PART

One

The Internet and the Router

It is impossible to write about routing protocols and the routing policies that affect the behavior of these routing protocols without a firm grasp of just what these twin tools are trying to accomplish. Routing *protocols* establish the global connectivity between routers that in turn establish the global connectivity that makes the Internet what it is today. Routing *policies* adjust and tune the behavior of the routing protocols so that this connectivity is made more effective and efficient.

Routers are the *network nodes* of the global public Internet, passing IP address information back and forth as needed so that every router that needs to knows when a new network (IP prefix) has been added anywhere in the world, or when a link or router has failed and so other networks might now be (temporarily) unreachable. Routers can dynamically route around failed links and routers in many cases, unless the destination network happens to be right there on the local router itself. Routers are network nodes in the sense that there are no users on the router itself that originate or read email (for example), although routers routinely take on a client or a server role (or both) for administrative purposes. Routers almost always just pass IP packet traffic through from one interface to another, input port to output port, all the while trying to make sure that the traffic is making progress through the network and moving one step closer to its destination.

The network that a great many routers find themselves attached to is, of course, the global, public Internet. This is not always the case, however, and there are still plenty of private router networks with no links to the Internet at all, sometimes for the sake of security, often just because connectivity to the Internet for this network is simply not needed or desired. Often local area networks (LANs) used in private organizations use routers to link departments, usually within the same building or office complex. This book will mention such private router networks only in passing, not because these networks are unimportant, but mainly because the role of routing policy is more critical when the global public Internet is involved than when connectivity between the Sales and the Marketing departments are the only issue. The emphasis on this book is on the global, public Internet.

The situation in the router world and on the modern Internet is complicated by considerations of dynamic host addresses, IP network address translation (NAT), and other features often used now for security purposes. The emphasis in this book will be on router use of publicly assigned IP address spaces. Again, the intent is not to downplay the significant role that dynamic host address configuration or NAT play in modern router networks, but just to make the main topics of routing protocol and routing policy behavior more understandable and less complex than they already are.

So this book starts off with a look at the role of the router as the platform of the routing protocols, and the history of the Internet that forms the context within which the routing policies operate.

CHAPTER

1

A Brief History of the Internet and Router

The days of conceiving the Internet as something to be mapped, grasped, understood, controlled, and so on are quite frankly gone. What exists instead in today's world of interconnected computers is a kind of *ISP grid net*, a haphazard, interconnected mesh of Internet service providers (ISPs) and related Internet-connected entities such as governments and learning institutions. But why introduce a new term when *Internet* is much more common and perfectly fine for most discussions of routers? Because only with an appreciation of the Internet as an ISP grid net can the important role of routing protocols and routing policies in today's Internet be understood. Talk of peers and aggregate summaries and backbones and access points and points of presence (POPs) make much more sense in the ISP grid net context than in the older context of a monolithic Internet.

The idea of the Internet as ISP grid net is shown in Figure 1.1. Large national ISPs, smaller regional ISPs, and even tiny local ISPs make up the grid net. In addition, pieces of the Internet act as exchange points for traffic such as CIX (Commercial Internet Exchange), FIX (Federal Internet Exchange), and NAPs (network access points). The precise role of the NAPs, CIX (now officially obsolete), and FIX will be explained later on in this chapter. They are included in the discussion to point out the overall and varied structure over time of what appears to be a unified Internet.

These Internet pieces are all chained together by a haphazard series of links with only a few rules, mostly of local scope (although there are important exceptions). NAPs, which are collections of routers where different ISPs can exchange traffic, are meshed with very high-speed links, and *Tier 1 ISPs* must have high-speed links to two (or more) NAPs. The smallest ISP can link to another ISP and thus allow their users to participate in the global, public Internet. Increasingly, linking between these ISPs is governed by a series of agreements known as *peering arrangements*. National ISPs may be peers to each other, but they view smaller ISPs as just another type of customer. Peering arrangements detail the reciprocal way that traffic is handed off from one ISP to another. Peers might agree to deliver each other's packets for no charge but bill non-peer ISPs for this privilege, since presumably the national ISP's backbone will be shuttling a large number of the smaller ISP's packets around but using the smaller ISP for the same purpose to a lesser degree. A few examples of Tier 1 ISPs, peer ISPs, and customer ISPs are shown in the figure.

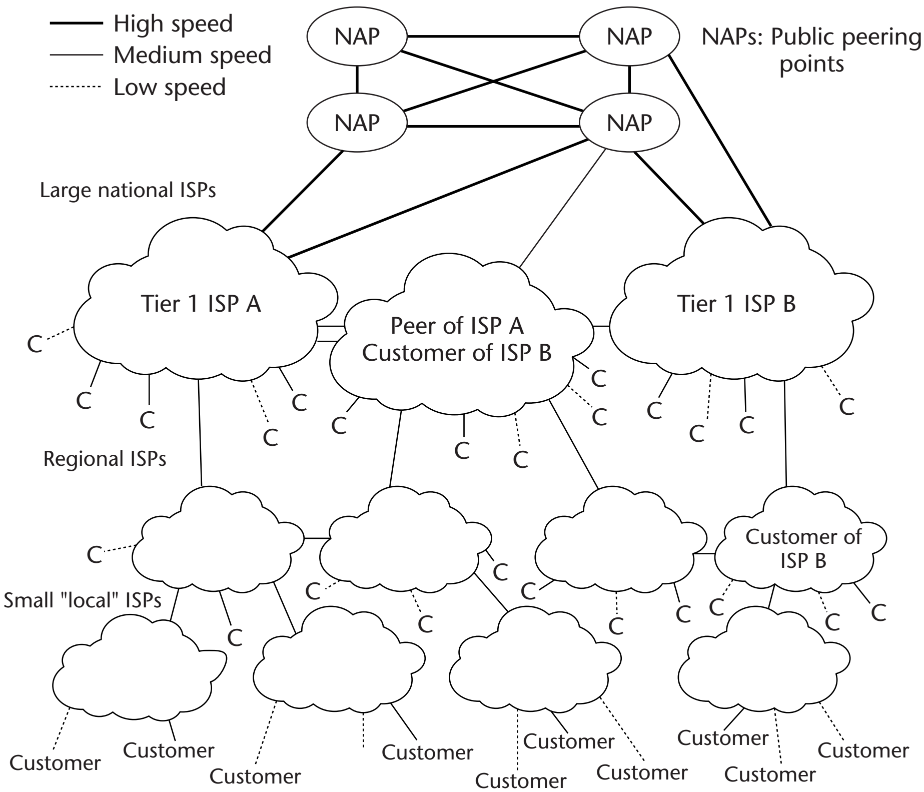

Figure 1.1 The ISP grid net.

At the bottom of Figure 1.1, millions of personal computers (PCs), minicomputers, and mainframes act as either clients, servers, or both on the Internet. These hosts—anything running Transmission Control Protocol/Internet Protocol (TCP/IP)—are usually attached by LANs and linked by routers to the Internet. These LANs are just shown as customers to the ISPs. Although all attached computers conform to this client/server architecture, many of them are strictly Web clients (that is, browsers) or Web servers (that is, Web sites) as the Web continues to take over more of the form and function of the Internet at large. Only at this bottom level is the term *customer* spelled out. At the other levels, members of each ISP's network are represented by just a *C*. For the sake of simplicity, Figure 1.1 ignores important details of the grid net such as the LANs and routers. However, it is important to realize that the clients and servers are on LANs and that routers are the network nodes of the Internet. The number of clients actually exceeds the number of servers many times over, but this is not apparent from the figure.

Moving up one level, the figure shows the thousands of ISPs that have emerged in the 1990s, especially since the Web explosion of 1993 to 1994. Usually, the link from the client user to the ISP is by way of a simple modem-attached, dial-up telephone line. In contrast, the link from a server to the ISP is most likely a leased private line, but there are important exceptions to this simplistic view. Although also not shown in the figure, a variety of Web servers may be within the ISP's own cloud network. For instance, the Web server on which an ISP's members may create and maintain their own Web pages would be located here.

A common practice in the networking field is to represent an ISP's (or any other type of service provider's) network as a cloud or oval. Sometimes the graphic actually *looks* like a cloud, but this practice only detracts from the figure's message in many cases, and networks are anything but light and fluffy. The use of the network cloud goes back to a telephony service provider data network known as the X.25 public packet-switching network, which shares many features with the Internet. The reason for the X.25 cloud was twofold. First, customers and users did not have to concern themselves with the details of the network in any way. Packets went into the cloud and emerged from the other side. Second, the cloud hid the fact that what was inside it was really exactly the same types of things that customers had on their own networks: network devices and links between them. There was no magic at all, just a network. X.25 was simply a public version of a private network, but with hidden details, packets, and economies of scale. In the same way today, ISPs condense their networks into clouds to hide the details of their actual network structures from customers (who do not need to know), competitors (who almost all want to know but should not), and hackers (who definitely should not know). Moving up again to a higher cloud layer, the smaller ISPs link into the large backbone of the national ISPs. Some may link in directly, whereas others are forced for technical or financial reasons to link in daisy-chain fashion to other ISPs, which link to other ISPs, and so on until an ISP with direct access to a NAP is reached. Note that direct links between ISPs, especially those with older Internet roots, are possible and sometimes common. In fact, the NAPs were once so congested that most major ISPs prefer to link to each other directly today, and so are peering directly to one another, bypassing the need to use the NAP hierarchy to deliver traffic.

The NAPs themselves are fully mesh-connected—that is, they all link directly to all other NAPs. Figure 1.1 shows only the general structure of the U.S. portion of the Internet. However, a large percentage of all inter-European traffic passes through the U.S. NAPs. Most other countries obtain Internet connectivity by linking to a NAP in the United States. Large ISPs routinely link to more than one NAP for redundancy. The same is true of individual ISPs, except for the truly small ones, which rarely link to more than one ISP, usually for cost reasons. Note also that peer ISPs often have multiple, redundant links between them.

Speeds vary greatly in different parts of the Internet. For the most part, client access is by way of low-speed dial-up telephone lines, typically at a speed 33.6 to 56 kilobits per second (Kbps). Servers are connected by medium-speed private leased lines, typically in the range of 64 Kbps to 1.5 megabits per second (Mbps). The high-speed backbone links between national ISPs run at higher speeds still, sometimes up to 45 Mbps. On a few, and between the NAPs themselves, speeds of 155 Mbps (known as OC-3c), 622 Mbps (OC-12c), 2.4 gigabits per second (Gbps) (OC-48c), and now even 10 Gbps (OC-192c) are not unheard of. Higher speeds are needed both to minimize large Web site page transfer latency times and to concentrate and aggregate traffic from millions of clients and servers onto one network.

Where did the ISP grid net come from? What happened to the Internet along the way? How did the routers and the protocols that run on these routers become so important to the Internet and Web? To answer these questions, we need to start at the beginning.

The Pre-Web Internet

A popular television commercial in the United States once switched back and forth between images of 1960s-era rock concerts and peace rallies and a small group of white-shirted, pocket-protected, glasses-wearing nerds trying to make a computer the size of a small car power up properly. The nerds consulted their slide rules, which were devices used by engineers to make calculations before there were electronic calculators, and finally managed to make the computer flash green lights in a satisfying manner. The commercial then ended in the present, and as the gray-haired and paunchy nerds labored with new equipment, a youthful engineer gazed in wonder at the slide rule that was found in a drawer. The point was, of course, that in 1969, while many people frolicked through the carefree 1960s, a few dedicated engineers were putting together the first sites for what would become the Internet. The commercial was full of obvious improbabilities, such as the stereotyped appearance of the group and the presence of a slide rule in a modern computer lab. Even by 1969, many engineers had already embraced the expensive and bulky laptop-sized electronic calculators that could only add, subtract, multiply and divide but were starting to appear on the market. But the Internet and computer networking in general are not all that old, and many network pioneers are still productively involved in all aspects of modern research and development. Despite the relative newness of the technology, the networking variations from the late twentieth century seem antiquated today. The Internet of 1990, for example, is in some ways as different from the modern Internet as an old World War I vintage biplane is from a modern jet fighter.

Of course, it is just as wrong, and just as right, to call the network built in 1969 the Internet as it is to call the contemporary ISP grid net the Internet. What the nerds had wrought, at the same time almost to the day that many other college students were happily rolling in the mud in upstate New York at the Woodstock Music and Arts Festival, was a U.S. government network called (in true federal government acronym fashion) the ARPANET, or Advanced Research Project Agency Network. ARPANET was funded in 1968 to perform research into packet-switching networks, and the network nodes were to be built by a company called Bolt, Beranek, and Newman (BBN). These network nodes were not called routers, or even gateways (the older Internet term for router). They were called interface message processors, or IMPs. Not everyone in government, even those who know about the BBN contract, was quite sure what was going on or even just what an "interface" was. The story goes that Senator Edward Kennedy, in whose home state of Massachusetts BBN was headquartered, sent a congratulatory message to BBN thanking them for the efforts to bridge religious differences with their new *interfaith* message processor.

ARPA itself had been created under the U.S. Department of Defense (DoD) to combat the perceived gap between the U.S. and Russian space programs. This gap was made painfully obvious to some when the Russians launched Sputnik, the first earth-orbiting satellite in 1957. The possibility of spying or even bombing from orbit became a real concern, and interservice rivalry between the Army and Navy over their own satellite plans slowed the U.S. response even further. Research into rocketry and related systems such as in-flight guidance at U.S. colleges and universities was slowed by a lack of communications between staff efforts to address problems. The answer, ARPA soon decided, to all these scattered efforts was closer coordination among agencies and institutions receiving ARPA funds under the DoD banner for research. Since many of the engineering issues that had been raised by then-current research were being addressed graphically on computers, it seemed plain to many that some form of computer network was needed to bring some semblance of order to these efforts.

The problem was that no one at the time had the slightest idea how a network for computers, as opposed to, say, telephones, should look and act. The early 1960s saw progress on this basic problem in the form of a series of papers. From 1961 to 1964, three crucial papers outlined the basic concepts. Leonard Kleinrock, at the Massachusetts Institute of Technology (MIT), examined packet switching using small parcels of data that came to be called datagrams; J.C.R. Licklider and W. Clark at MIT, explored the idea that computer communication could take place "online" in real time; and Paul Baran, at RAND, an important think tank, investigated the absolutely key concept that a network intended for national defense should have no central point of failure, or even a place where everything was controlled.

Implementation of these ideas started slowly. In 1965, two computers were linked with a 1,200 bits per second (bps) telephone line, pretty much state-of-the-art speed for the time. At least telephone giant AT&T had already invented the modulator-demodulator, or modem, for analog-digital conversion at Bell Laboratories sometime during World War II so that digital computer bits could flow over a standard analog telephone line. One story about the modem has it that Bell Labs wanted to demonstrate a new telephone system computer at a conference at Dartmouth College in New Hampshire. But in the early 1940s, the threat of sabotage and spying was considered too great to actually risk shipping the computer by truck from New Jersey. So the engineers devised the modem as a way for an engineer to sit at a teletype machine keyboard (those had been around since the early 1900s) at Dartmouth, type a command for the computer in New Jersey, and then see the output as it scrolled on the teletype machine's paper output. Its purpose served, the modems apparently went into a closet somewhere until they were needed again 20 or so years later.

Various plans for a full ARPANET were circulated over the next few years, until by December of 1969, four nodes were up and running. These were at the University of California at Los Angeles (UCLA, whose IMP was installed on August 30), Stanford Research Institute (SRI, whose IMP was connected on October 1), the University of California at Santa Barbara (UCSB, whose IMP was connected on November 1), and the University of Utah (linked soon after). A logical map of the initial four-node Internet appears in Figure 1.2. The computers linked were an IBM 360, a DEC PDP 10, an SDS Sigma 7, and an SDS 940. Ironically, the ARPANET was all ready to go after the space race was over, having been won by the United States in July of 1969 with the initial lunar landing.

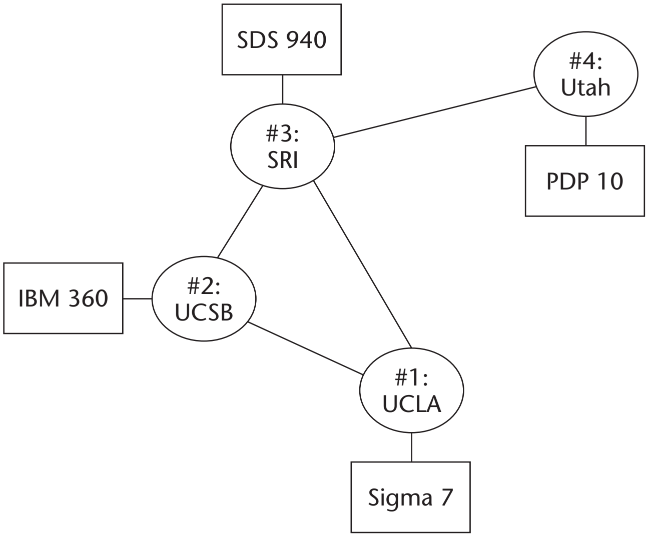

Figure 1.2 The original four-node ARPANET in 1969.

Note that the SRI IMP not only had to deliver packets to the attached SRI host but pass packets on to Utah. This forwarding aspect of traffic not for itself is the essence of network node and router operation. One of the most important features of this initial ARPANET was that the computers linked were from different vendors, as shown in Figure 1.2. So all four nodes used different operating systems, internal representations of data, and low-level languages. The function of the IMPs was to take the vendor-specific internals of the source data and translate them to a common "protocol" as the information flowed between the IMPs. Thus, each IMP only had to convert between two formats: the internal format of the host computer and the network format. This was much more important then than it is now. By the late 1970s, there were at least 10 major computer vendors in the United States alone, all with their own architectures and internals. An IMP, even with a whopping (for the time) 12 kilobytes (KB) of memory, could hardly be expected to understand and translate among them all. The format used on the network complied with the ARPANET Host-Host Protocol, which was soon replaced with the more robust Network Control Protocol (NCP), and later still by TCP/IP.

The early network pioneers, called the Network Working Group, were not even sure they were always doing what was expected of them. They had a mandate from Washington to create a computer network; that much was clear. But with the planners all the way back on the East Coast, the implementers were not taking any chances that they were somehow exceeding their authority of going beyond the strict terms of the contract between ARPA and BBN.

So right from the start, in April 1969, Stephen Crocker at UCLA decided to document implementation issues and Network Working Group decisions on how to solve them. Since many were convinced that some "pro from the East" would appear at some point and tell them exactly what to do, these messages back to the East Coast were titled "requests for comments." This practice gave birth to the famous series of Internet specifications, the RFCs, but unfortunately (or fortunately), no one ever showed up to take charge.

The newly born network spent many years more or less inventing itself. ARPANET was a hit from the start, but only among the groups under the ARPA umbrella. Digital lines (rare at the time) running at 56 Kbps came in 1970 to link BBN to UCLA, and MIT to Utah. Fifteen nodes were operational by the end of 1971, which was the planned target size, and the familiar email @ sign made its debut. International links came in 1973, to England by way of Norway, and in 1976 Elizabeth II, Queen of the United Kingdom, sent out an email from the first head of state at the Royal Signal and Radar Establishment.

This is not to say that the ARPANET did not have problems. In 1973, there was a famous lockup on Christmas Day when the Harvard IMP decided to tell all the other IMPs that it was zero hops away from every destination on the ARPANET. Naturally, all packets converged on Harvard, creating the first *black hole* in Internet history. And right from the start, it became obvious that pure packet streams could arrive out of sequence or not at all. Independent routing was robust and reliable when it came to reachability (if there's a way for a packet to get there, it will), but error-prone and "unreliable" when it came to basic service quality (no errors, then many packets missing; in sequence, then not . . .). So in 1974, these and other problems were addressed in a paper from Kahn and Vinton Cerf proposing an additional protocol layer to add some simple reliability to the packet shuffling through the IMPs on the ARPANET. This was to be called the *Transmission Control Program* (TCP).

ARPANET provided three key services to its users: email, remote computer access (to become Telnet), and file transfer (actually, file copy) across the network. In 1973, the most important use of the ARPANET appeared to be email: a study done that year showed that fully 75 percent of the traffic on the ARPANET was email. ARPANET quickly grew far beyond its initial 15-node vision, and by 1983 comprised 113 nodes.

The term *Internet* appears to have been introduced in 1982, once TCP (now Transmission Control Protocol) and IP (Internet Protocol) became the standard protocols for ARPANET. TCP and IP were originally intended to be meshed as one protocol layer, but during the development cycle in 1978, the decision was made to try to make TCP and IP independent. However, in many features the split was just not practical, so the designation TCP/IP reflected the close relationship between the two functions. Any collection of networks linked by TCP/IP formed an internet. Those that linked networks on the ARPANET formed the Internet according to some documents released around that time, although the ARPANET did not disappear officially until 1990.

However, the initial networks linked by the IMPs were just isolated minicomputers and mainframes with their associated terminals. It was not until the LAN came along in the early 1980s that the idea of a "network of networks" or internet was used in the modern sense of the word. So "network of networks" came to mean "a wide area network of local area networks," although there was and is no reason that wide area networks (WANs) could not form an internet in addition to LANs. In fact, this was done in 1982 when what was now the Internet was linked to another packet WAN. LANs came along when Bob Metcalf linked Xerox Alto computers (early PCs) with a coaxial cable snaking through the Xerox's Palo Alto Research Center (PARC) in the late 1970s. This network was soon standardized as *Ethernet*, although technically the standard version should be called an IEEE 802.3 LAN.

The first *exterior gateway protocol*, called without a sense of irony the Exterior Gateway Protocol (EGP), was used to link other WANs to the Internet through what was called a *gateway*. These gateways were another step along the evolutionary path to the modern router. Originally, all routing information was distributed using a single algorithm to each and every device on the Internet. So all network nodes knew every detail of the network. Even by the mid-1980s, this degree of housekeeping was proving to be beyond the capabilities of even high-end processors. Thus, the gateways allowed for interconnection of separate *routing domains* to the core Internet, and each piece of the Internet would run an interior gateway protocol (IGP) to provide full connectivity information within a specific routing domain. Between routing domains, however, only basic reachability information was needed so that each portion of the total Internet knew generally where things were. A driver leaving Chicago only needs to know that San Diego is in southern California somewhere. Only once across the California border are more precise directions needed.

Several other key developments and experiments matured during 1983 and 1984 and changed the world of networking forever. First and foremost, the cutover to TCP/IP from NCP began. Then the University of Wisconsin developed the concept of a *name server* so that users no longer had to look up the network address of computers and email recipients in bulky books. This work lead directly the Domain Name System (DNS) that could translate something like an email addressed to walterg@juniper.net to the proper IP address without user intervention. But beyond a doubt the most important development during that time period was the incorporation of TCP/IP into an open, standard, supported, powerful, inexpensive, multi-user operating system: UNIX. Once linking a computer to the Internet no longer required specialized hardware or software (almost every UNIX workstation had an Ethernet interface built in), the entire research community began to use UNIX and the Internet as an indispensable part of their work.

Now LAN interconnection over the Internet began in earnest, fueled by the ready availability of Berkeley UNIX (4.2 BSD) on desktop workstations. BSD took the BBN TCP/IP code and first merged it, and later essentially rewrote the code, into UNIX. BSD had the TCP/IP protocol preinstalled and instantly usable. In contrast, the "official" UNIX developed and distributed by AT&T Bell Labs had no integrated networking capability at all. But because of restrictions on AT&T as a regulated telephone company, UNIX could not be sold for profit by AT&T and was available to anyone who paid Bell Labs to ship it out. And these users in turn were free to add, subtract, or modify anything they wanted in UNIX, and even redistribute the result. BSD was a true distribution of UNIX independent of the official AT&T channels. It is important to realize that these new networking capabilities were at this time more or less restricted to UNIX users. The common IBM-architecture PC was in this timeframe marketed and sold as a standalone device that networking users tended to sneer at. And Apple computer users were always proud of their independence from standards or trends outside their own world.

UNIX today forms the core operating system for many types of computers, workstations, and even routers. Hardware vendors now had a choice. They could spend a lot of time and effort developing a new operating system from scratch, and then marketing it and teaching people to use the new commands and interface. Or they could just get a version of UNIX for practically nothing, adapt it for their own architecture, and then provide their users with a familiar command set and interface. Figure 1.3 traces the evolution of UNIX from its origins in 1969 to the many slightly different, yet always familiar, versions available today. Many users today encounter UNIX as Linux or FreeBSD. Even Mac OS 10 has strong elements of UNIX in it, to the extent that Mac OS 10 must spawn a Mac OS 9 process to run older, non-UNIX-oriented applications.

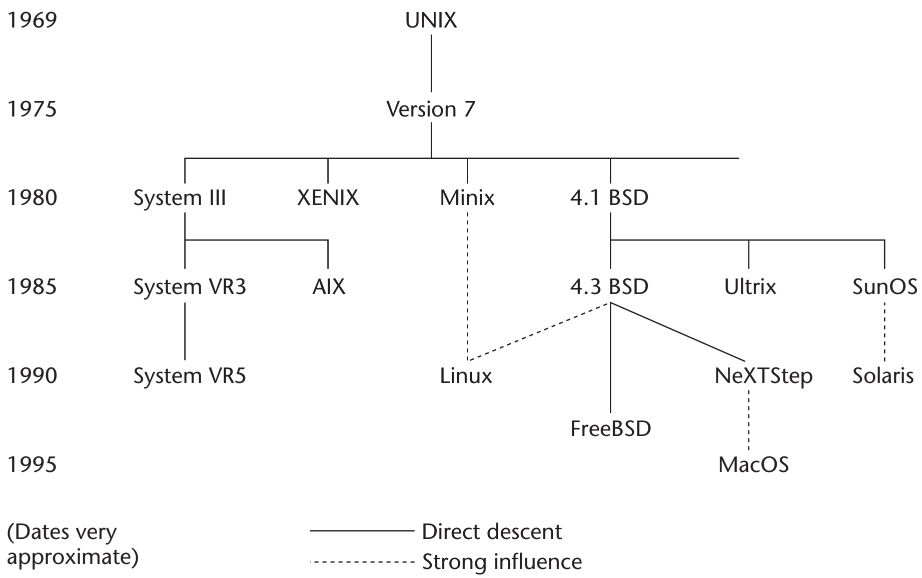

Figure 1.3 The evolution of UNIX.

In the same timeframe, ARPANET linked through a gateway to CSNET, another large network. All this linking made the U.S. Department of Defense nervous, since many military bases were now on the network as well. So in 1983, ARPANET split off the military bases on the network as MILNET, leaving only 45 of the 113 nodes behind. DNS registration began when symbolics.com was registered on March 15, 1985.

In 1986, the National Science Foundation (NSF) built a backbone for the Internet (NSFNET) with 56-Kbps lines to link supercomputer centers at universities. This began a new surge of universities and colleges onto the Internet, this time with no apparent government research ties. The routers of the day did a great job of routing around link failures just as planned—when they could. On December 12, 1986, an AT&T cable was broken between Newark, New Jersey, and White Plains, New York. All seven ARPANET trunk lines running to New England happened to run through that single cable, so that the region was effectively isolated from the rest of the world.

Early in the following year, NSF signed an agreement to manage the NSFNET backbone with Merit Network, Inc. Merit brought computer giant IBM and telephone company MCI on board, and together they founded Advanced Network Systems (ANS). The ANS dial-up network later became the basis for the America Online (AOL) network for PC modem users. ANS was the closest thing that the Internet ever had to a location the Internet was run from operationally. ANS was headquartered for years in Elmsford, New York, a tiny village about 30 miles north of New York City. And Senator Al Gore did indeed "invent" the Internet in a sense when he officially requested a national network for research and education. Why build something new when the Internet was lying around? After all, RFCs had reached 1,000, and attached hosts now topped 10,000.

The Internet was truly becoming global by 1988. NSFNET had links to Canada, Denmark, Finland, France, Iceland, Norway, and Sweden.

By the late 1980s, most of the major pieces of the Internet were in place. There were plenty of college students sending email, LANs, WAN links, routers (gateways), routing protocols, routing domains, the basic name service (DNS), and workstations running UNIX with TCP/IP built in. All that was missing was the ordinary people without a high degree of computer or network sophistication. With the Web, the people would come. And they would come in a rush.

The Web Comes to Town

Most people encountered the Internet in college. Internet links were usually subsidized and paid for by grants, so no one had to worry about paying for Internet access or connection time. The first commercial ISP came in 1990, when The World began charging customers to dial in to the Internet with a modem. Those without UNIX and TCP/IP (almost everyone with a PC or Mac) could use a terminal emulation application to communicate with a real host attached to the Internet. This caused some consternation at first, since the acceptable use policy (AUP) in place at the time mandated Internet use for research and education, not commercial purposes. But as government subsidies began to dry up, NSF lifted commercial restrictions on the Internet in March 1991. Now there was a scramble to figure out how to make money, or at least not go broke, linking pieces of the Internet together. Anybody with a router and a link could find a way to hook up to the Internet. But routers were very expensive.

The evolution of the Internet and the evolution of the router cannot be considered in isolation. As the Internet grew, routers had to evolve to keep up with the growth of traffic and routing table size. And more powerful routers in turn encouraged more Internet links, more networked applications, and so forth. When the Web explosion came, the router industry was ready, and one company even more so than others. That router company was Cisco.

As important as the University of California at Berkeley was for UNIX (BSD UNIX), Stanford University in nearby Palo Alto was even more important where the Internet was concerned. One Stanford engineering building, Margaret Jacks Hall, was the birthplace of industry leaders Silicon Graphics Incorporated (SGI), Sun Microsystems (Sun stands for Stanford University Network), and router "inventor" Cisco Systems. While Cisco did not really invent the router any more than anyone "invented" the Internet, Cisco did popularize the name *router* instead of gateway and basically defined what a router for the Internet should do. Routers at the time were also called *firewalls*, since routers could prevent many of the problems that surfaced when linking many LANs together with bridges. Bridging created one big LAN at the frame layer, and this meant that any problem on one LAN at one site could easily take down all of the sites linked by the bridge. Routers created one big network (or *internetwork*) at the packet layer, but all of the LANs were still separate at the frame layer. So LAN problems could be stopped by the packet-handling firewall created by the router.

By the early 1990s, there were 11 vendors of networking equipment that sold devices that had at least some routing capabilities. Some of these products were positioned as *enhanced bridges*, since most LAN administrators were more familiar and comfortable with the bridging features of these products and not all worried about full implementations of all routing protocols. But all dealt with IP packets at least at some level. These vendors were 3Com, Alantec, Bay (Wellfleet), Cabletron, Cisco, DEC, Hewlett-Packard, IBM, Proteon, Retix, and SMC. The most successful of these was to be Cisco, although when the Web hit town, the most powerful backbone IP routers were made by Wellfleet (soon to become, with Synoptics, Bay Networks). One dollar invested in Cisco in 1995 became eight dollars in 1998. By 1999, Cisco was the eighth most valuable company in the United States.

The Cisco AGS+ router was a marvel of engineering at the time, and yet today that fact is hard to believe. The CPU ran at 25 or 30 MHz and typically routed packets between four LAN and two serial WAN ports, although much larger numbers of ports were supported. A hard drive with a couple hundred megabytes was included as well. The AGS usually had about 4 MB of RAM, depending on configuration and position in the network, but 16 MB could be used as well. A single megabyte of RAM cost up to $250 at the time. But for the Internet of the day, the router worked: Packets went in and packets went out.

The Internet was not the only game in town. There were online services and bulletin board systems for ordinary citizens with PCs and Macs that were totally independent of the Internet. But these “proprietary” services such as CompuServe, Prodigy, and AOL were more concerned with keeping their users satisfied with their own content rather than unleashing them on the Internet. Besides, what was there on a government research network that everyday people would want? And pay for?

In the early 1990s, there were some other very good reasons why most people, even people with personal computers at work or at home, had no good reason to think about joining the Internet. People with PCs were just getting used to pointing and clicking at things with Windows (Mac users did not suffer from this sudden shock). The almost totally UNIX command-driven Internet was too different even from DOS to feel comfortable for most people. Computers did not come with modems built in, and external analog modems, which now ran at 9,600 bps, cost about $700 (the computer itself at the time could cost upwards of $6,000 for current technology, and early double-speed CD-ROM drives cost $1,000 or more). Most affordable modems ran at only 1,200 bps. And since there were few places to dial in for Internet access, the phone call itself was often long distance. Demand could quickly drive down the price for the hardware, but what could drive the demand for modems and computers, and routers?

The answer, of course, turned out to be the Web.

The Birth of the Web

Imagine life on the pre-Web Internet. UNIX systems ran TCP/IP and hooked up to a router. Server processes waited for client processes, run by users, to talk to them, usually when a user sat down at a monitor and keyboard and typed in a command like “mail.” Perhaps the user was a scientist at a research institution using the Internet to check his email. Today, there might be a response from a colleague who might be able to recommend a paper requested to advance a current research project. The email scrolls across the screen, and the colleague thinks there might be something on a computer on the Internet, but in another country. And the colleague cannot remember exactly where it was on the computer. No matter, the scientist can type in and run "telnet" as a client and log in as "guest" on the remote computer. Now the UNIX commands are run on the remote computer as if the scientist were a local user. Directories are listed and probed with Telnet commands. What luck! The title of one paper seems to be just the thing. Now the scientist can type in and run "ftp" as an "anonymous" user and again list the contents of the directory (however, the commands to list files in a directory with Telnet are different than those used in FTP). Assuming the scientist knows these commands, the scientist can get a copy of the paper transferred to his or her own machine. Unfortunately, the paper title was misleading and is of no help to the scientist at all.

What's wrong with this picture? Well, for one thing, this is a lot of work just to fetch a paper. The three programs are separate clients, and only Windows or a Mac or UNIX can actually run them all at once. Because the clients are three separate tools, the commands are different and they do not work with each other. The paper was found with Telnet, but Telnet cannot directly transfer the file (actually, there are other UNIX programs accessible through Telnet that could do this, but this is just an example).

The whole point is that with the Web site (server) and Web browser (client) in place, everything changes. The browser is a type of *universal client* that allows users to send email, access other computers, transfer files, view content, and so on all with the same familiar pointing and clicking instead of needing to remember arcane commands. Information at the Web site is organized according to what are still called *hypertext* links, although the links can lead to pictures, movies, or music.

In retrospect, the Web seems so obvious that it's a wonder it took so long to be invented. But in a world of expensive computers and slow link speeds, the Web seemed like a luxury rather than a necessity. Nevertheless, the reaction the first time people, especially people without a technical background in networks and computers, saw the Web was "I need that *now*!"

Not that inventing the Web was easy or fast.

The idea of hypertext was first used in 1981 in a book called *Literary Machines* by computer scientist Ted Nelson. Hypertext is the idea of following a nonlinear path through data and documents through a series of linked "nodes" of information.

This nonlinear concept can be traced back to 1945 to Vannevar Bush, who wanted to design a computer information database he called a *memex.* Memex computer users would be able to read data on a subject and from within that document be able to link to related information throughout the database. The user would follow a trail of links throughout the database that was not limited to a specific topic but allowed for connections across boundaries of classification. Information about "Wall Street" is just as likely to lead to "New York City" as "Stock Exchange." This whole concept became the basis for the Web. The very "web" structure sometimes makes for a frustrating experience, because following links can be confusing. If the objective of a "Wall Street" search is to find stock prices, "New York City" is not helpful. What does "New York City" have to do with all this?

Although many associate Web-like linking with the Internet, the idea would eventually have its first real application not on the Internet, but inside a computer. In 1987 Apple Computers released a software package called HyperCard that came bundled free with the Macintosh and its operating system. HyperCard not only allowed the linking of documents; it enabled users to create links to sounds and images within those documents. HyperCard restricted the links to the same system, but in some sense, the Web is just the old Mac HyperCard system (or Windows Help system) with links allowed to other computers.

The attempt to bring the concept of hypertext to the Internet began in the late 1980s when Tim Berners-Lee was employed as a software engineer at CERN, the European Particle Physics Institute, in Geneva, Switzerland. The high-energy physics community was spread throughout various universities and industries in Europe. The information and the research data this community of scientists produced spread over many computers and networks.

In March 1989, Berners-Lee proposed a way of linking text documents with other documents using *networked hypertext*. Working with Robert Cailliau, they eventually produced a design document in November 1990 that was a proposal for developing a system based on their idea of networked hypertext. "Hypertext is a way to link and access information of various kinds as a web of nodes in which the user can browse at will," the document stated. "Potentially, hypertext provides a single user-interface to many large classes of stored information, such as reports, notes, databases, computer documents, and online systems help."

This paper used the term *world wide web* without the capitals. After passing a few documents among computers, the pair quickly found that the documents needed a standard format that could be interpreted by each computer and still convey the information. So they developed a new language called the *Hypertext Markup Language (HTML)*, modeled after, and a subset of, a much more complex formatting language called the *Standard Generalized Markup Language (SGML)*. Today, HTML and SGML have been both essentially merged into a new tool called the *Extensible Markup Language (XML).*

The resulting Web software would be a mixture of Web servers (even a large group of Web servers was now just called a Web site) and Web browsers (the universal client terminology has become the Web browser). The browser is software that runs on the user's client computer and talks to the software running on the Web server to request certain files. These files could be any type representing the various resources a user needs to access. If written in HTML, the files can contain information the browser can display in addition to the names of files of related resources and their locations (the hypertext links).

The World Wide Web was used successfully at CERN by May 1991. The first browser developed was a *line-mode browser* that was really just an advanced version of a Telnet session, as shown in Figure 1.4. The term *line-mode* refers to the line-by-line output of this first browser. Berners-Lee then presented the World Wide Web to the world in December 1991 at the Hypertext '91 conference in San Antonio, Texas, where it caused somewhat of a sensation among the crowd, which included some influential Internet people. In January 1992, the line-mode browser was made available for downloading to anyone with an Internet connection.

Word of the World Wide Web quickly spread and was propagated through Internet discussion groups and conferences. In July 1992, the University of California at Berkeley became the birthplace of the first modern-looking browser. This graphical browser was developed by a postgraduate associate named Pei Wei. The browser, named Viola, was available on UNIX systems using X Windows, which is roughly the UNIX equivalent of Microsoft Windows. It was the first browser to introduce familiar and distinctive browser features such as distinguishing hypertext links by colors and underscoring. Viola's other key feature was the capability to allow simple mouse pointing-and-clicking to activate the hypertext links. But Viola remained only an interesting experiment, because Pei Wei had no interest in developing Viola further as an experiment or as a commercial product.

```
Welcome to the World-Wide Web
THE WORLD-WIDE WEB

For more information, select by number:
A list of available W3 client programs  [1]
Everything about the W3 project  [2]
Places to start exploring  [3]
The first International WWW Conference  [4]

This telnet service is provided by the WWW team
at CERN  [5]
1-5 Up, Quit, or Help
```

Figure 1.4 The original Web "browser."

The Web Explodes

Others, however, were extremely interested in the Web browser potential. They realized that as computing power grew, graphical Web browsers would not need the powerful UNIX operating system and X Windows to perform adequately. Even Macintoshes and Windows-based PCs would do just as well.

And so in February 1993, a new Web browser called Mosaic was released by the National Center for Supercomputing Applications (NCSA), part of the University of Illinois at Urbana-Champaign. Marc Andreessen, a graduate student at the University of Illinois at Urbana-Champaign, with the help of Eric Bina, a programmer at the NCSA, developed the new Web browser. It was originally designed to run in a UNIX environment (with X Windows). But because of its popularity, the Windows and Macintosh versions of Mosaic browser were released in the fall of 1993. This date can truly be thought of as the real beginning of the Web.

Mosaic was not just an innovative Internet application; it was *the* tool, the universal client that opened the portal to the Internet for ordinary people. Once exposed to the Web, many realized that HTML was not a complicated computer language, but rather a relatively simple formatting tool that even those who had previously only used a word processor could master in a short time. Today many word processors allow users to build pages within them and will generate the HTML code so that the users never even have to learn HTML at all. Users just keep using the word processor as they have all along.

Marc Andreessen left the NCSA to cofound a company called Mosaic Communications, which eventually became Netscape Communications Corporation with the release of the Netscape browser, a new redesign of Mosaic. Andreessen jokingly called Netscape a "combination of Godzilla and Mosaic," leading to the merged code name of "Mozilla."

The popularity of the Web caught many by surprise. Microsoft even held off the release of the Windows 95 operating system software so that it could incorporate its Internet Explorer Web browser into the package. To keep ahead of the pack, Netscape offered extensions to the HTML standards to enhance the interaction and function of its Web browser.

What was all the excitement about? Was the Web experience so different that people actually got excited about the Internet, or more specifically, the Web servers attached to the Internet that made up the Web? Well, yes. Without the Web the average PC or Macintosh user people would be unable to go online and buy a car, order a book, obtain hard-to-find forms, pay their bills, bid on goods and services, get software upgrades, find directions to out-of-the-way places, and many other everyday tasks that they now do.

The commercial potential of the Web was apparent very early. Even before the Web browser made its way onto the Macintosh and PC platforms, at least a few people could order pizza over the Internet using the Web.

By the end of 1994 there were several thousand Web sites scattered around the Internet world. You could listen to the Japanese national anthem, download sales information from a few high-tech companies who saw the promising future of this new tool, and (even then) see pictures that no one should show their mother. But there was one Web site that was so distinctive that many people went there just to look and marvel. This was arguably the first commercial Web site: www.pizzahut.com.

Legend has it that the software and hardware engineers working at SCO UNIX (Santa Cruz Operation, a leader in developing UNIX server software) in California had a real hunger for Pizza Hut pizza. Working long and late hours meant missed meals, and sending someone out for pizza meant juggling or stalled tasks while the rest of the teams worked. One of the key projects involved the World Wide Web. So the story goes that SCO got together with Pizza Hut and put a Web server into the local Pizza Hut store.

Now the software team could use their Web browser at SCO to access the www.pizzahut.com Web site by entering that URL (this used to be called "pointing your browser at . . ."). Once there, they filled in name, street address, and voice phone number (apparently, even software types were tempted to order deliveries to vacant lots). Note the use of the term *voice phone number* to prevent confusion with modem links. After a few more entries on other pages, the pizza would arrive. Oddly, this Santa Cruz Pizza Hut page could be accessed, and pizza could be ordered from anywhere in the Internet universe, but deliveries ordered outside of the local area were quite cold if and when they arrived.

Figure 1.5 shows the electronic storefront Pizza Hut Web page as it looked in 1994. Despite its rudimentary appearance today, which would be laughable if not for the page's pioneering status, there are many features that have become indispensable for e-commerce sites. First, note that the company logo is prominently displayed (the original has a glorious and then startling red roof). Note also the careful respect for trademarks (the ® signs) and the use of the company name in the URL. This was the first common use of the term "webmaster" (but with no capital W), and clicking on the link allowed users to send email to the Webmaster. Then there is the form to fill out, a feature that took many Web sites much longer to offer on their own pages. Of course, there is a hypertext link to SCO as well. How many people who went to look at Pizza Hut went to SCO as well? Probably plenty, and this directly lead to the kind of banner and sidebar advertising overkill seen today at many Web sites.

Other business-oriented Web sites were also around in 1994. For example, the Internet Shopping Network, at URL www.internet.net, offered lots of things for sale—among them clothes, kitchenware, and small appliances. But after some initial interest, most of these early efforts failed because of one simple thing: People knew that the Internet offered no security or privacy whatsoever. Anyone with a network monitor or the right software could see everything that went by, especially on the local LANs at each end, which just happened to be where all the clients and servers were.

Sending your credit card information over the Internet without any security features at all was deemed as irresponsible as leaving a credit card on your desk at work while you went to lunch. To be sure, there were people who routinely did both, and they thought nothing of it. These folks were variously called "trusting" or "suckers," depending on who was doing the talking.

For offerings like Pizza Hut, these considerations mattered little. You paid for the pizza on delivery. The stakes were relatively low. But for operations like the Internet Shopping Network, the stakes were potentially higher. How can you sell to people if they won't give you their credit card information over the Internet? You can't push cash money down the wire.

The Internet Shopping Network handled this by establishing a membership code for everyone who wanted to shop. Prospective shoppers had to call up by phone to give their names, addresses, and payment method information before they were allowed to buy at the Web site. Then at the site, they just entered their membership code and shopped till they dropped.

Unfortunately, there were many drawbacks to this plan. It struck many as odd that they had to call by phone to use the network. Why not just call in a catalogue order instead? It was hard to sent gifts also, since the only address shipped to was the member's address on file. There was no allowance for alternate payment methods, either, such as a second credit card. Naturally, all of these issues had to be addressed (and they were) before the Web became more attractive for general business use.

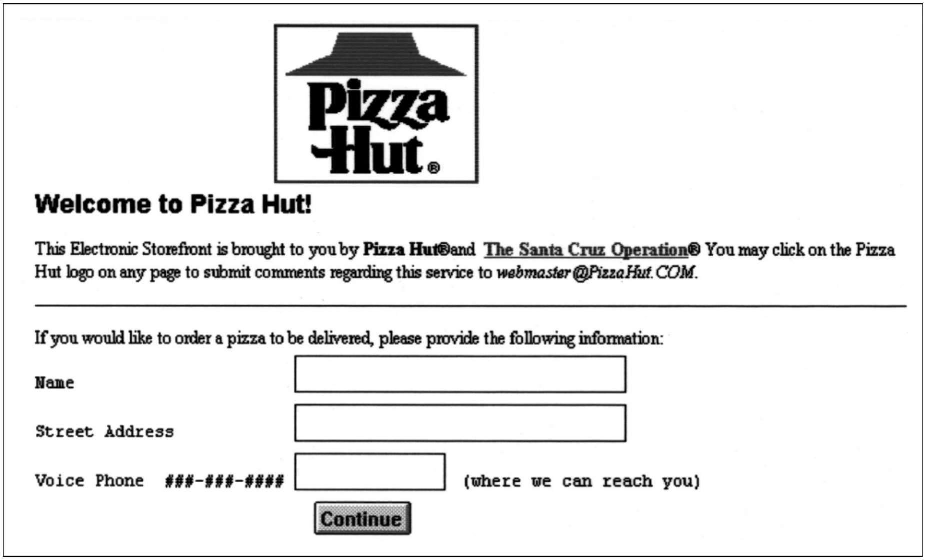

Figure 1.5 Web page from www.pizzahut.com, from 1994.

The Birth of the ISPs

In 1994, as the Web was having a huge impact on what the Internet meant to people, the federal support dollars for the NSFNET were phased out as planned. In 1993, the NSF deemed the then-experimental high-speed NSFNET network begun in 1984 a success, but stated that it was not the type of research network needed for the future. The NSF invested in a newer study of high-speed computing and networking and announced that "the Internet" would have to eventually fund itself commercially without government subsidies. The NSF continued funding for portions of the Internet, cutting funds 20 percent per year from 1994 until 1998 when they disappeared entirely.

The idea that the U.S. government should provide taxpayer dollars to encourage new technologies that might improve the life of citizens is outlined in the Constitution. The Internet did not set a precedent in this regard. For example, Samuel F. B. Morse, inventor of the telegraph, was an art professor, not a professional inventor (although he did explore more than the telegraph in this regard). Federal money was used to run the first telegraph line between Washington, DC, and Baltimore, Maryland. As a result, initial use of this line was free, but by the time the money ran out, it was obvious that the people had embraced the telegraph and money from users would support the new industry.

By the 1990s, it was also obvious that the Internet formed an important part of university and college life. Withdrawing the NSFNET subsidy just meant that funds for continued operations had to be gathered from the users of the network. To face this challenge, the Internet was restructured into a small number of network access points (NAPs) run for the most part by major telephone companies. Any part of the Internet could reach any other part of the Internet through these NAPs if the separate parts did not have direct peer connections to each other.

In addition to the NAPs, the Internet also came to include other types of peering points. The most important of these were the Commercial Internet Exchange (CIX) and the Federal Internet Exchange (FIX). The CIX and FIX were independent of the NAPs but also provide places for the ISPs that link to them to exchange traffic.

The history of the CIX is brief. In the days of the acceptable use policy, the Internet was supposed to be for "research and education" only. Use of the free Internet to make money was forbidden. As companies began to use the Internet for sales, marketing, and more, the line between what was acceptable and unacceptable became blurred. Is sending a sales brochure blatant commercial use of the Internet or still educational ("educating" people about the product?). Maybe to be safe, a company should send such commercial information over separate links and keep this traffic off of the global public Internet whenever possible.

So the CIX was formed to allow companies and service providers to send commercial content back and forth without worrying about violating the Internet spirit of research and education. ISPs or companies could exchange traffic at one of several CIX sites on the Internet, as long as they maintained links there, of course. CIX was disbanded in January 2002, its purpose obsolete in these days of the wide-open Internet. Today, going to www.cix.org (the old Web site) pops up the page for the U.S. Internet Service Providers Association. The new emphasis of the organization is on privacy, content restrictions, and intellectual property.

U.S. government networks exchanged traffic at two FIX points located on the east and west coasts of the United States. Government agencies such as NASA or the Department of Energy used these peering points in a fashion similar to the early days of the Internet. Today, peering points like CIX and FIX are mainly of historical interest.

Of course, any network entity providing Internet service (really just access to servers on the Internet) had to pay for the lines running to the NAPs and other charges as well. How would they get the money needed to pay for NAP connections? These organizations in many cases reached out beyond the university and college community and went public. Now calling themselves Internet service providers (ISPs), the smaller, regional portions of the Internet began advertising and signing up as many people as they could for Internet access. Some organizations were so successful that they grew into national ISPs very quickly, while others either accepted their smaller, regional role or at least were content with it.

It helped that there was already a precedent for such online services. In the 1980s, the availability of PCs and the appearance of these PCs in people's homes, as well as on their desktops at work, led to the appearance of various specialized networking companies or bulletin board services.

The popularity of the IBM PC in the early 1980s caused a revolution not only in business computer use but also in home computer use. For the first time, ordinary people at home had access to the enormous computing power previously available only to a select few with home computer terminals. And these select few had no real computing power in their homes. The computing power was still in the office mainframe or minicomputer; the terminal only provided access to the office computer over a telephone line. A person with a home PC could run applications and programs such as spreadsheets and word processors formerly restricted to a corporate environment. And when it came to word processing, spreadsheets, and even simple graphics, the humble PC often outperformed many mainframes and minicomputers, because of the rapid advances in technology. All that was missing was a network to tie together all of these new PC users, isolated in their homes in front of their screens.

By the mid-1980s, the PC community had a way to network their isolated PCs together. It appeared in the form of a variety of online services, some large, some small. The smaller ones consisted of a single PC running special "server" software similar in nature to Internet server software, but not based on the same protocols in most cases. A few telephone lines connected to modems allowed others to simply dial in with their terminal emulation software packages to become *clients*. This simple but powerful arrangement launched a cottage industry of bulletin board systems (BBSs) that lasted well into the 1990s and that still linger in some portions of the networking world. These small electronic bulletin boards usually addressed the needs of a small, contained group (a factory workers union local, for example) and offered meager features, such as email among users, as well as a common point to post files and messages. Many BBSs were staffed on a voluntary basis, and there was no charge to users beyond telephone calling time charges, which were usually minimal, except for users outside the local calling area.

The larger systems addressed the needs of the public at large for such services. Generally, the services were much the same, with the addition of some special user groups for those with shared interests, such as military history or houseplants. These larger online services benefited from special rules that the federal government laid down in the United States in 1984, specifically to encourage the growth of such services. The larger services had many points of presence (POPs) throughout the United States so that local users did not have to constantly dial long-distance numbers and pay high rates and the consequent high bills for access to these systems. Membership was typically based on a flat monthly rate and usage charges, although some offered totally flat-rate services, which users embraced wholeheartedly. Naturally, these online service providers were concentrated in major metropolitan areas, but there were some in almost every state in one place or another.

By the time the Web exploded onto the world in 1993 and 1994, the world of the large online service providers was filled by three companies: CompuServe (CSi), America Online (AOL), and Prodigy, arranged in order of their first public service offerings.

CompuServe began in 1982, virtually at the beginning of the PC industry itself. CompuServe quickly gained a reputation as the service of choice for technology-oriented PC enthusiasts. The more users cared about things such as operating system optimization, writing their own utility programs, or adding their own hard drive to their system, the more the users liked CompuServe. Oddly, for all its reputed sophistication, CompuServe was one of the last to realize that the Internet and Web were important to its users and lagged behind in offering easy and integrated Internet and Web access to its members.

America Online began offering services to the public at large in 1985. Their reputation was that their members did not need a great deal of PC sophistication or have to be a computer science wizard to access their various chat rooms and sample their offerings. AOL's aggressive marketing techniques turned some people off but certainly raised the consciousness of average PC users who at least tried the online service to see whether they would like it. Many did. AOL was among the first to see the Internet and Web as an adjunct service to their bulletin-board-based system, but not as a replacement to their "keyword" services.

Prodigy began as a joint effort between IBM and Sears in 1984 to encourage PC users to communicate online. Public service was launched in the United States in 1988, and Prodigy seemed determined to "out-AOL AOL" in terms of user-friendliness and services, such as home shopping and electronic newspapers. However, Prodigy alienated some with insistent advertising from sponsors (virtually absent on other service providers' systems) and suffered from some early negative publicity. While Prodigy saw that the long-term future of networking for ordinary people would revolve around advertising, sales, and marketing, CompuServe and AOL people at the time tended to treat these key aspects of the Internet today with disdain.

By the early 1990s, PC users in the United States were comfortable with the idea that an online service could put them in touch with other users with similar interests and offer them access to information that would otherwise be difficult to obtain. When the ISPs were faced with a loss of government funds, many of them saw a solution in simple repositioning. Rather than provide Internet access to colleges and universities, why not offer Internet access to anyone willing to pay the price? This involved installing TCP/IP client software on the user's PC, but software installation by this time was not the headache it once had been, and Windows 95 solved that problem by integrating TCP/IP into the operating system.

In the mid-1990s, three separate threads came together to encourage the expansion of Internet and Web access in the United States. The first was the idea that anyone with a PC could benefit from networking with other PC users through an online service. The second thread was that the ISPs were faced with having to make their operations financially self-supporting. The last thread was the Web itself, with its colorful graphics and multimedia in place of the stodgy commands of the recent past. The big online service providers, especially the Big Three of CSi, AOL, and Prodigy, were not slow to react. When the Web broke onto the world, Prodigy became the first of the Big Three to offer a Web browser as an integral part of its client software package in January of 1995. AOL quickly followed and went Prodigy one better. In 1995, AOL went out and bought the assets of ANS, the principal operations arm of the NSFNET Internet backbone. ANS got money and AOL got Web access for their users. CompuServe soon also realized that their members might want access to the Internet and the Web, but many had already decided not to wait and signed up with AOL or another ISP.

Today, there are still thousands of mostly small ISPs around the world. Only about 20 or 30 have an extensive national backbone of their own (the uncertainty is due to the lack of a standard definition of "national," "backbone," and "own" when it comes to ISPs). The others basically simply link to other ISPs, which might link to other ISPs, and eventually everything is linked together at the major NAPs. So the Internet is not one network but more like thousands of interconnected ISP networks, each with a large or small number of Web servers (Web sites) maintained by their individual members (customers) or members' organizations.

The Router's Role

So far, all well and good. The Internet took off when the Web hit town. The ISPs were right there to provide Internet (and Web) access for a steadily falling price. What has all this to do with routers, routing protocols, and routing policies?

As already mentioned, once the number of routes and routers reached a critical point, it was impossible for all routers to know all of the details of the Internet. So the Internet today consists of an increasing number of *routing domains*. Each routing domain has its own internal and external routing policies. The sizes of routing domains vary greatly, from only one IP address space to thousands. Each domain is called an autonomous system (AS). In many cases, an ISP has only one AS, but national or global ISPs might have several. For example, a global ISP might have one AS for North America, another for Europe, and a third for the rest of the world. Each AS must have a uniquely assigned AS number, although there can be various, logical "sub-ASs" called *confederations* or *subconfederations* (both terms are used) inside a single AS.

An AS forms a group of IP networks sharing a unified *routing policy framework*. A routing policy framework is a series of rules and guidelines used by the ISP to formulate the actual routing policies that are configured on the routers. Between different ASs, which are often administered by different ISPs, things are trickier. Careful coordination of routing frameworks and routing policies between ASs is needed to communicate complicated policies between different ASs.

Routes not only need to be advertised to another AS, but need to be accepted by the receiving AS. The decision on which routes to advertise and which routes to accept is determined by routing policy, of course. The situation is summarized in Figure 1.6, which shows the extremely simple exchange of routing information between two ASs that are peers. The routing information is transferred by the routing protocol running between the routers, usually BGP (Border Gateway Protocol, another reflection of the router as gateway).

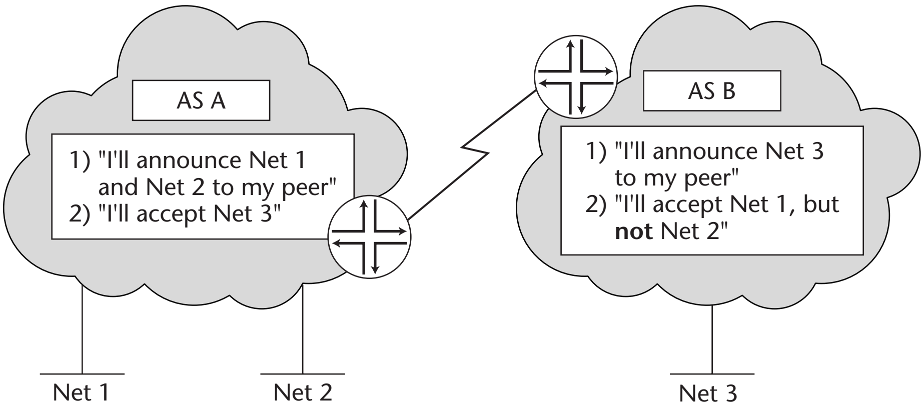

Figure 1.6 The effects of routing policy.

Usually, the exchange of routing information is bidirectional, but this is not always a given. In some cases, the routing policy might completely stifle the flow of routing information in one direction, either because of the routing policy of the sender (suppress the advertising of a route or routes) or the receiver (ignore the routing information from the sender). If routing information is not sent or accepted between ASs, peers or not, then hosts (clients or servers) in one AS cannot reach other hosts on the networks represented by that routing information in the other AS. This situation is shown in Figure 1.6 as well.

Financial considerations often play a role in routing policies. In the "old days" of federal subsidies this was not much of an issue, and there were always grants available for continuing support for the research and educational network. Now the ISP grid net has raised issues as ISPs installed POPs in many regions and countries. ISPs can have their own customers, but they can also be customers of other ISPs as well. Who pays whom, and how much?

Telephony solved this problem with a concept called *settlements*, where one telephone company usually bills the call originator and shares a portion of the amount with other telephone companies as an *access charge*. Access charges compensate the other telephone companies that carry the call for the loss of the use of their own facilities for the duration of the call. On the Internet, the issue becomes how should one ISP compensate (if at all) another ISP for delivering packets that originate on the other ISP. The issue is complicated by the fact that the call is now a stream of packets, and an ISP might just be a transit ISP for packets originating in one ISP's AS and destined for a third ISP's AS.

Nevertheless, ISP peers have tried three ways to translate the telephony settlements model to the Internet. First, there are very popular bilateral (between two parties) settlements based on the "call," usually defined as some aspect of IP packet flows. In this case, the first ISP, where the packet originates at a client, gets all of the revenue from the customer, but the first ISP shares some of this money with the other ISP where the server is located. Second, there is the idea of *sender keeps all* (SKA), where the flow of packets from client to server one way is presumably balanced by the flow of packets from client to server the other way, so each might as well just keep all of the customer revenue. Finally, there is the idea of *transit fees*, which are just settlements between one ISP to another, usually from smaller ISPs to larger (since this traffic flow is seldom symmetrical).

None of these methods have worked out well on the Internet. There are often many more than just two or three ISPs involved between client and server. There is no easy way to track and account for the packets that should constitute a call, and even TCP sessions leave a lot to be desired because a simple Web page load might involved many rapid TCP connections between client and server. It is often hard to determine the "origin," and packets do not follow stable network paths. Packets are often dropped, and it seems hardly fair to bill someone else for re-sent packets replacing those that might have been counted but not delivered in the first place. Finally, dynamic routing might not be symmetric: so-called hot-potato routing seeks to pass packets off to another ISP as soon as possible. So the path from client to server might pass through different ISPs than server-to-client replies.

The drawbacks of the telephony settlements model have resulted in a movement to more simplistic arrangements between ISP peers, which just means ISPs of roughly equal size. These are often called *peering arrangements* or just *peering*. But another issue has arisen because there is no strict definition of what a peer is or is not. In this book, *peering* will just mean that two ISPs are directly connected and have instituted some routing policies between them. Financially, there is often also a sender-keeps-all arrangement in place, so no money changes hands. However, this applies only to peers. An ISP that is not a peer of the others is in a sense just a *customer* of the other ISPs. And customers must pay for services rendered. Since there are typically no financial arrangements for peer ISPs providing transit services to a third peer, however, peer ISPs do not provide transit to a third peer ISP (unless, of course, the third peer ISP is willing to pay and become a customer of one of the other ISPs). This situation is shown in Figure 1.7.

All three ISPs are peers in the sense that they are roughly equal in terms of network resources. They might all be small or regional or national ISPs, but the point is the same. ISP A is peering with ISP B and ISP B is peering with ISP C, but ISP A has no peering arrangement (or direct link) with ISP C. So packet deliveries from hosts in ISP A to ISP B (and back) are allowed, as are packet deliveries from hosts in ISP C to and from ISP B. But ISP B has routing policies in place to prevent transit traffic from ISP A to and from ISP C through ISP B. Unless ISP A and ISP C are willing to peer with each other, or ISP A or ISP C is willing to become a customer of ISP B, there will be no routing information sent to ISP A or ISP B to allow these ISPs to reach each other through ISP B. The routing policies enforced on the routers in ISP B will make sure of this.

The ISP grid net, without a clearly defined hierarchy of natural peers, complicates this simplistic application of peering drastically. Peering is often a political issue. The politics of peering more or less began in 1997, when a large ISP informed about 15 other ISPs that their current easy-going peering arrangements would be terminated. New agreements for transit traffic were now required, the ISP said, and the former peers were effectively transformed into customers. As the trend spread among the larger ISPs, direct connections were favored over public peering points such as the NAPs or CIX. The ISPs with multiple, high-speed links to NAPs (Tier 1 ISPs) peered only with each other.

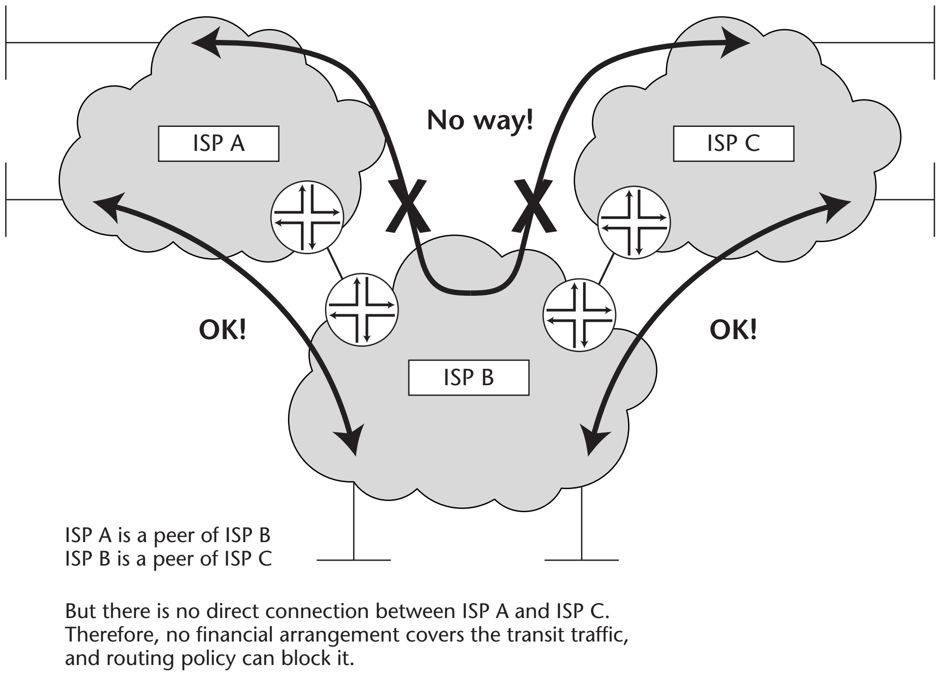

Figure 1.7 Peer ISPs do not provide free transit services.

Naturally, no ISP wants to be a customer of another ISP. All ISPs want to be peers, and peers of the biggest ISPs around. When it comes to peering, bigger is definitely better, so a series of mergers and acquisitions (although it is often claimed that there are really no mergers, only acquisitions) among the ISPs took place as each ISP sought to become a bigger peer than another. This consolidation has decreased the number of Tier 1 ISPs and reduced the number of potential peers considerably.

Today, potential partners for peering arrangements are closely scrutinized in several key areas. ISPs being considered for potential peerage must have high-capacity backbones, be of roughly the same size, cover key areas, have a good network operations center (NOC), have about the same quality of service (QoS) in terms of delay and dropped packets, and most importantly, exchange traffic roughly symmetrically. Because the work of the Internet routers is done delivering packets, nobody wants to peer with an ISP that supplies 10,000 packets for every 1,000 packets it accepts. Since servers, especially Web sites, tend to generate much more traffic than they consume, ISPs with tight networks with many server farms or Web hosting sites often have a hard time peering with anyone. This situation is shown in Figure 1.8.

ISPs without peering arrangements must rely on public exchange and peering points like the NAPs or (formerly) CIX for global connectivity. However, these exchange points are usually congested, which led to the rise of peering in the first place. And routing policies are still in place at the NAPs. A special device called the *route server* applies the routing policies in place between the ISPs that come together at these public peering points.

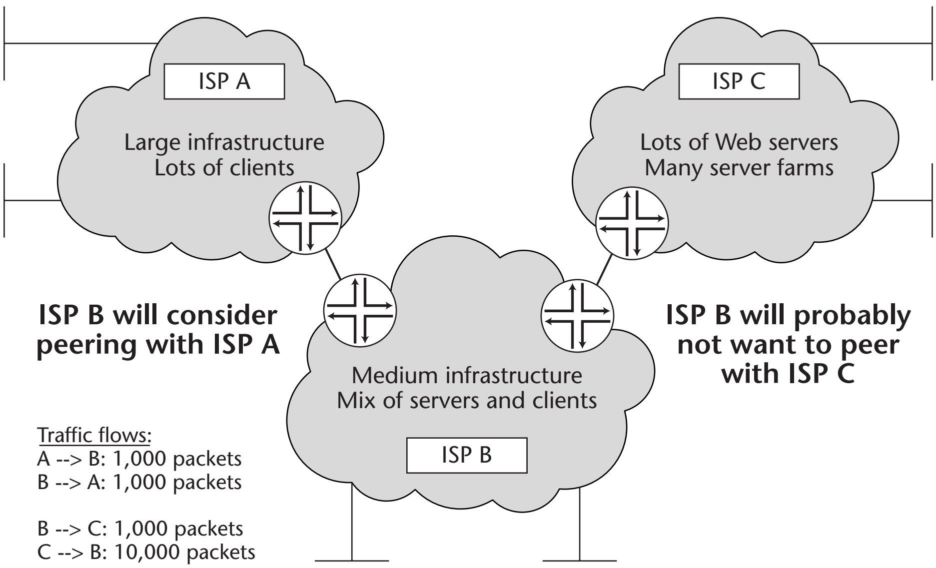

Figure 1.8 Good peer candidates and bad peer candidates.

Public peering arrangements are usually multilateral and applied to all ISPs that connect there, although there are bilateral arrangements in place here as well. The trend today is more toward private peering between pairs of peers. Exchanges are often overwhelmed by traffic, and it is almost impossible to enforce any QoS mechanisms in place on an ISP's network such as differentiated services for voice and video through a public exchange point.

Private peering can be done by simply installing a WAN link between routers at the AS borders of the two ISPs. Alternatively, peering can be done at a *collocation* site where the two peer's routers basically sit side by side. Both types of private peering are common.

The whole point of this chapter is that the Web explosion has created a need for new routing policies between peer ISPs. The Internet today has about 25 percent more routes (about 100,000) than there were *computers* attached to the Internet at the beginning of 1989 (about 80,000). Routing policy is necessary whether the peering is public or private, through a NAP or through a WAN link between routers. Routing information simply cannot be easily distributed everywhere all at once. Even the routing protocols play a role. When the initial NSFNET backbone was installed, only the Routing Information Protocol (RIP) routing protocol was mature enough to run on the backbone. Fully 80 percent of the traffic on the new NSFNET backbone was RIP routing updates, and so the quest for newer routing protocols such as Open Shortest Path First (OSPF) and Intermediate System-Intermediate System (IS-IS) began in earnest. Routing policies help IGPs such as OSPF and IS-IS distribute routing information within an AS more efficiently.

The Internet *must* be segmented into routing domains (the AS), and the flow of routing information between those domains must be controlled by routing policies to enforce the public or private peering arrangements in place between ISPs. Driven mainly by the use of the Web for commercial purposes, the Internet in 1990 was only 3 percent of the size of the Internet in 1996. The goal of this book is to detail the operation of the routing protocols and the creation of routing policies to enforce the aims of internal route distribution and peering between ASs to bring some order into this jungle of connections that is the Internet.