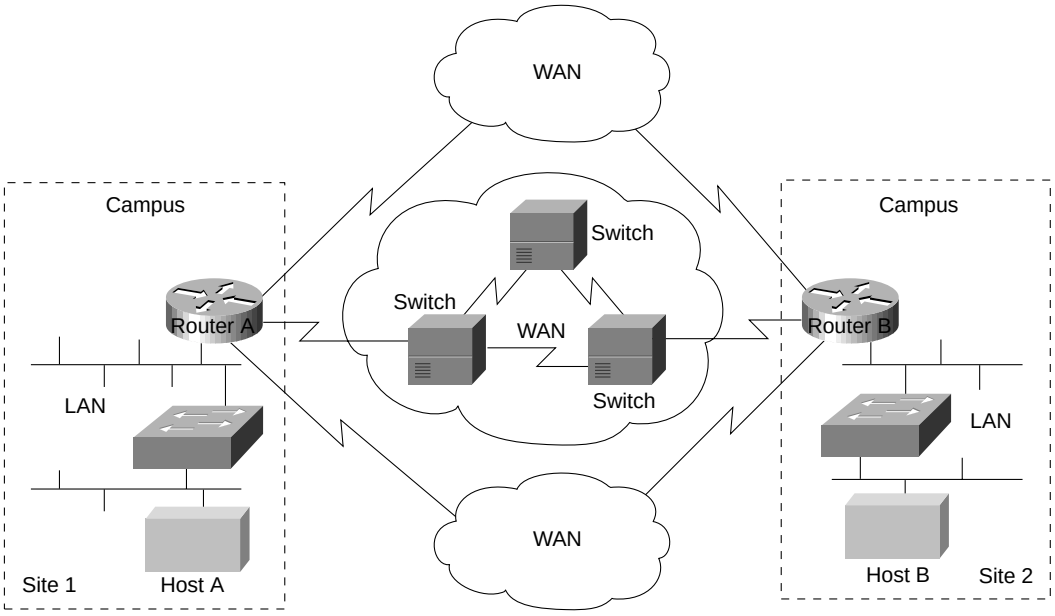
Introduction

Internetworking—the communication between two or more networks—encompasses every aspect of connecting computers together. Internetworks have grown to support vastly disparate end-system communication requirements. An internetwork requires many protocols and features to permit scalability and manageability without constant manual intervention. Large internetworks can consist of the following three distinct components:

- Campus networks, which consist of locally connected users in a building or group of buildings
- Wide-area networks (WANs), which connect campuses together
- Remote connections, which link branch offices and single users (mobile users and/or telecommuters) to a local campus or the Internet

Figure 1-1 provides an example of a typical enterprise internetwork.

Figure 1-1 Example of a typical enterprise internetwork.



Designing an internetwork can be a challenging task. To design reliable, scalable internetworks, network designers must realize that each of the three major components of an internetwork have distinct design requirements. An internetwork that consists of only 50 meshed routing nodes can pose complex problems that lead to unpredictable results. Attempting to optimize internetworks that feature thousands of nodes can pose even more complex problems.

Despite improvements in equipment performance and media capabilities, internetwork design is becoming more difficult. The trend is toward increasingly complex environments involving multiple media, multiple protocols, and interconnection to networks outside any single organization's dominion of control. Carefully designing internetworks can reduce the hardships associated with growth as a networking environment evolves.

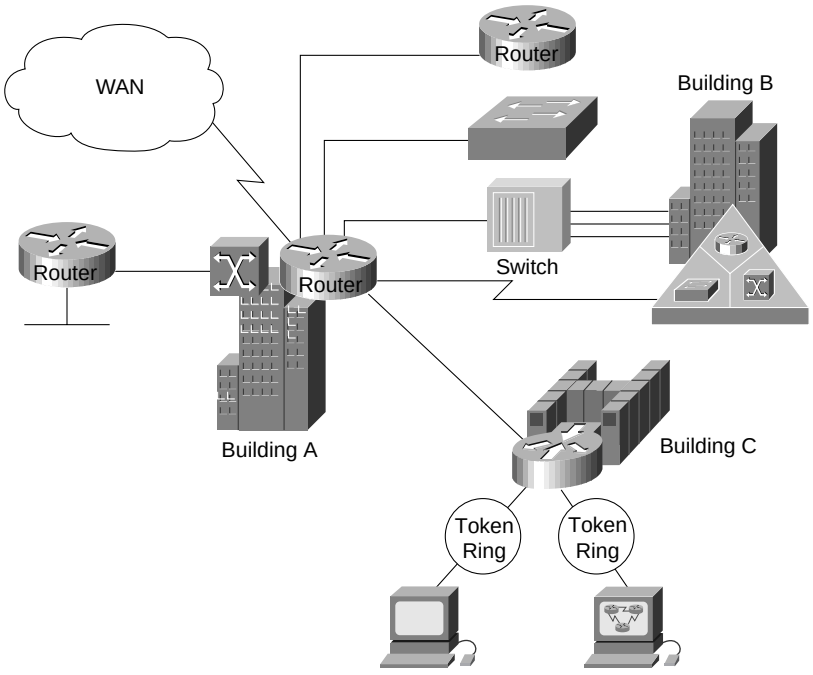
This chapter provides an overview of the technologies available today to design internetworks. Discussions are divided into the following general topics:

- Designing Campus Networks
- Designing WANs
- Utilizing Remote Connection Design
- Providing Integrated Solutions
- Determining Your Internetworking Requirements

Designing Campus Networks

A *campus* is a building or group of buildings all connected into one enterprise network that consists of many local area networks (LANs). A campus is generally a portion of a company (or the whole company) constrained to a fixed geographic area, as shown in Figure 1-2.

Figure 1-2 Example of a campus network.



The distinct characteristic of a campus environment is that the company that owns the campus network usually owns the physical wires deployed in the campus. The campus network topology is primarily LAN technology connecting all the end systems within the building. Campus networks generally use LAN technologies, such as Ethernet, Token Ring, Fiber Distributed Data Interface (FDDI), Fast Ethernet, Gigabit Ethernet, and Asynchronous Transfer Mode (ATM).

A large campus with groups of buildings can also use WAN technology to connect the buildings. Although the wiring and protocols of a campus might be based on WAN technology, they do not share the WAN constraint of the high cost of bandwidth. After the wire is installed, bandwidth is inexpensive because the company owns the wires and there is no recurring cost to a service provider. However, upgrading the physical wiring can be expensive.

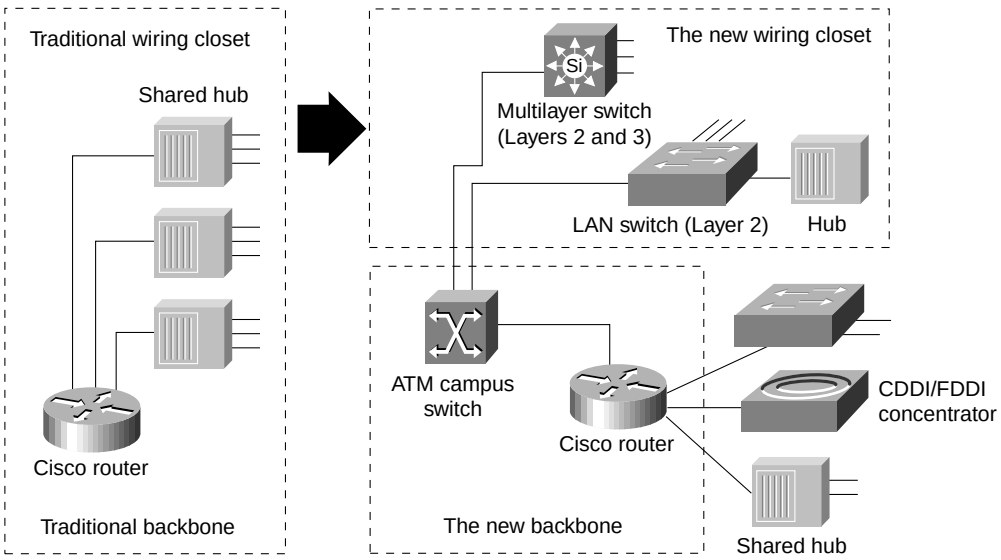
Consequently, network designers generally deploy a campus design that is optimized for the fastest functional architecture that runs on existing physical wire. They might also upgrade wiring to meet the requirements of emerging applications. For example, higher-speed technologies, such as Fast Ethernet, Gigabit Ethernet, and ATM as a backbone architecture, and Layer 2 switching provide dedicated bandwidth to the desktop.

Trends in Campus Design

In the past, network designers had only a limited number of hardware options—routers or hubs—when purchasing a technology for their campus networks. Consequently, it was rare to make a hardware design mistake. Hubs were for wiring closets and routers were for the data center or main telecommunications operations.

Recently, local-area networking has been revolutionized by the exploding use of LAN switching at Layer 2 (the data link layer) to increase performance and to provide more bandwidth to meet new data networking applications. LAN switches provide this performance benefit by increasing bandwidth and throughput for workgroups and local servers. Network designers are deploying LAN switches out toward the network's edge in wiring closets. As Figure 1-3 shows, these switches are usually installed to replace shared concentrator hubs and give higher bandwidth connections to the end user.

Figure 1-3 Example of trends in campus design.



Layer 3 networking is required in the network to interconnect the switched workgroups and to provide services that include security, quality of service (QoS), and traffic management. Routing integrates these switched networks, and provides the security, stability, and control needed to build functional and scalable networks.

Traditionally, Layer 2 switching has been provided by LAN switches, and Layer 3 networking has been provided by routers. Increasingly, these two networking functions are being integrated into common platforms. For example, multilayer switches that provide Layer 2 and 3 functionality are now appearing in the marketplace.

With the advent of such technologies as Layer 3 switching, LAN switching, and virtual LANs (VLANs), building campus networks is becoming more complex than in the past. Table 1-1 summarizes the various LAN technologies that are required to build successful campus networks. Cisco Systems offers product solutions in all of these technologies.

Table 1-1 Summary of LAN Technologies

LAN Technology	Typical Uses
Routing technologies	Routing is a key technology for connecting LANs in a campus network. It can be either Layer 3 switching or more traditional routing with Layer 3 switching and additional router features.
Gigabit Ethernet	Gigabit Ethernet builds on top of the Ethernet protocol, but increases speed ten-fold over Fast Ethernet to 1000 Mbps, or 1 Gbps. Gigabit Ethernet provides high bandwidth capacity for backbone designs while providing backward compatibility for installed media.
LAN switching technologies	
• Ethernet switching	Ethernet switching provides Layer 2 switching, and offers dedicated Ethernet segments for each connection. This is the base fabric of the network.
• Token Ring switching	Token Ring switching offers the same functionality as Ethernet switching, but uses Token Ring technology. You can use a Token Ring switch as either a transparent bridge or as a source-route bridge.
ATM switching technologies	ATM switching offers high-speed switching technology for voice, video, and data. Its operation is similar to LAN switching technologies for data operations. ATM, however, offers high bandwidth capacity.

Network designers are now designing campus networks by purchasing separate equipment types (for example, routers, Ethernet switches, and ATM switches) and then linking them together. Although individual purchase decisions might seem harmless, network designers must not forget that the entire network forms an internetwork.

It is possible to separate these technologies and build thoughtful designs using each new technology, but network designers must consider the overall integration of the network. If this overall integration is not considered, the result can be networks that have a much higher risk of network outages, downtime, and congestion than ever before.

Designing WANs

WAN communication occurs between geographically separated areas. In enterprise internetworks, WANs connect campuses together. When a local end station wants to communicate with a remote end station (an end station located at a different site), information must be sent over one or more WAN links. Routers within enterprise internetworks represent the LAN/WAN junction points of an internetwork. These routers determine the most appropriate path through the internetwork for the required data streams.

WAN links are connected by switches, which are devices that relay information through the WAN and dictate the service provided by the WAN. WAN communication is often called a *service* because the network provider often charges users for the services provided by the WAN (called *tariffs*). WAN services are provided through the following three primary switching technologies:

- Circuit switching
- Packet switching
- Cell switching

Each switching technique has advantages and disadvantages. For example, *circuit-switched* networks offer users dedicated bandwidth that cannot be infringed upon by other users. In contrast, *packet-switched* networks have traditionally offered more flexibility and used network bandwidth more efficiently than circuit-switched networks. *Cell switching*, however, combines some aspects of circuit and packet switching to produce networks with low latency and high throughput. Cell switching is rapidly gaining in popularity. ATM is currently the most prominent cell-switched technology. For more information on switching technology for WANs and LANs, see Chapter 2, "Internetworking Design Basics."

Trends in WAN Design

Traditionally, WAN communication has been characterized by relatively low throughput, high delay, and high error rates. WAN connections are mostly characterized by the cost of renting media (wire) from a service provider to connect two or more campuses together. Because the WAN infrastructure is often rented from a service provider, WAN network designs must optimize the cost of bandwidth and bandwidth efficiency. For example, all technologies and features used to connect campuses over a WAN are developed to meet the following design requirements:

- Optimize WAN bandwidth
- Minimize the tariff cost
- Maximize the effective service to the end users

Recently, traditional shared-media networks are being overtaxed because of the following new network requirements:

- Necessity to connect to remote sites
- Growing need for users to have remote access to their networks
- Explosive growth of the corporate intranets
- Increased use of enterprise servers

Network designers are turning to WAN technology to support these new requirements. WAN connections generally handle mission-critical information, and are optimized for price/performance bandwidth. The routers connecting the campuses, for example, generally apply traffic optimization, multiple paths for redundancy, dial backup for disaster recovery, and QoS for critical applications.

Table 1-2 summarizes the various WAN technologies that support such large-scale internetwork requirements.

Table 1-2 Summary of WAN Technologies

WAN Technology	Typical Uses
Asymmetric Digital Subscriber Line	A new modem technology. Converts existing twisted-pair telephone lines into access paths for multimedia and high-speed data communica- tions. ADSL transmits more than 6 Mbps to a subscriber, and as much as 640 kbps more in both directions.
Analog modem	Analog modems can be used by telecommuters and mobile users who access the network less than two hours per day, or for backup for another type of link.

Leased line	Leased lines can be used for Point-to-Point Protocol (PPP) networks and hub-and-spoke topologies, or for backup for another type of link.
Integrated Services Digital Network (ISDN)	ISDN can be used for cost-effective remote access to corporate networks. It provides support for voice and video as well as a backup for another type of link.
Frame Relay	Frame Relay provides a cost-effective, high- speed, low-latency mesh topology between remote sites. It can be used in both private and carrier-provided networks.
Switched Multimegabit Data Service (SMDS)	SMDS provides high-speed, high-performance connections across public data networks. It can also be deployed in metropolitan-area networks (MANs).
X.25	X.25 can provide a reliable WAN circuit or backbone. It also provides support for legacy applications.
WAN ATM	WAN ATM can be used to accelerate bandwidth requirements. It also provides support for multiple QoS classes for differing application requirements for delay and loss.

Utilizing Remote Connection Design

Remote connections link single users (mobile users and/or telecommuters) and branch offices to a local campus or the Internet. Typically, a remote site is a small site that has few users and therefore needs a smaller size WAN connection. The remote requirements of an internetwork, however, usually involve a large number of remote single users or sites, which causes the aggregate WAN charge to be exaggerated.

Because there are so many remote single users or sites, the aggregate WAN bandwidth cost is proportionally more important in remote connections than in WAN connections. Given that the three-year cost of a network is nonequipment expenses, the WAN media rental charge from a service provider is the largest cost component of a remote network. Unlike WAN connections, smaller sites or single users seldom need to connect 24 hours a day.

Consequently, network designers typically choose between dial-up and dedicated WAN options for remote connections. Remote connections generally run at speeds of 128 Kbps or lower. A network designer might also employ bridges in a remote site for their ease of implementation, simple topology, and low traffic requirements.

Trends in Remote Connections

Today, there is a large selection of remote WAN media that include the following:

- Analog modem
- Asymmetric Digital Subscriber Line
- Leased line
- Frame Relay
- X.25
- ISDN

Remote connections also optimize for the appropriate WAN option to provide cost-effective bandwidth, minimize dial-up tariff costs, and maximize effective service to users.

Trends in LAN/WAN Integration

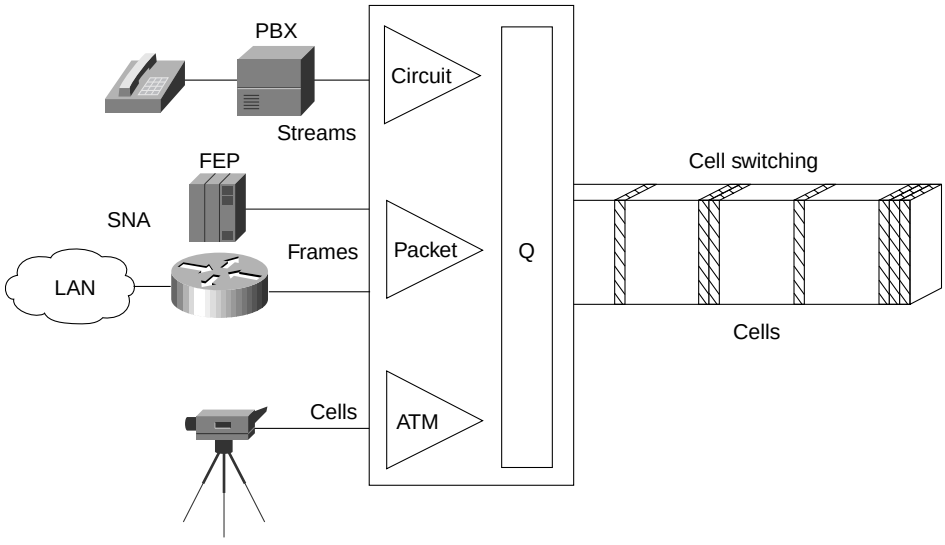
Today, 90 percent of computing power resides on desktops, and that power is growing exponentially. Distributed applications are increasingly bandwidth hungry, and the emergence of the Internet is driving many LAN architectures to the limit. Voice communications have increased significantly with more reliance on centralized voice mail systems for verbal communications. The internetwork is the critical tool for information flow. Internetworks are being pressured to cost less, yet support the emerging applications and higher number of users with increased performance.

To date, local- and wide-area communications have remained logically separate. In the LAN, bandwidth is free and connectivity is limited only by hardware and implementation costs. The LAN has carried data only. In the WAN, bandwidth has been the overriding cost, and such delay-sensitive traffic as voice has remained separate from data. New applications and the economics of supporting them, however, are forcing these conventions to change.

The Internet is the first source of multimedia to the desktop, and immediately breaks the rules. Such Internet applications as voice and real-time video require better, more predictable LAN and WAN performance. These multimedia applications are fast becoming an essential part of the business productivity toolkit. As companies begin to consider implementing new intranet-based, bandwidth-intensive multimedia applications—such as video training, videoconferencing, and voice over IP—the impact of these applications on the existing networking infrastructure is a serious concern. If a company has relied on its corporate network for business-critical SNA traffic, for example, and wants to bring a new video training application on line, the network must be able to provide guaranteed quality of service (QoS) that delivers the multimedia traffic, but does not allow it to interfere with the business-critical traffic. ATM has emerged as one of the technologies for integrating LANs and WANs. The Quality of Service (QoS) features of ATM can support any traffic type in separate or mixed streams, delay sensitive traffic, and nondelay-sensitive traffic, as shown in Figure 1-4.

ATM can also scale from low to high speeds. It has been adopted by all the industry's equipment vendors, from LAN to private branch exchange (PBX).

Figure 1-4 ATM support of various traffic types.



Providing Integrated Solutions

The trend in internetworking is to provide network designers greater flexibility in solving multiple internetworking problems without creating multiple networks or writing off existing data communication investments. Routers might be relied upon to provide a reliable, secure network and act as a barrier against inadvertent broadcast storms in the local networks. Switches, which can be divided into two main categories—LAN switches and WAN switches—can be deployed at the workgroup, campus backbone, or WAN level. Remote sites might use low-end routers for connection to the WAN.

Underlying and integrating all Cisco products is the Cisco Internetworking Operating System (Cisco IOS) software. The Cisco IOS software enables disparate groups, diverse devices, and multiple protocols all to be integrated into a highly reliable and scalable network. Cisco IOS software also supports this internetwork with advanced security, quality of service, and traffic services.

Determining Your Internetworking Requirements

Designing an internetwork can be a challenging task. Your first step is to understand your internetworking requirements. The rest of this chapter is intended as a guide for helping you determine these requirements. After you have identified these requirements, refer to Chapter 2, "Internetworking Design Basics," for information on selecting internetwork capability and reliability options that meet these requirements.

Internetworking devices must reflect the goals, characteristics, and policies of the organizations in which they operate. Two primary goals drive internetworking design and implementation:

- *Application availability*—Networks carry application information between computers. If the applications are not available to network users, the network is not doing its job.
- *Cost of ownership*—Information system (IS) budgets today often run in the millions of dollars. As large organizations increasingly rely on electronic data for managing business activities, the associated costs of computing resources will continue to rise.

A well-designed internetwork can help to balance these objectives. When properly implemented, the network infrastructure can optimize application availability and allow the cost-effective use of existing network resources.

The Design Problem: Optimizing Availability and Cost

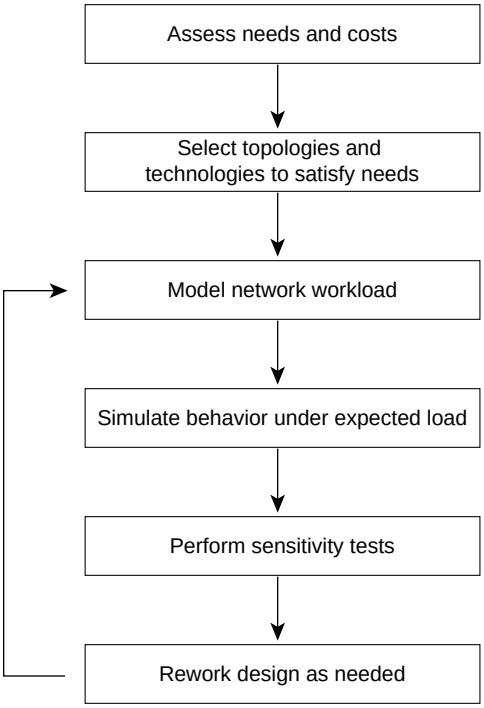
In general, the network design problem consists of the following three general elements:

- *Environmental givens*—Environmental givens include the location of hosts, servers, terminals, and other end nodes; the projected traffic for the environment; and the projected costs for delivering different service levels.
- *Performance constraints*—Performance constraints consist of network reliability, traffic throughput, and host/client computer speeds (for example, network interface cards and hard drive access speeds).
- *Internetworking variables*—Internetworking variables include the network topology, line capacities, and packet flow assignments.

The goal is to minimize cost based on these elements while delivering service that does not compromise established availability requirements. You face two primary concerns: availability and cost. These issues are essentially at odds. Any increase in availability must generally be reflected as an increase in cost. As a result, you must weigh the relative importance of resource availability and overall cost carefully.

As Figure 1-5 shows, designing your network is an iterative activity. The discussions that follow outline several areas that you should carefully consider when planning your internetworking implementation.

Figure 1-5 General network design process.



Assessing User Requirements

In general, users primarily want application availability in their networks. The chief components of application availability are *response time*, *throughput*, and *reliability*:

- Response time is the time between entry of a command or keystroke and the host system's execution of the command or delivery of a response. User satisfaction about response time is generally considered to be a *monotonic* function up to some limit, at which point user satisfaction falls off to nearly zero. Applications in which fast response time is considered critical include interactive online services, such as automated tellers and point-of-sale machines.
- Applications that put high-volume traffic onto the network have more effect on throughput than end-to-end connections. Throughput-intensive applications generally involve file- transfer activities. However, throughput-intensive applications also usually have low response-time requirements. Indeed, they can often be scheduled at times when response-time-sensitive traffic is low (for example, after normal work hours).
- Although reliability is always important, some applications have genuine requirements that exceed typical needs. Organizations that require nearly 100 percent up time conduct all activities online or over the telephone. Financial services, securities exchanges, and emergency/police/military operations are a few examples. These situations imply a requirement for a high level of hardware and topological redundancy. Determining the cost of any downtime is essential in determining the relative importance of reliability to your internetwork.

You can assess user requirements in a number of ways. The more involved your users are in the process, the more likely that your evaluation will be accurate. In general, you can use the following methods to obtain this information:

- *User community profiles*—Outline what different user groups require. This is the first step in determining internetwork requirements. Although many users have roughly the same requirements of an electronic mail system, engineering groups using XWindows terminals and Sun workstations in an NFS environment have different needs from PC users sharing print servers in a finance department.
- *Interviews, focus groups, and surveys*—Build a baseline for implementing an internetwork. Understand that some groups might require access to common servers. Others might want to allow external access to specific internal computing resources. Certain organizations might require IS support systems to be managed in a particular way according to some external standard. The least formal method of obtaining information is to conduct interviews with key user groups. Focus groups can also be used to gather information and generate discussion among different organizations with similar (or dissimilar) interests. Finally, formal surveys can be used to get a statistically valid reading of user sentiment regarding a particular service level or proposed internetworking architecture.
- *Human factors tests*—The most expensive, time-consuming, and possibly revealing method is to conduct a test involving representative users in a lab environment. This is most applicable when evaluating response time requirements. As an example, you might set up working systems and have users perform normal remote host activities from the lab network. By evaluating user reactions to variations in host responsiveness, you can create benchmark thresholds for acceptable performance.

Assessing Proprietary and Nonproprietary Solutions

Compatibility, conformance, and interoperability are related to the problem of balancing proprietary functionality and open internetworking flexibility. As a network designer, you might be forced to choose between implementing a multivendor environment and implementing a specific, proprietary capability. For example, the *Interior Gateway Routing Protocol* (IGRP) provides many useful capabilities, such as a number of features that are designed to enhance its stability. These include *hold-downs*, *split horizons*, and *poison reverse updates*.

The negative side is that IGRP is a proprietary routing protocol. In contrast, the integrated *Intermediate System-to Intermediate System* (IS-IS) protocol is an open internetworking alternative that also provides a fast converging routing environment; however, implementing an open routing protocol can potentially result in greater multiple-vendor configuration complexity.

The decisions that you make have far-ranging effects on your overall internetwork design. Assume that you decide to implement integrated IS-IS instead of IGRP. In doing this, you gain a measure of interoperability; however, you lose some functionality. For instance, you cannot load balance traffic over unequal parallel paths. Similarly, some modems provide a high level of proprietary diagnostic capabilities, but require that all modems throughout a network be of the same vendor type to fully exploit proprietary diagnostics.

Previous internetworking (and networking) investments and expectations for future requirements have considerable influence over your choice of implementations. You need to consider installed internetworking and networking equipment; applications running (or to be run) on the network; traffic patterns; physical location of sites, hosts, and users; rate of growth of the user community; and both physical and logical network layout.

Assessing Costs

The internetwork is a strategic element in your overall information system design. As such, the cost of your internetwork is much more than the sum of your equipment purchase orders. View it as a total cost-of-ownership issue. You must consider the entire life cycle of your internetworking environment. A brief list of costs associated with internetworks follows:

- *Equipment hardware and software costs*—Consider what is really being bought when you purchase your systems; costs should include initial purchase and installation, maintenance, and projected upgrade costs.
- *Performance tradeoff costs*—Consider the cost of going from a five-second response time to a half-second response time. Such improvements can cost quite a bit in terms of media selection, network interfaces, internetworking nodes, modems, and WAN services.
- *Installation costs*—Installing a site's physical cable plant can be the most expensive element of a large network. The costs include installation labor, site modification, fees associated with local code conformance, and costs incurred to ensure compliance with environmental restrictions (such as asbestos removal). Other important elements in keeping your costs to a minimum will include developing a well-planned wiring closet layout and implementing color code conventions for cable runs.
- *Expansion costs*—Calculate the cost of ripping out all thick Ethernet, adding additional functionality, or moving to a new location. Projecting your future requirements and accounting for future needs saves time and money.
- *Support costs*—Complicated internetworks cost more to monitor, configure, and maintain. Your internetwork should be no more complicated than necessary. Costs include training, direct labor (network managers and administrators), sparing, and replacement costs. Additional cost that should be included is out-of-band management, SNMP management stations, and power.
- *Cost of downtime*—Evaluate the cost for every minute that a user is unable to access a file server or a centralized database. If this cost is high, you must attribute a high cost to downtime. If the cost is high enough, fully redundant internetworks might be your best option.
- *Opportunity costs*—Every choice you make has an opposing alternative option. Whether that option is a specific hardware platform, topology solution, level of redundancy, or system integration alternative, there are always options. *Opportunity costs* are the costs of *not* picking one of those options. The opportunity costs of not switching to newer technologies and topologies might be lost competitive advantage, lower productivity, and slower overall performance. Any effort to integrate opportunity costs into your analysis can help to make accurate comparisons at the beginning of your project.
- *Sunken costs*—Your investment in existing cable plant, routers, concentrators, switches, hosts, and other equipment and software are your *sunken costs*. If the sunken cost is high, you might need to modify your networks so that your existing internetwork can continue to be utilized. Although comparatively low incremental costs might appear to be more attractive than significant redesign costs, your organization might pay more in the long run by not upgrading systems. Over reliance on sunken costs can cost your organization sales and market share when calculating the cost of internetwork modifications and additions.

Estimating Traffic: Work Load Modeling

Empirical *work-load modeling* consists of instrumenting a working internetwork and monitoring traffic for a given number of users, applications, and network topology. Try to characterize activity throughout a normal work day in terms of the type of traffic passed, level of traffic, response time of hosts, time to execute file transfers, and so on. You can also observe utilization on existing network equipment over the test period.

If the tested internetwork's characteristics are close to the new internetwork, you can try extrapolating to the new internetwork's number of users, applications, and topology. This is a *best-guess* approach to traffic estimation given the unavailability of tools to characterize detailed traffic behavior.

In addition to passive monitoring of an existing network, you can measure activity and traffic generated by a known number of users attached to a representative test network and then extrapolate findings to your anticipated population.

One problem with modeling workloads on networks is that it is difficult to accurately pinpoint traffic load and network device performance as functions of the number of users, type of application, and geographical location. This is especially true without a real network in place. Consider the following factors that influence the dynamics of the network:

- *The time-dependent nature of network access*—Peak periods can vary; measurements must reflect a range of observations that includes peak demand.
- *Differences associated with type of traffic*—Routed and bridged traffic place different demands on internetwork devices and protocols; some protocols are sensitive to dropped packets; some application types require more bandwidth.
- *The random (nondeterministic) nature of network traffic*—Exact arrival time and specific effects of traffic are unpredictable.

Sensitivity Testing

From a practical point of view, sensitivity testing involves breaking stable links and observing what happens. When working with a test network, this is relatively easy. Disturb the network by removing an active interface, and monitor how the change is handled by the internetwork: how traffic is rerouted, the speed of convergence, whether any connectivity is lost, and whether problems arise in handling specific types of traffic. You can also change the level of traffic on a network to determine the effects on the network when traffic levels approach media saturation. This empirical testing is a type of *regression* testing: A series of specific modifications (tests) are repeated on different versions of network configurations. By monitoring the effects on the design variations, you can characterize the relative resilience of the design.

Note Modeling sensitivity tests using a computer is beyond the scope of this publication. A useful source for more information about computer-based network design and simulation is A.S. Tannenbaum, *Computer Networks*, Upper Saddle River, New Jersey: Prentice Hall, 1996.

Summary

After you have determined your network requirements, you must identify and then select the specific capability that fits your computing environment. For basic information on the different types of internetworking devices along with a description of a hierarchical approach to internetworking, refer to Chapter 2, "Internetworking Design Basics."

Chapters 2–13 in this book are technology chapters that present detailed discussions about specific implementations of large-scale internetworks in the following environments:

- Large-scale Internetwork Protocol (IP) internetworks
  - Enhanced Interior Gateway Routing Protocol (IGRP) design
  - Open Shortest Path First (OSPF) design
- IBM System Network Architecture (SNA) internetworks

  - Source-route bridging (SRB) design
  - Synchronous Data Link Control (SDLC) and serial tunneling (STUN), SDLC Logical Link Control type 2 (SDLLC), and Qualified Logical Link Control (QLLC) design
  - Advanced Peer-to-Peer Networking (APPN) and Data Link Switching (DLSw) design
- ATM internetworks
- Packet service internetworks
  - Frame Relay design
- Dial-on-demand routing (DDR) internetworks
- ISDN internetworks

In addition to these technology chapters there are chapters on designing switched LAN internetworks, campus LANs, and internetworks for multimedia applications. The last 12 chapters of this book include case studies relating to the concepts learned in the previous chapters.