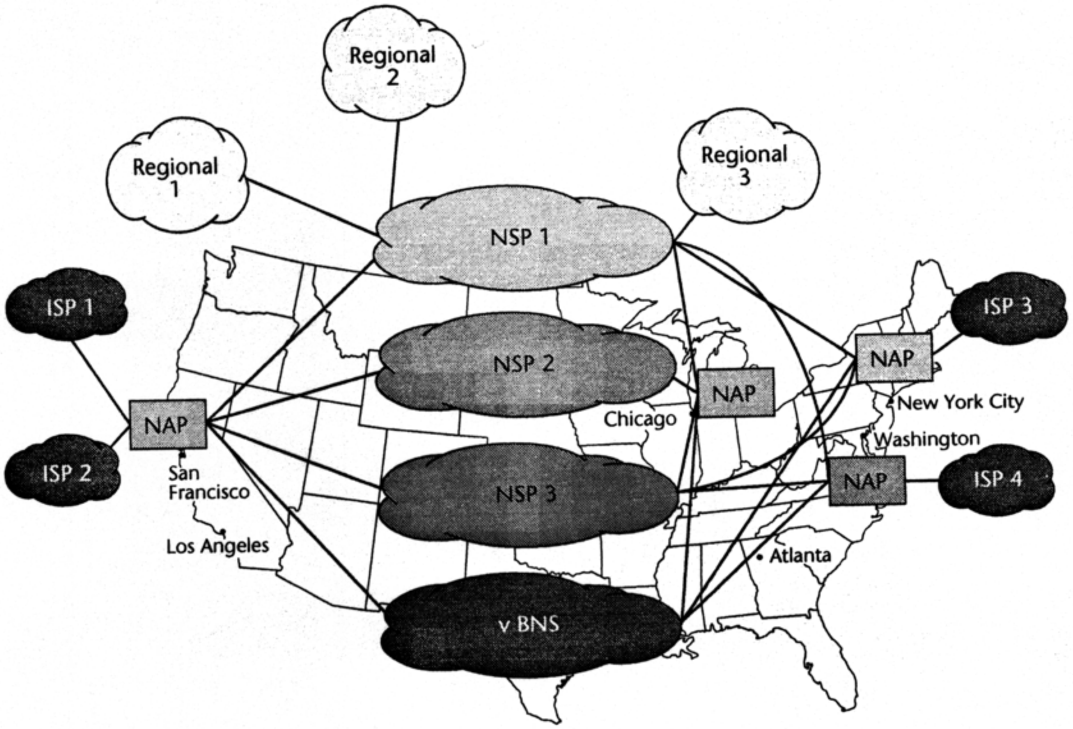

Figure 14-1: The architecture of the Internet, according to NSF Solicitation 93-52.

14.1.2 Network Access Points

Solicitation 93-52 sought to create at least three network access points, or NAPs, located in the San Francisco Bay area, the Chicago area, and the New York City area, as well as additional NAPs to the extent that funding might permit. The NAPs were expected to operate at the data link layer of the OSI Reference Model; they might be implemented as a shared high-speed LAN or as an SMDS or ATM service providing shared interconnection among NSPs/ISPs.

Each NAP would provide interconnection among NSPs and ISPs for purposes of traffic interchange—known in the trade as *shared interconnection,* and sometimes referred to as *public peering.* In an *interconnection,* or *peering,* relationship, Internet providers agree to accept traffic destined for one another's respective customers but not necessarily for other third parties. In a *customer,* or *transit,* relationship between Internet providers, by contrast, the transit provider typically agrees to accept traffic destined for any point in the global Internet.

The solicitation resulted in the NSF's supporting four NAPs. Those NAPs, along with their sponsoring firms, are shown in Table 14-1.

Table 14-1: Network Access Points and Sponsoring Firms

NAP	Sponsoring Firms
Chicago	Ameritech/BellCore
New York[1]	Sprint/San Diego Supercomputer Center
San Francisco	PacTel/BellCore
Washington, D.C.	Metropolitan Fiber Systems (MFS), in a facility called MAE East

1. New York is something of a misnomer: Sprint responded to the NSF's request with a proposal to locate its NAP in a facility situated some 90 miles southwest of New York, across the Delaware River from Philadelphia.

In the flurry of acquisitions that followed the Telecommunications Act of 1996, the identity of a number of these organizations is changing or has changed. BellCore has been acquired by SAIC; PacTel has been acquired by SBC; Ameritech is in the process of being acquired by SBC; and MFS was acquired by WorldCom, which subsequently became MCI WorldCom as a result of its acquisition of MCI.

The NSF's program was basically successful, even though things did not work out exactly as initially envisioned. The Washington NAP, better known as MAE East, and the Sprint NAP evolved into full-fledged national and global traffic interchange points; the Chicago and San Francisco NAPs, however, have been problematic since their inception. Initially, they suffered from technical problems associated with attempting to drive large volumes of IP traffic over still immature ATM switching products; subsequently, these NAPs have lacked a critical mass of large national backbone ISPs. As a result, they function, in practice, in the role of regional traffic concentration points. Quite a few additional regional public interchange points have come into existence in the United States in recent years, although they do not appear to play a strong role in the overall traffic flows of the Internet.

Meanwhile, the lack of a stable public peering point on the West Coast of the United States was, briefly, problematic for the NSP/ISP community. In practice, the large national backbone ISPs soon converged on a facility called MAE West, operated jointly by MFS and NASA.

As a result, the work originally envisioned for the NAPs can be viewed, for most purposes, as being performed by MAE East, MAE West, and the Sprint NAP. All three of these were based in the recent past on Fiber Distributed Data Interface (FDDI) LAN technology, and all three augmented the FDDI with high-speed LAN switches (gigaswitches). The gigaswitches support full-duplex FDDI, thus offering, in theory, 100Mbps of input and output simultaneously between each pair of routers.

By 1996, all of these facilities were suffering from inadequate capacity. In 1998, MCI WorldCom upgraded its MAE facilities in Washington (MAE East), San Jose (MAE West), and Dallas to offer modern ATM switches as a high-capacity alternative to the FDDI/gigaswitch architecture.

As we shall soon see, a largely unanticipated consequence of the evolution of the Internet away from the pure NSF 93-52 model has been a migration of interconnection traffic away from these three locations and into private arrangements among the NSPs. These *direct interconnections* are sometimes referred to as *private peering.*

14.1.3 Transit Service among Regional ISPs

Eight large regional ISPs went through a competitive bidding process, which resulted in a series of contracts for *transit* service—where transit represents carrying of data to other ISPs—being awarded to MCI. MCI connected the NSF-sponsored regional ISPs to one another and carried their traffic to other NSPs and to the NAPs. The arrangement worked well, in general, but it has largely been phased out today. The original regional ISPs have outgrown their need for transit services; however, most backbone ISPs now offer transit services, which continue to play a huge role in the public Internet today.

14.1.4 The Very High Speed Backbone Network Service (vBNS)

The NSF also awarded the vBNS to MCI. MCI runs routers over an infrastructure of ATM switches to interconnect a number of NSF-sponsored research institutions, including the supercomputer centers at Cornell University, Pittsburgh, San Diego, NCAR, and NCSA. The vBNS was initially operated at OC-3 speeds (155Mbps) and was subsequently upgraded to OC-12 (622Mbps).

The program has generally worked as intended; nonetheless, there are issues. First, it serves only a portion of the institutions to which connectivity is offered. (Thus, it is not available to all of Cornell University). Second, the original intent of using the vBNS as a testbed for research into high-speed internetworking may be inappropriate for a production network used to accomplish real work.

More recently, some of the functions of the vBNS have been subsumed by two newer initiatives: the *Next-Generation Internet (NGI),* and the *Abilene* project of *Internet* 2. The NGI is a U.S. government–sponsored initiative to provide a high-speed private internetwork among a number of major U.S.

agencies. Abilene is a high-speed private internetwork interconnecting members of UCAID, a consortium of research universities and private industry. Abilene has benefited from significant "in-kind" contributions of circuits and equipment from Qwest, Cisco, and Nortel.

Meanwhile, the original vBNS contract will soon expire. MCI WorldCom intends to replace it with a privatized next-generation vBNS network based on packet-over-SONET at OC-48 speeds.

14.1.5 The Router Arbiter (RA) Project

The Router Arbiter project was intended to conduct research into routing and to provide a database of Internet topology and policies, in order to enhance the stability, robustness, and manageability of the Internet. The Router Arbiter project also produced statistics about the Internet as a whole (see Section 14.6).

14.2 Structure of the Internet Today

Traffic interchange among backbone ISPs is fundamental to the operation of this system. As previously noted, in an *interconnection* relationship, Internet providers agree to accept traffic destined for one another's respective customers but not necessarily for other third parties. This is different from a *customer,* or *transit,* relationship between Internet providers, where the transit provider typically agrees to accept traffic destined for any point in the global Internet.

Historically, interconnection was called *peering,* in order to imply that traffic was interchanged among providers that were similar in size and capability. Over time, it came to be recognized that peers need not be similar in size; rather, what was important was that there be comparable value in the *traffic interchanged.*

I find it convenient to think of today's Internet as comprising two kinds of ISP: backbone ISPs and all others. The backbone ISPs interconnect with all other significant ISPs by means of a full set of interconnection relationships. Other ISPs may have some interconnection relationships, or they may not, but they have a significant dependency on a customer or transit relationship to one or more backbone ISPs. (Thus, the ability to reach all Internet destinations without the need for a transit relationship—sometimes called, somewhat inexactly, *default-free* status—is a strong indicator that an ISP should be viewed as a backbone ISP.) This yields the complex, Medusa-like structure shown in Figure 14-2.

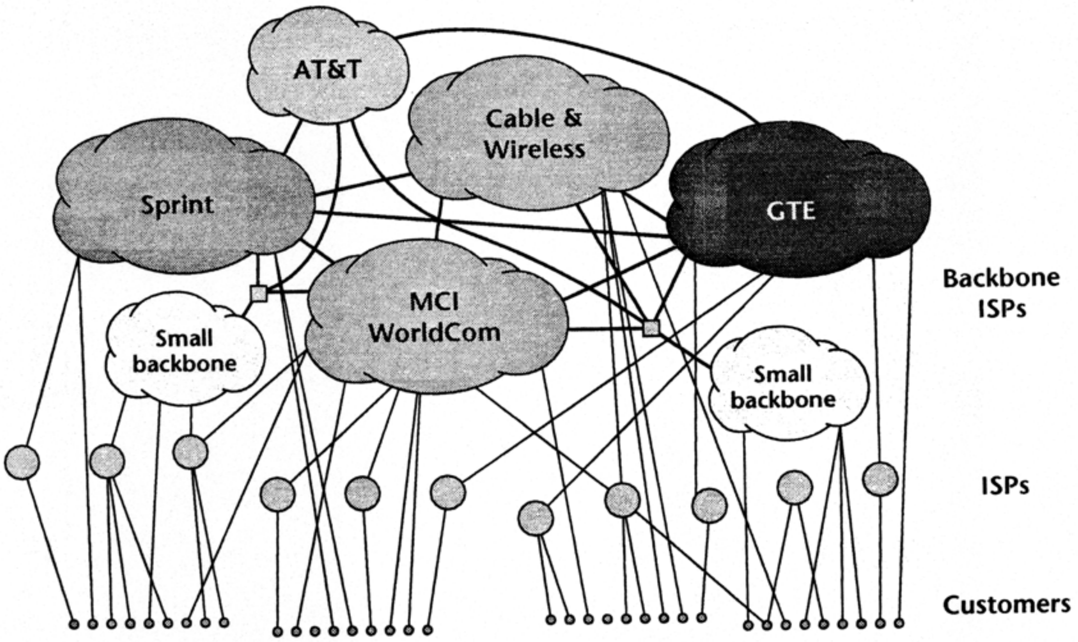

Figure 14-2: Present-day Internet structure: Backbone ISPs and other ISPs.

By these criteria, Cable & Wireless (formerly internetMCI), MCI WorldCom (including UUnet, ANS, and other previously independent ISPs), Sprint, AT&T, and GTE (including the former BBN and Genuity) should clearly be viewed today as backbone ISPs. Somewhere between six and perhaps thirty other ISPs could also be viewed as backbone ISPs. The vast remainder are dependent on the backbone ISPs for global interconnectivity.

14.3 Direct and Shared Interconnections (Public and Private Peering)

Over the past few years, there has been a marked trend away from the use of the shared interconnection points on the part of the large backbone ISPs. In absolute terms, there is still significant traffic growth at MAE East, MAE West, and the Sprint NAP; however, the growth is not commensurate with the growth in overall traffic in the Internet. Thus, these facilities are losing "market share" in terms of the number of bits that flow through them.

The trade press has been perplexed by this trend and has occasionally presented it as if it were a predatory tactic on the part of the large backbone ISPs. In fact, simple economic and technical considerations drive the move to direct interconnections (private peering). For that matter, the migration to private peering is, for the most part, neutral in its impact on smaller ISPs—which indeed have severe problems with the affordability of interconnections but not because of the upsurge in direct interconnections among large backbone ISPs.

Even our way of thinking about shared interconnection has evolved. Historically, we referred to the NAPs and MAEs as *public peering points.* This was something of a misnomer, as, in almost all instances, peering was not public! Interconnection (peering) at a shared interconnection point would be established as a bilateral business relationship between two backbone ISPs, often at no direct cost to either party, based on the shared perception that both parties benefited from that interconnection.

In 1993, it might have seemed that the NAP-based architecture shown in Figure 14-1 could be expanded indefinitely. As each component was outgrown, it would simply be upgraded. When the NAP shoe started to chafe, the NAP provider might simply buy a shoe of larger size.

Things have not played out quite the way that they were expected to. First, LAN technology did not keep pace with the growth in Internet traffic levels. As traffic at the three major interconnects grew, the shared-medium FDDI (with 100Mbps of bandwidth to *all* ISPs present) was upgraded several times but never seemed to be able to keep pace with traffic. More recently, the largest MAE facilities have migrated to ATM; nonetheless, the main reason that the Internet as a whole has not long since collapsed under its own weight is that the major backbone ISPs have diverted most of their traffic away from the shared interconnection points.

The second main factor in this migration is that the economics of interconnection work differently as the bandwidth ramps up. Under the original concept of a NAP, a single T-3 connection from each NSP would carry all of the peering traffic for that NSP for a major portion of the United States. The NAPs made economic sense because they allowed multiple peers to share resources. They also provided economies of scale, because operating a single shared T-3 circuit was far more cost-effective than operating one or more T-1 circuits to each of several peers.

Today, if a given backbone ISP has enough traffic to another backbone ISP to fill a T-3 circuit, there is no incentive to use a shared interconnect for that traffic. Indeed, in light of current Internet traffic levels, use of a public interconnect

- Would introduce the risk of overloading the shared interconnect, to the detriment of both backbone ISPs and possibly also of third parties

- Offers no additional economic incentives in the form of economies of scale
- Potentially complicates future upgrades, in that three parties are involved (two backbone ISPs and the shared interconnect facility manager) instead of two (just the two backbone ISPs)

For a large backbone ISP, the natural tendency is to use

- Private interconnects to those backbone ISPs with which one has a lot of traffic to interchange
- Shared interconnects to interchange traffic with smaller backbone ISPs and, possibly, with small ISPs

This yields a system that looks like Figure 14-3. The majority of interconnections among providers may continue to take place at shared interconnection points; however, the preponderance of traffic already flows across private peering interfaces, and this tendency is sure to accelerate in the coming years.

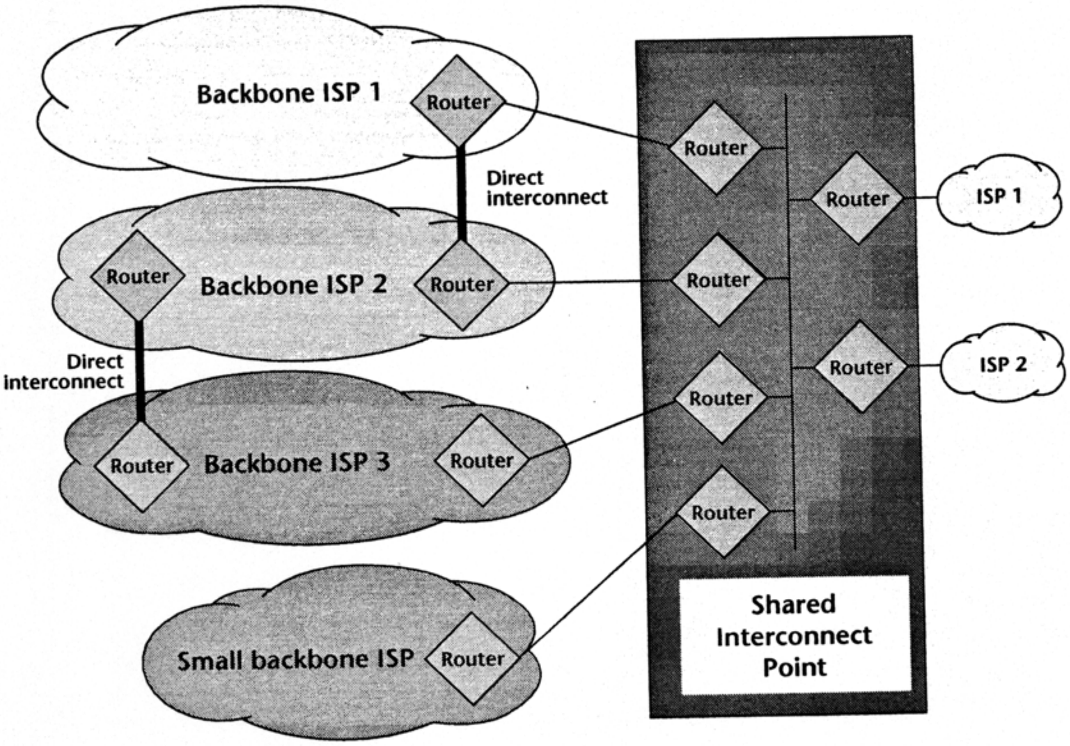

Figure 14-3: Direct and shared interconnections (private and public peering).