Telecommunications Essentials

The Complete Global Source for Communications Fundamentals, Data Networking and the Internet, and Next-Generation Networks

Lillian Goleniewski

Addison-Wesley

Boston • San Francisco • New York • Toronto • Montreal
London • Munich • Paris • Madrid
Capetown • Sydney • Tokyo • Singapore • Mexico City

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and Addison-Wesley, Inc. was aware of a trademark claim, the designations have been printed with initial capital letters or in all capitals.

Lido Telecommunications Essentials® is the registered trademark of The Lido Organization, Inc.

The author and publisher have taken care in the preparation of this book, but make no expressed or implied warranty of any kind and assume no responsibility for errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of the use of the information or programs contained herein.

The publisher offers discounts on this book when ordered in quantity for special sales. For more information, please contact:

Pearson Education Corporate Sales Division
201 W. 103rd Street
Indianapolis, IN 46290
(800) 428-5331
corpsales@pearsoned.com

Visit AW on the Web: www.aw.com/cseng/

Library of Congress Cataloging-in-Publication Data

Goleniewski, Lillian.
Telecommunications essentials : the complete global source for communications fundamentals, data networking and the Internet, and next-generation networks / Lillian Goleniewski.
p. cm.
Includes bibliographical references and index.
ISBN 0-201-76032-0
1. Telecommunication. I. Title.

TK5101 G598 2002
621.382—dc21 2001053752

For information on obtaining permission for use of material from this work, please submit a written request to:

Pearson Education, Inc.
Rights and Contracts Department
75 Arlington Street, Suite 300
Boston, MA 02116
Fax: (617) 848-7047

ISBN 0-201-76032-0
Text printed on recycled paper

1 2 3 4 5 6 7 8 9 10—CRS—0504030201
First printing, December 2001

Table 8.2 LAN Technologies and Cabling Requirements

Technology	Type of Cable
	Ethernet (10Mbps)
10Base5	Thick coax
10Base2	Thin coax
10BaseT	2-pair UTP
10BaseFL	2 strands of multimode optical fiber
	Fast Ethernet (100Mbps)
100BaseTX	2-pair Cat 5 UTP
100BaseT4	4-pair Cat 3 UTP
100BaseT2	2-pair Cat 3 UTP
100BaseFX	2 strands of multimode optical fiber
	Gigabit Ethernet (1Gbps)
1000BaseSX	Short-wavelength multimode optical fiber
1000BaseLX	Long-wavelength single-mode optical fiber
1000BaseCX	Coax patch cable
1000BaseT	4-pair Cat 5 or Cat 5e UTP
1000BaseTX	2-pair Cat 6 (currently a TIA draft proposal)

The prevailing standard in the world is Ethernet. It generally appears as Fast Ethernet or Gigabit Ethernet in the backbone, connecting together individual Fast Ethernet or 10Mbps Ethernet LAN segments (see Figure 8.3).

LAN Access Methods

The third main LAN characteristic is the access methods, which are involved in determining who gets to use the network and when they get to use it. There are two main approaches: token passing and carrier-sense multiple-access/collision detect (CSMA/CD).

Telecommunications Essentials

The Complete Global Source for Communications Fundamentals, Data Networking and the Internet, and Next-Generation Networks

Lillian Goleniewski

Addison-Wesley

Boston • San Francisco • New York • Toronto • Montreal
London • Munich • Paris • Madrid
Capetown • Sydney • Tokyo • Singapore • Mexico City

The author and publisher have taken care in the preparation of this book, but make no expressed or implied warranty of any kind and assume no responsibility for errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of the use of the information or programs contained herein.

The publisher offers discounts on this book when ordered in quantity for special sales. For more information, please contact:

Pearson Education Corporate Sales Division
201 W. 103rd Street
Indianapolis, IN 46290
(800) 428-5331
corpsales@pearsoned.com

Visit AW on the Web: www.aw.com/cseng/

Library of Congress Cataloging-in-Publication Data

Goleniewski, Lillian.
Telecommunications essentials : the complete global source for communications fundamentals, data networking and the Internet, and next-generation networks / Lillian Goleniewski.
p. cm.
Includes bibliographical references and index.
ISBN 0-201-76032-0
1. Telecommunication. I. Title.

TK5101 G598 2002
621.382—dc21 2001053752

ISBN 0-201-76032-0
Text printed on recycled paper

1 2 3 4 5 6 7 8 9 10—CRS—0504030201
First printing, December 2001

Chapter 11

Next-Generation Network Services

This chapter investigates traditional Internet services, as well as new generations of applications and the network platforms that support those applications. It discusses virtual private networks (VPNs), security, various uses of Voice over IP (VoIP) networks, and developments in the streaming media arena and emerging applications.

■ Traditional Internet Applications

Traditional Internet applications are called *elastic applications* because they can work without guarantees of timely delivery. Because they can stretch in the face of greater delay, they can still perform adequately, even when the network faces increased congestion and degradation in performance. The following are the most widely used elastic applications:

- **E-mail**—The most widely used of the Internet applications, generating several gigabytes of traffic per month, is e-mail. Because there is a standardized convention for the e-mail address—*username@domainname*—various companies can interoperate to support electronic messaging.
- **Telnet**—Telnet is one of the original ARPANET applications. It enables remote login to another computer that's running a Telnet server and allows

you to run applications present on that computer, with their outputs appearing in the window on your computer.

- **File Transfer Protocol (FTP)**—FTP allows file transfers to and from remote hosts. That is, you can use FTP to download documents from a remote host onto your local device.
- **The World Wide Web**—Key aspects that identify the Web are the use of the uniform resource locator (URL) and the use of Hypertext Transfer Protocol (HTTP)—the client/server hypermedia system that enables the multimedia point-and-click interface. HTTP provides hyperlinks to other documents, which are encoded in Hypertext Markup Language (HTML), providing a standardized way of displaying and viewing information contained on servers worldwide. Web browsers are another important part of the Web environment—they interpret the HTML and display it, along with any images, on the user's local computer. The ease of using the Web popularized the use of URLs, which are available for just about any Internet service. (The *URL* is the syntax and semantics of formalized information for location and access of resources via the Internet. URLs are used to locate resources by providing an abstract identification of the resource location.)

So, how are people actually using the Internet? Greenfield Online conducts polls online. In November 2000, Greenfield Online (www.greenfieldonline.com) reported that the Internet was being used as follows:

- 98% of Internet users went online to check their e-mail.
- 80% of Internet users were checking for local information, such as movie schedules, weather updates, or traffic reports.
- 66% of Internet users were looking for a site that provided images and sounds.
- 66% of Internet users wanted to shop at sites that provided images of the products they were interest in.
- 53% of Internet users downloaded some form of a large file.
- 37% of Internet users listened to Internet radio.

This shows that increasingly there is interest in imagery, multimedia, and entertainment-type aspects of the Internet. These are advanced real-time applications that are highly sensitive to timely data delivery. Therefore, any application that includes VoIP, audio streaming, video streaming, or interactive multimedia needs to be addressed by the administration of Quality of Service (QoS). The lack of control over QoS in the public Internet is preventing the deployment of these new applications at a more rapid pace.

Today's flat-rate pricing for Internet access is compatible with the Internet's lack of service differentiation, and it is partially responsible for that structure as well. The main appeal of a flat-rate pricing scheme is its simplicity. It means predictable fees for the users, and it means providers can avoid the administrative time and cost associated with tracking, allocating, and billing for usage. It also gives companies known expectations for payments, facilities planning, and budgeting. However, as QoS emerges within the Internet, the ability to differentiate services will result in differentiated pricing, thereby allowing revenue-generating service levels and packages—and that's extremely important. As we've discussed several times so far in this book, developments in optical networking and in wireless networking are providing more and more bandwidth. Hence, the cost—the cents per minute that you can charge for carrying traffic—is being reduced. If network operators are going to continue to make money in the future, they will need to do so through the administration of value-added services, differentiated performance, and tiered pricing. Therefore, the QoS aspect is very important to the materialization of new revenue-generating services.

Key applications from which service providers are expected to derive revenues include e-commerce, videoconferencing, distance learning and education networks, Webcasting, multiplayer gaming, unified messaging, call centers, interactive voice response, and IP-based centrex systems. Evolving next-generation networks—such as VPNs, VoIP and Packet over IP, streaming audio and video, multimedia collaboration, network caching, application hosting, location-based online services, software downloads, and security services—are introducing a variety of Class of Service (CoS) and QoS differentiators.

VPNs

A big driver of interest in VPNs is that customers increasingly need to communicate with people outside their enterprise, not just those inside the enterprise. As mentioned in Chapter 6, "Data Communications Basics," in the 1980s, about 80% of the information that was used within a given address of a business came from within that address. Only 20% was exchanged outside the walls of that location. Today, the relationship has reversed. As much as 80% of information exchanged is with points outside a given business address.

Another reason for interest in VPNs is that customers want to quickly and securely change their access points and needs as changes occur in their businesses. Many strategic alliances and partnerships require companies to exchange messages quickly. Some of these are temporary assignments—for example, a contractor

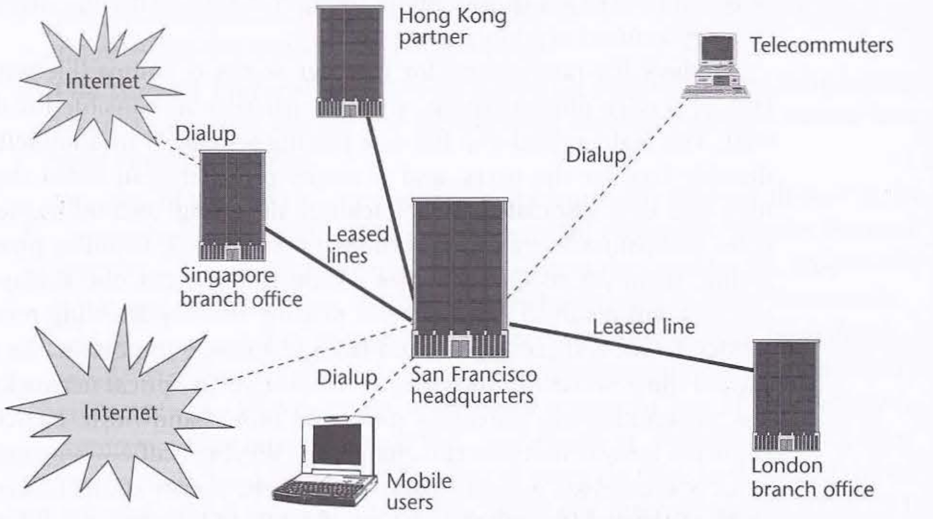

Figure 11.1 An enterprise network based on leased lines

building out a fiber-optic loop or an applications developer building a new billing system—that might last a few months, during which time the individuals involved need to be incorporated into the network. Leased lines are infamous for requiring long waits for provisioning—often 6 months to 18 months! VPNs allow rapid provisioning of capacity where and when needed.

What we see emerging is a requirement for networks that can be very quickly provisioned and changed in relationship to organizational structures. This results in a steady migration of traffic away from the traditional networks, based on leased lines (see Figure 11.1), to public networks. As a result, we're seeing a steady growth in the pseudoprivate realm of the VPN (see Figure 11.2). A *VPN* is a logical network that isolates customer traffic on shared service provider facilities. In other words, the enterprise's traffic is aggregated with the traffic of other companies. VPNs have been around for quite some time—since X.25 closed user groups on the packet-switched network, and with the AT&T Software-Defined Network (SDN) on the circuit-switched networks. A VPN looks like a private network, but it runs across either the public circuit-switched network or public packet-switched data networks. Thus, VPNs are not just a solution within the IP realm—a VPN is a concept, not a specific set of technologies, and it can be deployed over a wide range of network technologies, including circuit-switched networks, X.25, IP, Frame Relay, and ATM.

A VPN uses a shared carrier infrastructure. It can provide additional bandwidth on demand, which is an incredible feat, as compared to the weeks that it normally takes to add bandwidth to dedicated networks. Carriers build VPNs with

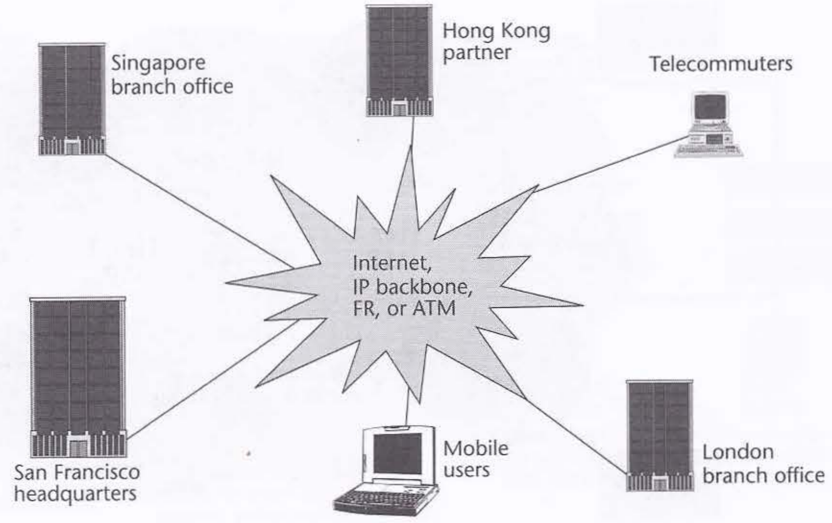

Figure 11.2 An enterprise network using a VPN

advanced survivability and restoration capabilities, as well as network management tools and support, so that QoS can be considered and service-level agreements (SLAs) can be administered and met.

Two basic VPN deployment models exist: customer based and network based. In *customer-based VPNs,* carriers install gateways, routers, and other VPN equipment on the customer premises. This is preferred when customers want to have control over all aspects of security. In *network-based VPNs,* the carrier houses all the necessary equipment at a point of presence (POP) near the customer's location. Customers that want to take advantage of the carrier's VPN economies of scale prefer this type of VPN.

VPN Frameworks

Contemporary VPNs can be described as belonging to one of two categories: the Internet-based VPN and the provisioned VPN.

Internet-Based VPNs

In an *Internet-based VPN* (see Figure 11.3), smaller ISPs provide local access services in defined geographical regions, requiring an enterprise to receive end-to-end services from multiple suppliers. An Internet-based VPN uses encryption to create a form of closed user group, thereby isolating the enterprise traffic and providing acceptable security for the enterprise across the public shared packet network. However, because

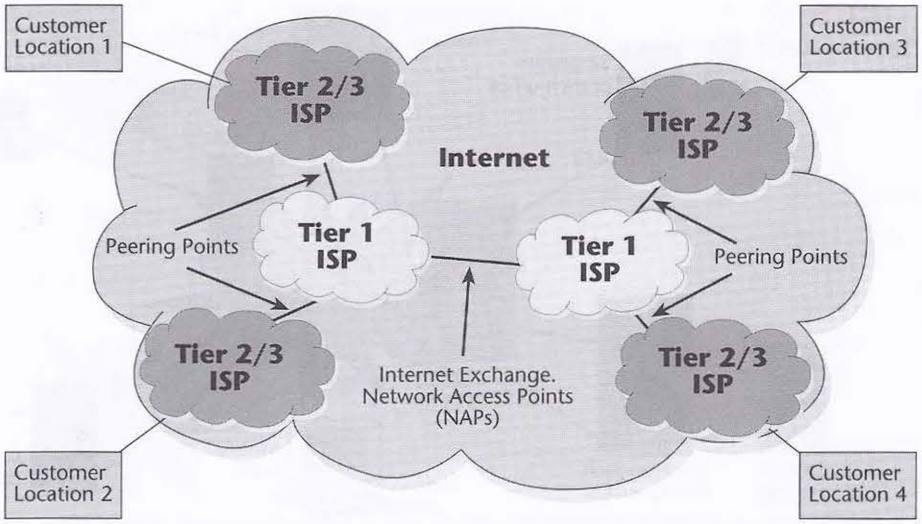

Figure 11.3 An Internet-based VPN

it involves multiple ISPs in the delivery of the VPN, the performance is unpredictable. The biggest problem of having multiple suppliers is the inability to define and meet consistent end-to-end bandwidth or performance objectives.

Figure 11.4 shows what is involved in providing an Internet-based VPN. The customer would have on the premises a wide variety of servers that dish up the corporate content, the finance systems, the customer service systems, and so on. A VPN is responsible for the encapsulation of the information and hence the security aspects. *Remote Authentication Dial-in User Services* (RADIUS), an authentication and access control server, is used for purposes of authenticating whether a user is allowed access into the corporate resources. The RADIUS server connects to a firewall, which is used to determine whether traffic is allowed into or out of the network. The router selects the optimum path for the messages to take, and the circuit physically terminates on a channel service unit/data service unit (CSU/DSU). A private line interfaces with the Internet provider's POP. From that point, the VPN either uses the public Internet that's comprised of multiple ISPs, or it relies on IP backbones provided by a smaller group of providers. Users who are working on mobile devices would have laptops equipped with the client and VPN services necessary for encapsulation and the administration of security.

Provisioned VPNs

VPNs rely on the capability to administer preferential treatment to applications, to users, and so on. The public Internet does not support preferential treatment

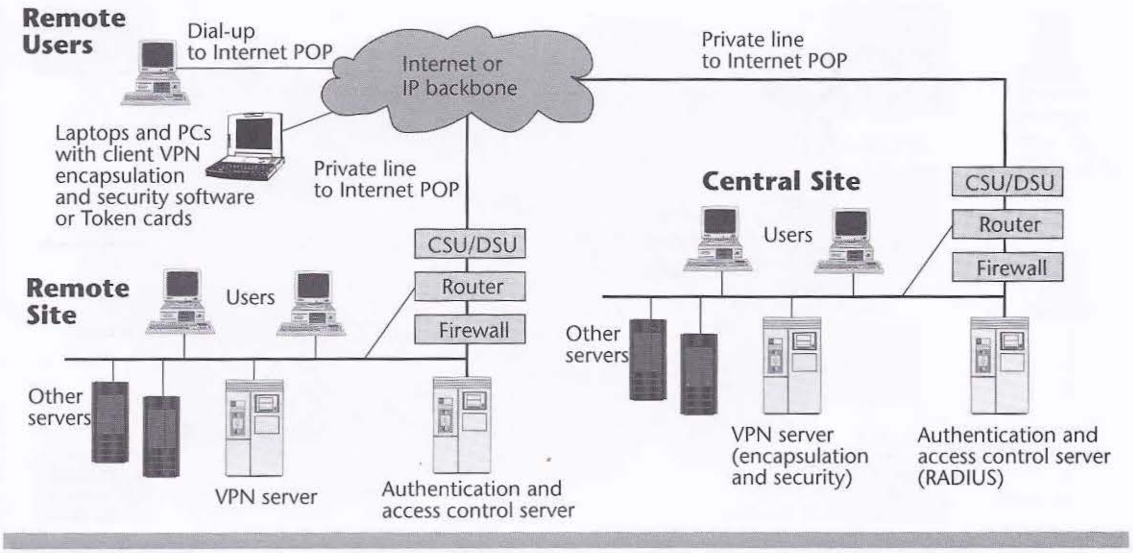
Remote Users
Dial-up to Internet POP
Internet or IP backbone
Private line to Internet POP
Laptops and PCs with client VPN encapsulation and security software or Token cards
Private line to Internet POP
Central Site
Users
CSU/DSU
Router
Firewall
Remote Site
Users
CSU/DSU
Router
Firewall
Other servers
Other servers
VPN server
Authentication and access control server
VPN server (encapsulation and security)
Authentication and access control server (RADIUS)

Figure 11.4 The parts of an Internet-based VPN

because it is subject to delay, jitter, and loss; it is therefore unsuitable for next-generation services that require high performance. In most cases, to accommodate business customers that are interested in such advanced services and who demand SLAs, the underlying transport is really Frame Relay or ATM. These Frame Relay and ATM VPNs offer greater levels of QoS and can fulfill the SLAs that customers and vendors agree to. They do, however, require that the customer acquire an integrated access device (IAD) to have on the premises, which can increase the deployment cost significantly. IADs enable the enterprise to aggregate voice, data, and video traffic at the customer edge.

A *provisioned VPN* (see Figure 11.5) is a packet-switched VPN that runs across the service provider's backbone, generally using Frame Relay or ATM. This type of VPN is built on OSI model Layer 2 virtual circuits, such as those used by Frame Relay, ATM, or Multiprotocol Label Switching (MPLS), and it is provisioned based on customer orders. Virtual circuits based on predetermined locations create closed user groups and work well to carve out a VPN in a public shared network, by limiting access and usage to the provisioned VPN community. However, encryption is still required to securely protect the information from theft or modification by intruders.

The provisioned VPN is differentiated from the IP VPN by its ability to support multiple protocols and by the fact that it offers improved performance and management. These VPNs are characterized as having excellent performance and security, but the negative is that a single vendor offers both reach and breadth in terms of service offerings.

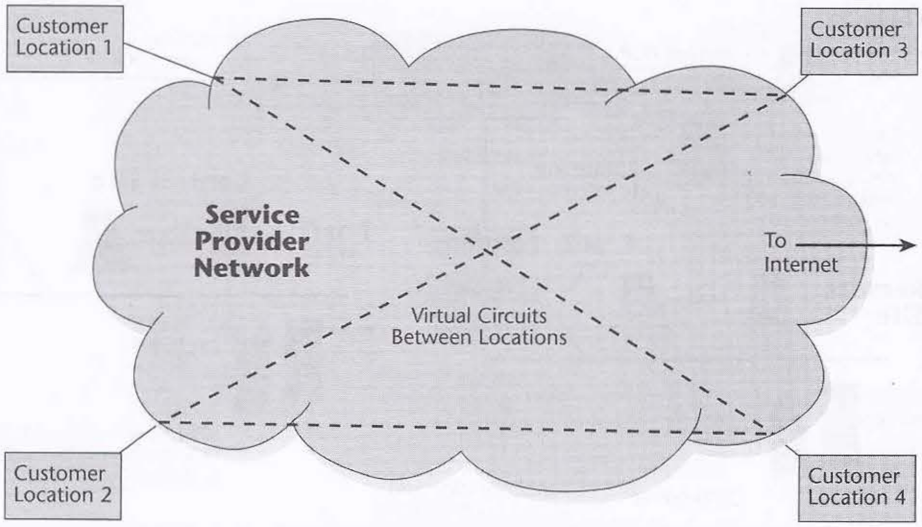

Figure 11.5 A provisioned VPN

Figure 11.6 shows what the equipment would like look at a customer premise in support of a Frame Relay- or an ATM-based VPN. The customer would have an IAD that would allow voice and data to be converged at the customer premise. The IAD would feed into the data communications equipment, over which a circuit would go to the service provider's POP. At the service provider's POP would be a multiservice access device that enables multiple protocols and interfaces to be supported and that provides access into the service provider's core network, which would be based on the use of Frame Relay or ATM. To differentiate Frame Relay- and ATM-based VPNs from Internet-based VPNs, service providers stress that multiple protocols are supported and that they rely on the use of virtual circuits or MPLS labels to facilitate the proper path, thereby ensuring better performance and providing traffic management capabilities.

To further differentiate Frame Relay- or ATM-based VPNs from regular Frame Relay or ATM services, additional functions—such as packet classification and traffic isolation, the capability to handle multiple separate packet-forwarding tables and instances of routing protocols for each customer—reside at the edge.

VPN Applications

A VPN is an architecture, a series of products and software functions that are tied together and tightly calibrated. Managing a VPN entails dealing primarily with two issues: security policies and parameters and making sure that applications function within the latency requirements.

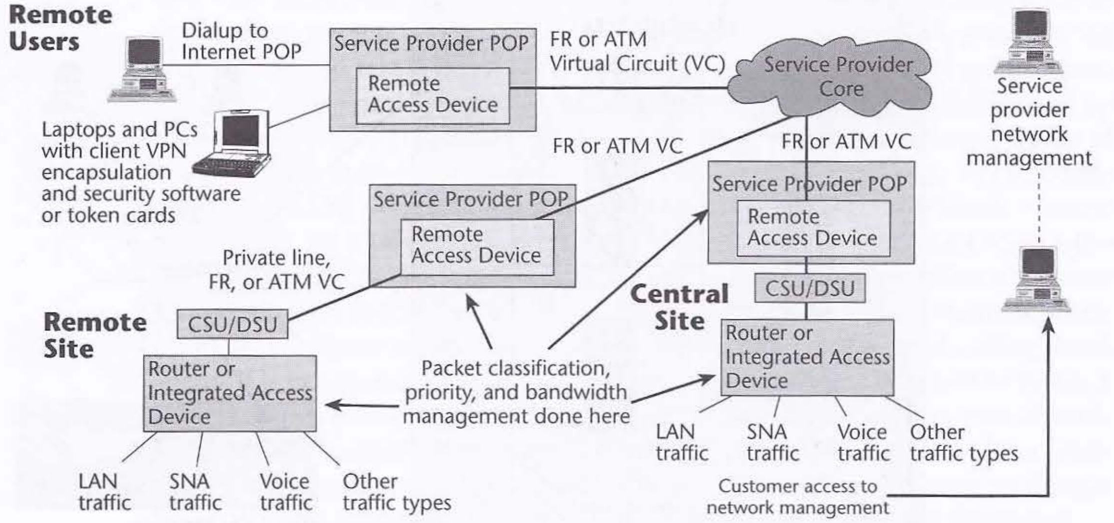
Remote Users
Dialup to Internet POP
Service Provider POP
Remote Access Device
FR or ATM Virtual Circuit (VC)
Service Provider Core
Service provider network management
Laptops and PCs with client VPN encapsulation and security software or token cards
FR or ATM VC
FR or ATM VC
Service Provider POP
Remote Access Device
Service Provider POP
Remote Access Device
Private line, FR, or ATM VC
Central Site
CSU/DSU
Remote Site
CSU/DSU
Router or Integrated Access Device
Packet classification, priority, and bandwidth management done here
Router or Integrated Access Device
LAN traffic
SNA traffic
Voice traffic
Other traffic types
LAN traffic
SNA traffic
Voice traffic
Other traffic types
Customer access to network management

Figure 11.6 A Frame Relay- or ATM-based provisioned VPN

VPN applications provide maximum opportunities to save money and to make money—by substituting leased lines with Internet connectivity, by reducing costs of dialup remote access, and by stimulating new applications, using extranets. These savings can be substantial. According to TeleChoice (www.telechoice.com), in the realm of remote access, savings over customer-owned and maintained systems can range from 30% to 70%; savings over traditional Frame Relay services can range from 20% to 60%; savings over leased lines or private lines can range from 50% to 70%; and savings over international private lines can be up to 90%.

It is important to be able to effectively and easily manage the VPN environment. You need to consider the capability to track the tunnel traffic, the support for policy management, the capability to track QoS, the capability to track security infractions, and the support for public key certificate authorities (CAs).

The one-stop-shopping approach to VPNs—managed VPN services—is designed to lock in users and to reduce costly customer churn, but with this approach, interoperability is very restricted. Managed VPNs provide capabilities such as IP connection and transport services, routers, firewalls, and a VPN box at the customer site. Benefits of this approach include the fact that it involves a single service vendor, SLAs, guaranteed latency and bandwidth, and the security of traffic being confined to one network. Approximately one-third of VPN users opt for such a managed service.

There are three major applications of VPNs—intranets (that is, site-to-site VPNs) remote access, and extranets—which are examined in the following sections.

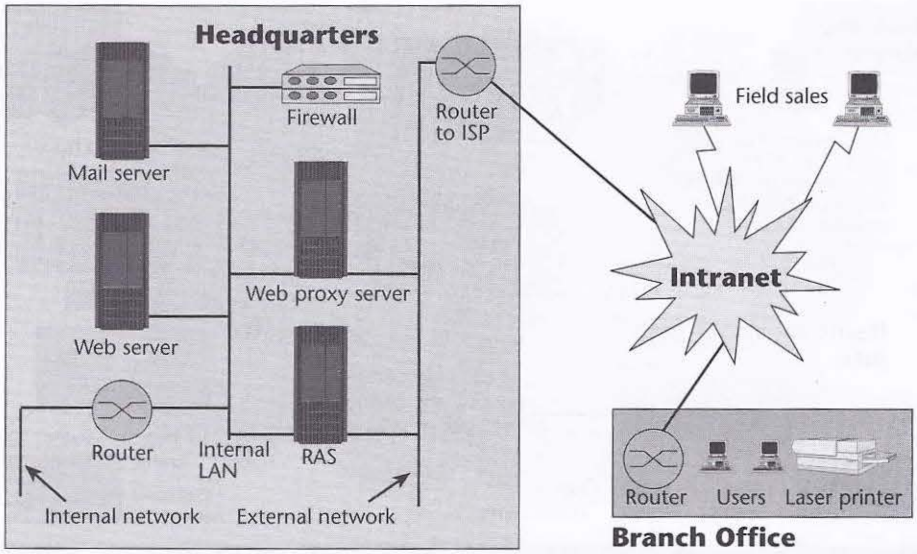

Figure 11.7 An intranet-based VPN

Intranet VPNs

Intranet VPNs are site-to-site connections (see Figure 11.7). The key objective of an intranet VPN is to replace or reduce the use of leased-line networks, traditional routers, and Frame Relay services. The cost savings in moving from private networks to Internet-based VPNs can be very high, in the neighborhood of 50% to 80% per year. Remember that Internet-based VPNs allow less control over the quality and performance of applications than do provisioned VPNs; this is a bit of a deterrent, and many clients still want to consider the Frame Relay- or ATM-based ATMs, which would provide better QoS. The savings might drop a bit, but the cost of a provisioned VPN would be substantially less than the cost of using leased lines.

There are a few key barriers to building out more intranets based on VPNs:

- No standardized approach to encryption
- Variance between vendors' products, which leads to interoperability problems
- Lack of standards regarding public key management
- Inability of today's Internet to provide end-to-end QoS

Remote Access VPNs

The most interesting and immediate VPN solution for most customers is the replacement of remote access servers. VPN remote access implementations can save customers from 30% to 70% over traditional dialup remote access server deployment. Remote access servers provide access to remote users, generally via analog plain old telephone service (POTS) lines, or, perhaps, ISDN connections, including dialup protocols and access control for authentication (administered by the servers). However, a remote access server requires that you maintain racks of modems, the appropriate terminal adapters for ISDN services, or DSL-type modems for DSL services. You also need remote access routers, which connect remote sites via a private line or public carriers and provide protocol conversion between the LANs and WANs. To have an internal implementation of remote access, you have to acquire all these devices, as well as the talent to maintain them.

If an enterprise needs remote access connections outside local calling areas, and/or if it needs encrypted communications, it is generally fairly easy to justify a VPN service over an enterprise-based remote access server. The initial cost of hardware for a VPN approach is about 33% less than the cost of hardware for a traditional dialup remote-access server deployment. The customer also saves on charges for local access circuits, and costly toll and international charges are eliminated.

By virtue of supporting a greater range of customers, a service provider that offers VPN-based remote access is more likely to support a wider variety of broadband access options, including xDSL, cable modems, and broadband wireless. VPN-based remote access also reduces the management and maintenance required with modem banks and remote client dial-in problems. For these reasons, remote access represents the primary application for which customers turn to VPNs. Figure 11.8 shows an example of remote access VPN.

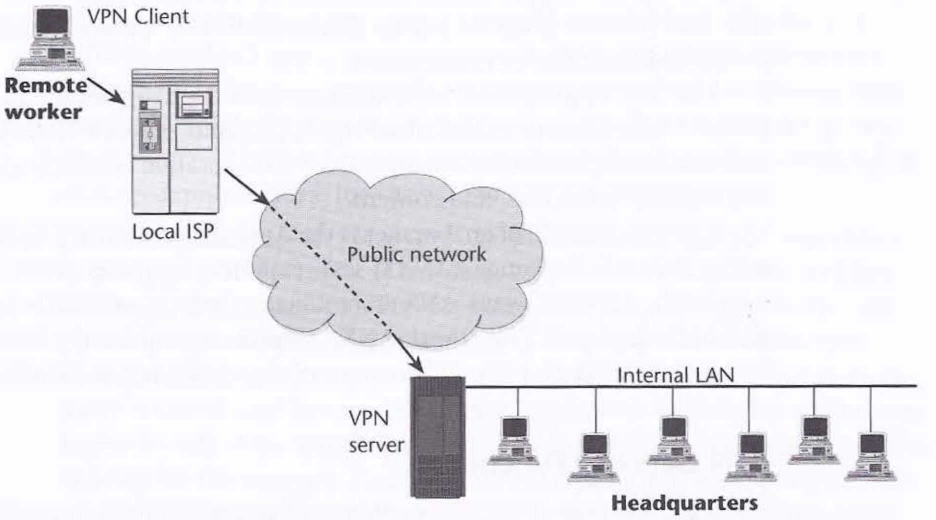

Figure 11.8 A remote-access VPN

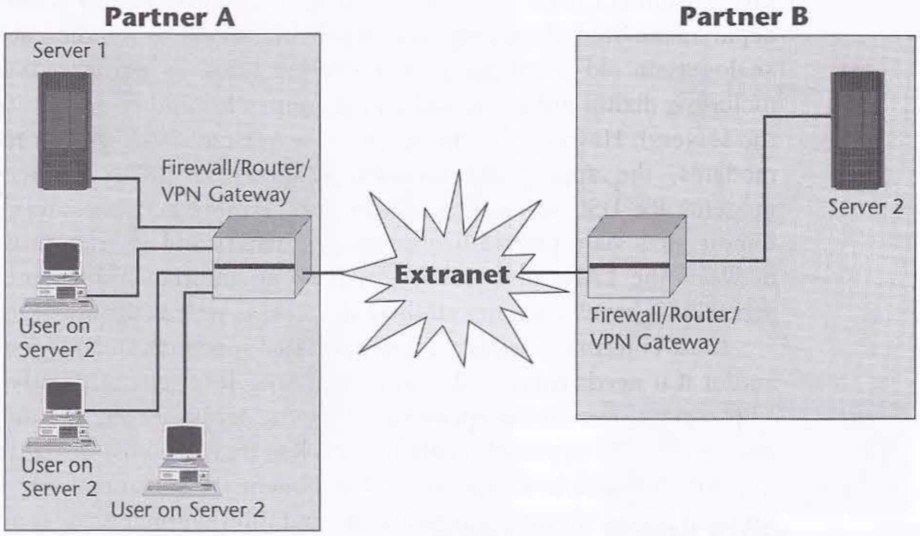

Figure 11.9 An extranet-based VPN

Extranet VPNs

Extranet VPNs allow an external organization to have defined access into an enterprise's internal networks and resources (see Figure 11.9). There are three major categories of extranets: supplier extranets, which focus on speeding communications along the supply chain; distributor extranets, which focus on the demand side and provide great access to information; and peer extranets, which create increased intraindustry competition.

The key applications for extranets include distribution of marketing and product information, online ordering, billing and account history, training policy and standards, inventory management, collaborative research and development, and e-mail, chat, news, and content.

A prime example of an extranet is the Automotive Industry Action Group's Automatic Network Exchange (ANX). This extranet comprises some 50,000 members worldwide. In many ways ANX is producing de facto standards for how extranets should be deployed. Check with ANX (www.anxo.com) for the latest information on how extranets are evolving and how one of the world's largest extranets is performing.

VPN Gateway Functions

The main purpose of the VPN gateways that are required to enable VPNs is to set up and maintain secure logical connections, called *tunnels*, through the Internet. Key functions of VPN gateways include packet encapsulation, authentication, message integrity, encryption, key exchange and key management, as well as firewalling, network address translation, access control, routing, and bandwidth management. The following sections describe these functions in detail.

Tunneling Protocols

Tunneling is a method of encapsulating a data packet within an IP packet so that it can be transmitted securely over the public Internet or a private IP network. The remote ends of the tunnel can be in one of two places: They can both be at the edges of the service provider's network, or one can be at the remote user's PC and the other at the corporate boundary router. Between the two ends of the tunnel, Internet routers route encrypted packets as they do all other IP traffic.

Three key tunneling protocols are needed in VPNs:

- **Point-to-Point Tunneling Protocol (PPTP)**—PPTP was developed by Microsoft, 3Com, and Ascend, and it is included in Windows 95, Windows 98, Windows Me, Windows NT, Windows 2000, and Windows XP. PPTP is a Layer 2 protocol that can work in a non-IP enterprise environment, which is one of its strengths for customers that use multiple protocols rather than using only IP. PPTP provides low packet overhead and good compression, but its weaknesses are on the security front: It does not provide encryption or key management in the published specification, and it essentially relies on the user password to generate keys. But all implementations of PPTP include Microsoft Point-to-Point Encryption (MPPE).
- **Layer 2 Tunneling Protocol (L2TP)**—The IETF promotes L2TP, which is a merger between PPTP and Cisco's Layer 2 Forwarding (L2F) protocol. L2TP is another Layer 2 protocol that can work in a non-IP enterprise environment. L2TP is used primarily by service providers to encapsulate and carry VPN traffic through their backbones. Like PPTP, it does not provide encryption or key management in the published specification (although it does recommend IPSec for encryption and key management).
- **IP Security (IPSec)**—IPSec is an IETF protocol suite that addresses basic data integrity and security. It covers encryption, authentication, and key exchange. IPSec involves a 168-bit encryption key, although the key size can vary, depending on the capabilities of each end of the connection. Recent drafts address encapsulating the secured payload, the key management protocol, and key creation. IPSec emphasizes security by authenticating both ends of the tunnel connection, negotiating the encryption protocol and key for the encrypted session, and encrypting and decrypting the session establishment data. However, IPSec is restricted to IP environments, each user is required to have a well-defined public IP address, and IPSec cannot run on networks that use network address translation.

Benefits and Evolution of VPNs

The main benefit of VPNs as compared to leased lines or Frame Relay is cost savings. VPNs also optimize environments with IP; they have less overhead than Frame Relay, and tunneling protocols may eliminate the need for proprietary encapsulation of protocols. Provisioned VPNs also have the additional benefits of Frame Relay and ATM in the administration of virtual circuits and QoS. VPNs also provide the capability to support dialup access, and greater redundancy is achieved in the network by virtue of meshed nets. Also, VPNs do not necessarily demand a digital fiber infrastructure end-to-end.

VPNs are undergoing an evolution, and various parameters still need to be addressed. Among those are the QoS guarantees. Effective traffic prioritization is at the heart of QoS, and current mechanisms that are available include Differentiated Services (DiffServ), Class-Based Queuing, Common Open Policy Service (COPS), and Multiprotocol Label Switching (MPLS). (These mechanisms are covered in Chapter 10, "Next-Generation Networks.") Other areas of evolution in VPNs are tiering of VPN services (that is, bandwidth tiering and different policy management), the capability to support autoprovisioning, and the emphasis on security.

QoS and security are the two most important considerations in administering VPNs, so uptimes, delays, and SLAs need to be structured. For example, QoS guarantees could be structured to promise 100% premises-to-premises network availability and a maximum latency guarantee of 80 milliseconds. Some vendors offer separate SLAs for dedicated and remote access. For dedicated access, the SLA offers an availability guarantee of 99.9% and a maximum latency guarantee of 125 milliseconds. On remote access SLAs, a busy-free dial availability guarantee of 97% is stipulated, and the latency guarantee specifies an initial modem connection speed of 26.4Kbps at 99%.

■ Security

Security is very important to the proper operation of VPNs. This section describes the available security mechanisms and the anticipated developments in the realm of standards for encryption and key management.

Firewalls

A *firewall* is typically defined as a system or a group of systems that enforces and acts as a control policy between two networks. It can also be defined as a mechanism used to protect a trusted network from an untrusted network—usually while still allowing traffic between the two. All traffic from inside to outside and vice versa must pass through the firewall. Only authorized traffic, as defined by the

Viruses

Virus is a term that's used broadly to refer to a program that is designed to interfere with computers' normal operations. The tab for all computer viruses in 1999 was, shockingly, greater than US$12 billion.

The term *virus* can be used more narrowly to refer to programs that move from one file to another and can be transmitted to other PCs via an infected file. They generally don't seek out the Internet or e-mail to spread.

Another type of virus is the *worm,* such as the Love Bug. Worms make use of a LAN or the Internet (especially via e-mail) to replicate and forward themselves to new users.

Finally, a *Trojan horse* hides within another program or file and then becomes active when someone opens the unwitting host.

A big part of administrating security involves managing viruses. The fact that we can deploy such functionality on a proxy server is very attractive.

local security policy, is allowed to pass through it. The system itself is highly resistant to penetration. A firewall selectively permits or denies network traffic.

There are several variations of firewalls, including these:

- A firewall can use different protocols to separate Internet servers from internal servers.
- Routers can be programmed to define what protocols at the application, network, or transport layer can come in and out of the router—so the router is basically acting as a packet filter.
- Proxy servers can be used to separate the internal network users and services from the public Internet. Additional functions can be included with proxy servers, including address translation, caching, encryption, and virus filtering.

Authentication

Another aspect of security is the authentication of users and access control, which is commonly handled by RADIUS. RADIUS servers are designed to block unauthorized access by remote users. RADIUS provides authentication, authorization, and accounting, and it relies on Challenge Handshake Authentication Protocol (CHAP) to authenticate remote users, which means that there's a back-and-forth dialogue to verify a user's identity. In fact, RADIUS makes use of CHAP, which uses a three-way handshake to periodically verify the identity of the peer throughout the connection. The server sends a random token to the remote workstation. The token is then encrypted, by using the user's password, and sent back to the server. The server performs a lookup to see whether it recognizes the password. If the values match, the authentication is acknowledged; if the values do not match, the connection is terminated. Because a different token is provided each time a remote user dials in, CHAP provides robust authentication.

Encryption

The best way to protect electronic data is to use encryption—that is, to encode data so as to render a document unreadable by all except those who are authorized to have access to it. The content of an original document is referred to as *plain text*. When encryption is applied to the document, the plain text is scrambled, through the use of an algorithm and a variable or a key; the result is called *ciphertext*. The key is a randomly selected string of numbers. Generally, the longer the string, the stronger the security.

There are two major categories of encryption algorithms: symmetric and asymmetric (also called *public key encryption*).

Symmetric Encryption

In symmetric encryption, the sender and the receiver use the same key or machine setup. There are two approaches to encoding data using symmetric encryption: block cipher and streaming cipher. With the block cipher approach, the algorithm encodes text in fixed-bit blocks, using a key whose length is also fixed in length. With the streaming cipher approach, the algorithm encodes the stream of data sequentially, without segmenting it into blocks. Both of these techniques require a secure method of reexchanging keys between the participants.

Symmetric encryption algorithms include the following:

- **Data Encryption Standard (DES)**—DES was developed in the 1970s and is very popular in the banking industry. It is a block cipher that encodes text into fixed-bit blocks, using a 56-bit key. DES is being replaced by the Advanced Encryption Standard (AES).
- **Triple DES (3DES)**—3DES is 168-bit encryption that uses three 56-bit keys. 3DES applies the DES algorithm to a plain text block three times.
- **Rivest Cipher 4 (RC4)**—RC4 is a streaming cipher technique; a stream cipher adds the output of a pseudorandom number generator bit by bit to the sequential bits of the digitized plain text.
- **Blowfish**—Blowfish is a 64-bit block code that has key lengths of 32 bits to 448 bits. Blowfish is used in more than 100 products, and it is viewed as one of the best available algorithms.

- **International Data Encryption Algorithm (IDEA)**—IDEA, developed by ETH Zurich, is free of charge for noncommercial use. It is viewed as a good algorithm and is used in Pretty Good Privacy (PGP) and in Speak Freely, a program that allows encrypted digitized voice to be sent over the Internet.
- **Twofish**—Twofish, developed by Bruce Schneier of Counterpane Internet Security, is very strong, and it was one of the five initial candidates for the AES.

According to the National Institute of Standards and Technology (NIST), it would take 149 trillion years to crack the U.S. government's AES, which uses the Rijndael algorithm and specifies three key lengths—128 bit, 192 bits, and 256 bits. In comparison, DES, which uses a 56-bit key, would take only a matter of hours using a powerful computer, but, of course, this is totally dependent on the speed of the hardware used for cracking the code; a typical desktop PC would require much more than a few hours to crack a 56-bit DES key.

Asymmetric Encryption

Key encryption requires a secure method for exchanging keys between participants. The solution to key distribution came, in 1975, with Diffie and Hellman's public key cryptography scheme. This permits the use of two keys, one of which can be openly published and still permit secure encrypted communications. This scheme later became known as *asymmetric key cryptography.*

Asymmetric cryptography can be used for authentication. After encrypting a signature by using a private key, anyone with access to the public key can verify that the signature belongs to the owner of the private key. As shown in Figure 11.10, the following are the steps in public key encryption:

1. User A hashes the plain text.
2. User A encrypts that hash value with a private key.
3. User A encrypts the plain text with user B's public key.
4. User B decodes the cipher text with the private key.
5. User B decodes the hash value, using User A's public key, thereby confirming the sender's authenticity.
6. User B compares the decrypted hash value with a hash value calculated locally on the just-encrypted plain text, thereby confirming the message's integrity.

Public key management involves the exchange of secrets that both ends use to produce random short-term session keys for authenticating each other. It is a method of encrypting data by using two separate keys or codes. The sender uses a public key that is generally provided as part of a certificate issued by a CA to

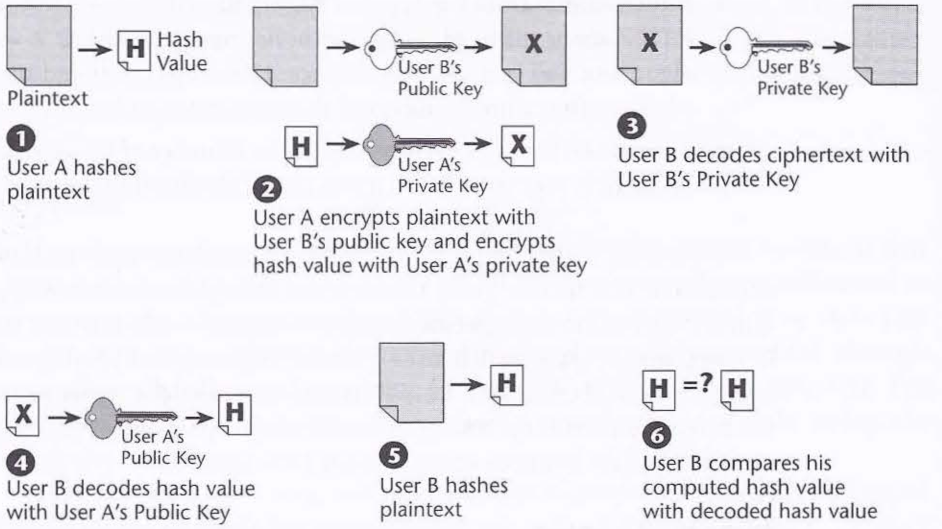
Hash Value
Plaintext
H
❶
User A hashes plaintext
User B's Public Key
X
H
User A's Private Key
X
❷
User A encrypts plaintext with User B's public key and encrypts hash value with User A's private key
X
User B's Private Key
❸
User B decodes ciphertext with User B's Private Key
X
User A's Public Key
H
❹
User B decodes hash value with User A's Public Key
H
❺
User B hashes plaintext
H =? H
❻
User B compares his computed hash value with decoded hash value

Figure 11.10 Encryption and authentication

scramble data for transmission. The receiver then uses a unique private key to decrypt the data upon receipt. The CA is an entity that, like a bank, is government regulated. It issues certificates that contain data about individuals or enterprises that has been verified to be authentic. In essence, the CA vouches for the authenticity of other parties so that their communications are secured.

Message authentication verifies the integrity of an electronic message and also verifies that an electronic message was sent by a particular entity. Before an outgoing message is encrypted, a cryptographic hash function—which is like an elaborate version of a checksum—is performed on it. The hash function compresses the bits of the plain-text message into a fixed-size digest, or hash value, of 128 or more bits. It is then extremely difficult to alter the plain-text message without altering the hash value.

Message authentication mechanisms include Message Digest-5 (MD5) and Secure Hash Algorithm-1 (SHA-1). MD5 hashes a file of arbitrary lengths into 128-bit value. SHA-1 hashes a file of arbitrary length into 160-bit value; it is more processor intensive but it renders higher security.

Public key management provides a secure method for obtaining a person's or an organization's public key, with sufficient assurance that the key is correct. There are three main public key algorithms: RSA (named for its creators, Rivest, Shamir, and Adelman), Diffie-Hellman, and PGP. RSA is 22 years old, and its security derives from the difficulty of factoring large prime integers. Diffie-Hellman is used mostly for exchanging keys; its security rests on the difficulty of computing discrete algorithms in a finite field, generated by a large prime number. PGP, which is a commercial product sold by Network Associates, was created in 1991. It is one of the most popular public key exchange (PKE) schemes.

Without a functioning universal public key infrastructure, we cannot reliably and easily acquire certificates that contain public keys for persons or organizations we want to communicate with. Standards are emerging, including Public Key Infrastructure (PKI), IETF Public Key Infrastructure X.509 (PKIX), Simple PKI (SPKI), and Public-Key Cryptography Standards (PKCS).

PKI is a system that provides protocols and services for managing public keys in an intranet or an Internet environment—it involves distributing keys in a secure way. PKI secures e-business applications such as private e-mail, purchase orders, and workflow automation. It uses digital certificates and digital signatures to authenticate and encrypt messages and a CA to handle the verification process. It permits the creation of legally verifiable identification objects, and it also dictates an encryption technique to protect data transmitted over the Internet. Trusted PKI suppliers include Entrust and VeriSign. PKI technology is now moving from pilot testing into the real world of e-commerce. Web browsers such as Microsoft Internet Explorer and Netscape Navigator include rudimentary support for PKI by providing an interface into a computer's certificate store, and browsers often include the certificates for some top-level CAs, so that the users can know, incontrovertibly, that the roots are valid and trustworthy.

IKE is the key exchange protocol used by IPSec, in computers that need to negotiate security associations with one another. A security association is a connection between two systems, established for the purpose of securing the packets transmitted across the connection. It supports preshared keys, which is a simplified form of key exchange. It does not require digital certificates. Every node must be linked to every other node by a unique key, and the number of keys needed can grow out of control; for example, 2 devices need 1 key, and 8 devices need 28 keys. New versions of IKE generate new keys through a CA. Legal and political problems will most likely delay widescale use of IKE.

One of the biggest hurdles e-commerce companies face is confirming the identity of the parties involved. Ensuring identity requires an encrypted ID object that can be verified by a third party and accepted by a user's browser. Personal digital IDs contained in the user's browser accomplish this. Historically, these client certificates have been used to control access to resources on a business network, but they can also contain other user information, including identity discount level or customer type. Third parties (that is, CAs) guarantee these types of certificates. The user's browser reads the server certificate, and if it's accepted, the browser generates a symmetric session key, using the server's public key. The server then decrypts the symmetric key, which is then used to encrypt the rest of the transaction. The transaction is then signed, using the user's digital ID, verifying the user's identity and legally binding the user to the transaction.

Digital Certificates

Digital certificates, based on the ANSI X.509 specification, have become a de facto Internet standard for establishing a trusting relationship using technology. Digital certificates are a method for registering user identities with a third party, a CA (such as Entrust, UserTrust, or VeriSign). A digital certificate binds a user to an electronic signature that can be trusted like a written signature and includes authentication, access rights, and verification information. CAs prepare, issue, and manage the digital certificates, and they keep a directory database of user information, verify its accuracy and completeness, and issue the electronic certificates based on that information. A CA signs a certificate, verifying the integrity of the information in it.

By becoming their own digital CAs, service providers can package electronic security with offerings such as VPN and applications services. Vendors that provide the technology required to set up as a CA include Baltimore Technologies (in Ireland), Security Dynamics Technologies, and Xcert.

Server certificates ensure Internet buyers of the identity of the seller's Web site. They contain details about the Web site, such as the domain name of the site and who owns it. Third parties, such as Thawthe in South Africa, then guarantee this information. Sites with server certificates post the CA, and Internet browsers accept their certificates for secure transactions.

There are still many security developments to come and there is a bit of unsettlement in this area. Standards need to be defined and formalized before e-commerce will truly be able to function with the security that it mandates. For now, these are the types of mechanisms that are necessary to ensure that your data remains with you.

■ VoIP

VoIP has been drawing a lot of attention in the past couple years. This section covers the types of applications that are anticipated for VoIP, as well as what network elements are required to make VoIP work and provide similar capabilities to what we're used to from the PSTN.

VoIP Trends and Economics

Although VoIP calling is used for billions of billed minutes each year, it still represents a very small percentage of the market—less than 5% overall. According to Telegeography (www.telegeography.com), 40% of VoIP traffic originates in Asia and terminates in North America or Europe; 30% travels between North America and Latin America; one-third of U.S. international VoIP traffic goes to Mexico, with future volume increases predicted for calling to China, Brazil, and India, and the rest moves among the U.S., Asia Pacific, and Western European regions. It is important to closely examine who will be using this and what carriers or operators will be deploying these technologies. Probe Research (www.proberesearch.com) believes that by 2002, 6% of all voice lines will be VoIP. This is still rather minor, given the fact that some have been saying that VoIP would have replaced circuit-switched calling by now. Piper Jaffray (www.piperjaffray.com) reports that minutes of communication services traveling over IP telephony networks will grow from an anticipated 70 billion minutes and 6% of all the PSTN traffic in the year 2003 to over a trillion minutes by the year 2006. In the United States alone, the PSTN is handling some 3.6 trillion minutes of traffic monthly.

Although VoIP has a very important place in telecommunications, it's important to realize that it is not yet taking over the traditional circuit-switched approach to accommodating voice telephony. The exciting future of VoIP lies in advanced and interesting new applications, an environment where voice is but one of the information streams comprising a rich media application. Many expect that sales of VoIP equipment will grow rapidly in the coming months and years. Part of the reason for this growth is that the network-specific cost for VoIP on dedicated networks is quite a bit lower than the cost of calls on circuit-switched networks—about US 1.1 cents per minute as compared with US 1.7 cents per minute. Using VoIP to carry telephony traffic greatly reduces the cost of the infrastructure for the provider, but at the expense of possibly not being able to maintain QoS. Potential savings are even greater if VoIP is implemented as an adjunct to data network.

Another factor encouraging customers to examine VoIP is the use of shared networks. Because IP emphasizes logical rather than physical connections, it's easier for multiple carriers to coexist on a single network. This encourages cooperative sharing of interconnected networks, structured as anything from sale of wholesale circuits to real-time capacity exchanges. Also, VoIP can reduce the barriers to entry in this competitive data communications world. New companies can enter the market without the huge fixed costs that are normally associated with the traditional circuit-switched network models. Furthermore, because IP telephony will enable new forms of competition, there will be pressure to better align government-controlled prices with underlying service costs. International VoIP services are already priced well below the official rates and some of VoIP's appeal is that it eliminates the access charges interexchange carriers normally have to pay to interconnect to the local exchange carrier. In the United States, these charges range from US 2 cents to US 5 cents per minute.

Advantages of VoIP

The key benefits of VoIP are cost savings associated with toll calls, enhanced voice services, and creative and innovative new applications. The key concerns related to

Regulations Related to VoIP

It's one thing to approach telephony on the Internet such that the incumbent is protected from competition with other voice telephony services on the Internet. But stating that voice on the Internet should not be allowed would be to cut your own throat. All the exciting new applications on the Internet do involve the use of multimedia applications, and voice is part of that overall stream. So, we have to be very careful about what we're regulating—whether it's voice, which is increasingly part of a larger application set, or whether it's traditional voice telephony.

VoIP are voice quality compared to that in today's PSTN; the cost of QoS to ensure the same quality as in the PSTN; security; the current lack of compelling applications; and regulatory issues, such as whether voice will be allowed on the Internet and whether voice will be treated as an altogether different environment—as a converged, integrated application.

VoIP Applications

VoIP includes any set of enabling technologies and infrastructures that digitize voice signals and transmit them in packetized format. Three major network architectures can be used in support of VoIP applications:

- Voice-over intranets, which could be based on leased lines, Frame Relay, ATM, or VPNs
- Voice-over extranets, which could also be based on leased lines, Frame Relay, ATM, or VPNs
- Voice over the public Internet

The following sections discuss some of the key issues related to VoIP applications.

IP Long-Distance Wholesale

So far, the most compelling business case for VoIP has been in IP long-distance wholesale, where there are clear financial benefits and low barriers to entry. Early pioneers in this area include iBasis, ITXC, and Level 3, which predominantly offer IP services to domestic and international carriers, but also offer services to corporations and other service providers. What the customers gain by doing business in this fashion is a reduction in cost associated with carrying their traffic over expensive toll or international transit links.

In IP long-distance wholesale, the voice service levels must match those of the PSTN. End customers of the international carriers expect to perceive the same voice

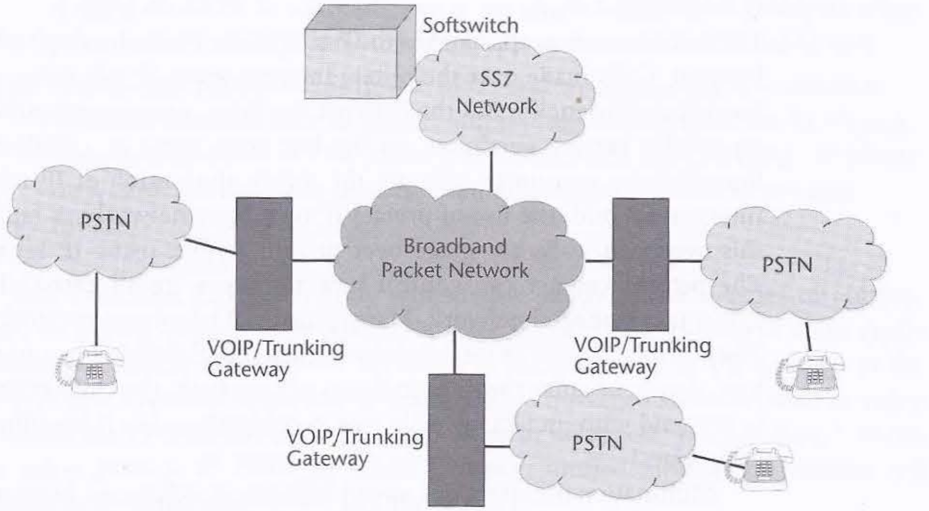

Figure 11.11 A converged long-distance network

quality throughout. How can providers guarantee that when it's almost impossible to control QoS over the public Internet? Even in the case of IP backbones, QoS depends on the underlying architecture used. The solution lies in smart management of packet latency, to ensure circuit-like behavior inside the IP network. For example, iBasis developed a proprietary routing algorithm that monitors performance on the Internet; when it detects that congestion levels may affect the quality of the voice, it switches the calls over to the circuit-switched network, thereby ensuring that customers experience the high quality that they expect end-to-end.

The IP long-distance wholesale environment takes advantage of a converged voice/data backbone by using trunking gateways to leverage the PSTN (see Figure 11.11). This allows support and processing of voice calls. The trunking gateways enable connection of the data network to the PSTN, to support long-haul carrying of the switched calls. In addition, switching services can be added to the data networks through the use of softswitches. (The functions and types of softswitches and gateways that make up the new public network are discussed later in this chapter.)

These are main issues in selecting providers of IP long-distance wholesale:

- **Voice quality versus bandwidth**—How much bandwidth do you use to ensure the best quality?
- **Connecting to the customer**—How many services need to be supported (voice, data, dialup modem, fax, ISDN, xDSL, cable modem)?
- **Maintaining voice quality**—As bandwidth becomes constrained, how do you maintain the voice quality?

IP Telephony

There are two main approaches to IP telephony. First, there's IP telephony over the Internet. Calls made over the public Internet using IP telephony products provide great cost-efficiencies. But the Internet is a large, unmanaged public network, with no reliable service guarantee, so the low costs come at a trade-off. International long-distance consumer calls are the major application of IP telephony over the Internet. Second, the use of private IP telephony networks is rapidly emerging. In this approach, calls are made over private WANs, using IP telephony protocols. The network owner can control how resources are allocated, thereby providing QoS and a managed network. Many private IP telephony networks are being built. They enable an enterprise to take advantage of its investments in the IP infrastructure. Again, because this is a single-owner network, the QoS issues are much easier to contend with; in fact, a single-owner network makes it possible to contend with QoS issues!

Multinational enterprises spend billions of dollars on international voice services each year, so the savings that IP telephony offers is compelling. The cost benefit of running voice services over a private IP network is on the order of 20% or more savings on international long distance, as compared to using traditional voice services. Private IP transport platforms will be increasingly deployed, therefore, as an enterprisewide telephony option.

Recent deployment of IP local exchange products, coupled with low-bandwidth, high-quality voice compression, creates a solid foundation for extending business telephone service to telecommuters at home or on the road. The efficiencies of IP packet technology, coupled with the ITU G.723.1 voice compression standards at 6.4Kbps, enable road warriors and small office/home office workers to have a complete virtual office over a standard 56Kbps Internet modem connection to the office. The really great feature of this environment is that your current location is your office and your IP phone rings wherever you are. However, this requires an IP local exchange—a carrier-class product that resides in the service provider network and provides PBX-like telephony service to multiple business and telecommuter customers. It also requires a softswitch (that is, call-agent software) that's used for purposes of managing call processing functions and administration. Also, end-user services are delivered via IP Ethernet phones or analog telephones that use Ethernet-to-analog adapters.

There are three major categories of IP phones:

- **POTS phone**—The advantage of the POTS phone is high availability and low price. The disadvantage is that it has no feature buttons and the required Ethernet-to-analog adapter is quite costly.
- **Soft phone**—A soft phone is software that runs on the user's PC and graphically resembles a telephone. Its advantage is low price. Its disadvantage is that

it relies on the PC sound card, and it can create volume level problems when you switch between it and other applications that use the PC sound card.

- **IP Ethernet phone**—This device looks and works just like a traditional multiline business display phone, and it plugs into an Ethernet RJ-45 jack. It's priced similarly to PBX phones, at US$300 and up. Emerging "IP phone on a chip" technologies promise dramatically lower prices in the near future.

The evolution of IP telephony will involve many different types of applications, including long-distance wholesale voice services; the support of voice applications for campus or enterprise networks in bringing VoIP to the desktop in the form of new advanced applications that involve converged streams (such as video conferencing or multimedia in the establishment of remote virtual offices); Internet smart phones; IP PBXs; IP centrex service; unified messaging; Internet call waiting; and virtual second-line applications.

VoIP Enhanced Services

Another approach to supporting voice services is to look toward enhanced services. There are two categories of enhanced services:

- **Transaction-oriented services**—These services include Click-N-Call applications, interactive chat, Surf-With-Me, videoconferencing, and varieties of financial transactions.
- **Productivity-enhancing services**—These services include worldwide forwarding, multiparty calling, a visual second line, unified messaging, collaboration, access to online directories, visual assistance, CD-quality sound, personal voice response, and video answering machines.

The key to enhanced services is not cost savings, but cost savings are realized through toll bypass, QoS differentiation, the capability to support remote access, and the capability to create new forms of messaging. Because of the cost savings and features available, the use of enhanced services will grow by leaps and bounds over the next several years.

VoIP is part of a larger trend toward innovative voice-enabled Internet applications and network interactive multimedia. This trend includes various facilities to enhance e-commerce, customer service, converged voice and visual applications, new intelligent agents and various forms of bots, and e-calling campaigns. These sorts of advanced services make it possible to gain greater value from the IP investments that have been made, and at the same time, they create interesting new revenue streams with altogether new businesses.

We'll see VoIP applications increasingly used in a number of ways. VoIP applications will be included on Web-based call centers as automatic call-backs from customer service-based phone numbers entered into a Web page; as multiparty conference calls, with voice links and data sharing, initiated also from a Web page; and in the process of reviewing and paying bills. The key is to blend rich, Internet-based content with a voice service. An example of an emerging application that illustrates such innovation is online gaming. InnoMedia and Sega Enterprises are integrating InnoMedia Internet telephony into Sega Dreamcast game consoles to allow game players worldwide to voice chat with each other while playing games. This device can also be used to cost-effectively place calls in more than 200 countries through InnoSphere, InnoMedia's global network. For example, the rate from the United States to Hong Kong will be US2 cents per minute, from the United States to the United Kingdom it will be US5 cents, and from the United States to Japan, Australia, and most of Europe, it will be US9 cents.

Another example of an interesting new VoIP application is Phonecast, a media network of Internet-sourced audio channels for news, entertainment, and shopping, available to telephones. Created by PhoneRun and WorldCom, Phonecast is modeled after television and radio broadcasting, and it allows callers to create a personal radio station and direct it by using simple voice commands. This is the first of a series of innovative content and service partnerships, assembled to form a comprehensive voice-portal product line.

VoIP Service Categories

There are several main VoIP service categories:

- **Enterprise-based VoIP**—In enterprise-based VoIP, whether for the LAN or WAN, specialized equipment is required at the customer site.
- **IP telephony service providers**—These providers are generally involved in toll-bypass operations. They do not require specialized equipment at the customer site, but they may require additional dialing procedures to gain access to the network. Currently, multistage dialing is one of the problems we still face: You have to dial a seven- or eight-digit number to gain access to your ISP, and then you have to dial a string of digits for the authentication code, and then you have to dial the string of digits corresponding to the number you want to reach. Single-stage dialing will remedy this situation in the very near future.
- **Converged service providers**—These companies will bundle together voice, data, and video services.
- **Consumer VoIP**—Consumer VoIP is generally geared toward consumer connections over the public Internet.

VoIP Network Elements

VoIP may seem like rocket science compared to conversations, but the concept is really quite simple: Convert voice into packets for transmission over a company's TCP/IP network. Two characteristics determine the quality of the VoIP transmission: latency and packet loss. Latency is the time it takes to travel from Point A to Point B. The maximum tolerance for voice latency is about 250 milliseconds, and it's recommended that the delay be less than 150 milliseconds. Small amounts of packet loss introduce pops and clicks that you can work around, but large amounts of packet loss render a conversation unintelligible. With too much packet loss, you would sound like you were saying "Da dop yobla bleep op bop," because little packets with much of your conversation would have been lost in congestion and could not be retransmitted while working within the delay requirements of voice. Hence, packet loss with VoIP can cause big chunks of a conversation to be lost. (We will talk about ways to resolve that a little later in this chapter.)

VoIP gateways have allowed IP telephony applications and new, innovative VoIP applications to move into the mainstream. Other features that have helped the development of VoIP are Internet telephony directory, media gateways, and softswitches, as well as telephony signaling protocols.

VoIP Gateways

VoIP gateways bridge the traditional circuit-switched PSTN and the packet-switched Internet. Gateways overcome the addressing problem. A couple years ago, for two VoIP users to communicate, they had to be using the same software, they had to have sound cards and microphones attached to their PCs, and they had to coordinate a common time during which both would be online in order to engage in a VoIP session. Gateways have made all that unnecessary, and now the only requirement is that you know the user's phone number. Phone-to-PC or PC-to-phone operation requires the use of only one gateway. Phone-to-phone operation requires two gateways, one at each end.

VoIP gateway functionality includes packetizing and compressing voice; enhancing voice quality by applying echo cancellation and silence suppression; dual-tone multifrequency (DTMF) signaling support (that is, touch-tone dialing); routing of voice packets; authentication of users; address management; administration of a network of gateways; and the generation of call detail records that are used to create bills and invoices.

To place a call over a VoIP network, the customer dials the number the same way as on a traditional phone. The edge device, the VoIP gateway, communicates the dialed number to the server, where call-agent software—that is, a softswitch—determines what is the appropriate IP address for that destination call number and returns that IP address to the edge device. The edge device then converts the voice signal to IP format, adds the given address of the destination node, and sends the signal on its way. If enhanced services are required, the softswitch is called back into action to perform the additional functions. (The softswitch is also referred to as a Class 5 agent because it behaves like a local exchange or a Class 5 office.)

There are two primary categories of VoIP gateways:

- **Gateways based on existing router or remote access concentrator (RAC) platforms**—The key providers here include the traditional data networking vendors, such as 3Com, Cisco, Lucent, and Motorola. As incumbent equipment suppliers to ISPs, the data networking vendors are capturing the largest percentage of these sales. They represented the majority of VoIP gateway sales through 2000 because ISPs were buying gateways at a fast rate based on the significant wholesale opportunity available to larger carriers.
- **Server-based gateways**—These are designed from the ground up to support VoIP. Key providers of server-based gateways include telecommunications vendors, as well as companies specifically designed for this business; Clarent, Ericsson, Lucent, NetSpeak, Nortel, Nuera, and VocalTec are among the vendors involved. These gateways will overtake router and RAC solutions as incumbent carriers deploy more server-based gateways with extensive call server and signaling capabilities.

More and more merger and acquisition activities will lead to blended solutions, causing the distinction between the different types of gateways to blur. RAC- and router-based gateways will take on more enhanced call-server characteristics as a result. The market segments for the two categories, then, are composed of the following:

- **Enterprise VoIP gateways**—These gateways are customer premise equipment deployed between a PBX and a WAN device, typically a router, to provide call setup, call routing, and conversion of voice into IP packets and vice versa.
- **VoIP routers**—Voice cards perform packetization and compression functions and are inserted into a router chassis. The router then directs the packets to their ultimate destination.
- **IP PBXs**—An IP PBX is an infrastructure of distributed telephony servers that operates in packet-switched mode and offers the benefits of statistical multiplexing and IP routing. We are still in the early days for IP PBXs, although they are beginning to emerge as a viable alternative. A key concern is reliability. (IP PBXs are discussed in more detail later in this chapter.)
- **Service-provider VoIP gateways**—These are used to aggregate incoming VoIP traffic and route the traffic accordingly. The role is analogous to that of the local exchange. Challenges include the local loop competition among the incumbent carriers, quality concerns, shortage of product, interoperability

issues, the lack of hot-swappable and redundant support, and the lack of Network Equipment Building Systems (NEBS) compliance.

- **VoIP access concentrators**—VoIP cards fit into an existing dial access concentrator.
- **SS7 gateways**—SS7 gateways are critical to enabling us to tap into the intelligence services that enhance so much of the telephony activity on the PSTN.

There are many gateway vendors. All gateway vendors share the need for digital signal processors and embedded software solutions that provide for silent suppression, echo cancellation, compression and decompression, DTMF signaling, and packet management. Therefore, another very important part of this equation is the component vendors. Manufacturers of VoIP equipment need to continue to make quality improvements in the underlying technology. This includes addressing interoperability between different gateway vendors' equipment; improving the tradeoffs between cost, function, and quality; and introducing single-stage dialing and the ability to dial from any telephone.

Internet Telephony Directory

An Internet telephony directory is a vital piece of the VoIP puzzle, so this section talks a little bit about the IETF Request for Comment 2916, also known as ENUM services. ENUM services convert telephone numbers into the Internet address information required to support all forms of IP-enabled communication services, including real-time voice, voicemail, fax, remote printing, and unified messaging. In other words, ENUM is a standard for mapping telephone numbers to IP addresses. DNS translates URLs to IP addresses, and EMUM uses the DNS to map a PSTN phone number (based on the E.164 standard) to the appropriate URLs.

ICANN is considering three proposals for the .tel domain. The applicants are NetNumber, which currently runs the Global Internet Telephony Directory (an implementation of ENUM that is used by IP-enabled platforms to convert standard telephone numbers into Internet address information), Number.tel, and Telnic based in the United Kingdom. The ITU is trying to advance an implementation of the IETF ENUM standard under the domain e164.arpa. In this implementation, control of telephone number addressing on the Internet would be distributed to the more than 240 national public network regulatory bodies that administer telephone numbers for the PSTN.

Media Gateways

Media gateways provide seamless interoperability between circuit-switched, or PSTN, networking domains and those of the packet-switched realm (that is, IP, ATM, and Frame Relay networks). They interconnect with the SS7 network and enable the handling of IP services. They're designed to support a variety of telephony signaling protocols. Media gateways are designed to support Class 4, or toll-switch, functions, as well as Class 5, or local exchange, services. They operate in the classic public network environment, where call control is separate from media flow. They support a variety of traffic—including data, voice, fax, and multimedia—over a data backbone. Enhanced applications of media gateways include network conferencing, network-integrated voice response, fax serving, network, and directory services.

As shown in Figure 11.12, media gateways fit between the access and core layers of the network, and they include several categories: VoIP trunking gateways, VoIP access gateways, and network access service devices. They provide service interconnection or intercarrier call handling. The trunking gateways interface between the PSTN and VoIP networks, terminating trunks associated with SS7 control links. These Time Division Multiplexed trunks carry media from an adjacent switch in the traditional circuit-switched network, and the adjacent switch generally belongs to another service provider. (Depending on the agreements between service providers, these are also referred to as cocarrier trunks or feature group D trunks.) The trunking gateways manage a large number of digital virtual circuits. The access gateways provide traditional analog or ISDN interfaces to the VoIP networks; they are devices that terminate PSTN signaling and media, and they connect to PBXs, as well as to traditional circuit switches, such as the Class 5 and Class 4 offices. With network access servers, you can attach a modem to a telephone circuit

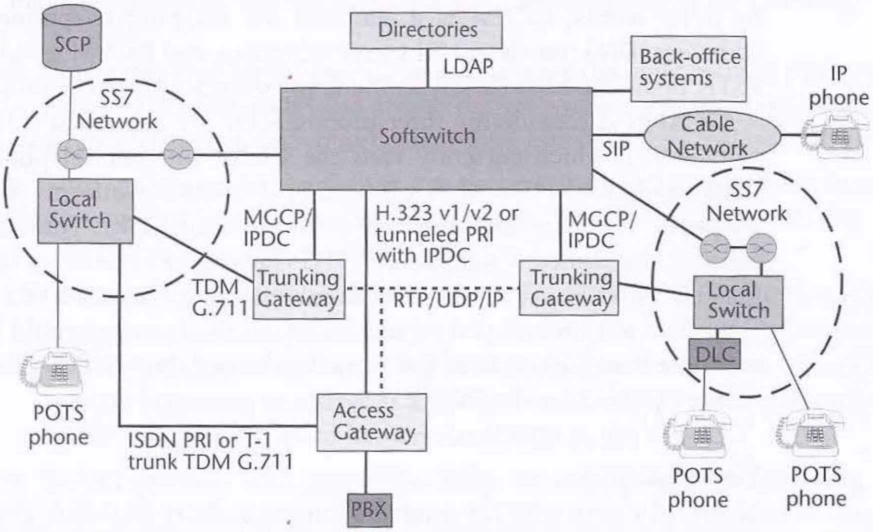

Figure 11.12 VoIP network architecture

and provide data access to the Internet, so that you can attain managed modem service by using cocarrier trunks.

VoIP Softswitches

Call-control intelligence is outside the media gateways and VoIP gateways; it is, instead, handled by a *softswitch*, also referred to as a *media gateway controller* or *call agent*. The softswitch implements the service logic. It controls external trunking gateways, access gateways, and remote access servers. Softswitches run on commercial computers and operating systems, and they provide open application programming interfaces.

A softswitch is a software-based, distributed switching and control platform, and it controls the switching and routing of media packets between media gateways, across the packet backbone. Softswitches provide new tools and technologies to build services in a more productive Internet-based service creation environment. Operators are advised to adopt a "service separation" strategy and to distribute applications throughout the network, avoiding the monolithic closed system that is similar to the circuit-switched environment. We can use application servers to partition enhanced telecommunications services and to determine what interface protocol to select for facilitating interoperability between the softswitches and the applications servers.

The softswitch functionally controls the voice or data traffic path by signaling between media gateways that actually transport the traffic (see Chapter 10). The gateway provides the connection between an IP or ATM network and the traditional circuit-switched network, acting a lot like a multiprotocol cross-connect. The softswitch ensures that a call's or a connection's underlying signaling information—automatic number identifiers, billing data, and call triggers—are communicated between the gateways. Softswitches must reuse intelligent network services through an open and flexible directory interface, so they provide a directory-enabled architecture with access to relational database management systems, and to Lightweight Directory Access Protocol (LDAP) and Transaction Capabilities Applications Part (TCAP) directories. Softswitches also offer programmable back-office features, along with advanced policy-based management of all software components.

The softswitch is a very important element in the new public network. It is what enables the media and trunking gateways to communicate with the underlying infrastructure of the PSTN and thereby to draw on the service logic needed to support telephony activities. In addition, softswitches will be able to reach to new application servers on which new generations of applications have been designed for new versions of enhanced services.

Telephony Signaling Protocols

New generations of signaling and IP telephony control protocols are emerging, and their purpose is to control the communication between the signaling gateway and IP elements. Since the early days of exploring the nature of VoIP and creating devices to enable it, a number of telephony signaling protocols have been considered. Some of the contenders have been H.323, Internet Protocol Device Control (IPDC), Signal Gateway Control Protocol (SGCP), Multimedia Gateway Control Protocol (MGCP), Multimedia Gateway Control (MEGACO), Session Initiation Protocol (SIP), and IP Signaling System 7 (IPS7). Many of those contenders have combined, so this section focuses on the ones that have the strongest presence and potential today.

H.323 The ITU H.323 version 2 specification is based on ISDN standards and limited to point-to-point applications. Version 2 requires multipoint control units (MCUs) to manage multiple sessions. H.323 version 2 provides much of the foundation for exchange of voice and fax messages. The advantage of H.323 is that it is the most mature of the telephony signaling protocols, so many vendors offer it and vendor interoperability is good. On the other hand, H.323 is not as robust as some of the newer entrants, so other protocols on the horizon might eclipse H.323 before too long.

MCGP Bellcore and Level 3 merged their respective SGCP and IPDC specifications into MCGP. In MCGP, softswitches provide the external control and management, so MCGP is becoming a good way to connect an IAD to a gateway.

MEGACO MEGACO is also called H.248 and it is another emerging ITU standard. MEGACO describes how the media gateway should behave and function.

SIP SIP (IETF Request for Comment 2543) is an application-layer control, or signaling protocol, for creating, modifying, and terminating sessions with one or more participants. SIP is used to set up a temporary session, or call, to the server so that the server can execute the necessary enhanced service logic. These sessions may include Internet multimedia conferences, Internet telephony, or multimedia distribution. Linking caller ID to Web page content can link the status of a mobile phone with instant messaging. Members in a session can communicate via multicast or via a mesh of unicast relations, or by a combination of these. This is increasingly popular as the protocol between softswitches and application servers.

LDAP LDAP is the standard directory server technology for the Internet. LDAP enables retrieval of information from multivendor directories. In fact, LDAP 3.0 provides client systems, hubs, switches, routers, and a standard interface to read and write directory information. Directory-oriented services best suited for an LDAP lookup include unified messaging, free phone (that is toll-free number translation), calling name service, and Internet phone number hosting. Remember that as the Internet moves forward, it must connect with the underlying intelligence in the PSTN.

IPS7 The SS7 network acts as the backbone for the advanced intelligent network. SS7 provides access to all the advanced intelligent network features, allows for efficient call setup and teardown, and interconnects thousands of telephony providers under a common signaling network. The capability to communicate with the SS7 network is essential for all service providers. It gives next-generation local exchange carriers access to an existing base of service features, and it ensures that packet-based telephony switching gateways can support key legacy service and signaling features. The interconnection between a legacy circuit switch provider, such as the incumbent local exchange carrier, and a competitive local exchange carrier operated over a packet backbone would include the gateway switch to packetize and digitize the voice coming from the Class 5 office, and the SS7 gateway to provide access into the underlying intelligent network infrastructure. (Chapter 5, "The PSTN," discusses SS7 and next-generation gateway switches in more detail.)

Next-Generation Standards and Interoperability

Next-generation network standards are widely deployed across the globe and are generating billions of dollars in service revenue. Packet-enabled intelligent networks will enhance the revenue stream with new technology to provide intelligent networking services, such as local-number portability, carrier selection, personal numbers, free phone, prepaid call screening, call centers, and voice VPNs. End-to-end, next-generation networks function as seamlessly interoperating wholes; they consist of the legacy-based circuit-switched network, with its underlying SS7 and service logic delivering today's enhanced features, as well as a packet-based network for transport efficiencies that can also be served by new-generation IP servers and enhanced applications, for features we haven't yet thought of.

There are a few key groups to be aware of in the area of standards and interoperability for next-generation networks. There's iNOW!, which stands for Interoperability NOW!, and its members include Ascend, Cisco, Clarent, Dialogic, Natural MicroSystems, and Siemens. These members will interoperate also with Lucent and VocalTec, as well as each other. iNOW! advocates interoperability and certification based on H.323.

The Technical Advisory Committee (TAC), formed by Level 3 Communications, includes 3Com, Alcatel, Ascend, Cisco, Ericsson, Level 3, and others.

The International Softswitch Consortium is focused on enabling softswitch technology and applications on an IP infrastructure. This group advocates interoperability and certification based on H.323, SIP, MGCP, and Real-Time Transfer Protocol (RTP). It is working to develop and promote new standardized interfaces for portable applications, which ride on top of an IP-based softswitch network. The International Softswitch Consortium has more than 68 member companies.

Finally, the Multiservice Switching Forum (MSF) is an open-membership organization committed to developing and promoting implementation agreements for ATM-capable multiservice switching systems. The goal of MSF is to develop multiservice switching with both IP- and ATM-based services, and its founding members are WorldCom, Cisco Systems, Telcordia, AT&T, Alcatel, Lucent, British Telecom, Fujitsu Network Communications, Lucent Technologies, Nortel Networks, Siemens, Telecom Italia, Telia AB, and Qwest.

IP PBXs

IP PBXs are in the very early stages, and they will present some benefits as well as some challenges. Companies can take advantage of IP-based intranets that have been set up between headquarters and remote locations to cost-effectively integrate voice and data traffic. The key strength of IP PBXs is their capability to network over existing IP networks. Because the information is programmed into the phone, phones can be relocated by simply unplugging and moving them. It is also easier to network over existing IP WANs, as long as there is adequate bandwidth to support voice traffic.

Among the challenges to the convergence of IP PBXs is that we expect them to provide high reliability and high availability, which we always require with telephony. Telephony-grade servers are classified as fault tolerant when they achieve 99.99% (that is, four nines) survivability. The standard for most PBX voice systems is 99.999% (that is, five nines), so four nines is quite a bit less than what we're accustomed to. The industry is slowly embracing Windows NT and Windows 2000 for core call processing, but some feel that these products are not reliable enough in their current form. To be fair, research on NT stability and security shows that almost always the problems are a result of poor or improper administrative procedures, not a result of problems in the operating system itself. As NT administrators have gained operational experience, the reliability and security of Windows-based data centers has improved. In summary, customer concerns include security, reliability, survivability, operability, maintainability, and accountability.

Another important issue related to IP PBXs is power distribution. PBXs have internal power distribution plants to support processing memory and internal interface circuit cards. All analog and proprietary digital telephones are line powered by the centralized PBX, using standard unshielded twisted-pair (UTP) wiring. Larger PBXs often have redundant power conversion and distribution elements throughout the cabinet design, and fluctuations in power—such as spikes and surges—are also regulated by the PBXs. Although there are Ethernet switches that can deliver power to the desktop via Category 5 cabling, they are just being introduced, and standards have not yet been developed for this.

Voice quality is another big issue with IP PBXs. The voice quality delivered over an IP PBX has to match that of the PSTN, so VoIP systems will need to meet stringent technical requirements to manage delay and echo, which are affected by the amount of compression and the type of codec used, as well as by the QoS capabilities of the underlying transport network. Voice quality will become a new performance variable, with various levels available and reflected in the pricing of services.

Another issue related to IP PBXs is network QoS. Voice QoS must remain adequate when it shares the network with bandwidth-intensive data applications. Packet loss must be minimized, and latency must be reduced. We still need to figure out how much voice traffic the data network can accept before voice, data, and video start to degrade.

Features and functionality are other issues. PBXs in general have 400 to 500 features, whereas IP phone systems provide only about 100 features.

There are also issues surrounding distance limitations. Fast Ethernet Category 5 cabling is limited to distances of 330 feet (100 meters), whereas PBXs support analog phone extensions over UTP at up to 2 miles (3.5 kilometers) and proprietary digital phones at up to 1 mile (1.5 kilometers).

Another issue is the lack of management systems. Systems designed to accommodate moves, adds, and changes, as well as troubleshooting, need to be developed.

There are also security questions (for example, Will voice over the LAN demand encryption of voice traffic?) and issues related to legacy voice investments (that is, enterprises generally protect investments in their existing equipment). Finally, we face a lack of a really compelling value proposition. However, PBXs are migrating toward a future in which IP-based packet transport will replace circuit-switched Time Division Multiplexing.

This market is poised for major change in the next several years. PBXs are migrating toward telephony server models, in which a nonproprietary platform will perform the call control and feature provisioning. But these are still the early days. Through integration, we will eventually have a much more cost-efficient platform for network services.

The Future of VoIP

VoIP is very important, and it's part of a larger application set that enables the integration of voice, video, data, and images. It is the early days for VoIP as well. Today, VoIP accounts for only a very small amount of global voice traffic. With VoIP we face issues of interoperability, scalability, and the number of features that can be supported. We face issues of whether the incumbents are motivated to replace all Class 5 exchanges with next-generation telephony. IP QoS is still immature, as these are the early days.

■ Multimedia on the Internet: Streaming Media

There's such a wide variety of applications for multimedia on the Internet that we're only just beginning to consider where and how we can use visualization and other sensory information. There are three major categories of multimedia on the Internet: communications applications (including VoIP, video telephony, and video and multimedia conferencing applications), computer applications (including interactive rich media, videomail, and streaming audio, video, and media content), and entertainment applications (including broadcast video, video-on-demand, and network games). This section concentrates on streaming media, and in Chapter 15, "The Broadband Home and HANs," we'll explore the fantastic future of smart devices and sensory networks.

Streaming Media Trends

Most streaming audio and video on the Internet today (for example, music, ads) is entertainment or consumer oriented. Many companies are now also beginning to use streaming media as a business tool. For example, I offer e-learning solutions on a streaming basis.

The appeal of streaming media is evident: Audio and video grab people's attention and can quickly present information that is easy to absorb and retain. Streaming media also allows for novel ways to reach clients, employees, and prospective customers. Audio and video are highly effective in sales and marketing, advertising, corporate communications, motivation, training, instruction, and customer support. Businesses realize gains in revenues and greater efficiencies and decreased costs for information delivery by turning to streaming media. By getting audio and video content in front of an audience, you can charm that audience, but it is costly and can be difficult. Downloading big files is time-consuming. Video file sizes run into the tens of megabytes, and a 5-minute video can be as large as 55MB. Audio files are often several megabytes.

Streaming is the solution to the problem of downloading large media files. Using a streaming media player, such as Apple's QuickTime, Microsoft's Windows Media Player, or RealNetworks's RealAudio and RealVideo, a user can play audio and/or video within seconds after the first bits of the stream hit the user's computer. These players support both live Internet broadcasts and video-on-demand, in which the streaming server keeps a copy of the content so that clients can request it at any time. Millions of people access some form of streaming content—audio or video—every day, and the number of streams available on the Internet is growing phenomenally. With all these people using streaming media and so many streaming media offerings available, streaming search technology is an emerging requirement in next-generation networks.

Streaming Media Applications

There are many applications for streaming media. Streaming media can be used as a novel way to reach and communicate with employees, clients, and partners. And streaming media is becoming more and more necessary where you need to respond to regulations that require a full disclosure or a method for informing a very wide audience. Another key application is virtual roadshows, such as pre-IPO presentations to potential investors. Product demonstrations are a very strong application of streaming media (for example, launches and rollouts of new products, virtual education and training). Some of the key customers for streaming media at this point are entertainment, financial institutions, health care, and education.

With streaming media, the content provider must digitize the content and set up a server that is specific to the client. The provider's total hardware and software costs are typically only in the thousands of dollars per streaming server. All the client needs is the player software, which is free or very inexpensive, and a sound card. Companies that offer Web hosting and streaming media services, however, need to make sure that they address network latencies, bandwidth management, digital rights management, billing systems, ad insertion, player licenses, and storage space. Again, we are in the early stages with streaming media, but it is certainly going to be a very important area.

Streaming Media on the Internet

Streaming media on the Internet today suffers because of restricted bandwidth (which makes the video jerky), poor reliability in the network (resulting in missing frames or dropouts in the audio), a lack of QoS in the network (which causes various types of distortions or artifacts in the video and audio), and packet loss at Internet peering points (ranging up to 40% during peak traffic hours, which is when key problems occur). These factors are being addressed, however; therefore, streaming media has great potential in the business realm.

Businesses have spent billions of dollars on streaming media, and they are expected to spend billions each year in the next several years. Businesses are seeking content conversion or capture, hardware and software infrastructure, network access and transport services, and other services, such as installation and support. Entertainment and consumer-oriented uses of streaming media are also huge. Streaming media is becoming fundamental to the way corporations, as well as individuals, communicate. Software and service providers are addressing three basic problems: delivery, performance monitoring, and content management.

Streaming Media Delivery

Edge caching has gained considerable momentum as a solution to the peering point problem. With edge caching, Web content is duplicated on a machine close to the end user the first time the user requests the content. Subsequent requests for this content, then, are satisfied from the nearby machine. This improves the speed and reliability of access because it avoids the Internet backbone and its peering points. Providers of edge caching include CacheFlow, InfoLibria, Inktomi, Network Appliance, and Novell.

In addition to edge caching, other techniques can be applied, such as hop-by-hop retransmission. This minimizes latency and increases the usefulness of retransmission for real-time broadcasts. With hop-by-hop retransmission, an intermediate device retransmits, so the retransmission travels a shorter path over a fewer number of hops and is therefore less delayed.

Application-layer multicasting is another technique. It ensures that just one stream goes across the backbone whenever possible. It is similar to IP multicasting, but it occurs at the application layer. FastForward Networks provides such solutions.

Streaming Media Performance Monitoring

Streaming content is vulnerable to fluctuations in bandwidth and QoS. The user's experience is closely correlated with metrics such as throughput, jitter, and dropped packets. Streaming applications require performance monitoring systems and services that track such measurements. Key providers in this realm include Mercury Interactive and WebHancer.

Streaming Media Content Management

Content management is another issue that is related to streaming media. Companies that regularly stream content need a system for making the content searchable and navigable by creating metadata that indexes it. Metadata is then stored on an application server, and the video is stored on a video server. Multimedia data search systems are automated software tools that analyze video, comparing each frame to known images and computing image similarity. They also create and index a voice-to-text transcription. Convera Corporation (which is a joint venture between Intel and Excalibur Technologies), Taalee, Virage, and WordWave are key providers of multimedia data search systems.

ISPs will look to offer new types of visually enabled services as a way to make up for reduced connection-fee revenue. Other organizations will look for ways to blend video with Web sites or portals, to improve the way they disseminate information, and to provide enhanced customer interaction. We can expect to see visually enabled call centers, visually enabled help desks, visual virtual meeting rooms, visual chat rooms, and visually enabled e-commerce.

Real-time interactive visual communications have been available for some time, but to date only niche markets (for example, distance learning, telemedicine, and corporate video conferencing) have seen their benefits and adopted their use.

Broadband Internet access with QoS is required for streaming media, and as it becomes more widely available in the near future, we will see many more adopters of this technology. Chapter 10 further discusses the demands of real-time visual streams and related QoS issues.

For more learning resources, quizzes, and discussion forums on concepts related to this chapter, see www.telecomessentials.com/learningcenter.