InterOffice Memo

To: Paul Maritz, Jim Allchin, Brad Silverberg, Tom Evslin, Dave Thompson

CC: John Ludwig, Tod Nielsen, Bob Muglia, Ty Carlson, Aaron Contorer, Steven Sinofsky, Dave Cutler

From: J. Allard

Date: January 25, 1994

Subject: Windows: The Next Killer Application on the Internet

With every passing minute, 2 new systems become wired. A new, connected network appears every 40 minutes. In 1993 more than 25 books appeared on the shelves of major bookstore chains, and over 2,600 articles in popular periodicals such as New Yorker, Time and PC Week can be cited. You can even pick up a periodical from the local newsstand dedicated entirely to this very hot topic: The Internet.

The three fundamental building blocks of the Internet include users, infrastructure, and information that I collectively refer to as the *infostructure*. The Internet is very well aligned with our corporate vision of *information at your fingertips* and serves as an effective infostructure to increase product group productivity and defray support costs. Microsoft is very well positioned to grow our Systems marketshare in some strategic areas and win the mindshare of millions with incremental effort in our present product plans. This memo summarizes a decade of Internet history and technologies, identifies some recent trends, and explores what it will take to make Windows the next killer application on the Internet.

Overview

The Internet represents a decade-plus old network that has very rapidly grown to over 16 thousand interconnected networks. Current surveys indicate an excess of 2 million connected nodes servicing approximately 25 million users in 137 countries[1]. Originally a defense department project (ARPANet), the Internet has grown to include thousands of academic, research, and commercial organizations. In the last year, the Internet has *doubled* in size, with a growth rate in excess of 5% per month. In 1994, commercial network providers such as AT&T, MCI, ANS and PSI are enhancing their services to offer Internet connectivity via modem, ISDN and even coax, allowing this vast infrastructure to extend easily and inexpensively to the home, small business, and K-12 schools.

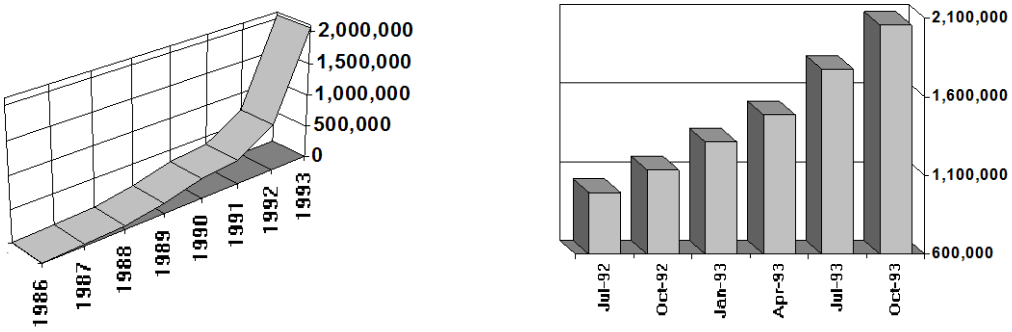

Figure 1: Internet Growth (Total Systems) 1986-1993. Source: SRI International

[1]Source: SRI International. Note: these figures represent systems which are not only connected, but are also visible worldwide at the time of census. This census is unable to account for the many organizations which connect via a "gateway" to the Internet - such as Microsoft. Microsoft offers a prime example of huge low-side variations in this data - several thousand employees are able to regularly share e-mail with Internet users, access Usenet and FTP files. However, Microsoft contributes only a few dozen systems to the above census figures. Of the 137 countries with Internet access, only some 60 have direct connections to the Internet at this time.

Virtually every network'able operating system in the world is represented somewhere on the Internet. This is made possible by universal support for the TCP/IP and related internetworking protocols. Although TCP/IP provides the basic framework, the real end-user appeal takes shape in the form of the applications, information and connectivity capabilities the Internet has to offer. The past years represent the initial "explosion" of systems connecting to the Internet, driven by the desire of global connectivity and e-mail access. This boom has sparked the development of new efforts in the community improving accessibility and usability as the user-base expands beyond computer experts. I refer to these efforts loosely as "Internet navigation tools" or "second generation information tools".

Many of these second generation tools allow users to navigate the Internet as a virtual library of information. Although a user session may span a number of systems, on different networks, in many countries, many new tools and protocols do their best to shield the user from such detail. Instead, the user navigates through hierarchies of information pointers (links), arranged by subject matter. From an end-user's perspective this is akin to browsing a bookstore - initially seeking information by subject, then by title, and finally by page. Architecturally, the concept might be more appropriately paralleled with a single application designed to allow a user to browse CompuServe, DowJones, America Online, and Genie without requiring the selection of a specific service - from the user's perspective they simply flow between related topics, while the tool manages connections between the appropriate services in the background. Of course, an application of this nature would be very difficult to develop given the use of proprietary protocols and the physical presence of these information services on multiple, disjoint networks. One of the greatest strengths of the current Internet infrastructure is that a single system can provide multiple, simultaneous connections to different information providers with a single physical link.

Two aspects of the Internet have stimulated generalized information tools that behave like the one I describe above:

- Thousands of organizations and information sources are globally connected on a common "universal highway system" - *given an on ramp, all destinations are reachable with the same vehicle.*
- Protocols used on the Internet are developed as collaborative, open efforts between interested producers and consumers of the information. This methodology results in protocols which consider implications such as international use and heterogeneous systems - *all cars and drivers on the information highway obey international law.*

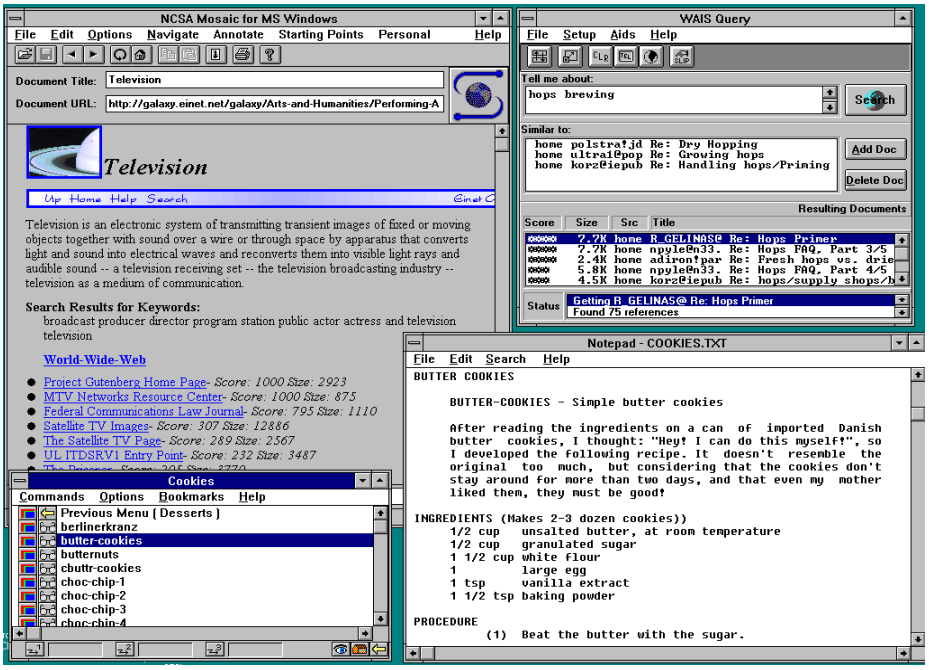

Figure 2: Surfing the Internet with Windows-based Navigation Tools

The Internet model facilitates worldwide access to a wealth of information from virtually all modern operating systems. Users are required to establish a connection to the infrastructure (generally via a service provider), and then acquire the appropriate tools before they can begin to "surf the infostructure".

Services Available on The Internet

I classify the primary interactive services available on the Internet into three categories: simple messaging, file transfer, and information discovery and retrieval. (This section may read a little "dry" but identifies some terms and trends while providing the necessary foundation for opportunity exploration. If you are familiar with these capabilities, jump to page 9).

Simple Messaging

E-mail: familiar to us all, two basic protocols deliver mail to the millions of users on the Internet: SMTP[2] is used primarily as the mail routing and delivery protocol. POP[3] is a relatively new protocol generally used by individual users (e.g., users with non-permanent Internet connections) to download their mail to and from a connected SMTP store. POP is used in a fashion similar to the Microsoft campus Xenix mail protocol. Rich document embedding (graphics, sound, etc) is facilitated by MIME[4], basic binary file transfers are often facilitated by ASCII encoding of the binary files (uuencoding).

Apart from typical person-to-person messaging, a valuable use of e-mail not unique to the Internet is the use of global technical and discussion aliases. Such aliases are frequently used to facilitate international discussion or to hold "electronic meetings" at very low cost both within global organizations as well as between them. Much of the work done by the TCP/IP networking team involves the communication of ideas and information over several global e-mail aliases, some of which are "cross posted" to Usenet as well (see discussion below). As these global discussion aliases are maintained by individual sites, there is no telling how many of them there are. Most predictions indicate the number of these aliases is in the 10s of thousands given the traffic patterns measured on the NSFNet (figure 7) relative to the Internet user base. Microsoft offers SMTP mail connectivity via a MS-DOS based gateway product, but no direct SMTP or POP access from the MS-Mail client. Several third parties offer SMTP and POP-based mail clients for MS-DOS and/or Windows, generally with their TCP/IP product offerings.

Usenet: Usenet is a distributed public bulletin board system consisting of over 6,000 different professional, technical and recreational discussion groups (more commonly, "newsgroups"). The topics include everything from aeronautics*(sci.space.aeronautics)* to computer programming *(comp.sys.win32. programming)* to geology*(sci.geo.geology)* to zymurgy*(rec.crafts.brewing)*. The groups are organized using a hierarchical naming scheme designed to provide additional context for the nature of discussion. Administrators may filter which newsgroups will be carried locally and may choose to add local which will only be exchanged with other consenting sites. Global newsgroups are created by a public electronic nomination and election process.

NNTP[5]is used to transfer messages between the news servers scattered around the Internet, as well as to deliver the messages to the "newsreader" client application software. Many Microsoft employees are familiar with a set of these discussion groups offered via hexnut, wingnut, BBS1 and BBS2 ITG services. Both character-mode and graphical newsreader clients are available for many platforms by commercial vendors and in the public domain.

File Transfer

FTP[6]: FTP is probably the most widely used Internet application - well, at least it makes the most noise - FTP accounts for over 40% of all the interactive Internet traffic in a given day (23% of all traffic on the Internet). FTP is a client-server protocol that provides users with rudimentary browsing of remote file systems and the ability to transfer ASCII and binary data files between systems. Security is weak at best (passwords are passed in cleartext), so the most prevalent use of the FTP protocol is on "anonymous FTP servers". Anonymous FTP refers to guest access to FTP servers, users connect to public servers providing only an e-mail address, and receive (generally) read-only access to specific volumes of the remote system. Although the standard FTP interface is a character mode application, many graphical clients have been developed to make FTP look more like File Manager. Windows NT provides a simple FTP client and a scalable, integrated FTP server as part of the base product. Clients are available on virtually all operating systems supporting the TCP/IP protocols, the thousands of infostructure FTP servers are dominated by Unix, and announce themselves to their users as such.

---

[2]*RFC 821: Simple Mail Transfer Protocol*, 1982, August; Postel, J.B.
[3]*RFC 1225: Post Office Protocol: Version 3*,1991, May; Rose, M.T.
[4]*RFC 1341: MIME (Multipurpose Internet Mail Extensions):Mechanisms for Specifying and Describing the Format of Internet Message Bodies*,1992 June; Borenstein, N.; Freed, N.
[5]*RFC 977: Network News Transfer Protocol*, 1986 February; Kantor, B.; Lapsley, P.
[6]*RFC 959: File Transfer Protocol*, 1985 October; Postel, J.B.; Reynolds, J.K.

Gopher*[7]*: Named after the University of Minnesota's athletic mascot, the Internet gopher project was inspired by the inability of FTP to span multiple servers or filesystems. From the gopher protocol specification:

> **gopher:** n. **1.** Any of various short tailed, burrowing mammals of the family Geomyidae, of North America. **2.** (Amer. colloq.) Native or inhabitant of Minnesota: the Gopher State. **3.** (Amer. colloq.) One who runs errands, does odd-jobs, fetches or delivers documents for office staff. **4.** *(computer tech.) software following a simple protocol for burrowing through a TCP/IP internet.*

Gopher documents (and services) may reside on many servers, running a very simple client-server protocol. The gopher interface was designed to present users with a virtual filesystem containing document objects, directory objects, and searching capabilities across multiple servers with the click of a mouse. Directory objects unify the virtual filesystem - although a user views all directory objects as equals, some directory objects traverse the filesystem to which the user has most recently connected. In other cases, invoking a directory object result in a redirection to another server elsewhere on the Internet. Directory objects that provide these redirections are inserted as "links" manually by gopher server administrators.

Beyond offering a filesystem view, most server implementations allow administrators to provide very descriptive text for the objects with few naming constraints. On a well-managed gopher server, the directory c:*home**jallard* might appear as *J. Allard's home directory*, and the file *dhcp.doc* might appear as *DHCP ConneXions Article (Microsoft Word format).* As you might imagine, this capability is a big plus for new users of Internet tools. The gopher protocol also provides simple document type information allowing the client tool to choose to display the document once retrieved, or to invoke a more suitable viewer of the data such as Paintbrush or WinWord. Since this concept is very useful in the PC industry, the designers have recently expanded the protocol (gopher+[8]) to include richer document type descriptions and other minor enhancements to correct oversights in the original protocol.

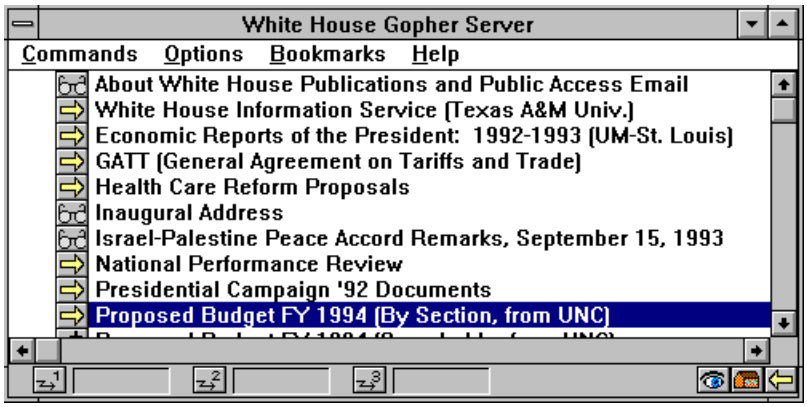

Figure 3: Using Gopher to Explore the online White House Library

The University of Minnesota maintains a directory with links to all gopher servers which independent administrators have chosen to register by e-mail with the University. Most gopher clients provide a link to this directory of directories from their main window, allowing users to readily locate gopher information servers as they join the global Internet. At the time of this writing, 1189 gopher servers appear in the list (approximately 6% support the gopher+ protocol). This list provides a single link to the "root" of each organization's gopher "filesystem" which help guide the user to their specific destination. Gopher clients are readily available for Unix, X Windows, Macintosh, Next, OS/2, MVS, MS-DOS, and Windows in the public domain. As with FTP and messaging servers on the Internet, nearly all gopher servers are sponsored on systems running a variant of Unix.

[7]*The Internet Gopher Protocol: a Distributed Document Search and Retrieval Protocol*, Spring 1992; Bob Alberti, Farhad Anklesaria, Paul Lindner, Mark McCahill, Daniel Torrey, University of Minnesota

[8]*Gopher+: Proposed Enhancements to the Internet Gopher Protocol*, Summer 1992; Bob Alberti, Farhad Anklesaria, Paul Lindner, Mark P. McCahill, Daniel Torrey, University of Minnesota

Information Discovery and Retrieval (Second Generation Tools)

Archie[9]: The archie service was inspired by the incredible number of FTP servers which appeared on the Internet during the 80's. Archie is a simple service that contacts a configured set of chosen FTP servers periodically (using the basic FTP protocol) and builds indexes based on filenames stored on the servers. Users generally access archie via terminal services, although there is also a simple client-server protocol that can be used to query known archie servers on the Internet from specialized client software. Archie is provides no content-based searching mechanisms whatsoever, although it is a useful tool to locate one or more FTP locations for a particular file.

For example, a query on "pkzip*" (PKZIP is a popular PC compression utility) results in a list of FTP servers, and the respective location on each which maintain copies of the PKZIP software. As archie only accesses the directory listings on the servers, and not the information itself, archie offers no content based indexing or query capabilities. BunyIP Information Systems (an Internet services company) has recently announced plans to extend the archie capabilities to further include indexing of gopher menu objects in the second half of 1994.

WAIS[10]: The Wide Area Information Servers system was developed in cooperation by **Apple Computer**, **Dow Jones** and Thinking Machines Corp. (TMC) to provide a full-text information retrieval system for the Internet. The WAIS protocol is based on the draft ISO standard Z39.50 designed for bibliographic retrieval, with some necessary extensions for full-text searching. On the Internet, the role of WAIS servers is to generate inverted indexes on local documents and execute full-text searches on these indexes. TMC maintains a WAIS server which acts as a database of databases, allowing users to locate information sources (servers) on the Internet which may contain information relevant to their query. Although the WAIS protocol does not restrict the use of complex (e.g., boolean) queries, most WAIS servers today only supports simple keyword matching, forcing the user to refine queries by adding additional keywords.

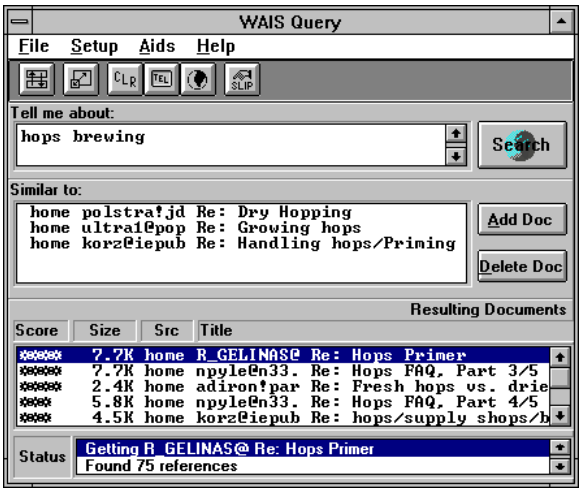

Figure 4: Efficient Zymurgy Research using WAIS Technology

A user's initial WAIS journey may be a little rough, but once accustomed to the interface, topic-based research can be made very efficient. Let's say we've been a little out of touch with Bill Clinton's plans for healthcare reform and want to learn more. We start at the directory of servers and scan for Clinton. This query gives us two databases on the Internet which would be appropriate: *Clinton-Speeches* and *White House Papers*. We choose to limit future queries to these two databases. We then provide the keywords *health* and *taxes* and WAIS connects to the two servers offering these databases, queries them, and merges the results. We are returned a total of 75 documents by title, sorted by how effectively our keywords matched the document text. We can then choose to either refine our query by adding additional keywords, or choose to browse the documents by title. Although there are several hundred WAIS servers on the Internet today, topic coverage is actually somewhat limited for non-technical topics. TMC provides free source code to implement WAIS servers, although I have yet to see a non-Unix implementation. Several client programs exist, the only Windows-based client was developed by the US Geological Survey (illustrated above).

---

[9] *Archie - An Electronic Directory Service for the Internet*, Peter Deutsch, McGill University.
[10] *An Executive Information System for Unstructured Files: Wide Area Information Servers*, November, 1991; Brewster Kahle, Harry Morris, Franklin Davis, Kevin Teine, Clare Hart, Robin Palmer, Thinking Machines.

World Wide Web (WWW or W3): The World Wide Web refers collectively to a set of information servers which run HTTP[11], servicing HTML[12]compatible information files on the Internet. Users navigate "the Web" through a graphical point-and-shoot interface not unlike the Microsoft Multimedia Viewer. Rich documents are presented with embedded sounds and inline graphics with links to other related topics. Advanced capabilities such as MPEG animations are typically not supported directly by Web clients, the clients simply invoke an appropriate viewer. Links provide references to other WWW documents which can be located on any other server in the Web. The Web has received the most attention of any of the new technologies due to the nature of its simple user interface, and the quality and organization of the servers that have recently come to be part of the infostructure. Some examples:

- **Novell** offers all of their product documentation and support documents
- The **NASDAQ** Financial Executive Journal is available from the Cornell Law School
- The **1994 Winter Olympic Games** schedule and results are online and "live" courtesy of Oslonett in Norway
- The **SEC's EDGAR** system which manages all 10-K and 10-Q reports will be offered via the Web at no charge to the public by year end, currently only available at $39/hour and $15/document through private systems
- O'Reilly and Associates (a popular Unix book publishing company) has established a commercial Web server online with corporate advertising support, Internet-related monthly articles and Doonesbury cartoons
- The **Library of Congress** server is presently hosting a Vatican art collection exhibit
- ***Wired*** publishes a "Web'd" version of their popular magazine (2 months after hits the stands)
- The University of Georgia offers a graphical tour of their greenhouses on a Botany department server

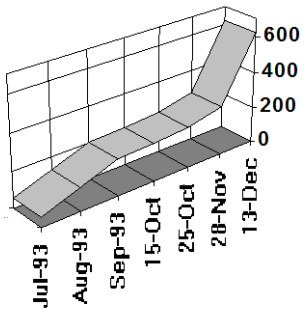

Figure 5: World Wide Web Server Growth July -December 1993. Source: Matthew Gray[13]

[11]*HTTP (HyperText Transfer Protocol): A Protocol for the Retrieval and Manipulation of Textual and Hypermedia Information*, Tim Berners-Lee, CERN

[12]*Hypertext Markup Language (HTML): A Representation of Textual Information and MetaInformation for Retrieval and Interchange*, Tim Berners-Lee, CERN. (Note: HTML is a variant of the ISO/SGML text markup language).

[13]Matthew is a Physics student at MIT which has developed the "Web Wander" program - an automated knowbot which visits different Web sites looking for other Web references in the corresponding documents it locates. Because it cannot locate servers which do not have links from known servers, does not recurse through documents, and cannot account for gateways, the data can be considered accurate (e.g., there are at least 623 Web servers), but incomplete.

Given the ease-of-use of most Web client interfaces, and the recent excitement and press surrounding the Web, it's no wonder that its use and presence have grown at such incredible rates. Although it's very difficult to measure the exact number of Web information services, the Internet has been picking more than one new Web server a day since June - as of December '93, over 600 World Wide Web servers on reachable on the Internet. This statistic does not reflect the number of Web servers which are used as internal document distribution servers in many universities and research centers.

The most popular Web client is considered to be NCSA's *Mosaic* which is available for X Windows, the Macintosh and now Windows 3.x. PCWeek recently dedicated 4 columns to the public domain Mosaic program for Windows (unfortunately, the screenshot demonstrated a connection to the **Novell** WWW Internet server...). The NCSA team has well to measure the number of Mosaic downloads via FTP servers which number close to 700,000 by year-end 1993. Considering the software is likely to be distributed by sites not reporting to NCSA and that a number of alternative Web clients are available, *it is fair to assume that over 1,000,000 end users have access to the Web*, and that this number is growing rapidly. Although available for Windows, the current implementation is considerably weaker than its Unix counterpart.

Figure 6: NCSA's Mosaic for Windows browsing O'Reilly's InfoCenter

The Future of Navigation Tools

The following graph indicates the ratio of *packets* measured between the different interactive protocols discussed in this memo, offering some sense for the current utilization of various interactive *protocols* on the Internet today. These results must be taken with a grain of salt; given the dynamics of these protocols these ratios may present a poor picture of actual use of particular *applications*. For example, an automated FTP script used for daily replication may generate 100,000 packets in a minute while an hour long end-user interactive research session with WWW on particle physics could yield as few as 1,000.

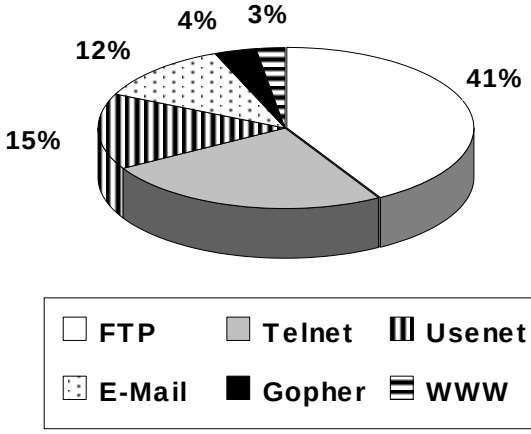

Figure 7: Internet "Interactive" Traffic by Protocol, December 1993. Source: NSFNet[14]

This diagram does help to demonstrate that the Internet has not completely undergone the shift to this new information based focus. The growth of gopher and WWW servers has been explosive since the protocols were discovered, and new information based applications and protocols are appearing on the horizon. *Over the next year we will see a huge number of new information servers come to the Internet based on gopher and Web technologies*, we can expect to see some work will be done in both protocols to offer simple yet valuable enhancements.

The biggest change facing the Internet in the next two years is the introduction of commercial services - initially in the areas of electronic shopping and publishing. Commercial servers based on WWW are inevitable, at least 3 organizations have announced commercial servers with commercial backing from vendors such as **Land's End**, **LL Bean** and **Victoria's Secret**. A number of businesses are already promoting their products using gopher and Web technology, some even accepting electronic orders today. Today, you can visit the *Virtual Record Store* to find and order a CD, visit *Quantum Books* without flying to Cambridge and browse the latest technical titles, buy a list of government surplus properties for sale from *Counterpoint Publishing*, and even make a discreet stop at *JT Adult Toy Store* without leaving your home.

The potential for Internet profit has inspired work on commerce protocols (the Internet Mercantile Protocol) to make electronic shopping easier and integrated with these navigation tools. Video/audio conferencing and shared whiteboard applications are also under development to enhance workgroup productivity over wide distances. The quarterly IETF engineering meetings have been multicast for over a year on the Internet using experimental software at over 100 sites worldwide. This experimental technology provides 3 frame/sec video and bi-directional audio capabilities for over a year. The software has been developed and runs exclusively on SunOS, **Sun MicroSystems**' variant of Unix.

[14]Source: NSFNet (The National Science Foundation Network). Although this data specifically accounts for the traffic patterns on NSFNet for December 1993, this federally funded backbone is believed by most to best represent the traffic patters on the Internet as it carries such a large volume of public traffic. Note also that for purposes of this graph, all other protocols such as NFS, X Windows, and remote process execution have been eliminated to show the relative use of the "interactive" Internet protocols discussed in this memo.

Opportunities

I have identified three immediate opportunities that coincide with our present business focus and the current Internet swell:

- Reduction of Support Costs
- Mind- and Marketshare
- Enhanced Research Capabilities

A fourth, implicit opportunity deserves a brief mention here - our efforts in this global infostructure provides a unique testbed for us to explore. Although Al Gore and Company's NII (National Information Infrastructure) efforts appear to be making some initial headway, *the Internet provides an incredible opportunity for Microsoft to effectively explore large scale public networks from many levels*: customer needs, technical challenges, quality of service issues, electronic commerce and information browsing technologies. *The Internet is version 1.0 of the National Information Infrastructure*.

Support Cost Reduction

The Internet offers an extensive user base that is already substantially larger than the combined users of **CompuServe**, **America Online** and **Prodigy** services. With a well-focused, low-cost investment, it is possible to effectively leverage the infrastructure to offer support to the millions of users connected to the global Internet. The Microsoft Internet FTP server (ftp.microsoft.com) represents a solid first step in the area of Internet-based customer support that has been extremely popular with both end-users as well as the press:

> "Technological migration takes time, but **Novell** is really behind the eight ball on the Internet. Internet interlopers logging into **Univel** find that its FTP server runs **SunOS** 4.1, and not Novell's year-old Unixware. Meanwhile, plugging into **Microsoft**'s FTP server revealed that Microsoft has a **Windows NT** FTP Server up and running."[15]

ftp.microsoft.com is a Windows NT FTP server, *brought online the same day Windows NT shipped to customers*. The server offers simple file transfer services to users of Microsoft products, about 70,000 files are available for a total of 500 Mb of information for customers to access. We've recently begun measuring usage of this server, which is charted below in figure 8. *Since November 1, 1993 we've serviced 25,000 users and 75,000 files per week on average*, the only cost being the management of an NT server, and the time of program and product managers putting up information for retrieval.

Approximately 50,000 users have downloaded the *MS-DOS 6.2 StepUp* program since it went online, over 5,000/week on average, compared to only 26,000 from **CompuServe**, 3,000 on **Genie**, and 1,500 on **America Online**. Consider the COGS associated with making this upgrade available in EggHead, or a customer via the postal service on a floppy to be about $5. Putting *StepUp* on the Internet represents a quarter million dollar savings in two months. All that was necessary was a simple file copy to a server also shared on the corporate network (\\gowinnt\ftp). Our efforts in this area barely scratch the surface of what is possible.

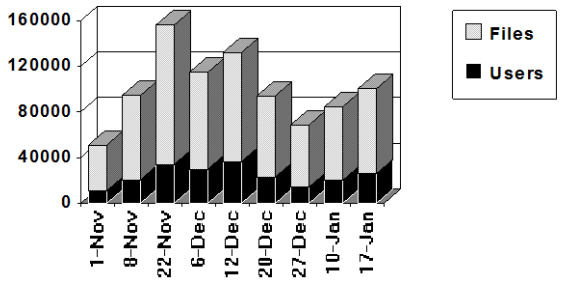

Figure 8: Microsoft Internet FTP Server Usage: November 1, 1993-Jan 10,1994

[15]Computer Reseller News, November 8, 1993

Mind- and Marketshare

The operating system that has received the credit for driving the Internet infostructure is unquestionably Unix. Nearly every FTP, Gopher and WWW server on the Internet is running some variant of Unix[16]. Initial client and server availability for any of these protocols have been developed exclusively on Unix as well. The "best Internet navigation tools" are the domain of the X Windows system on **Unix** and to the **Apple Macintosh** running MacTCP. The word on the Internet is that if you want to bring a server online, buy a Unix. If you want to get a cool Internet client, use Unix or better yet - buy a Macintosh. The industry standardization of the Windows Sockets API[17] has increased the availability of the tools for the Windows 3.x platform, but the "Internet==Unix" attitude is still dominant in this community, the quality of the Windows-based tools is very poor. Of the 500 highly technical engineering professionals that attended the last IETF[18] meeting, I encountered exactly 2 Sun SparcBooks running **SunOS**, one Intel/**Windows** system, and over 40 **Macintoshes** (oh yes, my laptop was the only one running NT).

The work done by our team position both Daytona and Chicago very well to shatter these Unix-biased perceptions. With built in TCP/IP, dialup connectivity via SLIP[19]/PPP[20] and Windows Sockets support, we will deliver millions of "Internet-ready" systems to the masses in 1994. Assuming that Web, Gopher and FTP systems are being added as information servers at the rate of 100s/month in 1994, there is a clear opportunity to influence mindshare and establish market presence. *The "Internet information server" is an specific market where Windows NT falls shy to Unix today.* If teachers, journalists, researchers and hobbyists are looking for the best system to connect to the Internet, the answer should be Chicago or Daytona. *The "Internet explorer system" is a specific area where Windows falls shy to Macintosh today.* By providing the technology to drive these exciting new services and to connect to the Internet easily, we stand to increase mind- and marketshare in areas which we have fallen short in the past. We have the opportunity to start displacing these highly visible Unix systems by adding the power and ease of Windows NT to administrators of current Internet servers.

I should point out that the dominance of Unix and Macintosh in these areas is a direct result of their historical presence in the academic community. With the exception of **Apple**'s involvement with the WAIS technology, **Apple**, **Novell**, and **AT&T** have contributed near-zero to these efforts that has enhanced the value and presence of their systems on the Internet. Most Internet users are far from married to the Mac or Unix; by taking a leadership approach in the integration of Internet connectivity, messenging and exploration with the Windows family, we will win new users and current users will be more willing to migrate away from these competitive platforms. *A title from a recent PC Week article: "Setting up the Web servers: Installing Unix is the toughest part"*

Research and Development

All of the facts referenced in this document were acquired in a few hours using the Internet tools FTP, Gopher, and WWW from my office[21] on a private Internet tap (a service not available to the average Microsoft employee). Even with our terrific library resources, excellent research staff, and access to Internet e-mail, it would have taken a considerable amount of time to assemble the data for this memo without direct access to these Internet resources. In fact, much of the data provided here is available exclusively in the form of automated reports and discussions on the Internet. All of the product and research teams at Microsoft have a great deal to gain by having ubiquitous access to the Internet. *Today, it is probably fair to consider Internet connectivity a competitive advantage in the software industry. Tomorrow it will be a measurable disadvantage if we're not wired.*

---

[16]Not entirely true, in the last month, I have been pleasantly surprised in connecting to 5 Windows NT FTP servers on the Internet in the course of research.

[17]*Windows Sockets: An Open Interface for Network Programming under Microsoft Windows Version 1.1*, January 1993; Martin Hall, Mark Towfiq,Geoff Arnold, David Treadwell, Henry Sanders.

[18]Internet Engineering Task Force - the body of engineering professionals which define the Internet protocols.

[19]*RFC 1055: Nonstandard for transmission of IP datagrams over serial lines: SLIP.* 1988 June; Romkey, J.L.

[20]*RFC 1331: The Point-to-Point Protocol (PPP) for the Transmission of Multi-protocol Datagrams over Point-to-Point Links.* 1992 May; Simpson, W.

[21]Note that due to limitations in the Windows-based Internet navigation tools, I generally choose to use my Sun Sparcstation to access the Internet. The Unix/X-Windows based which provides a much more mature environment for Internet navigation and information processing. As an Internet user, the Unix-centric perceptions I mention in this memo hold for myself much of the time given the present state of available technologies and tools for Windows and Windows NT.

A Phased Approach

Our increased participation in the Internet will yield results, however we must exercise caution as we become more active in this infostructure. Although Microsoft has built a fair amount of respect in the TCP/IP community, our global presence on the Internet is hardly secure. From a recent discussion on Web authoring tools on the Internet:

> >>>I think your idea is great, although I frankly think that the whole HTML concept needs to import some ideas from
> >>>(gasp) Microshaft Word.
> You don't want anything of the sort. The correct solution to the problem is...

In order to build the necessary respect and win the mindshare of the Internet community, I recommend a recipe not unlike the one we've used with our TCP/IP efforts: embrace, extend, then innovate. ***Phase 1 (Embrace):*** all participants need to establish a solid understanding of the infostructure and the community - determine the needs and the trends of the user base. Only then can we effectively enable Microsoft system products to be great Internet systems. ***Phase 2 (Extend):*** establish relationships with the appropriate organizations and corporations with goals similar to ours. Offer well-integrated tools and services compatible with established and popular standards that have been developed in the Internet community. ***Phase 3 (Innovate):*** move into a leadership role with new Internet standards as appropriate, enable standard off-the-shelf titles with Internet awareness. *Change the rules: Windows become the next-generation Internet tool of the future*.

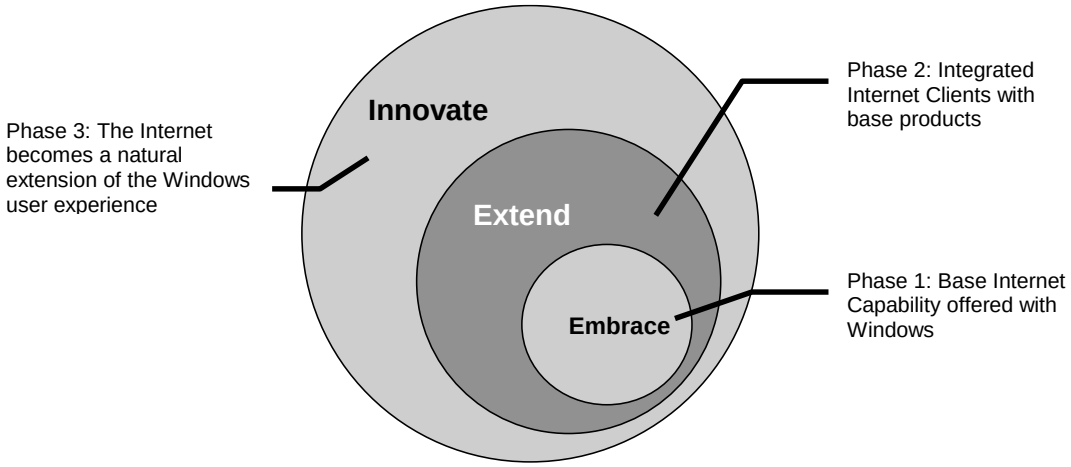

Figure 9: Peeling the Internet Onion

Although we are years behind many of our competitors on the Internet, our agility and creativity will allow us to catch up quickly. Be aware that the Internet presents a strange culture with a deep-rooted history that we cannot safely ignore and still achieve these goals. One concern that must be considered as we press forward is efficient communication between the groups within Microsoft that these efforts impact. There is a lot of potential synergy between these groups which we can exploit to make these visions become realities, although the potential for disaster is also present. The obvious teams to be involved include (but are not limited to) PSS, Workgroup Applications (Mail and Messaging), the TCP/IP internetworking team, DRG, and the Marvel folks. Each of these teams have unique perspectives and expertise to share contributing to a formula for success.

Phase 1: Base Internet Capabilities (Embrace)

We *must* deliver Internet-ready systems to customers in 1994. The biggest challenge in meeting this goal will be the extensive interoperability testing necessary to successfully establish Windows as an Internet-ready system. Although the development team is on track to deliver the necessary base technologies, we ensure 100% compatibility with the 100+ commercial Internet providers. Our in-the-box support for easy, compatible TCP/IP support is well ahead of **Apple**, **IBM**, and especially **Novell**'s positioning and pricing models, and sets the stage for our future endeavors.

Although offering Internet-compatible systems with Chicago and Daytona put the initial stake in the ground, poor application support can tarnish our image in the near term. The current state of public domain Windows-based tools is very sad. We need to inspire the development of high-quality Win32 Internet applications and begin winning the mindshare of those which have formerly embraced Unix as the standard Internet platform. Once the standard tools and services that have been available on Unix and Macintosh for years become widely available on Windows NT, those interested in offering Internet services will have a very attractive alternative to Unix. We can further enhance our image by solving customers' bootstrapping problem; by making quality Internet tools easily available to Windows customers, we will shorten users' paths to access to the infostructure.

The best way to establish respect in the Internet community is to become more active in it; we need to beef up our Internet support and participation through PSS. The FTP server experiment went very well, given the late-night efforts of those wanting to make it happen. It's apparent that more effective management of the server and increased internal communication will help make this resource much more valuable and allow us to further defray support costs. We also need to consider Usenet-based customer support and other services such as gopher, WAIS and WWW to deliver both pre- and post-sales information effectively to customers as companies like **Novell** and **Digital** have.

As important as bringing customers to Microsoft via the Internet, we need to provide the product groups access to the rich technical resources it has to offer. From specifications, to analysis, to algorithms, every product group within Microsoft stands to gain a great deal from being able to access these resources readily. A number of technical barriers face us in offering desktop connectivity right now, however setting up public access terminals will go a long way toward enhancing our corporate understanding of the Internet as well as increase our research capacity.

Weaving the Internet into our networking marketing messages will establish a more visible presence and get us up on the Internet "wave". The stability and use of the FTP server is testimony that Windows NT can handle mission-critical 24x7 service to huge numbers of concurrent users, and an excellent message to share. Our efforts in compatibility testing our products with 100 Internet providers ensures our commitment to Internet-based support and Internet-compatible products needs to be communicated and leveraged.

Action Items:

- Deliver easy-to-use, Internet support with Chicago and Daytona: TCP/IP, Windows Sockets, & SLIP/PPP
  - *Need commitment from Chicago for TCP/IP PPP development work*
  - *Ensure interoperability of Microsoft dialin offerings with all major Internet providers, offer Internet provider list with ResKit*
  - *Ensure interoperability of Microsoft TCP/IP offerings with all existing Windows Internet tools*
- Ensure the availability of high-quality Internet navigation tools and servers
  - *Sponsor a development contest for Win32-based Internet tools*
  - *Deliver winning Internet tools to customers via the Daytona/Chicago ResKits and on the Internet itself*
- Manage the Microsoft FTP server more efficiently, establish guidelines for use
  - *Dedicate PSS resources to the efficient management and administration of the Internet server*
  - *Develop a PSS requirements document detailing the needs of Microsoft customers w/ Internet access*
- Weave the Internet into Marketing Messages
  - *Daytona/Chicago press releases/Interop+NetWorld demos*
- Enhance employee access to the Internet as a research tool (kiosks)
  - *Provide Windows NT Internet systems in the library, EBC, etc. for public use*
  - *Examine extending desktop access through the SIAM project, allow creation of Internet-wide aliases for technical discussions*
- Develop a cross-team plan for the next level of Internet interoperability

Phase 2: Integrated Internet Clients and Servers (Extend)

Once we deliver core Internet connectivity technologies with the base system, it becomes easier to exploit the infostructure to win the mind- and marketshare of the Internet community. Providing the base connectivity technologies will enable Windows to become less of a stranger to users the Internet quickly and will accelerate development efforts in research/academia. We also buy the necessary time to shake out these new technologies, understand what users want from the Internet, and understand the trends more effectively. In this phase, we make Windows the *choice* system of the Internet by embracing the current technology and integrating these standard tools and services with our base products.

Although the strategy for this phase requires participation and input from many Microsoft teams there are some obvious opportunities for integration to exploit. Internet mail and Usenet should be integrated with our Capone and EMS products. The Explorer is a natural fit with technologies like FTP and gopher which present users with filesystem-like organization of rich documents. Web viewers are not totally unlike our WinHelp or Multimedia viewer tools. When the next generation of Windows include support for standard Internet protocols, the Internet becomes a natural extension of the Windows end-user experience, and the favored way to explore. With the interest in information systems and global connectivity in the K-12 area, the Internet carrot can help us grow the Windows market where the **Macintosh** is the favored choice today. Unless **Novell** repositions, these efforts, coupled with TCP/IP in-the-box will infringe on their extremely successful *LAN Workplace for DOS* TCP/IP product.

Integration will help us win the mindshare of new Internet clients, but in order to win the mindshare of current Internet users and server administrators, we will need to round out our server-side offerings to include other technologies like gopher, and WWW. The content-based query technology in Cairo is a natural fit to service WAIS and Z39.50 queries over the Internet as well.

Many of these efforts would appear to overlap with the goals of the Marvel project, I view the two efforts as complementary. It is very important for Marvel to determine a plan of action for Internet connectivity. Although it is possible to build a parallel network with some simple gateways (such as **CompuServe** and **America Online** offer), the Internet is *already* a global village with an incredible user base. With 2 million hosts growing at 5+% per month, MarvelNet is going to have a hard time competing with the growth of the Internet. I believe the most practical approach is to embrace these users and to use the wire that they're already on to acquire their business. Providing access in this fashion will be viewed by many users as competitive benefit over other information providers that only offer basic gateway functionality to the Internet.

Marvel *should (at least) be the central provider of any Microsoft support content* (e.g., updates, KnowledgeBase, and so forth). We should provide customers with the choice to access this information by either dialing in to the private MarvelNet directly, or to come in via the Internet. When the *MS-DOS 6.2 upgrade* was available on **CompuServe**, but not on the Internet, we received hundreds of e-mail messages demanding Internet access to the upgrade as well. Many customers love our Internet support, and barely tolerate CompuServe. We need to use this to our advantage to offer support to the widest range of customers while eliminating the overhead of maintaining multiple support areas.

Action Items (additional action items will shake out of the cross-team plan recommended in the previous phase):

- Complement existing support to include integrated support for standard protocols
  - *Offer integrated client support with the Explorer, Capone, Viewer and others as appropriate*
  - *Round out Windows NT server offering to include support for gopher and WWW*
  - *Integrate WAIS/Z39.50 query technology with Cario content indexing capabilities*
  - *Investigate WWW authoring capabilities with next version of Multimedia Toolkit or as WinWord macros*
- Leverage new technologies to provide support for Microsoft customers w/ Internet access
  - *Provide gopher access to ftp.microsoft.com*
  - *Index and format KnowledgeBase articles for simple browsing via WWW*
- Enhance employee access to the Internet as a research tool
  - *Bring Internet research capabilities to users' desktops with firewall/SIAM extension technology (productize this)*
- Consolidate PSS and Marvel support by offering Internet access to Marvel
  - *Free services should be available via FTP and gopher*
  - *Full Marvel services should be available using the Internet as the wire*
  - *Get Marvel team involved with ongoing commerce efforts such as IMP (the Internet Mercantile Protocol)*

Phase 3: Windows as the Global Infostructure Explorer (Innovate)

As Windows penetrates the Internet, we advance beyond simple integration to a point when *Windows becomes the next killer application for the Internet*. Specialized tools like gopher, FTP and WWW become old hat on the Internet quickly, the press and excitement shifts from specialized tools and Unix to Windows and its strong capabilities on the information highway. The enormous momentum generated by our integration efforts and increased presence feeds the Windows foothold in the Internet, as users begin to take advantage of the native system features to manage and navigate the infostructure. There are a number of activities that can be imagined:

The existing Domain Name System is becomes quickly eclipsed by the Cairo directory service; providing universal resource location services without the need for specialized, arcane applications like archie. Users of the SMTP mail infrastructure take advantage of the advanced features of EMS while remaining backward compatible with their current Internet messaging solutions. Distributed information on the Internet is indexed by Cario and browsed using the Explorer across thousands of information servers worldwide without the need of tools like WAIS and Mosaic. *Windows becomes the global infostructure explorer.*

The hottest content server on the Internet? Microsoft Marvel. It's already integrated with all of these services and has a huge content base and commercial vendor support. Perhaps most importantly, *Marvel solves the difficult electronic commerce problem well ahead of the Internet: the ability to securely purchase goods in a networked environment*. Cairo becomes the clear platform choice for new information providers (information providers are everyone everywhere, not just Dow Jones anymore, but every grade school in America), offering the most flexibility in both management and technology. OLE Windows applications are now Internet-enabled - people schedule meetings across the Internet and share ideas using Microsoft application products on Windows. No more specialized tools, no more need for Unix to carry the infostructure. Access to the Internet is a natural extension of the base Windows system as well as Microsoft applications.

Action Items:

- Ensure tight integration of Cairo technologies to provide simple migration
  Integrate DNS technology with Cario directory service
  Integrate WAIS/Z39.50 query technology with Cario content indexing capabilities
  Ensure OLE network capabilities and resource location scale well enough to satisfy Internet use
- Promote the use of Cairo as the infostructure super-server
  Seed universities with the necessary software to displace FTP and Web servers with Cario
  Develop simple migration tools to upgrade info servers from Unix to Cairo
- Enable workgroup applications for Internet use
  Have DRG promote the "right way" to develop Internet-ready tools for Windows to ISVs
  Ensure that Microsoft applications use OLE/DFS/Windows Sockets to guarantee Internet compatibility

Potential Risks

Wile pursuing even the most conservative opportunities which I present in this memo, these efforts are not entirely without risk. Here I identify some of the potential hazards we can expect to encounter.

Microsoft/Internet Culture Clash

One of the biggest challenges facing Microsoft's success in the Internet community is acceptance and respect. Although we have an incredible amount of respect in the commercial software business, the Internet has been founded on public domain protocols and products which generally included source availability at no charge. It has been only recently that vendors have suggested profiting from the Internet by selling the browsing tools and technologies, and offering commercial services on the Internet itself. The information and software has been free for 15 years, we need to be careful to embrace the current technologies and community before we attempt to reshape it.

Development/Testing - Interoperability Issues

We need to be very sensitive to the fact that many of the protocol "specifications" for these Internet tools were written after the tools themselves. The development of the Windows NT FTP server experience illustrated this point too effectively. Although we were able to develop a working server in 2 man-months, over 3 months of concerted interoperability testing were necessary in order to account for all of the interpretations made by the authors of the several dozen FTP clients. This process was accelerated greatly by putting the server online and having 1,000s of users test it for us by connecting, however the lack of history and involvement in the Internet community made this effort more difficult. Server testing is nice, because end-users can do it for us. Client-side testing poses new challenges. Without an extensive beta cycle, it will be difficult to ensure full interoperability of Internet clients with existing server implementations. My recommendation to ship Internet clients with the Resource Kit(s) in the initial stage and hold off putting them in the product allows us to guarantee they are rock-solid interoperable and supportable.

Support

Internet support provides a number of challenges and an incredible number of variables. Imagine a Chicago customer using a third party TCP/IP product, a public domain gopher client, dialing in via an ISDN provider and not being able to connect to their University gopher server. They will call us and ask whose problem it is and how to resolve it. By providing high-quality, interoperable solutions in-the-box we can eliminate a number of these variables. However, we need to establish a plan to support these customer situations effectively, perhaps exclusively through the Internet on public forums. This is another argument for offering well-tested, interoperable clients in the resource kits before integrating them into the core product.

Address Space/Routing Table Explosion

The Internet was never intended to grow to the size that it has. The designers of the original TCP/IP protocols did not take into account all of the scalability issues. As such the Internet is facing some scaling problems that need to be addressed. The short story is, the Internet backbone routers are being forced to maintain 30 new routing entries a day as new networks join the party. With 16,000 networks to route between and 30 new networks a day joining in, efficient routing is becoming more and more difficult for the Internet routers that bind the Internet. Further, the original addressing scheme and assignment policy was very lax, leading to very poor utilization of the 32-bit IP address space. Current predictions estimate the life expectancy (given current growth rates) of the Internet address space to be ~5 years. If the growth curve takes another dramatic upswing, this has the potential to become problematic in less time. Several engineering groups within the IETF have been chartered to design solutions to this pressing problem which are referred to as IPng, or Internet Protocol, next generation. The IETF has organized a 17 member board to provide guidance in this area to which I was appointed. This gives Microsoft an opportunity to protect our corporate interests by steering the proposals to consider requirements of longer-term developments that we'd like to explore. Some are of the opinion that the Internet will just blow up, that the IETF cannot solve this problem and that NII will be far enough along to eclipse the Internet altogether. I do not share this belief; I think that *the Internet will become a leaf network and the primary source of information for the initial phases of the NII*, eventually it will be overhauled, but it's the only universal game in town today.

Summary

Microsoft has stood quiet during the initial explosion in this infostructure growth. Driven primarily by academic and corporate desire for e-mail access and global connectivity, millions of users have become wired so they may access this huge source of information and the interactive communication capabilities. This incredible growth has spawned the development of second-generation navigation tools that is already fueling the next boom in infostructure; at-home, K-12, and commercial access. As the user base expands and begins to adopt the new Internet protocols such as World Wide Web, the quantity and accessibility of the information available on the Internet grows without bound (and the dependency on Unix and Macintoshes deepen). The Internet provides the opportunity for us to capture this user base by positioning Windows as the ideal information system for their needs.

The Internet infostructure can deliver an enormous blow to our Systems competitors, provided that we are successful in getting this influential and growing community to first accept Windows. By embracing current technologies available on the Internet, we position Windows as the choice system for interactive Internet services and prepare for the shift to the native IAYF technologies offered by Cairo and other Microsoft products. At the same time, we enhance our customer service, expand the Marvel user base, increase our research potential and reduce our support costs by leveraging the infostructure ourselves.

"On the Internet, nobody knows you're a dog."

Source: *The New Yorker*, July 5, 1993