After reading this chapter, you should be able to perform the following tasks:

- Describe the traditional use of networks for both voice and data traffic.
- Describe the new-world network model that integrates both voice and data into a single network.
- Describe the Cisco network design model in general terms.
- Briefly describe Cisco's AVVID technology and the standards used for implementing AVVID.

CHAPTER 1

Merging Voice and Data Networks

Tremendous changes have rocked the computer networking industry over the past decade. Those changes have included widespread networking in businesses, such frequent networking in homes. As a result, Microsoft Windows 98 Second Edition now comes with built-in support for small networks (DHCP server and NAT), the World Wide Web, and the phenomenon of integrated voice and data. There is also a whole host of wireless technologies such as Cisco Systems Aironet, infrared links between Personal Digital Assistants (PDAs) and PCs, cell phone modems, and the JetCell product from Cisco Systems that permits cell phones to register with a campus network and uses the "in-house" voice over IP network to service those cell phone calls. This chapter provides a brief history of how we got where we are today, some of the tools we use for voice-over services, a brief explanation of network design, and where Cisco's products fit in the network, both from a network design perspective and from a customer perspective.

Traditional Networks

Let's look at a very brief history of networks in general. First we'll look at telephone, or voice, networks, and then we'll look at data networks. *Traditional networks* are networks that use the voice network to transmit voice, video streams, and data. Frequently a carrier's network is delivered to the end customer as if it were three distinct networks, one for each of the services listed above. More recently it has become common to share the service provider's network but still segregate the network into three discrete networks at the point of ingress to the customer's site. After we briefly look at traditional networks, we'll see how new ideas and technologies are combining these networks into a single network—from desktop through the network to the desktop at the opposite end of a call, video conference, or PC application. This makes a single network infrastructure, protocol management, and addressing structure possible.

Voice Networks

The phone system began as a system based on copper wires from Point A to Point B. For several decades the current on the wires was used to transmit only analog voice calls. In the mid-20th century, technologies were developed to transmit calls via digital networks. Today nearly all calls are handled within the network cloud by digital transmission techniques. Most home and small-business service is still analog, from the home handset to the central office (CO). For many readers this is probably the system with which you grew up. In Europe and other areas of the world, the Integrated Services Digital Network (ISDN)—a digital service that provides support for integrated voice, video, and data—is the predominant service for all but the smallest businesses and homes.

Analog Voice Networks

We'll discuss analog technology in much greater detail in Chapter 2, "Introduction to Analog Technology." For now, think of the old black-and-white movies where the operator used to pull cables from the switchboard and plug them into the proper jack to complete a call. What the operator was doing was providing a continuous copper circuit from one end of the call to the other. The entire circuit transmitted the voice as an analog wave, similar to the transmission of sound (as electricity) from a modern stereo receiver or audio/video receiver to the speaker. A dedicated pair of copper wires was required for each call because of the nature of an analog call in the old phone network. The advent of frequency division multiplexing (FDM) permitted a single analog circuit to carry several voice calls, still using analog technology, and typically over a coaxial cable. Eventually techniques were devised to permit the voice stream to be digitized and sent as a digital signal.

Digital Voice Networks

The gradual replacement of the analog voice network with a digital voice network began in the United States and other more developed countries in the 1960s. Many of these changes were required in order to carry not only greater volumes of voice traffic, but also data from the then-new computer industry. In the digital telephone network, calls from homes are typically digitized when they reach the local exchange carrier's (LEC's) CO. In a modern business the digital line can extend all the way to the private branch exchange (PBX), a kind of miniature CO that serves a private organization, and on to the phone handset itself. In these networks, the only analog voice stream is from the mouth to the microphone of the handset. We'll cover digital voice networks in detail in Chapter 3, "Introduction to Digital Voice Technology," and PBXs in Chapter 12, "Old-World Technology: Introduction to PBXs."

Data Networks

The widespread use of the phone network for personal communications led to the use of the same infrastructure for data networks. In the early days of data communications, the protocols used to communicate between machines were all proprietary. As the computer communications industry grew, the need for different machines to interwork with each other required common communications protocols.

In this section, we discuss traditional data networks so that you can compare them to new-world data networks, which are covered later in this chapter. Even though we categorize these networks as traditional as if they are no longer common, there are several places in the world where they are still the rule. One of the greatest challenges for those converting from traditional networks to new-world networks is making sure that the current network is well understood. Using a U.S.-centric model for a global network will not provide an optimal solution based on providers' local offerings and may lead to suggesting solutions that simply are not available in some parts of the world.

Network Cloud Terminology

The terms LEC, IXC, "tail circuit" or "last mile," and "local loop" are all U.S.-centric. In many countries the local telecommunications services have not been deregulated and there is only one carrier—the government. In that case, the terms are irrelevant and may be confusing to readers from outside the United States. It is not intended to be confusing—far from it. Because of the size and familiarity many readers have with the U.S. market, these terms are used to help distinguish the differences where they exist. In those countries that do not have a deregulated telecommunications market, all services are obtained through the local Post, Telephone, and Telegraph (PTT) agency.

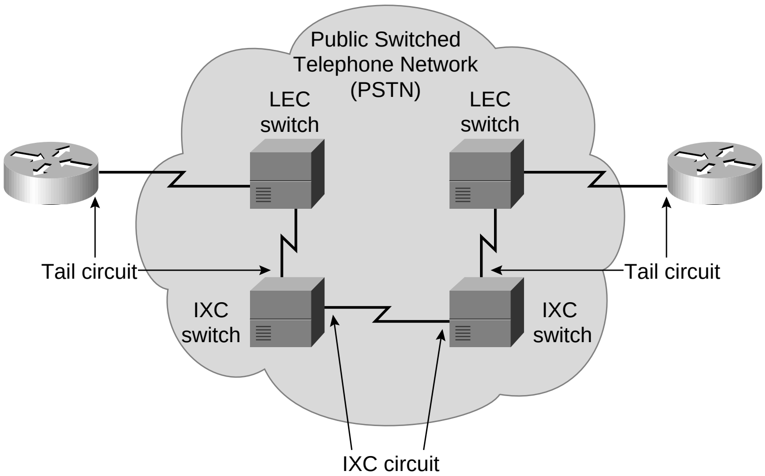
Leased-Line Networks

In a leased-line network, the customer typically leases all of the services from one or more service providers. Leased-line networks may be simple point-to-point networks connecting two points, running a simple Layer 2 protocol such as HDLC and a single Layer 3 protocol such as IP. A leased-line network may also consist of a network running multiple protocols and connecting multiple sites, some in point-to-point configurations and some in a series topology called multipoint. A leased-line network typically consists of three components:

- The local circuit, frequently called the "tail-circuit" or the "last mile," from the customer's local site to the inter-exchange carrier's (IXC's) local point-of-presence (POP) via the LEC's CO.
- A circuit through the IXC's network to the remote site.
- Another local circuit at the remote site. See Figure 1-1 for an illustration of these network components.

In a site that does not cross tariff boundaries or in an area where competitive local exchange carriers (CLECs) provide service, all of the traffic may stay on the same network throughout. Leased-line networks typically come in two flavors, point-to-point and multi-drop.

Figure 1-1 *A Typical Leased-Line Network*



Point-to-Point Networks

A point-to-point network is just what it sounds like: a network with only two connections, one at each end. Point-to-point networks provide a specified amount of bandwidth, which is available whether it is used or not. An example of a common implementation of point-to-point networks that offer great error recovery and flexibility is the X.25 public data networks. These are discussed in more detail in the sidebar "X.25 Networks" later in this chapter. Point-to-point connections are frequently replaced with Frame Relay networks, as described in the "Frame Relay Data Networks" section of this chapter. The earliest days of point-to-point networks required the same kind of equipment on each end of the link. For the past few years, however, Point-to-Point Protocol (PPP) has provided a mechanism for heterogeneous networking. The PPP suite provides a rich set of features, discussion of which is beyond the scope of this book.

Multi-Drop Networks

Multi-drop networks were a common implementation using proprietary protocols such as IBM's Synchronous Data Link Control (SDLC) protocol and Digital Equipment Corporation's Digital Data Communications Message Protocol (DDCMP), as well as standardized protocols such as the High-Level Data Link Control (HDLC) protocol's Multilink Procedures (MLP) extension.

These networks permitted the phone company to bridge connections within their CO or POP so that a single frame sent from a primary station (typically the host or the host's front-end processor, or FEP) would be transmitted to all secondary stations. The frame included an address that one of the secondary stations recognized as its own, and that station would process the frame; all other stations on the multi-drop link would discard the frame.

Multi-drop networks were characteristic primarily of the proprietary mainframe and minicomputer worlds. Many enterprises have replaced these multi-drop networks with Frame Relay networks where Frame Relay is available. In many areas, the lack of Frame Relay circuits requires continued use of multi-drop networks.

Network Design

This book is not solely or primarily about network design, but we do discuss network design as it relates to converged voice and data networks in Chapter 13, "Network Design Guide." This section is meant to serve as a general review for those who have studied network design and a gentle introduction for those who have not studied this subject.

Cisco has a well-developed and highly respected design model that comprises three layers:

- Access layer
- Distribution layer
- Core layer

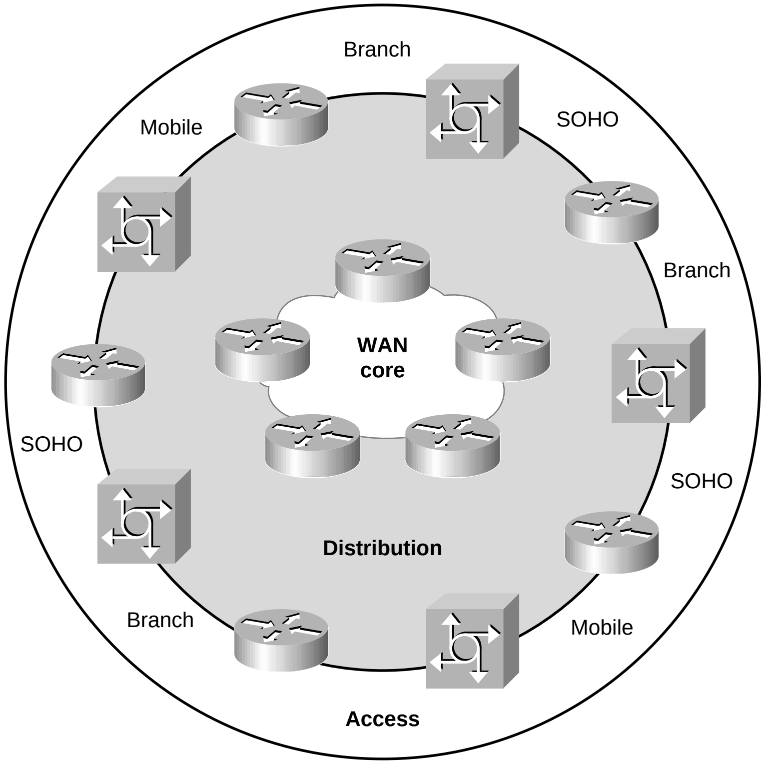
Each layer has specific characteristics and is used to perform specific tasks within the overall network architecture. The model is illustrated in Figure 1-2. We briefly discuss these characteristics and tasks in the following sections.

The Access Layer

The access layer is the layer that is closest to the users and provides them access to the resources on the network, hence the name. The primary task of the access layer is to provide a point of ingress/egress to users of the network. Characteristics of the access layer are:

- It uses switched or shared LAN technology in a campus environment.
- It uses remote access technologies (leased line, Frame Relay, ISDN, and asynchronous dial) in a WAN environment.
- It may provide for deterministic fail-over in mission-critical environments.

Figure 1-2 *Hierarchical Wide-Area Design*



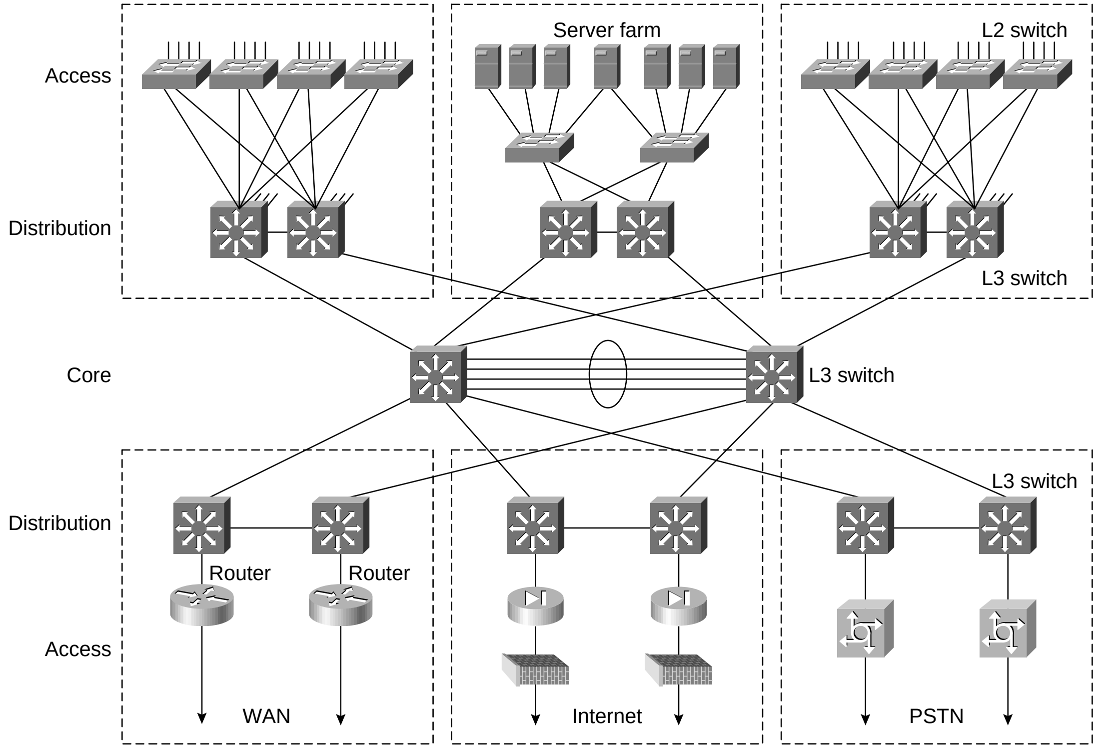
The Distribution Layer

The distribution layer is the layer of the design model that provides for the aggregation of traffic from the access layer and serves as the point of ingress to the core layer. The distribution layer may have several characteristics depending upon the specific features implemented. Among these are:

- Policies for routing, quality of service (QoS), or security
- Address summarization
- Broadcast or multicast domain boundaries
- Media translations (for example, between 10BaseT Ethernet and 100BaseT or Gigabit Ethernet)
- Routing redistribution between different routing protocols
- Bandwidth aggregation from the low- and medium-speed access links into higher-speed backbone links

This list is just a subset of characteristics that might be considered to apply to the distribution layer. Frequently the foundation of a distribution layer area is a multilayer switch, and for that reason they are called “distribution blocks” or “switch blocks.” Figure 1-3 illustrates the use of distribution blocks in a network and the layers to which each series of devices belongs. See the Cisco Press book titled *Cisco LAN Switching* for more information on the concept of distribution blocks.

Figure 1-3 *Enterprise Campus Network Design Model*



The Core Layer

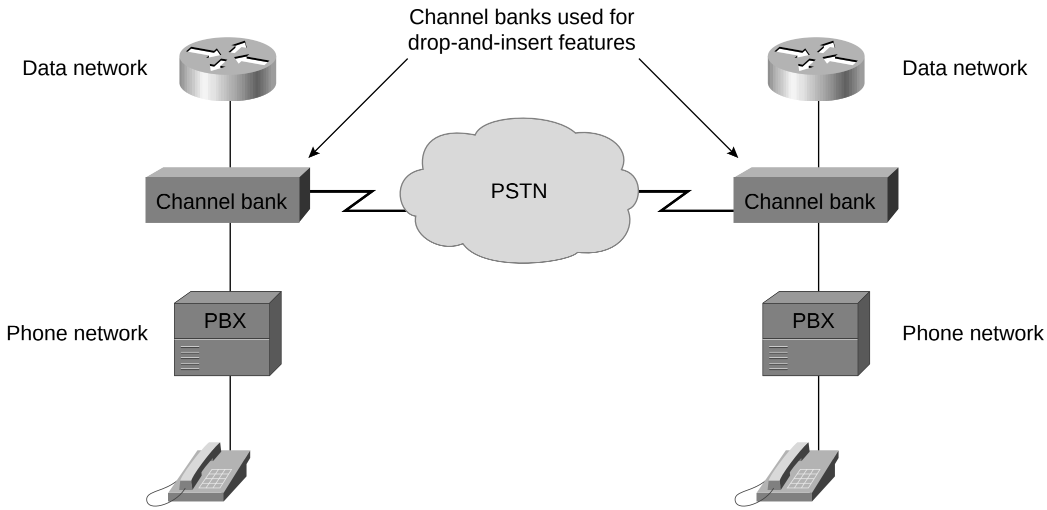
The purpose of the core layer is to perform high-speed switching between one distribution area and another distribution area. Typically no filtering or route summarization is done here; this layer simply switches frames or cells quickly from one port to another port.

Sharing Circuits with Drop-and-Insert Devices

From the 1970s through the 1990s it was common to have both a voice network and a data network on entirely separate infrastructures in an enterprise. Many enterprises still use this kind of network, since the total bandwidth of their data and voice traffic will fit within one T1 or E1 circuit (see Chapter 3 for more information on T1 and E1 circuits).

To better use the bandwidth, they insert a channel bank between the telephone company's (telco's) circuit termination point and their telephone and computer equipment. The channel bank takes the input of a single trunk from the telephone company and sends some channels to the phone equipment and others to the data equipment. This type of deployment is illustrated in Figure 1-4. The function that channel banks provide is called drop-and-insert because the user can drop channels from, say, the voice stream and insert them into the data stream. Even with this kind of network, though, when telephone calls aren't being made there is unused (but paid-for) bandwidth. Likewise, unused bandwidth that is allocated to the data stream cannot be used during periods of heavy telephone circuit usage.

Figure 1-4 *A Shared Network Using Drop-and-Insert Devices*



New-World Networks

The solution to the problems posed by traditional networks was to develop a new way to network voice and data, as described in the following sections.

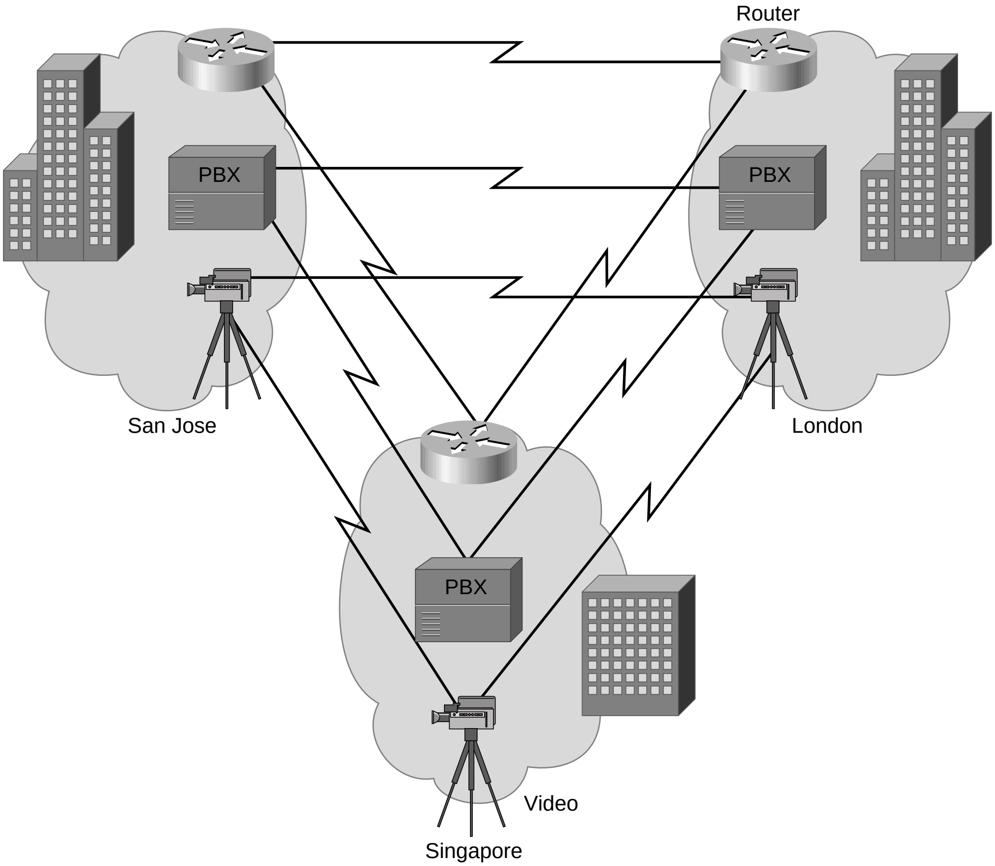
Circuit Switching, Packet Switching, and Virtual Circuits

The traditional telephone system uses switches throughout the network to connect the two endpoints together. Once a call—irrespective of whether it is a voice call or a data call—is established, there is a continuous path from one end to the other that is dedicated solely to that call. This is the meaning of the term "circuit switching"; it refers to switching a shared infrastructure from one dedicated use to another dedicated use. This term is applied equally to either voice circuits or data circuits. Since many voice calls consist of "dead time" when neither party is speaking, there is a tremendous waste of resources in a circuit-switched network. On the other hand, since the circuit is entirely dedicated to that one call there is no need for any of the techniques we will discuss later to provide what we call *quality of service (QoS).* In a packet-switched network the network is shared between users and between applications, so a single user may have several applications running that use the local network but are switched on an individual packet-by-packet basis to their ultimate—and frequently different—destinations.

In the paragraph above we discussed circuit switching versus packet switching. Another concept we build on extensively in this book is that of "virtual circuits." We have probably all heard of virtual reality, but what is a virtual circuit? Remember that circuit switching allocates an entire circuit to either a voice or data call, while packet switching switches packets over a shared infrastructure. Using a shared infrastructure to build a circuit that appears to be a permanently allocated or dedicated circuit between two points is what we are referring to when we speak—or write, as the actual case may be—about "virtual circuits." Virtual circuits are of two types, either permanent or switched, and are encountered in X.25, Frame Relay, and ATM environments. A permanent virtual circuit (PVC) is one that is established between two points and left in place in perpetuity—or until there is a technical problem with the infrastructure. A switched virtual circuit (SVC) is one that is built on demand and destroyed or torn down when the need for the circuit is over. All of the transports we have encountered—X.25, Frame Relay, and ATM—can make use of SVCs. It is most common to see them in X.25 and ATM environments, although they are becoming more common in some service providers' networks.

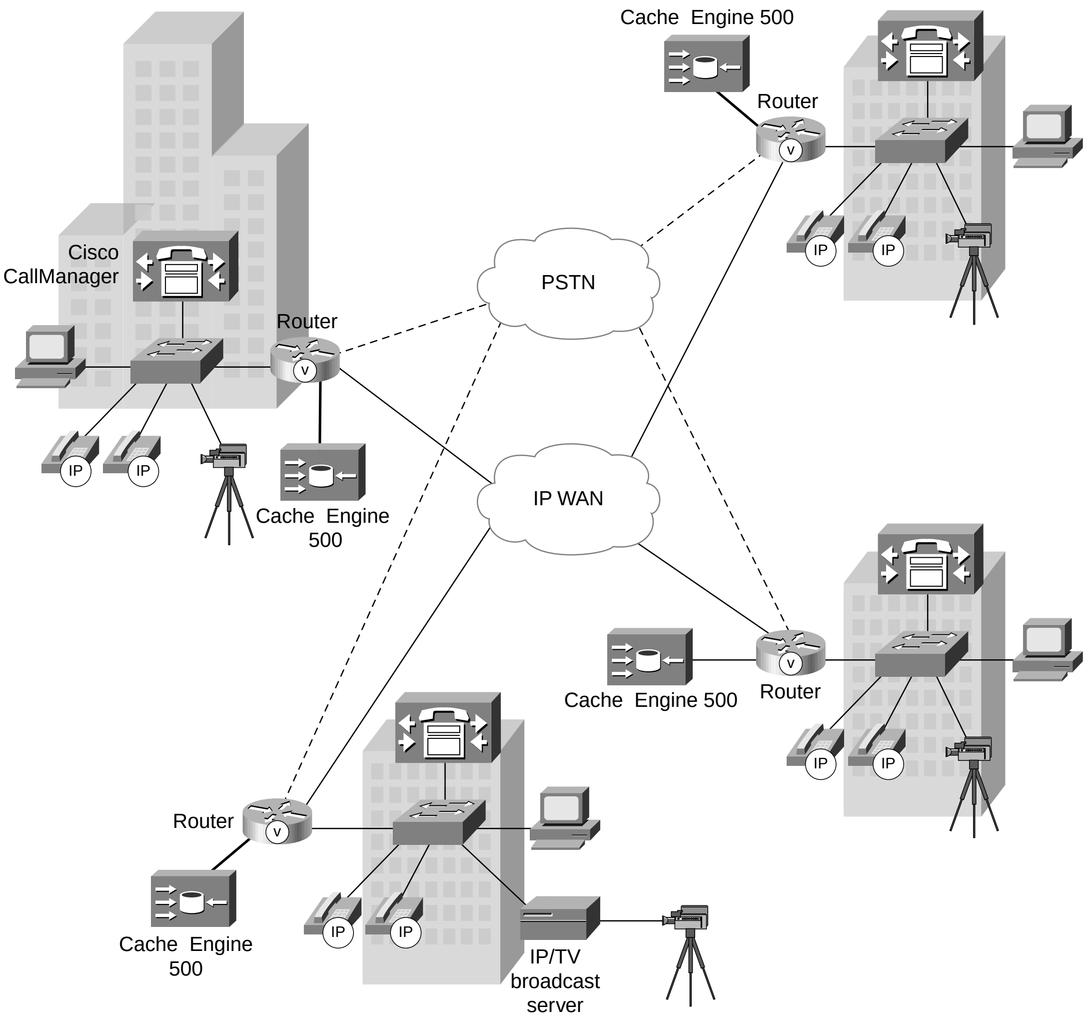
This chapter opened by noting that tremendous changes have rocked the computer networking industry recently. This is best illustrated by the technology reviewed in this section. In "old world" communications, an enterprise typically had a network dedicated to voice, another to data, and perhaps even a third dedicated solely to video links, as shown in Figure 1-5.

Figure 1-5 *Separate Data, Voice, and Video Networking*



Now voice, video, and integrated data are coming together into a single network. Cisco refers to its vision for this as the Architecture for Voice, Video, and Integrated Data (AVVID). An enterprise that deploys an AVVID network can reduce costs and increase productivity through the use of a single network for all of its communications needs. This network is standards-based, using technologies such as Frame Relay, IP, and other IETF-based standard protocols, and International Telecommunications Union (ITU) standards for signaling and voice compression. An implementation of this type of network is illustrated in Figure 1-6. We'll look at each of these subjects briefly now, and in detail in later chapters of this book.

Figure 1-6 *A Converged Network for Voice, Video, and Integrated Data*



Frame Relay Data Networks

Frame Relay networks were first deployed in the early 1990s, but did not really become deployed in large numbers until the mid-1990s. Prior to the advent of Frame Relay, most large switched networks were based on the X.25 family of standards. The X.25 protocol is an extremely flexible technology, but each link is of a fixed speed. The X.25 protocol is a high-overhead protocol (see sidebar), where Frame Relay has a very low overhead. Frame Relay fixed the overhead issue and the fixed-speed issue.

X.25 Networks

The X.25 protocol is a protocol built to operate over circuits that are sometimes unreliable; X.25 has multiple layers of error checking and sequencing. These reliability features require processes that take time and induce delay into the end-to-end path. Many implementations of X.25 were limited to what today would be called "low-speed" links—56 kilobits per second (kbps) or 64 kbps. Nothing in the X.25 standard limits X.25 to such low speeds. In fact, the standard specifically states that the interface may operate up to 2 megabits per second (Mbps). One of the editors of this book worked in a network where X.25 was carried internally at speeds of 8 Mbps! The X.25 protocol is unsuitable for an integrated service network because of its typically slow speed and high overhead requirements.

Frame Relay permits a customer to subscribe with a carrier for a certain amount of bandwidth, but with the ability to "burst" over the subscribed amount for short periods of time. This behavior is well suited to the short and bursty nature of much of the traffic seen in client/server environments, such as the Internet. Another benefit of Frame Relay is that the customer does not need a dedicated circuit from each site to each of the other sites. The Frame Relay cloud takes care of interconnecting sites according to the customer's preferred design. Frame Relay can be deployed in full-mesh networks, partial-mesh networks, or star configurations. For a fuller discussion of the important elements of Frame Relay network design, see *Cisco Internetwork Design* or *Building Cisco Remote Access Networks*.

AVVID

AVVID includes four building blocks: basic infrastructure, such as routers and switches; call processing, used to resolve dialed numbers to gateways or other end-devices; applications, such as call control; and clients, such as IP telephones, H.323 videoconferencing equipment, and PCs. AVVID uses Internet Protocol as the common transport for the converged network, permitting common QoS, policy, routing, management and security services to be deployed throughout a network.

One of the key aspects of AVVID is the unified messaging component that permits e-mail, voice mail, and faxes to be routed to the appropriate device. Voice mail, for example, may be forwarded as a WAV file as an attachment to an e-mail. AVVID leverages efficient network design with appropriate technology to provide higher availability at a lower total cost of ownership.

Techniques for Combining Voice and Data

Earlier in this chapter we discussed how voice and data have been carried in the past. We looked at separate networks and a single network with dedicated portions of that network carrying either voice or data. New-world telephony treats voice and video just like any other data. The data stream is converted to a digital signal, compressed, and packetized for transport by the IP suite of protocols. We can establish quality of service for the IP packets that carry voice, video, or other time-sensitive traffic (such as whiteboarding), and in this case the network differentiates between that time-sensitive data and traditional application data. The old-world phone system has distinct steps required to establish and tear down a call. Those signaling functions must be replicated in new-world telephony. AVVID supports open standards to provide those signaling functions.

Voice Compression and Packetization

The old-world telephone system dedicates 64 kbps for each phone call, even during periods of silence. The new-world telephone system uses low-bandwidth voice-compression standards such as the ITU-T's G.729 and G.729a for phone calls requiring as little as 11.2 kbps during periods of burst and minimal bandwidth during periods of silence. As new standards are developed and implemented, that bandwidth will fall. The voice input is first digitized using open standards, then packetized into small data packets to avoid delay and jitter.

Delay and Jitter

Delay is when the packet takes so long to arrive at the other end that the speaker at that end may interpret the delay as silence on the other person's part. *Jitter* occurs when packets arrive with varying amounts of delay. Delay is normal, but long delay is annoying. Varying amounts of delay create jitter. Delay and jitter are common in traditional networks, but several steps are taken to reduce their impact. De-jitter buffers are built into switches in the network to ensure there is enough of the bit-stream buffered to resynchronize the bit-stream if necessary. The network is engineered to make sure that delay does not become a degrading issue (like using multiple satellite hops from end to end). Delay and jitter are discussed more in Chapter 3 and are covered in detail in *Voice over IP Fundamentals* from Cisco Press. You can use various QoS features to avoid or compensate for delay and jitter; these will be addressed in Chapter 10, "Configuring Cisco Access Routers for VoIP."

It is also possible to transmit voice directly over OSI Layer 2 without the benefit of IP. The compression and encapsulation procedures are the same as when the data is transmitted over IP; the only changes are the encapsulation layer and the specific mechanisms used to provide QoS. Irrespective of the layer, the principles of encapsulation, fragmentation, and queuing are used in providing QoS.

Signaling Protocols

There are several kinds of signaling in the old-world telephone system. The new-world telephone system must support the old system for interoperation with legacy systems. The new-world paradigm must also provide signaling systems that are standards-based and that are designed to operate over IP networks. There are several protocols designed to do this, including the ITU-T protocol H.323, and the IETF protocols Media Gateway Control Protocol (MGCP) and Session Initiation Protocol (SIP). These protocols are discussed in more detail in Chapter 10. Traditional protocols such as ITU-T Q.931 and ITU-T Q.700 (Common Channel Signaling System 7, or just SS7) are also supported. Connectivity between traditional PBXs was provided via tie-lines that used E&M signaling, which is still supported in several Cisco platforms. There are also protocols that permit signaling between proprietary old-world systems, such as the European Computer Manufacturers Association (ECMA) QSIG, based on the ITU-T's Q.931. QSIG is discussed briefly in Chapter 12; for a more thorough treatment of this protocol, see *Voice over IP Fundamentals*. British Telecommunications (BT) developed a proprietary protocol in the 1980s to connect PBXs known as the Digital Private Network Signaling System (DPNSS), which is being replaced by QSIG in many installations.

Cisco's Role in New-World Networking

Cisco Systems plays several roles in the arena of new-world telephony. It is a manufacturer and is active in many standards bodies that define the standards for the new world. Cisco's equipment fits into all three of the design layers discussed earlier in this chapter. In this section we briefly look at the technologies Cisco is deploying and where they fit, both in the design model and from a customer perspective.

Cisco's Role from a Design Viewpoint

Recall the three-layer design model discussed earlier: access, distribution, and core. Each of these layers plays a role in new-world networking.

The Access Layer

The access layer is the layer that provides entry to the network for users. At the access layer, one would see technology that either integrates traditional telephony—for example, connecting to a legacy PBX or handset—or that is based on open standards—for example, Cisco's IP Phone or SoftPhone or another standards-based telephony device. There are also collaboration tools like Microsoft's NetMeeting and similar products that permit simultaneous voice, video, and whiteboarding.

Integrating Traditional Telephony

Cisco equipment that integrates traditional telephony has interfaces for devices that use T1/E1, ISDN, traditional FXO/FXS connections, or tie-lines between PBXs using an E&M interface. These are the legacy interfaces we have inherited from nearly 100 years of analog telephony. The access device, such as a Cisco 2610, would have handset(s) connected directly to it to provide the ability to call from one location to another. There may be a line to the local CO to provide access to the PSTN.

New-World Telephony

New-world networking devices such as the IP Phone act as a DHCP client to request an IP address or be statically configured locally with a permanent IP address. They use IETF and ITU-T standards throughout to provide quick and easily deployed solutions to problems. These devices connect directly to a switch and may even take their DC power from the switch. This pure IP environment permits great flexibility and granularity of service, management, policy, QoS, and security services throughout the network.

The Distribution Layer

The distribution layer is the point at which some of the policy, QoS, and other services mentioned above may be implemented. At this layer, one will find either Cisco routers or Catalyst switches with an embedded router to forward packets to their destination. If necessary, the class of service (COS) may be changed to a more appropriate class to provide the QoS necessary for real-time services. Features such as Random Early Detection (RED), Weighted RED (WRED), or other "flavors" of RED may be used to "throttle back" reliable connections such as HTTP to provide bandwidth for voice and video.

The Core Layer

The core layer is designed to forward packets as quickly as possible. In this environment, one will probably find Catalyst switches, GSRs, and LS1010s—depending on the network size, complexity, and distribution-layer speeds—to quickly forward frames to the appropriate distribution layer. In a service provider's core or a very large enterprise core, one may find such switches as IGX, MGX, and/or BPXs.

Cisco's Role from a Customer Viewpoint

The customer will not only need to look at the basic design of the network as outlined above, but will also need to determine specific applications. The customer type will determine the Cisco equipment deployed in the network.

Enterprise Telephony

Enterprise telephony can be thought of as a purely internal system. Say the Widget Company, for example, decides to cut down on its long-distance bills for interoffice calls. It also wants a unified messaging system and the ability to conduct videoconferencing or online training. This customer will want to implement the appropriate access technology, depending on whether they will coexist with legacy equipment or prefer to switch to new-world technology. They may or may not need to reconfigure their distribution and core layers, depending upon their design strategy in their old-world network.

Service Provider Telephony

Many service providers are adopting Cisco's AVVID technology to replace the legacy, proprietary equipment that limits their business flexibility. Cisco has a wide range of equipment that is only briefly covered or not covered by this text. These include all the equipment previously mentioned as well as routers that implement interactive-voice response (IVR) and permit authentication via TACACS+ or RADIUS for prepaid long-distance service. Also included in this expanded view of Cisco's product offerings are the large-scale switches such as the IGX, BPX, and MGX. These switches provide the ability to send voice, video, and data using ATM as the transport, at interface speeds of up to OC-12 (622 Mbps) with backplane switching speeds in the several gigabits per second range.

Summary

In this chapter we have looked briefly at the development of both voice and data networks. From completely separate networks they evolved through shared networks with dedicated bandwidth and now into fully converged networks. Different technologies and design strategies have been deployed throughout the years, leading us to the scalable and easily replicated three-layer hierarchical design model. The old-world network was built on dedicated lines and proprietary protocols; the new-world network is built on a shared or converged network using open standards. These standards replicate the functions of the old-world proprietary standards, with the benefit of interoperability with legacy equipment during periods of coexistence and migration. Cisco Systems provides a wide range of products to meet the needs of enterprises and service providers, from access to core.

Review Questions

The following questions should help you gauge your understanding of this chapter. You can find the answers in Appendix A, "Answers to Review Questions."

1 Name the three components of a leased-line network.

2 Name the two types of leased-line networks.

3 Name the three layers of the Cisco network design model and briefly describe their functions.

4 What are the chief benefits of implementing an AVVID network?

5 What is a commonly implemented ITU-T standard used for voice compression?

6 What are two business markets that can benefit from implementing Cisco's AVVID technology?