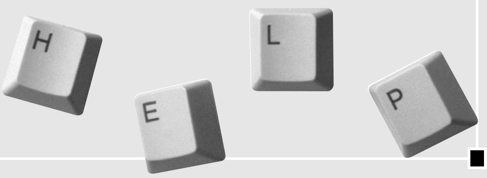
5674ch28.qxd_jd 9/24/03 9:57 AM Page 677

28

Voice Over Internet Protocol



Voice over Internet Protocol (VoIP) is a technology to use the Internet Protocol instead of phone-switching technology for voice, facsimile, and message services. The typical consumer probably hasn't heard of VoIP yet. However, the technology is already commonly in use. The main example would be the dominance of inexpensive prepaid telephone cards on the market. VoIP is the technology in use for almost all the prepaid telephone long-distance cards. It's how their per-minute charges are so competitive, compared with traditional long-distance service.

Until just a few years ago, every telephone call utilized the traditional Public Switched Telephone Network (PSTN). With PSTN, a long-distance call first goes through a local telephone company, then to the long-distance provider (for a connection fee) in an all-analog voice signal. Large telephone switches are in place to make the transfer. With VoIP prepaid telephone cards, you pick up your phone and dial a special number, then the long-distance number. The route that the call takes differs from that of traditional PSTN. VoIP simplifies the system with an advanced combination of hardware and software. The analog call is converted to digital. The digitized call is compressed and translated into IP packets for transmission over a private IP network or the Internet. The call bypasses the traditional route through the expensive telephone switch at a telephone company's central office. At the other end of the wire, the process is reversed: The digital signal is converted back to an analog signal, and it's again routed through regular telephone equipment.

Like many nascent technologies, VoIP was originally considered a cute novelty. But it's gaining momentum as a viable technology. VoIP holds several solid advantages, both in the home and in business, over traditional PSTN networks. A few include:

- Cost Reduction—VoIP eliminates traditional circuit-switched networks and reduces the associated per-minute long-distance fees. VoIP reduces the amount of hardware needed (infrastructure overhead) by converging voice and data networks. Network efficiency improves with shared equipment. Excess bandwidth, rarely exploited, can be fully utilized. In short, it's a bargain.
- Simplicity—A single piece of equipment supports both voice and data communications. Less hardware means less cost.
- Advanced Applications—Like PSTN, basic telephony and facsimile are the core elements of VoIP. However, because VoIP uses a compressed, packetized digital format, the possibility for advanced multimedia (and multiservice) applications is limitless. A few possible applications include: Web-enabled call centers, collaborative white boarding, remote telecommuting, and personal productivity applications.

The History of VoIP

VoIP had been talked about long before Vocaltec, Inc. released Internet Phone Software in 1995. This software was designed to run on a home PC (486/33 MHz) with sound cards, speakers, microphone, and modem. The software compressed the voice signal, translated it into voice packets, and shipped it out over the Internet. The technology worked as long as both the caller and the receiver had the same equipment and software. Although the sound quality was nowhere near that of conventional equipment at the time, this effort represented the first IP phone. By converting analog voice into compressed digital IP packets, the software enabled PC-to-PC Internet telephony.

It's tempting to view the advent of VoIP as a singular technological event. In reality, however, VoIP is an evolutionary extension of decades-long communications and network technology progress. The highly reliable telephone network has been in a state of constant evolution for more than 100 years.

Today, billions of calls traverse the world's phone systems every day with little human intervention. This hasn't always been the case. In the early 1900s, each call was switched manually by a live telephone company switchboard operator and again with private board operators hired by each company. An important development for the telephone company network was the automation of the call switching function. With the invention of the Private Branch Exchange (PBX), companies were able to cut their payrolls from many in-house operators to just a handful of receptionists.

The next big step forward for the telephone network was in the early 1960s, with the introduction of pulse code modulation (PCM) technology. PCM addressed the inherent signal problems of transmitting voice in the analog world by converting the analog signal to binary 0s and 1s. This reduced the distance and interference noise on the line. (In fact, on transatlantic calls, one could hear the waves as transient noise moving over the cable.) The binary voice signal became as clear at the receiving end of the line as at the sending end. In the 1980s, Time Division Multiplexing (TDM) became a popular way to deliver analog voice over digital networks. With TDM, many analog channels were digitized and allocated a specific time slot. The technology allowed different speeds for each channel and supported traffic aggregation. TDM transformed analog voice to digital over switched networks, laying very important groundwork for VoIP, because voice could now be viewed as data.

TDM provided the major advantage of putting more voice channels onto a single line. For example, two pair of copper wires could now support a T1 line, or 24 channels. The downside of TDM was that allocated channel bandwidth couldn't be dynamically reassigned when not in use. It was for this reason that a different way was sought to transmit voice and data over a single network—a significant challenge because of their differing natures.

Data is sporadic. It comes in bursts. Voice is time-sensitive. If a large data file takes up system bandwidth, voice packets can drop out or be delayed, a condition considered unacceptable for voice transmissions. It would be very difficult to follow a conversation in which large chunks of a sentence are missing.

The Asynchronous Transfer Mode (ATM) technology solved this problem in the early 1990s. ATM is a high-bandwidth, low-delay, packetlike switching technique. It has the quality of service (QoS) built in to integrate voice and data. Although many of today's international backbones utilize ATM, it has proven to be too costly for most corporate environments.

At the same time, however, many companies had started to integrate different types of data on their networks with low-cost frame relay, wide area networks (WANs). IT managers added intelligence to their networks to support different applications and protocols to utilize the same pipe. This became a universal method to transport a variety of applications. IP spread rapidly. Data and telecommunications managers could see the promise of IP as a universal transport mechanism.

Voice is the latest core function making its way into the IP world. It's only logical because to an IP network, voice is just another application. The PBX can be a highly reliable application server. It seems natural to converge traditional voice with existing data networks.

In the years since VoIP was introduced, a growing list of technology providers have begun to offer PC telephony software. There is a spate of gateway manufacturers entering the market. Until recently, VoIP provided PC-to-PC telephony primarily over intranets typically found in a business environment. With the introduction of gateway infrastructure outfitted with VoIP technology, users can now look forward to the widespread proliferation of Internet telephony. VoIP infrastructure is the real on-ramp to the Internet itself.

How Does VoIP Work?

It's simple: Voice becomes another data application running over the IP network. The PBX becomes a large server.

Technically, a VoIP call takes place in the following way:

A user picks up the handset, which signals an off-hook condition to the signaling application portion of the VoIP in the network. Then the session application portion of VoIP issues a dialtone and waits for the remote client to dial the number. When a number is dialed, the numbers are stored by the session application. After a sufficient number of digits are accumulated, the number is mapped to an IP host through the dial plan mapper. The IP host has a direct connection to either the destination telephone number or a PBX to complete the call.

Next, the H.323 session protocol (a standard discussed a little later in this chapter) establishes a transmit and receive channel over the IP network. If a PBX is handling the call, it will forward the call to the destination telephone. The VoIP system then starts a coder-decoder compression scheme for both ends of the connection. The remote client's communication is with Real-Time Transport Protocol/User Datagram Protocol/Internet Protocol (RTP/UDP/IP) as the protocol stack (also discussed further later in this chapter).

Call progress indications cut through the voice path as soon as an end-to-end audio channel is established. Additionally, any signaling that the voice port detects, such as inband dual-tone multifrequency (DTMF), is captured by the session application at both ends of the connection and carried over the IP network in Real-Time Control Protocol (RTCP), using the RTCP application-defined (APP) extension mechanism.

Finally, when either remote user hangs up, Resource Reservation Protocol (RSVP) reservations are undone if they are used, and the session ends. Each remote terminal is now idle and waits for the next off-hook state to trigger another call setup.

The Special Requirements of Voice

Though voice is just another data application in an IP system, it nevertheless has some distinct requirements. Some unique characteristics set voice apart. These include all of the following:

- Voice is a real-time application. This means data packets must be processed as they happen in the real world. They do not have the same tolerance for delay and packet loss as other data applications.
- Data traffic sometimes happens at unpredictable intervals, or "bursts." Voice, on the other hand, follows a consistent flow and is more predictable.
- Data transmissions are generally "asymmetrical" because file transfers are typically much larger on the download side than they are on the upload side. Conversely, voice transmissions are typically "symmetrical" because the rate of transfer is almost the same in both directions.
- An industry concern is the reliability in a migration from traditional PSTN-based voice to packetized voice. It is a critical application (and one people have come to expect rock-solid reliability from). Any voice net-

work, traditional or digital, must provide nearly 100% reliability of uninterrupted service. Voice service must be accessible and predictable to provide the continuity human conversation depends on. In a digital voice network, this continuity requires that digital voice packets arrive in the same order in which they were transmitted and be assembled in proper order. Otherwise, it would not be an accurate representation of the source analog input.

How IP Gateways Packetize Voice

The first step in transforming voice into an IP application is to collect raw analog voice data. To simplify design and efficiently utilize bandwidth, most packet systems transmit a constant number of samples (these typically span several milliseconds for each data packet). This series of samples is usually referred to as a *frame*. The size of the frame is often determined by the voice compression algorithm in use. In fact, most popular voice compression algorithms require a predetermined size of the input frame. Likewise, they produce a predetermined output frame. The time the system uses to create voice packets from raw analog voice creates latency. Latency is defined as the average time for a packet to work its way across the network. A typical voice system will gather at least 8 milliseconds of voice data before transmitting the packet. That 8 millisecond latency is introduced into the real-world conversation between two people.

In the world of VoIP, echo cancellation is the next step after raw voice samples have been collected and packetized. The echo cancel function removes the echo (inherent to the analog equipment used in the phone network). In addition, echo cancellation removes the perception of an echo created by the latency introduced when raw analog voice is digitized and broken into packets.

When the echo cancellation procedure is finished, the frame is compressed. This process typically falls within the special domain of a digital signal processor (DSP), which is a specialized processor optimized to handle real-time signals in the digital domain. The DSP may compress the frame in ratios ranging from 1:1 to as much as 10:1, depending on the protocol being used. In addition to compression, the DSP minimizes additional latency caused by the time used to compress the packet.

The VoIP system collects several frames of data before sending a voice packet. This is due to the series of protocol headers added to every packet. When the system has collected enough data for a packet, the DSP adds the headers at the beginning of the data to indicate the destination and the type of packet. It is then put in line for transmission.

How long do all these steps take? The whole process from raw analog voice to a packetized, transmission-ready compressed voice takes no more than 20 milliseconds. VoIP is virtually transparent.

Goals for VoIP Implementation

The goal of the implementation of a VoIP system is quite simple. It is to add telephone-calling capabilities to IP-based networks. Then it interconnects them to the public telephone network and to private voice networks to maintain voice quality standards and preserve the features users expect from their telephones.

A VoIP system should have specific characteristics:

- Voice quality should be comparable to what is available using the PSTN, even over networks with variable levels of QoS.
- The underlying IP network must meet strict performance criteria, including minimal call refusals, network latency, packet loss, and disconnects. This goal should be met even during congested periods or when multiple users must share network resources.
- Call control must make the telephone process transparent so that callers are unaware of the technology they are using.
- PSTN/VoIP service interworking should involve gateways between the voice and data network environments.
- System management, security, addressing, and accounting should be provided, preferably consolidated with the PSTN operation support systems (OSSs).

The Challenge of VoIP

Speech quality should be at least equal to the PSTN (usually referred to as "toll-quality" voice). Some experts argue that a *cost-versus-function-versus-quality* trade-off should be applied.

QoS usually refers to the fidelity of the transmitted voice and facsimile documents. QoS can also be applied to network availability (i.e., call capacity or level of call blocking), telephone feature availability (conferencing, calling number display, etc.), and scalability (any-to-any, universal, expandable).

The quality of sound reproduction over a telephone network is fundamentally subjective, although standardized quality measurements have been developed by the International Telecommunications Union (ITU). These measures represent special challenges in implementing VoIP and can profoundly impact QoS. Challenges include:

- Delay—Two problems result from high end-to-end delay in a voice network. These are echo and talker overlap. Echo becomes a problem when the round-trip delay is more than 50 milliseconds. Because echo is a significant quality problem, VoIP systems must address the need for cancellation. Talker overlap (the problem of one caller stepping on another talker's speech) becomes significant if the one-way delay becomes greater than 250 milliseconds. The end-to-end delay budget is therefore the major constraint and driving requirement for reducing delay through a packet network.
- Jitter (delay variability)—Jitter is the variation in interpacket arrival time as introduced by the variable transmission delay over the network. Removing jitter requires collecting packets and holding them long enough to allow the slowest packets to arrive in time to be played in the correct sequence, which causes additional delay.
- Packet Loss—IP networks cannot guarantee that packets will arrive at all, much less in the right order. Packets will be dropped under peak loads and during periods of congestion (caused, for example, by link failures or inadequate capacity). However, because of the time sensitivity of voice transmission, normal TCP-based retransmission schemes are not suitable. Approaches used to compensate for packet loss include interpolation of speech by replaying the last packet and sending of redundant information. Packet losses greater than 10% are generally intolerable.

All of these technical conditions do present challenges. However, the major benefits of VoIP are fueling the industry to work through standards and technology advancement to make them nonfactors for the user.

Standards

VoIP represents a potential for users to enjoy advanced services such as unified messaging, Web-based call centers, *follow me anywhere* services, and other forward-looking applications. The key to any technology is standards. VoIP currently has a glut of standards.

Traditional voice service has a strong set of international standards to specify and clarify design principles, communication processes, test procedures, and environmental conditions. Since H.323, the first VoIP standard, was ratified in 1996 by the ITU, a host of standards and variations of those standards have emerged from the ITU, as well as from other standards bodies. In addition, there are many more VoIP standards in the works. The following is an overview of the major standards governing the way in which traditional circuit-switched networks and packet-switched networks interact.

H.323

This standard was originally created for established multimedia applications. H.323 is part of a broad family of standards developed by ITU. It describes how audio, video, and data communications take place among terminals, network equipment, and services on IP networks. By 1997, H.323 was accepted as the prevailing VoIP network standard, a position it still enjoys today. H.323 allowed designers to get products to market by giving them a platform for packet network communications and interoperability among early vendor equipment. Because it was originally developed for multimedia applications, H.323 does burden VoIP systems with unnecessary overhead. Nevertheless, its wide use makes it the de facto choice for interoperability among VoIP equipment. The standard describes four major functions of networked communications:

1. Terminals—These are LAN client terminals that enable two-way communication.
2. Gateways—Gateways are designed for real-time, two-way communication between H.323 terminals on a network and other ITU terminals residing on a switched-based network or on another H.323 gateway.
3. Gatekeepers—Within a given zone, gatekeepers are the nexus for calls, providing services to endpoints.
4. Multipoint control units (MCUs)—MCUs function as endpoints for three or more terminals and gateways, enabling multipoint conference communication.

Session Internet Protocol

Session Internet Protocol (SIP) is relatively new on the VoIP standards scene. SIP is an application-layer control protocol that makes up for many of H.323's inherent faults. The standard, developed by the Internet Engineering Task Force (IETF), addresses the call setup and teardown, error handling, and interprocess signaling that are functions of every point-to-point connection. It also changes and terminates multimedia sessions, including conferences, Internet telephony, distance learning, and other applications. SIP enables VoIP gateways, client endpoints, PBXs, and other systems to communicate over packet networks from an equipment perspective.

Compared with H.323, SIP is a simpler protocol with less overhead. The minimum number of message exchanges to set up a call between two endpoints is three. SIP locates the recipient of a call, ensures that the equipment is congruent with the caller's equipment, then allows other protocols to take care of other functions, such as data transfer and security. In addition, SIP differentiates itself from H.323. SIP distributes much of the call management and routing among different areas of the network. Like H.323, SIP has morphed into several related protocols. Furthermore, SIP is a text-based and largely free-formatted protocol, making it easy to debug protocol implementations and to add new features to the protocol to meet new industrial demands.

Media Gateway Control Protocol

Also created by the IETF, Media Gateway Control Protocol (MGCP) is a proposed standard to convert audio signals on the PSTN to data packets that traverse the Internet. The protocol is based on an architecture to move call-control intelligence away from the gateway for processing by external call-control or call agents. MGCP allows media gateways to communicate.

Megaco/H.248

Megaco/H.248 is a new protocol born of a joint effort between the ITU (ITU-T Study Group 16) and the IETF (Megaco is work group of the IETF). Functionally, the proposed standard enables a control of media gateways. Megaco/H.248 is designed to succeed MGCP, adding peer-to-peer interoperability and ensuring a way to control IP telephone devices operating in a master/slave manner. The standard breaks the H.323 gateway function into separate subcomponents. It also determines the protocols employed by each communication component.

So Why All the Standards?

With all of the established VoIP standards and a veritable plethora of others in the works, the obvious questions arise: Why all the standards? Are they all necessary? Probably. First and foremost, standards deliver on the promise of interoperability. This is the ability for myriad VoIP-enabled products to communicate successfully. Interoperability, in turn, delivers an open platform for which many different solutions providers can create innovative products. Perhaps most important, standards and interoperability together give consumers a greater choice in products and technology. They lower the cost of ownership through free-market competition.

Overall, standards are good. Yet this still leaves the question of why there are so many standards, each seemingly with its own unique merits and faults. As discussed earlier, H.323 was originally established for multimedia applications. It was utilized for VoIP because of its maturity and the lack of a perfected standard for voice over packet communications. Nevertheless, despite shortcomings for voice, it enabled designers to get to market quickly with early VoIP products. It provided a vehicle for packet network communications and interoperability among early VoIP vendor equipment. By 1997, H.323 had achieved preeminence as the prevailing standard, a position it still maintains today.

H.323 served communication efforts among early VoIP equipment well. However, early equipment interfaced primarily with other packet-based VoIP equipment within an intranetworking environment, but VoIP systems must interface with legacy POTS (plain old telephone system) phones—the PSTN—and their accompanying standards were not well served by H.323. VoIP requires voice detecting and transmission via the packet network. It also requires tones, which are part of a body of incumbent switched-network standards initiated at the transmitting device.

Today, most VoIP systems don't employ IP from one end to the other. Accordingly, it is necessary to have a standard way of transmitting DTMF tones, as well as to provide a communication mechanism for other named telephony signal events to be transmitted in RTP packets. Request for Comment (RFC) 2833, created by the IETF and formally named *RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals*, is part of the standards track of the Networking Working Group at the IETF. It describes how to carry and format tones and events in an RTP packet. With separate payloads, RFC 2833 ensures a common way to avoid voice encoder payloads and insure redundancy for accurate reception.

Many VoIP applications will use RFC 2833. In Internet phones, it can emulate DTMF functionality. This alleviates the need to create legacy tones. In the receiving phone, it eliminates tone detection. As mentioned previously, VoIP gateways utilize DTMF relay support from RFC 2833.

In RTP, the packet network serves in place of the circuit-switched network. There is another possible application to benefit from the derivative standard. Here, RFC 2833 offers a way to perform transparent communication of telephony signals and events between Class 5 switches and IP networks or endpoints. One real-world example would be to allow MSOs (multisystem operators, such as a cable provider which also offers VoIP) to deploy services more quickly and cost-effectively than through use of the standard in IP telephony. Legacy equipment provides an interface between the Packet Cable network and the Class 5 switch through an IP terminal communicating the necessary GR303 signaling.

Further Standard Evolution—Vocoders

Today, most VoIP applications use voice codecs (also called *vocoders*), which were created for existing digital telephony applications. Like H.323, they are robust and proven. They allow designers to get interoperable products to market quickly. The downside, however, is that they are limited to the telephony signal band of 200–300 MHz with 8-kHz sampling. This worked fine when VoIP quality goals were to achieve the same levels as existing switched networks. However, this level of quality is far below real-world, face-to-face communication. It requires a much wider bandwidth of 50–7,000 Hz.

Just as the industry addressed previous limitations with VoIP service, it is now developing new wideband codec standards designed to push the technology to higher levels of service. The ITU G.722.1 vocoder delivers 24-Kbps and 32-Kbps data rates, and it enables wideband performance with a 16-kHz sampling rate. It is currently being used in some IP applications. Meanwhile, another wideband standard, the Adaptive Multi-Rate Wideband (AMR-WB) in the 50- to 7000-Hz bandwidth, has been jointly developed by the ITU-T and the Third-Generation Partnership Project (3GPP)/ETSI. Recently approved by the ITU and referred to as G.722.2, the standard further improves voice quality over G.722.1, and it enables seamless interface between VoIP systems and wireless base stations.

Ongoing Standards Development

Some may see the ongoing development of new VoIP standards as an alphabet soup and a cumbersome way to negotiate the development of a new technology. In reality, however, the process of continually creating new and more evolved standards is the single most influential force in the promotion and proliferation of a new technology. Standards allow manufacturers to get to market quickly with open systems; they create across-the-board interoperability; and they reduce the consumer's cost of ownership.

VoIP in the Enterprise

Within an enterprise or business communication environment, there are some distinct advantages of combining voice, fax, data, and multimedia traffic onto a single multipurpose network. Lower recurring transmission charges, reduced long-term network ownership costs, and the ability to implement a host of new and powerful voice-enabled applications all are compelling arguments. They have sparked interest in the business world for VoIP technology. At the same time, however, IT leaders have several reasons to move cautiously into the largely uncharted world of converged networks and packetized voice. Quality of voice calls on the data network, stability of VoIP solutions, and the consequences of being prematurely locked into a given architecture are areas of concern within the enterprise. This section will first look at the advantages of VoIP in the enterprise, then it will give some consideration of the risks.

There may be much discussion of the hows and whens of converging voice with other data services in the enterprise. However, the benefits of doing so are very convincing and an idea already being proven by early adopters in the enterprise.

Lower Transmission Charges

Reduced monthly phone charges are one clear advantage of converging voice calls with the corporate data network. For some companies, this reason may be the most compelling of all, depending on several factors, including the volume of calls within the company and the distance between company offices. Companies with offices overseas stand to reap the biggest rewards by eliminating international long-distance charges, especially to countries with monopolistic long-distance markets. In addition to intracompany calls, savings can also be realized for extracompany calls. First the call is routed to a destination outside of the company over the corporate network to the nearest remote office. It then interfaces with the PSTN. As an example, a company with offices in Los Angeles and Tokyo could route calls to and from each office over the corporate network, then hand the call off to a local carrier for another destination outside of Tokyo but still within the country of Japan. In this case, not all but a significant portion of long-distance charges have been avoided.

Economically speaking, it is attractive to use data networks for voice for a couple of reasons. First, data networks nearly always offer spare capacity; second, voice typically requires little bandwidth when compared with other data transactions. Compression of voice data typically makes it possible to integrate the new service into the existing network without additional capacity investments.

Reduced Cost of Ownership

The lower cost associated with VoIP systems is not just relegated to lower monthly bills. Converged data networks also reduce the recurring cost of owning two separate networks, one for voice and one for data. With any network, there are both human and equipment costs associated with purchase, implementation, maintenance, software licensing, and traffic monitoring. Personnel costs have always been a major concern for businesses, and now more than ever it is cost-intensive to attract and retain qualified IT personnel. Convergence of voice service with other data services gives companies an effective way to streamline their networks and make the most of their IT manpower. This in turn gives VoIP adopters a distinct advantage in the marketplace.

Powerful New Converged Applications

Saving money is an important way for companies to promote the interests of their employees and shareholders. Beyond this, companies exist to be profitable, win in their markets, and better serve their customers. By providing a host of new converged voice and data applications, such as Web-enabled call centers, unified messaging, and real-time collaboration, VoIP allows companies to serve their customers better, giving them additional advantages in their markets.

As an example, Web-enabled call centers present a way to overcome the communication problems inherent in turning site visitors from browsers into buyers. Historically, Web-based interaction between buyer and seller has been problematic at best. However, Web-based call centers, allow a customer with a question simply to click on a hyperlink to initiate a conversation with a live customer service agent who can answer questions and move the customer to a purchase decision more quickly.

Real-time multimedia/audio conferencing, distance learning, and embedding voice links into electronic documents are other coming applications. And these are just the beginning. There are many VoIP-enabled applications yet to be thought of.

Voice Quality

Most companies view any degradation in the voice quality provided by current switched networks as a significant concern. Within a data network, packets move around in a somewhat nonlinear fashion, and sometimes they are even lost or dropped. This is especially true in Ethernet networks that make up the most common enterprise computing environments. With most data applications, this doesn't present much of a problem, because Ethernet or IP error correction readily compensates for these events. Voice data, where there must be an efficient, real-time packet flow throughout the network, is much more sensitive to these problems.

Reliability

Anyone who has worked with computers and network data applications understands the frustrations of having the system or the network down. At the same time, most take for granted that when they pick up their telephone receivers, there will always be dialtone, along with reliable, uninterrupted service. VoIP must achieve close to this same level of reliability before it as seen as a wholesale replacement for switched networks.

Technology Adoption Issues

As businesses move to adopt converged voice and data networks, there is the concern that they will prematurely buy into the wrong technology. First, there is the concern over "buyer's remorse," the idea that a superior solution will become available just as the company has invested significant dollars in another solution. Second, the more far-reaching concern is that buy-in to a particular solution will result in a long-term commitment to a particular (possibly inferior) architecture that may limit the ability to choose services and management tools.

The continuing evolution of VoIP standards, as discussed earlier in this chapter, will in many ways alleviate this concern over time.

Risk vs. Reward

The reality with VoIP and converged networks is simply put: They *are* coming, and the companies that implement them in an effective way will be the ones to benefit first from the new services and cost savings they enable. The companies that integrate VoIP technology will enjoy significant competitive advantages.

The Elements of a VoIP System

Every VoIP system is made up of four essential functions: hard and soft clients, communication servers, client and trunk media gateways, and application servers. Together, these building blocks are spread over a telephony or business-grade IP network, providing the necessary levels of reliability, voice quality, and traffic management.

Client Side

For end users in the enterprise, the VoIP client is their personal way of interfacing with the network. Traditional switched networks offer one primary interface, which is typically the phone sitting on the user's desk. VoIP client equipment, however, may take several forms, including a dedicated IP telephone (which may look very similar to an ordinary phone), a wireless LAN telephone, or a PC running client software. In general, fixed-function devices, such as the IP phone, are much more reliable, while PC-based systems offer more flexibility.

The number and types of client equipment, which will grow with market demand, roughly fall into two categories: thin clients and thick clients. As their name implies, thin clients feature less inherent intelligence, relying primarily on network intelligence provided by the communication server to initiate communications and manage feature operation. Likewise, thick clients are more highly integrated and use built-in intelligence to serve the user.

Communication Servers

The heart of every VoIP system in the enterprise is the communication server. It provides the control that allows call establishment across the network. How much intelligence a particular call server integrates depends on whether thick clients, such as PCs, or thin clients, such as IP phones, are being used as endpoints. The functional purpose of a communication server is to coordinate address translation and handle call signal processing, call setup, management, resource management, and admission control in an IP network. The state of active calls and accompanying logs are maintained by communications servers, and with thin clients, H.323 clients, and some SIP clients, they also record state tracking. Signaling is also typically the domain of communication servers because a network may support multiple protocol stacks.

With integrated application programming interfaces, communication servers can integrate application servers into a VoIP system, and they are easily replicated for high availability and networked for scalability. The two most common types of communication servers are server-based and purpose-built.

Server-Based Communication Servers

Software is the underlying technology for server-based communication servers. The software components reside on industry-standard computing platforms, such as Windows or UNIX servers, delivering call processing and resource management functions for both media gateways (the next VoIP functional element) and VoIP clients. Advantages of server-based implementations include easy integration of third-party software; disadvantages include susceptibility to outside hackers.

Purpose-Built Communication Servers

In the enterprise, purpose-built communication servers are typically integrated into a multifunctional PBX running an embedded operating system in a closed environment. The advantages of purpose-built servers are that they are scalable and secure, and through integration in existing equipment, they allow maximum return on equipment investment.

Media Gateways

Media gateways come in two varieties: client-side and trunk-side. They serve as liaisons between VoIP packet data and analog or digital T1 voice trunks and analog or digital telephone set interfaces. A media gateway's purpose is to provide media mapping and/or transcoding functions between the IP network and circuit-based networks. Specific functions include compression, silence suppression, and echo cancellation. In addition, media gateways are responsible for H.323 or SIP on the VoIP side and other signaling required on the client or trunk side.

The three common implementations of client- and trunk-side media gateways are:

1. Standalone—Used in networks that utilize purpose-built communication servers, standalone media gateways provide client- and/or trunk-side media gateway functionality. They can be implemented anywhere, and they can expand without regard to underlying network architecture.
2. Network-based—These media gateways are consolidated with other network devices, such as routers and access devices. The clear benefit of a network-based media gateway is that there are fewer systems to support, and its functions can be coupled with routing, bandwidth, and traffic management aspects of the network. Reliability is an issue, however, due to the fact that they require the same software fixes and upgrades as other network functions.

3. Integrated media gateways—Typically, integrated media gateways are found in association with a PBX. Their feature set is closely coupled with the PBX and includes: call routing, trunk selection, and telephony class-of-service capabilities. As an advantage, they maintain the reliability and scalability of the PBX, but their disadvantage is also tied to the PBX, in that they may not be as modular as a standalone or network-based gateway.

Application Servers

Application servers act as a bridge between the VoIP and legacy worlds. They support a host of services and applications, such as unified messaging, conferencing, and collaborative multimedia services, regardless of whether the client uses VoIP or legacy voice systems. They can reside anywhere on the network to maintain application performance balance and optimal network conditions.

The Future of VoIP

As with any new technology, widespread adoption depends on a complex mix of market needs and market resources set against the new application's potential to transform the user experience. Customers must see the clear advantages of the new technology over their present systems; they must have financial incentive to adopt the new technology; and they must feel confident that it will make them more productive, save money, or fulfill other pressing business or personal requirements. More and more every day, VoIP is living up to these tough criteria. Still, as illustrated earlier, there are important issues that the industry must continue to address. The continual evolution of complementary standards and robust product offerings are just a few.

So where is the VoIP market today, and where is headed? Currently, the enterprise, in the form of corporate intranets and commercial extranets, holds the most immediate promise. VoIP is particularly attractive in these environments because IP-based infrastructures allow operators to decide who can and cannot use the network.

The VoIP gateway is another factor propelling the proliferation of packet voice technology. Gateway functionality, which once resided on a PC-based platform, has now migrated to robust embedded systems, each able to process hundreds of simultaneous calls. For corporations, the economies of scale from this integration will allow them cost-effectively to deploy large numbers of VoIP connections that merge data, voice, and video into integrated networks. In fact, the reduced expenses and competitive cost advantages associated with integrated IP networks will be the most compelling factor for many companies.

Outside of the corporate intranet setting, commercial extranets are also pushing VoIP acceptance. These carefully engineered IP networks are already delivering voice and fax over the Internet to customers. As an example, many of the calling cards available at the local convenience store checkout stand are really private IP extranets that earn a profit on the margins that data networks enjoy over traditional switched networks.

VoIP on the Internet

Realistically, it will be a few years before VoIP is a widespread phenomenon in the home. VoIP products that use the Internet will, for the time being, be niche markets that can deal with the varying performance levels. In the next three to five years, however, VoIP on the Internet will make significant strides in the wake of two important developments:

- Backbone bandwidth and access speeds will increase by several orders of magnitude, made possible by the deployment of IP/ATM/synchronous optical network (SONET), cable modems, and digital subscriber lines (DSL).
- The Internet will become tiered, requiring users to pay for the particular services they need or want.

Fax over IP (FoIP) products and services over the public Internet will proliferate more quickly than voice and video, mainly because it is more cost-effective and less technologically challenging. In fact, it is believed that most corporations over the next couple of years will take their fax traffic off the PSTN and move it to the Internet and corporate intranet.

Summary

VoIP's adoption in the industry and the home is happening incrementally, yet inexorably. By the end of the decade, videoconferencing with data collaboration will be the standard way to communicate in the workplace. Video cameras in the workplace are now paving the way for inexpensive equipment in the home, a development that should propel packet video and accompanying packet voice into the mainstream. When this happens, VoIP will become interwoven in the fabric of daily life.