired their grievances before the U.S. House Committee on Science on March 12, 1992. Following the public hearing, ANS and CIX reached a trial interconnection agreement and decided to discuss how ANS should charge for network backbone service (Perlman, 1995).

Appendix C: Characteristics Of The Internet Protocol Version 6 (IPv6)

IPv6[83] is a new version of the Internet Protocol. It has been designed as an evolutionary, rather than revolutionary, step from IPv4. Functions that are generally seen as working in IPv4 were kept in IPv6. Functions that do not work or are infrequently used were removed or made optional. A few new features were added where the functionality was felt to be necessary.

The Important Features Of IPv6 Include:

- **Type of service capabilities: It is probably the most critical change from the previous protocol. IPv6 has a new capability to enable the labeling of packets belonging to particular traffic "flows" for which the sender has requested special handling, such as non-default type of service or "real-time" service.**

- Expanded addressing and routing capabilities: The IP address size is increased from 32 bits to 128 bits providing support for a much greater number of addressable nodes, more levels of addressing hierarchy, and simpler auto-configuration of addresses.

- The scaleability of multicast routing is improved by adding a "scope" field to multicast addresses: A new type of address, called a "cluster address" is defined to identify topological regions rather than individual nodes. The use of cluster addresses in conjunction with the IPv6 source route capability allows nodes additional control over the path their traffic takes.

---

[83] Source: Bradner, S. and A. Mankin (1995), RFC 1752.

- Simplified header format: Some IPv4 header fields have been dropped or made optional to reduce the common-case processing cost of packet handling and to keep the bandwidth overhead of the IPv6 header as low as possible in spite of the increased size of the addresses. Even though the IPv6 addresses are four times longer than the IPv4 addresses, the IPv6 header is only twice the size of the IPv4 header.

- Support for extension headers and options: IPv6 options are placed in separate headers that are located in the packet between the IPv6 header and the transport-layer header. Since most IPv6 option headers are not examined or processed by any router along a packet's delivery path until it arrives at its final destination, this organization facilitates a major improvement in router performance for packets containing options. Another improvement is that unlike IPv4, IPv6 options can be of arbitrary length and not limited to 40 bytes. This feature plus the manner in which they are processed, permits IPv6 options to be used for functions that were not practical in IPv4. A key extensibility feature of IPv6 is the ability to encode, within an option, the action which a router or host should perform if the option is unknown. This permits the incremental deployment of additional functionality into an operational network with a minimal danger of disruption.

- Support for authentication and privacy: IPv6 includes the definition of an extension that provides support for authentication and data integrity. This extension is included as a basic element of IPv6 and support for it will be required in all implementations. IPv6 also includes the definition of an extension to support confidentiality by means of encryption. Support for this extension will be strongly encouraged in all implementations.

- Support for auto configuration: IPv6 supports multiple forms of auto configuration, from "plug and play" configuration of node addresses on an isolated network to the full-featured facilities offered by DHCP.

- Support for source routes: IPv6 includes an extended function source routing header designed to support the Source Demand Routing Protocol (SDRP). The purpose of SDRP is to support source-initiated selection of routes to complement the route selection provided by existing routing protocols for both inter-domain and intra-domain routes.

- Simple and flexible transition from IPv4: The IPv6 transition plan is aimed at meeting four basic requirements:

    I. Incremental upgrade: Existing installed IPv4 hosts and routers may be upgraded to IPv6 at any time without being dependent on any other hosts or routers being upgraded.

    II. Incremental deployment: New IPv6 hosts and routers can be installed at any time without any prerequisites.

    III. Easy Addressing: When existing installed IPv4 hosts or routers are upgraded to IPv6, they may continue to use their existing address. They do not need to be assigned new addresses.

    IV. Low start-up costs: Little or no preparation work is needed in order to upgrade existing IPv4 systems to IPv6, or to deploy new IPv6 systems.

10. Glossary

56k Line

A digital phone-line connection (leased line) capable of carrying 56,000 bits-per-second. At this speed, a Megabyte will take about 3 minutes to transfer. This is 4 times as fast as a 14,400 bps modem.

ADN

(Advanced Digital Network). Usually refers to a 56 Kbps leased-line.

ARPANET

(Advanced Research Projects Agency Network). The precursor to the Internet. Developed in the late 60's and early 70's by the US Department of Defense as an experiment in wide-area-networking that would survive a nuclear war.

ASCII

(American Standard Code for Information Interchange). This is the de facto world-wide standard for the code numbers used by computers to represent all the upper and lower-case Latin letters, numbers, punctuation, etc. There are 128 standard ASCII codes each of which can be represented by a 7 digit binary number: 0000000 through 1111111.

Backbone

A high-speed line or series of connections that forms a major pathway within a network. The term is relative as a backbone in a small network will likely be much smaller than many non-backbone lines in a large network.

Bandwidth

How much “stuff” you can send through a connection. Usually measured in bits-per-second. A full page of English text is about 16,000 bits. A fast modem can move about 15,000 bits in one second. Full-motion full-screen video would require roughly 10,000,000 bits-per-second, depending on compression.

Baud

In common usage the “baud rate” of a modem is how many bits it can send or receive per second. Technically, “baud” is the number of times per second that the carrier signal shifts value—so a 2400 bit-per-second modem actually runs at 300 baud, but it moves 4 bits per baud (4 x 300 = 1200 bits per second).

BBS

(Bulletin Board System). A computerized meeting and announcement system that allows people to carry on discussions, upload and download files, and make announcements without the people being connected to the computer at the same time. There are many thousands of BBS’s around the world, most are very small, running on a single IBM clone PC with 1 or 2 phone lines. Some are very large and the line between a BBS and a system like CompuServe gets crossed at some point, but it is not clearly drawn.

Bit

(Binary digit). A single digit number in base-2, in other words, either a 1 or a zero. The smallest unit of computerized data. Bandwidth is usually measured in bits-per-second.

BITNET

(Because It’s Time Network). A network of educational sites separate from the Internet, but e-mail is freely exchanged between BITNET and the Internet. Listservs, the most popular form of e-mail discussion groups, originated on BITNET. BITNET machines are IBM VM machines, and the network is probably the only international network that is shrinking.

Bps

(Bits-Per-Second). A measurement of how fast data is moved from one place to another. A "28.8 modem" can move 28,800 bits per second.

Browser

A client program (software) that is used to looking at various kinds of Internet resources.

Byte

A set of Bits that represent a single character. Usually there are 8 Bits in a Byte, sometimes more, depending on how the measurement is being made.

Client

A software program that is used to contact and obtain data from a Server software program on another computer, often across a great distance. Each Client program is designed to work with one or more specific kinds of Server programs, and each Server requires a specific kind of Client. A "Web Browser" is a specific kind of Client.

Cyberspace

Term originated by author William Gibson in his novel "Neuromancer", the word Cyberspace is currently used to describe the whole range of information resources available through computer networks.

Dial-up Connection

A temporary, as opposed to dedicated, connection between machines established over a standard phone line.

Domain Name

The unique name that identifies an Internet site. Domain Names always have 2 or more parts, separated by dots. The part on the left is the most specific, and the part on the right is the most general. A given machine may have more than one Domain Name but a given Domain Name points to only one machine. Usually, all of the machines on a given Network will have the same thing as the right-hand portion of their Domain Names, e.g.

gateway.gbnetwork.com

mail.gbnetwork.com

www.gbnetwork.com

and so on. It is also possible for a Domain Name to exist but not be connected to an actual machine. This is often done so that a group or business can have an Internet e-mail address without having to establish a real Internet site. In these cases, some real Internet machine must handle the mail on behalf of the listed Domain Name.

E-mail

(Electronic Mail). Messages, usually text, sent from one person to another via computer. E-mail can also be sent automatically to a large number of addresses (Mailing List).

E-mail Address

The domain-based or UUCP address that is used to send electronic mail to a specified destination. For example an editor's address is "gmalkin@xylogics.com".

Ethernet

A 10-Mbps standard for Local Area Networks (LANs), initially developed by Xerox, and later refined by Digital, Intel and Xerox (DIX). All hosts are connected to a coaxial cable where they contend for network access using a Carrier Sense Multiple Access with Collision Detection (CSMA/CD) paradigm. See also: 802.x, Local Area Network, token ring.

FAQs

(Frequently Asked Questions). FAQs are documents that list and answer the most common questions on a particular subject. There are hundreds of FAQs on subjects as diverse as Pet Grooming and Cryptography. FAQs are usually written by people who have tired of answering the same question over and over.

FDDI

(Fiber Distributed Data Interface). A standard for transmitting data on optical fiber cables at a rate of around 100,000,000 bits-per-second (10 times as fast as Ethernet, about twice as fast as T-3).

Federal Networking Council (FNC)

The coordinating group of representatives from those federal agencies involved in the development and use of federal networking, especially those networks using TCP/IP and the Internet. Current members include representatives from DOD, DOE, DARPA, NSF, and NASA.

File Transfer

The copying of a file from one computer to another over a computer network.

Finger

An Internet software tool for locating people on other Internet sites. Finger is also sometimes used to give access to non-personal information, but the most common use is to see if a person has an account at a particular Internet site. Many sites do not allow incoming Finger requests, but many do.

Fire Wall

A combination of hardware and software that separates a LAN into two or more parts for security purposes.

FIX

Federal Information Exchange is one of the connection points between the American governmental internets and the Internet.

FTP

(File Transfer Protocol). A very common method of moving files between two Internet sites. FTP is a special way to login to another Internet site for the purposes of retrieving and/or sending files. There are many Internet sites that have established publicly accessible repositories of material that can be obtained using FTP, by logging in using the account name “anonymous”, thus these sites are called “anonymous FTP servers”.

Gateway

The technical meaning is a hardware or software set-up that translates between two dissimilar protocols, for example Prodigy has a gateway that translates between its internal, proprietary e-mail format and Internet e-mail format. Another, sloppier meaning of gateway is to describe any mechanism for providing access to another system, e.g. AOL might be called a gateway to the Internet.

Gopher

A widely successful method of making menus of material available over the Internet. Gopher is a Client and Server style program, which requires that the user have a Gopher Client program. Although Gopher spread rapidly across the globe in only a couple of years, it is being largely supplanted by HyperText, also known as WWW (World Wide Web). There are still thousands of Gopher Servers on the Internet and we can expect they will remain for a while.

Header

The portion of a packet, preceding the actual data, containing source and destination addresses, error checking and other fields. A header is also the part of an electronic mail message that precedes the body of a message and contains, among other things, the message originator, date and time.

High Performance Computing and Communications (HPCC)

High performance computing encompasses advanced computing, communications, and information technologies, including scientific workstations, supercomputer systems, high speed networks, special purpose and experimental systems, the new generation of large scale parallel systems, and application and systems software with all components well integrated and linked over a high speed network.

Host

Any computer on a network that is a repository for services available to other computers on the network. It is quite common to have one host machine provide several services, such as WWW and USENET.

HTML

(HyperText Markup Language). The coding language used to create HyperText documents for use on the World Wide Web. HTML looks a lot like old-fashioned typesetting code, where you surround a block of text with codes that indicate how it should appear, additionally, in HTML you can specify that a block of text, or a word, is “linked “ to another file on the Internet. HTML files are meant to be viewed using a World Wide Web Client Program, such as Mosaic.

HTTP

(HyperText Transport Protocol). The protocol for moving HyperText files across the Internet. Requires a HTTP client program on one end, and an HTTP server program on the other end. HTTP is the most important protocol used in the World Wide Web (WWW).

HyperText

Generally, any text that contains “links” to other documents—words or phrases in the document that can be chosen by a reader and which cause another document to be retrieved and displayed.

Internet

(Upper case I) The vast collection of inter-connected networks that all use the TCP/IP protocols and that evolved from the ARPANET of the late 60's and early 70's. The Internet now (July 1995) connects roughly 60,000 independent networks into a vast global internet.

internet

(Lower case i) Any time you connect 2 or more networks together, you have an internet—as in inter-national or inter-state.

Internet Architecture Board (IAB)

The technical body that oversees the development of the Internet suite of protocols. It has two task forces: the IETF and the IRTF. "IAB" previously stood for Internet Activities Board.

Internet Engineering Task Force (IETF)

The IETF is a large, open community of network designers, operators, vendors, and researchers whose purpose is to coordinate the operation, management and evolution of the Internet, and to resolve short-range and mid-range protocol and architectural issues. It is a major source of proposals for protocol standards which are submitted to the IAB for final approval. The IETF meets three times a year and extensive minutes are included in the IETF Proceedings.

Internet Protocol (IP)

The Internet Protocol, defined in STD 5, RFC 791, is the network layer for the TCP/IP Protocol Suite. It is a connectionless, best-effort packet switching protocol. See also: packet switching, Request For Comments, TCP/IP Protocol Suite.

Internet Relay Chat (IRC)

A world-wide "party line" protocol that allows one to converse with others in real time. IRC is structured as a network of servers, each of which accepts connections from client programs, one per user.

Internet Society (ISOC)

The Internet Society is a non-profit, professional membership organization which facilitates and supports the technical evolution of the Internet, stimulates interest in and educates the scientific and academic communities, industry and the public about the technology, uses and applications of the Internet, and promotes the development of new applications for the system. The Society provides a forum for discussion and collaboration in the operation and use of the global Internet infrastructure. The Internet Society publishes a quarterly newsletter, the Internet Society News, and holds an annual conference, INET. The development of Internet technical standards takes place under the auspices of the Internet Society with substantial support from the Corporation for National Research Initiatives under a cooperative agreement with the US Federal Government.

Interoperability

The ability of software and hardware on multiple machines from multiple vendors to communicate meaningfully.

IP Number, IP Address, or Internet Address

Sometimes called a "dotted quad". A unique number consisting of 4

parts separated by dots, e.g.

165.113.245.2

Every machine directly connected to the Internet has a unique IP number—if a machine does not have an IP number, it is not really on the Internet. Most machines also have one or more Domain Names that are easier for people to remember. Domain Name allows the translation from numbers to letters. So, when people find information and share it with others on the Internet they do it by means of addresses (sort of like postal addresses when you mail a letter). Any piece of information which can be found on the Internet has an address. Addresses are used to define the location of a computer connected to the Internet, or a file or piece of information on a computer connected to the Internet. Generally speaking, there are several types of addresses: E-mail, URL, and newsgroup.

E-mail addresses look like this: s.omeone@some.where. You need an e-mail address in order to send and receive e-mail.

URLs (Uniform Resource Locators) are the addresses used to identify various types of Internet sites including Web sites. If you are using Netscape to connect to the Internet, the address or URL of the site you are currently connected to is listed in the Location/Netsite box near the top of your Netscape window. URLs start with HTTP, https, FTP, or gopher. If you know the Internet address, you can type it in the Location/Netsite box at the top of your Netscape window and press Enter. Examples of URLs are:

HTTP//www.anyweb.site

FTP//down.upload.afile

gopher://want.to.findfile

Newsgroup addresses are used to identify newsgroups. Examples of newsgroup addresses are:

rec.music.folk

alt.pets.hamsters

rec.sport.golf

ISDN

(Integrated Services Digital Network). Basically a way to move more data over existing regular phone lines. ISDN is rapidly becoming available to much of the USA and in most markets it is priced very comparably to standard analog phone circuits. It can provide speeds of roughly 128,000 bits-per-second over regular phone lines. In practice, most people will be limited to 56,000 or 64,000 bits-per-second.

ISP

(Internet Service Provider). An institution that provides access to the Internet in some form, usually for money.

Kilobyte

A thousand bytes. Actually, usually 1024 (2^10) bytes.

LAN

(Local Area Network). A computer network limited to the immediate area, usually the same building or floor of a building.

Latency

The time it takes for a packet to cross a network connection, from sender to receiver.

Layer

Communication networks for computers may be organized as a set of more or less independent protocols, each in a different layer (also called level). The lowest layer governs direct host-to-host communication between the hardware at different hosts; the highest consists of user applications. Each layer builds on the layer beneath it. For each layer, programs at different hosts use protocols appropriate to the layer to communicate with each other. TCP/IP has five layers of protocols; OSI has seven. The advantages of different layers of protocols is that the methods of passing information from one layer to another are specified clearly as part of the protocol suite, and changes within a protocol layer are prevented from affecting the other layers. This greatly simplifies the task of designing and maintaining communication programs.

Leased-line

Refers to a phone line that is rented for exclusive 24-hour, 7 -days-a-week use from your location to another location. The highest speed data connections require a leased line.

Listserv

The most common kind of maillist, Listservs originated on BITNET but they are now common on the Internet.

Local Area Network (LAN)

A data network intended to serve an area of only a few square kilometers or less. Because the network is known to cover only a small area, optimizations can be made in the network signal protocols that permit data rates up to 100Mb/s.

Login

> Noun or a verb. Noun: The account name used to gain access to a computer system. Not a secret (contrast with Password).

> Verb: The act of entering into a computer system, e.g. "Login to the WELL and then go to the GBN conference."

Maillist (or Mailing List)

A (usually automated) system that allows people to send e-mail to one address, whereupon their message is copied and sent to all of the other subscribers to the maillist. In this way, people who have many different kinds of e-mail access can participate in discussions together.

Megabyte

A million bytes. A thousand kilobytes.

Modem

(MOdulator, DEModulator). A device that you connect to your computer and to a phone line, that allows the computer to talk to other computers through the phone system. Basically, modems do for computers what a telephone does for humans.

Mosaic

The first WWW browser that was available for the Macintosh, Windows, and UNIX all with the same interface. “Mosaic” really started the popularity of the Web. The source-code to Mosaic has been licensed by several companies and there are several other pieces of software as good or better than Mosaic, most notably, “Netscape”.

Netiquette

A pun on “etiquette” referring to proper behavior on a network.

Netizen

Derived from the term citizen, referring to a citizen of the Internet, or someone who uses networked resources. The term connotes civic responsibility and participation.

Netscape

A WWW Browser and the name of a company. The Netscape ™ browser was originally based on the Mosaic program developed at the National Center for Supercomputing Applications (NCSA).

Netscape has grown in features rapidly and is widely recognized as the best and most popular web browser. Netscape corporation also produces web server software.

Netscape provided major improvements in speed and interface over other browsers, and has also engendered debate by creating new elements for the HTML language used by Web pages—but the Netscape “extensions” to HTML are not universally supported.

The main author of Netscape, Mark Andreessen, was hired away from the NCSA by Jim Clark, and they founded a company called Mosaic Communications and soon changed the name to Netscape Communications Corporation.

Network

Any time you connect 2 or more computers together so that they can share resources, you have a computer network. Connect 2 or more networks together and you have an internet.

Newsgroup

The name for discussion groups on Usenet.

NIC

(Networked Information Center). Generally, any office that handles information for a network. The most famous of these on the Internet is the InterNIC, which is where new domain names are registered.

Node

Any single computer connected to a network.

National Science Foundation (NSF)

A U.S. government agency whose purpose is to promote the advancement of science. NSF funds science researchers, scientific projects, and infrastructure to improve the quality of scientific research. The NSFNET, funded by NSF, is an essential part of academic and research communications. It is a highspeed "network of networks" which is hierarchical in nature. At the highest level, it is a backbone network currently comprising 16 nodes connected to a 45Mb/s facility which spans the continental United States. Attached to that are mid-level networks (regionals) and attached to the mid-levels are campus and local networks. NSFNET also has connections out of the U.S. to Canada, Mexico, Europe, and the Pacific Rim. The NSFNET is part of the Internet.

Packet

The unit of data sent across a network. "Packet" a generic term used to describe unit of data at all levels of the protocol stack, but it is most correctly used to describe application data units.

Packet Switching

The method used to move data around on the Internet. In packet switching, all the data coming out of a machine is broken up into chunks, each chunk has the address of where it came from and where it is going. This enables chunks of data from many different sources to co-mingle on the same lines, and be sorted and directed to different routes by special machines along the way. This way many people can use the same lines at the same time.

POP

Two commonly used meanings: “Point of Presence” and “Post Office Protocol”. A “Point of Presence” usually means a city or location where a network can be connected to, often with dial-up phone lines. So if an Internet company says they will soon have a POP in Belgrade, it means that they will soon have a local phone number in Belgrade and/or a place where leased lines can connect to their network. A second meaning, “Post Office Protocol” refers to the way e-mail software such as Eudora gets mail from a mail server. When you obtain a SLIP, PPP, or shell account you almost always get a POP account with it, and it is this POP account that you tell your e-mail software to use to get your mail.

Port

3 meanings. First and most generally, a place where information goes into or out of a computer, or both. E.g. the “serial port” on a person computer is where a modem would be connected.

On the Internet “port” often refers to a number that is part of a URL, appearing after a colon (:) right after the domain name. Every service on an Internet server “listens” on a particular port number on that server. Most services have standard port numbers, e.g. Web servers normally listen on port 80. Services can also listen on non-standard ports, in which case the port number must be specified in a URL when accessing the server, so you might see a URL of the form:

gopher://peg.cwis.uci.edu:7000/

which shows a gopher server running on a non-standard port (the standard gopher port is 70).

Finally, “port” also refers to translating a piece of software to bring it from one type of computer system to another, e.g. to translate a Windows program so that is will run on a Macintosh.

Posting

A single message entered into a network communications system.

EX: A single message “posted” to a newsgroup or message board.

PPP

(Point to Point Protocol). Most well known as a protocol that allows a computer to use a regular telephone line and a modem to make TCP/IP connection and thus be really and truly on the Internet.

Protocol

A formal description of message formats and the rules two computers must follow to exchange those messages. Protocols can describe low-level details of machine-to-machine interfaces (e.g., the order in which bits and bytes are sent across a wire) or high-level exchanges between allocation programs (e.g., the way in which two programs transfer a file across the Internet).

RFC

(Request For Comments). The name of the result and the process for creating a standard on the Internet. New standards are proposed and published on line, as a “Request For Comments”. The Internet Engineering Task Force is a consensus-building body that facilitates discussion, and eventually a new standard is established, but the reference number/name for the standard retains the acronym “RFC”, e.g. the official standard for e-mail is RFC 822.

Route

The path that network traffic takes from its source to its destination. Also, a possible path from a given host to another host or destination.

Router

A special-purpose computer (or software package) that handles the connection between 2 or more networks. Routers spend all their time looking at the destination addresses of the packets passing through them and deciding which route to send them on.

Server

A provider of resources (e.g., file servers and name servers).

Shell Account

Shell accounts connect the terminal emulation software (home computers) through a modem to an Internet host, which provide a gateway to the Internet. These type of accounts do not support graphics; it contains ASCII (American Standard Code for Information Interexchange) text only.

A computer, or a software package, that provides a specific kind of service to client software running on other computers. The term can refer to a particular piece of software, such as a WWW server, or to the machine on which the software is running, e.g."Our mail server is down today, that's why e-mail isn't getting out." A single server machine could have several different server software packages running on it, thus providing many different servers to clients on the network.

SLIP

(Serial Line Internet Protocol). A standard for using a regular telephone line (a "serial line") and a modem to connect a computer as a real Internet site. SLIP is gradually being replaced by PPP.

T-1

A leased-line connection capable of carrying data at 1,544,000 bits-per-second. At maximum theoretical capacity, a T-1 line could move a megabyte in less than 10 seconds. That is still not fast enough for full-screen, full-motion video, for which you need at least 10,000,000 bits-per-second. T-1 is the fastest speed commonly used to connect networks to the Internet.

T-3

A leased-line connection capable of carrying data at 44,736,000 bits-per-second. This is more than enough to do full-screen, full-motion video.

Talk

A protocol which allows two people on remote computers to communicate in a real-time fashion.

TCP/IP

(Transmission Control Protocol/Internet Protocol). This is the suite of protocols that defines the Internet. Originally designed for the UNIX operating system, TCP/IP software is now available for every major kind of computer operating system. To be truly on the Internet, your computer must have TCP/IP software.

Telnet

The command and program used to login from one Internet site to another. The telnet command/program gets you to the “login:” prompt of another host.

Terminal

A device that allows you to send commands to a computer somewhere else. At a minimum, this usually means a keyboard and a display screen and some simple circuitry. Usually you will use terminal software in a personal computer—the software pretends to be (“emulates”) a physical terminal and allows you to type commands to a computer somewhere else.

Terminal Server

A special purpose computer that has places to plug in many modems on one side, and a connection to a LAN or host machine on the other side. Thus the terminal server does the work of answering the calls and passes the connections on to the appropriate node. Most terminal servers can provide PPP or SLIP services if connected to the Internet.

Topology

A network topology shows the computers and the links between them. A network layer must stay abreast of the current network topology to be able to route packets to their final destination.

UNIX

A computer operating system (the basic software running on a computer, underneath things like word processors and spreadsheets). UNIX is designed to be used by many people at the same time (it is “multi-user”) and has TCP/IP built-in. It is the most common operating system for servers on the Internet.

URL

(Uniform Resource Locator). The standard way to give the address of any resource on the Internet that is part of the World Wide Web (WWW). A URL looks like this:

HTTP//www.matisse.net/seminars.html

or telnet://well.sf.ca.us

or news:new.newusers.questions

etc.

The most common way to use a URL is to enter into a WWW browser program, such as Netscape, or Lynx.

Usenet

A world-wide system of discussion groups, with comments passed among hundreds of thousands of machines. Not all Usenet machines are on the Internet, maybe half. Usenet is completely decentralized, with over 10,000 discussion areas, called newsgroups.

WAN

(Wide Area Network). Any internet or network that covers an area larger than a single building or campus.

Workstation

A general-purpose computer designed to be used by one person at a time and which offers higher performance than normally found in a personal computer, especially with respect to graphics, processing power and the ability to carry out several tasks at the same time.

WWW

(World Wide Web). Two meanings—First, loosely used: the whole constellation of resources that can be accessed using Gopher, FTP, HTTP, telnet, Usenet, WAIS and some other tools. Second, the universe of HyperText servers (HTTP servers) which are the servers that allow text, graphics, sound files, etc. to be mixed together.

11. Bibliography

Aboba, B. (1993). The On-line User's Encyclopedia: Bulletin Boards and Beyond. Reading: Addison-Wesley.

Almes, G., Ford, P., & Lothberg, P. (1992). Proposal for Global Internet Connectivity. Internet Engineering and Planning Group (IEPG). Working Document. June, 12. Distributed electronically. http://www.aarnet.edu.au/iepg/docs/IEPG-gix-proposal.html.

Almquist, P. (1992). Type Of Service In The Internet Protocol Suite. Tech. Rep. Network Working Group.

Ameritech. (1995). The Chicago NAP. Distributed Electronically. HTTP//ameritech.na.com/products/data/nap/The_Chicago_NAP.html

Association for Computer Machinery Proceedings. (1986). The ARPANET and Computer Networks. Proceedings: Conference of the History of Personal Workstations, Jan. 9-10, Palo Alto, California.

Bailey, J.P. (1995). Internet Economics. Tech. Rep. M.I.T. Distributed Electronically.

Berners-Lee, T., & Cailliau, R. (1990). World Wide Web: Proposal for a HyperText Project. European Particle Physics Laboratory (CERN).

Blake, J. D., & Tiedrich, L. T. (1994, June) The National Information Infrastructure Initiative and the Emergence of the Electronic Superhighway. Federal Communications Law Journal. Vol. 46, No. 3.

Bohn, R., Braun, H., Claffy, K.C., & Wolff, S. (1994). Mitigating The Coming Internet Crunch: Multiple Service Levels Via Precedence. Tech. Rep. Distributed Elextronically.

Bradner, S., Mankin, A. (1995). The Recommendation for the IP Next Generation Protocol. RFC 1752. Harvard University and Information Science Institute. January.

Brownlee, J.N. (1994). New Zealand Experiences with Network Traffic Charging. Tech. Rep. University of Auckland.

Cerf, V. (1993). How the Internet came to be. In B. Aboba (Ed.), The On-line User's Encyclopedia: Bulletin Boards and Beyond. Reading, Massachusetts: Addison-Wesley.

Cerf, V. (1995, November-December). Computer Networking: Global Infrastructure Driver for the 21st Century. On The Internet.

Cerf, V. (1989). Requiem for the ARPANET. In Users' Dictionary of Computer Networks. Bedford, MA.

Cocchi, R., Estrin, D., Shenker, S., & Zhiang, L. (1991). A Study of Priority Pricing in Multiple Service Class Networks. In Proceedings of Sigcomm '91. Distributed electronically.

Cocchi, R., Estrin, D., Shenker, S., & Zhiang, L. (1992). Pricing in computer Networks: Motivation, formulation, and example. Tech. Rep. University of Southern California.

Corporation for National Research Initiatives. (1995). Gigabit Testbed Initiative. Tech. Rep. Distributed electronically. Reston, Virginia.

Dacey, R. (1996). Personal interviews. Department of International Business at the University of Idaho.

Economides, N. S. (1994). The Economics of Networks. New York University. EARIE Conference, Chania, Greece. September. Distributed Electronically.

Economides, N. S., Woroch, G. A. (1992). Benefits and Pitfalls of Network Interconnection. Discussion Paper EC-92-31. New York University and GTE Laboratories. Distributed Electronically.

Economides, N. S., & Himmerlberg, C. (1994). Critical Mass and Network Size. Tech. Rep. Unknown.

Engineering & Operations Working Group (EOWG). 1995. Proposal for the Continued Interconnection of the Federal Agency Networks. Federal Network Council. Version 1.0. December 11. Distributed electronically. HTTP//www.fnc.gov/FEPG_proposal.html.

Faulhaber, G. .R. (1990). Pricing Internet: The efficient subsidy. Technical Report. Department of Public Policy & Management. Wharton School, University of Pennsylvania. Distributed electronically.

FIND/SVP. (1996). The American Internet User Survey. http://etrg.findsvp.com/features/newinet.html.

FNC. (1995). Federal Networking Council. Distributed electronically. HTTP//www.fnc.gov/.

Frizzier, K. D. (1995). NSFNET: Final Report 1987-1995. Merit Network, Inc., Ann Arbor.

Gilder, G. (1993, December). "The Coming of the Fibersphere", Forbes ASAP.

Gilster, P. (1994). The Internet Navigator. Second Edition. Wiley & Sons, Inc.: New York.

Gong, J., & Srinagesh, P. (1995). The Economics of Layered Networks. Draft. NII Initiatives, Bell Communications Research, Inc. Distributed Electronically.

Gwertzman, J. S. (1995). Autonomous Replication in Wide-Area Internetworks. Thesis in partial fulfillment of the honors requirements for the degree of Bachelor of Arts. Harvard College. Cambridge, Massachusetts. April 10. Distributed electronically. http://www.eecs.harvard.edu/~vino/web/push.cache/thesis.html

Hardy, H. E. (1993). The History of the Net. Master's thesis. School of Communications. Grand Valley State University. Allendale, Michigan. Distributed electronically. ftp://umcc.umich.edu/pub/users/seraphim/doc/nethist8.txt.

Hauben, M. & Hauben, R. (1996). Netizens: On the History and Impact of the Net. Distributed electronically. Columbia University. http://www.columbia.edu/~hauben/netbook/

Heart, F., McKenzie, A., McQuillian, J., & Walden, D. (1978). The Completion Report. Bolt Beranek Newman Report 4799.

Hoffman, D.L., Kalsbeek, W. D., & Novak, T. P. (1996). Internet Use in the United States: 1995 Baseline Estimates and Preliminary Market Segments. Project 2000 Working Paper. April 12. Vanderbilt University and University of North Carolina at Chapel Hill.

HPCC. (1995). High Performance Computing and Communications: Technology for the National Information Infrastructure. Blue Book. Distributed Electronically. HPCC.

HPCC. (1996). High Performance Computing and Communications: Foundation for America's Information Future. Blue Book. Distributed Electronically. HPCC.

Huston, G. (1994). Internet Service Provider Peering. Internet Engineering and Planning Group (IEPG). Distributed Electronically. http://www.aarnet.edu.au/iepg/settlements.html.

IEEE (1978, November). Applications of Information Networks. Proceedings of the IEEE. Vol. 66, No. 11.

ISOC. (1995). Internet Society. Distributed electronically. http://info.isoc.org:80/whatis/what-is-isoc.html.

Jackson, K. (1994). Commerce On The Net: What's The Holdup? In Open Systems Today. P. 162. October 31.

Kahn, R. (1995). The Role of Government in the Evolution of the Internet. In National Academy of Sciences, Revolution in the U.S. Information Structure.

Kanaley, R. (1994, 4 November). Internet Passes Quietly Into Private Hands. Philadelphia Inquirer. Distributed electronically. http://www.cs.umbc.edu/pub/news/umbc.cs/msg00287.html

Kleinrock, L. (1976, December). On Communications and Networks. IEEE Transactions on Computers. Vol. C-25 No. 12.

Krol, E. (1994). The Whole Internet User's Guide & Catalog. Second Ed. Sebastopol: O'Reilly & Associates.

Lasher, M. (1996). Personal Interviews. Assistant Director of Telecommunications Technology. Department of Computer Services, University of Idaho. February 16.

Lottor, M. (1992, January). Internet Growth (1981-1991). SRI International. RFC 1296. Distributed Electronically.

MacKie-Mason, J. K. and Varian, H.R. (1993). Some Economics of the Internet. Tech. Rep. University of Michigan.

MacKie-Mason, J.K. and Varian, H.R. (1994a). Economic FAQs about the Internet. Journal of Economic Perspectives, 8(3).

MacKie-Mason, J., & Varian, H. R. (1994b). Pricing the Internet. Working paper presented at the Kennedy School: Public Access to the Internet. May 26-27, 1993. University of Michigan.

MacKie-Mason, J., &Varian, H. R. (1994c). Pricing Congestible Network Resources. University of Michigan.

MacKie-Mason, J. & Varian, H. R. (1994d). Some FAQs about Usage-Based Pricing. Tech. Rep. Distributed electronically. University of California, Berkeley.

Malkin, G., & LaQuey, T. (1993). Internet Users' Glossary. RFC 1392. Network Working Group. January. gopher://nic.merit.edu:7043/0/introducing.the.internet/users.glossary.

Marine, A., Reynolds, J., & Malkin, G. (1994). FYI on Questions and Answers: Answers to Commonly asked "New Internet User" Questions. RFC 1594. Network Working Group.

McGarty, T., & Haywood, C. (1993). Internet Architectural and Policy Implications for Migration from High End User to Low End User. Presented at the Kennedy School, Harvard University. May 26 at Public Access to the Internet, Cambridge, MA.

MCI PR. (1995). MCI And The National Science Foundation Announce New Very High Speed Backbone Network Service To Connect Nation's Supercomputer Centers. April 24. Distributed Electronically. http://www.gov.mci.net/vBNS/press_release_vBNS.html

MCI. (1995). Very high speed BACKBONE NETWORK SERVICE (vBNS). Distributed electronically. http://www.gov.mci.net/vBNS/vBNS.html.

Mitchell, B., Srinagesh, P. (1995). Economic Issues. In UNIVERSAL ACCESS TO E-MAIL: Feasibility and Societal Implications. RAND Corporation. Distributed Electronically.

Murphy, J., & Murphy, L. (1994) Bandwidth Allocation by Pricing in ATM networks. Tech. Rep. EECS Department, University of California, Berkeley.

National Research Council. (1994). Realizing the Information Future: The Internet and Beyond. Computer Science and Telecommunications Board. National Academy Press: Washington, D.C.

Nicholson, W. (1995). Microeconomic Theory: Basic Principles and Extensions. Sixth Edition. Dryden Press. Fort Worth, Texas.

Nielsen Media Research. (1995). Nielsen Interactive Services. http://www.nielsenmedia.com.

North, D. C. (1993). Institutions, Transactions Costs and Productivity in the Long Run. Working Paper. Washington University, St. Louis. Distributed Electronically.

NSF 93-52. Network Access Point Manager, Routing Arbiter, Regional Network Providers, and very high-speed Backbone Network Service. Program Guideline. National Science Foundation: CISE/NCR. Distributed electronically.

NSF PR 93-52. New NFSNET Architecture. National Science Foundation Press Release. May, 1993. Distributed electronically. National Science Foundation.

NSF PR 95-37. NSFNET Backbone Decommissioned. National Science Foundation Press Release. May 15, 1995. Distributed electronically. National Science Foundation.

NSF. (1990). Acceptable Use Policy. National Science Foundation. Distributed electronically. ftp://nis.nsf.net/nsfnet/acceptable.use.policy

O'Reilly Research. (1995). Defining the Internet Opportunity. http://www.ora.com/www/info/research/users/index.html.

Parris, C., Keshav, S., & Ferrari, D. (1992). A framework for the study of pricing in integrated networks. Tech. Rep., International Computer Science Institute, Berkeley, March 1992.

Perez, M.A. (1995). Pricing the Internet: A Model and a Practical Implementation. Proceedings INET 95. Universidad Catolica de Chile.

Perlman, B. A. (1995). Pricing the Internet: How to Pay the Toll for the Electronic Superhighway. Center for Science and International Affairs (CSIA). Discussion Paper 95-01, Kennedy School of Government. Harvard University. March.

Plott, C. R., Wilkie, S. (1995, Summer). Local Competition and Universal Service. Jobs & Capital. Milken Institute for Job & Capital Formation. Vol. IV.

Pool, I. de S. (1983). Technologies of Freedom. Cambridge, Harvard University Press.

Pool, I. de S. (1990). Technologies Without Boundaries. Edited by Eli Noam. Cambridge, MA.

Quarterman, J. S. (1993). Internet Services and Access. Working Paper: Texas Internet Consulting. Distributed electronically.

Rickard, J. (1995). The Internet by the Numbers--9.1 Million Can't Be Wrong. BoardWatch Magazine. Volume IX: Issue 12 - ISSN:1054-2760. December.

Rams∅y, T. (1995). The Governing Structure Of The Internet: Coordination Without Central Control. Master's thesis. Alfred P. Sloan School of Management and the School of Engineering. Massachusetts Institute of Technology. Cambridge, Massachusetts.

Kahin, B. (1990). Commercialization of the Internet. Summary Report of the Science, Technology and Public Policy Program Workshop. RFC 1192. John F. Kennedy School of Government. Harvard University. March 1-3. Distributed electronically.

Schickele, S. (1993). The Economic Case For Public Subsidy Of The Internet. Presented at Public Access to the Internet Conference. JFK School of Government. Harvard University. May 26-27. Sonoma State University.

Shenker, S. (1993). Service Models and Pricing Policies for and Integrated Services Internet. Tech. Rep. Xerox Corporation.

Sprint. (1994). Sprint Network Access Point Handbook. Version 1.0. Distributed electronically. http://www.merit.edu/nsf.architecture/Sprint/NAP.handbook.html

Stahl, D. O., & Whinston, A. B. (1995). The Internet: A Future Tragedy of the Commons? The University of Texas at Austin. Conference on Interoperability and the Economics of Information Infrastructure. July 6-7, 1995. Rosslyn, VA

Staudt, D. A. (1996). Private e-mail. Division of Networking and Communications Research & Infrastructure. National Science Foundation.

Tyrlik, Z. J. (1996). Private E-mail. APK Net Ltd. (ISP): Network Manager.

Wenders, J. T., & Taylor, L. D. (1976, Autumn). Experiments In Seasonal-Time-Of-Day Pricing Of Electricity To Residential Users. The Bell Journal of Economics. Vol. 7, No. 2.

Wenders, J. T. (1977). Pricing in Regulated Industries: Theory and Application. Editor. Al Brower for Printing, Inc. Carefree, Arizona.

Wenders, J. T. (1987). The Economics of Telecommunications. Cambridge: Ballinger Publishing Company.

Wenders, J. T. (1996). Regular meetings. Thesis chairman. University of Idaho.

Zakon, R. H. (1993). Hobbes' Internet Timeline v2.3. Distributed Electronically. The MITRE Corporation.