White Paper

Voice Over IP 101

Understanding VoIP Networks

Juniper Networks, Inc.
1194 North Mathilda Avenue
Sunnyvale, CA 94089 USA
408 745 2000 or 888 JUNIPER
www.Juniper.net

Part Number: 200087-003 January 2006

Gary Greenberg
Consulting Engineer – VoIP Media and Communications

Stefan Brunner
Senior Network Security Consultant

Akhlaq A. Ali
Senior Marketing Engineer

Scott Heinlein
Solutions Marketing Manager

Contents

Introduction

Although Voice over IP (VoIP) has existed for several years, it has only recently begun to take off as a viable alternative to traditional public switched telephone networks (PSTN). Interest in VoIP has grown in part because the technology can help organizations reduce costs by using a single IP network to support both data and voice applications.

But cost is not the only factor driving VoIP's growth. Service providers are also attracted by VoIP's revenue potential. Operators can use their VoIP networks to rapidly deploy new value-added and high-margin applications and services.

Although VoIP is an attractive alternative to traditional PSTN voice services, deploying VoIP is not as simple as flipping a switch. Before implementing a VoIP solution, organizations must consider the following issues:

- What impact will deploying VoIP have on the organization?
- What functionality will the organization require from its VoIP network?
- How can an organization evaluate whether their VoIP solution will be able to provide high-quality voice services?

Organizations can choose from a variety of equipment and networking protocols to implement their VoIP solution. Just as in data networking, identifying the appropriate equipment and technology for the VoIP network depends heavily on an organization's business and technical requirements.

This paper describes the basic networking functions, components, and signaling protocols in VoIP networks. It explores the ramifications of deploying VoIP as well as the service considerations that drive specific equipment and technology choices. This paper is intended to provide organizations with a general understanding of VoIP, so that they will be better prepared to solve the more complex issues associated with deploying a secure and assured VoIP network.

Understanding Basic PSTN and VoIP Network Functions

Before diving into the details of VoIP networking components and technologies, it is important to understand the basic network functions that make voice services possible. This section describes how PSTN and VoIP networks use:

- *Database services* to locate endpoints and translate between the addressing schemes used in two (usually heterogeneous) networks
- *Signaling* to coordinate the actions of the various networking components needed to complete a call between two endpoints
- *Call connect and disconnect (bearer control)* mechanisms to transport audio content
- *Coder-decoder (CODEC)* operations to convert analog waveforms to digital information for transport

Database Services

PSTN and VoIP networks use database services to locate the endpoints for a given call and to translate between the addressing schemes used by two (usually heterogeneous) networks. These database services typically include:

- A call control database that contains the address mappings and translations for the call's endpoints (the end users' phones)
- A report function that generates transaction reports to support billing
- Additional logic that can provide network security—for example, the ability to prevent a specific endpoint from making international calls

Unlike the PSTN, which identifies endpoints by their phone number[1], VoIP networks identify endpoints by their IP address and port number. Some VoIP networks use a Domain Name System (DNS) to abstract addresses.

PSTN and VoIP networks couple these database services with call state control and signaling to coordinate the activities of their network elements.

Signaling

Signaling enables individual network devices to communicate with one another. Both PSTN and VoIP networks rely on signaling to activate and coordinate the various components needed to complete a call.

In a PSTN network, phones communicate with an analog Class 5 switch or traditional digital private branch exchange (PBX) for call connection and call routing purposes.

In a VoIP network, the VoIP components communicate with one another by exchanging IP datagram messages. The format of these messages may be dictated by any of several standard protocols. The most commonly used signaling protocols—H.323, Session Initiation Protocol (SIP), Media Gateway Control Protocol (MGCP), and H.248—will be described later in this paper.

Call Connection and Audio Transport Mechanisms

To complete a call, two endpoints must be able to open and sustain a communication session.

The public or private switches in the PSTN complete calls by connecting logical Digital Signal-0 (DS-0) channels through the network. Each DS-0 is a 64 kbps bi-directional channel that the PSTN dedicates exclusively to the communication session for the duration of the call. The PSTN uses Pulse Code Modulation (PCM) to represent analog audio frequencies, enabling the network to transmit the audio payload through these DS-0 channels as a digitally encoded pulse amplitude value.

Like the PSTN, VoIP networks use PCM to encode the audio payload. However, instead of transmitting the audio payload directly over a dedicated DS-0 channel, VoIP networks transport the audio payload using shared network resources. To complete a connection, VoIP networks place a set of one or more PCM samples, known as *frames,* into an IP datagram. The VoIP

[1] The phone numbers used in the PSTN are often referred to as E.164 numbers per ITU specifications.

solution formats the datagrams according to the Real-Time Protocol (RTP); then forwards them over a routed or packet-forwarded IP network. Because the IP network does not implicitly allocate resources to these RTP packets, ensuring high-quality VoIP communications can pose a significant challenge to service providers and enterprises. The issues associated with ensuring VoIP service quality will be discussed later in this paper.

CODEC Operations

The fourth basic network function is the process the network uses to convert analog waveforms to digital information. Both PSTN and VoIP solutions use CODECs or voice coder-decoder (VOCODERs) to do this. The process that achieves this conversion is complex and well beyond the scope of this paper. For the purposes of this discussion, it is sufficient to say that there are many ways to transform an analog voice signal—all of which are governed by industry standards and most of which are based on PCM.

Each encoding scheme has its own history and merit. Each has its own bandwidth needs based on its compression capabilities. Table 1 lists some of the more important encoding standards covered by the International Telecommunications Union (ITU). Notice the tradeoff that the standards make between encoding efficiency, reduced bandwidth consumption, and increased conversion delay.

Table 1: ITU Encoding Standards

ITU Standard	Description	Bandwidth (Kbps)	Conversion Delay (ms)
G.711	PCM	64	< 1.00
G.721	ADPCM	32, 16, 24, 40	< 1.00
G.728	LD-CELP	16	~ 2.50
G.729	CS-ACELP	8	~ 15.00
G.723.1	Multirate CELP	6.3, 5.3	~ 30.00

Understanding VoIP Solution Components

Although VoIP networks take a different approach to fulfilling the four primary networking functions, the major components of VoIP network ultimately deliver very similar functionality to that of a PSTN. Consequently, VoIP networks can perform all of the same tasks that the PSTN does.

A typical VoIP network has five major components:

- VoIP phones, consoles and other devices, which end users use to initiate and receive VoIP calls
- The Call Processing Server/PBX, which manages all VoIP control connections
- One or more Media/PSTN-to-VoIP Gateways, which convert voice content for transport over the IP network
- The IP network, which transports the audio payload

- One or more Session Border Controllers (SBCs), which control real-time, session-based traffic at the signaling (call control) and transport layers as it crosses network borders and network domains

VoIP Phones, Consoles and Other End-User Devices

End users can initiate and receive VoIP calls using a variety of VoIP phones and consoles. VoIP phones can either be hardware-based devices that resemble traditional phones or software-based devices, known as *softphones*.

Softphones offer the same basic features as hardware-based VoIP phones, but are typically run on notebook computers. Softphones are primarily targeted to mobile users, who often use them to connect to the corporate network over a Secure Sockets Layer (SSL) based virtual private network (VPN). Once the VPN connection has been established, the mobile users can make and receive calls through the corporation's PBX as if they were at the office.

VoIP consoles are applications that offer certain control characteristics. Consoles usually include a softphone. They can also interact with a VoIP phone or a legacy phone via a voice gateway.

Call Processing Server/PBX

At the heart of a VoIP phone system is the call processing server/PBX that manages all VoIP control connections. Typically software-based, these VoIP call processing servers can be deployed as a single server, cluster of servers, or a server farm with distributed functionality.

VoIP communications require the network to transmit two types of traffic:

- The voice stream itself, known as voice media or the VoIP payload
- Traffic associated with the signaling mechanism used to establish calls, which is categorized as control traffic

Call processing servers generally do not handle VoIP payloads, except for a few situations. As Figure 1 shows, in a typical VoIP solution, the VoIP payload flows in a peer-to-peer fashion between the VoIP terminals in the network. The VoIP terminals determine the flow of the VoIP payload traffic, while the call processing servers negotiate those flows within the control messages.

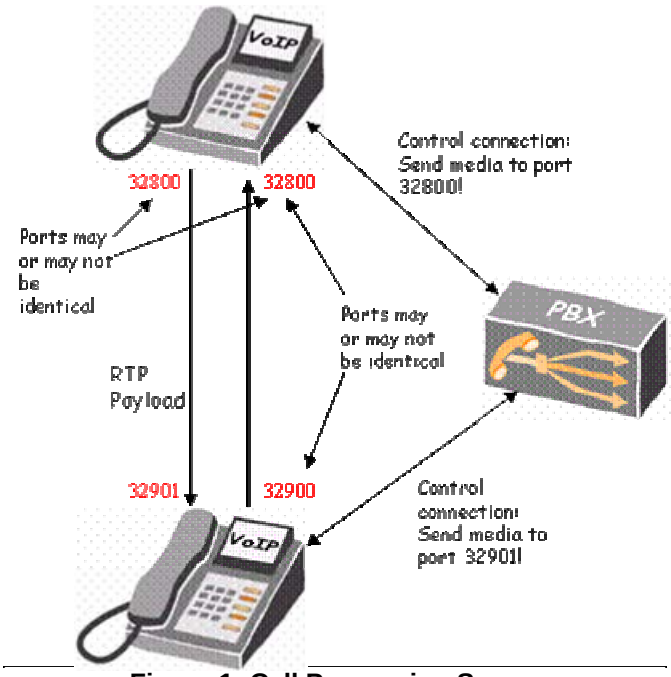

Figure 1: Call Processing Server

Media/VoIP Gateways/Gatekeepers

The VoIP network uses media gateways to perform the traditional CODEC functions—the analog-to-digital conversion of voice traffic—and to create voice IP packets. The media gateway provides the necessary interface for transporting voice content over the IP network and is the source of VoIP bearer traffic for that IP network. Typically, the media gateways convert each conversation or call into a single IP session that can be transported by a RTP that runs over User Datagram Protocol (UDP). Media gateways also provide optional features, such as analog and/or digital voice compression, echo cancellation, silence suppression, and statistics gathering.

Media gateways exist in several forms. For example, media gateways could be a dedicated telecommunication equipment chassis, or even a generic PC running VoIP software. In the case of an IP phone, the media gateway function resides in the handheld device.

Media gateway features and services can include some or all of the following:

- Trunking gateways that interface between the PSTN network and a VoIP network. Such gateways typically manage a large number of digital circuits
- Cable modem/cable set-top boxes, xDSL devices, broadband wireless devices and other residential gateways that provide a traditional analog interface to a VoIP network
- Small-scale (enterprise) VoIP gateways and other access media gateways that provide a traditional analog or digital PBX interface to a VoIP network

The industry sometimes uses the terms “gateway” and “gatekeeper” interchangeably. This reflects the fact that technology advances now enable the same device to perform the gatekeeping functions—Call Admission Control and bandwidth management—as well as the analog-to-digital

conversion and other voice processing tasks traditionally associated with media gateways.

The IP Network

IP networks have traditionally been used to transport data services for end users content with best-effort IP network transport. IP networks are also used to transport premium data services, such as those depended upon by a business enterprise, which demand better-than-best-effort transport. But voice requires real-time transport and is extremely sensitive to latency, packet loss and jitter. If an IP network is to carry both voice and data, it must be able to distinguish between these different traffic types and prioritize service delivery accordingly.

Class of service (CoS) ensures that packets of a specific application are given priority. The ability to prioritize traffic according to its CoS is a minimum requirement for real-time VoIP applications, because it enables the IP network to keep voice services from being affected by other traffic flows. Although the IP network must not reintroduce the inefficiencies of the dedicated PSTN circuits, the IP network's ability to support a virtual circuit or tunnel that enforces a differentiated service class is critical to delivering an assured user experience for voice services.

Session Border Controllers

Session Border Controllers (SBCs) control real-time, session-based traffic at the signaling (call control) and transport layers as it crosses network borders and between network domains. The SBC simultaneously supports VoIP signaling protocols associated with session management, such as SIP, H.323, H.248, and MGCP, as well as RTP and Real-time Transport Control Protocol (RTCP) flows associated with voice, video, or multimedia content. By correlating the signaling and media planes, SBCs are able to address the VoIP challenges of security, service assurance and quality of service (QoS), as well as interworking and regulatory compliance. Because the SBC can inspect the content of the session planes, it can also police bandwidth and isolate problems by network domain. These are all critical features—without them, service providers would not be able to meet service level agreements (SLAs) and would risk losing revenue and customers. Fundamentally, SBCs allow real-time, peer-to-peer services that require QoS and enhanced privacy. Voice is just one of many services that rely on SBCs to travel across a network infrastructure that was designed to—and is better suited to—support the asynchronous client server-oriented traffic flow common to the Internet.

Understanding VoIP Signaling Protocols

The signaling protocols implemented in a VoIP solution determine the features and functionality available on that network, as well as the way in which the VoIP solution components interact with one another. There are a variety of VoIP protocols and implementations, with a wide range of features that are currently deployed.

This section describes the signaling protocols commonly deployed in the field to establish VoIP connections and to signal control information between VoIP network elements, including:

- H.323, the ITU standard for establishing VoIP connections

- SIP, the Internet Engineering Task Force (IETF) standard for establishing VoIP connections
- MGCP, the first protocol developed by the IETF to signal control information between VoIP network components
- H.248, the protocol both the IETF and ITU use to signal control information between VoIP network elements

Although the ITU and the IETF govern the protocols used to deliver voice and other multimedia services over packet-based networks, the VoIP signaling protocols described in this section are all relatively new. Their definitions leave substantial room for interpretation. This is particularly true for SIP, the protocol the IETF uses to establish VoIP connections. Some vendors have implemented proprietary schemes to fill apparent gaps in the protocols or to add vendor-specific functionality. As a result, VoIP solutions often require interworking between vendor implementations.

Establishing VoIP Connections with H.323

The ITU's choice for establishing VoIP connections, H.323, is a packet-based multimedia communication system. The H.323 specifications define various signaling functions, as well as media formats related to packetized audio and video services.

Generally speaking, the H.323 standards were the first to classify and solve multimedia delivery issues over LAN technologies. However, as IP networking and the Internet became prevalent, many Internet RFC standard protocols and technologies were developed based on earlier H.323 ideas.

As Figure 2 illustrates, H.323 networks contain three primary solution components:

- Call Processing Servers, which provide call routing and communication to VoIP gateways and end user devices
- Media gateways, which serve both as the H.323 termination endpoint and interface with non-H.323 networks, such as the PSTN
- Gatekeepers, which function as a central unit for call admission control, bandwidth management and call signaling

Although gatekeepers are not required elements in H.323, they can increase the network's overall scalability by separating call control and management functions from the gateways.

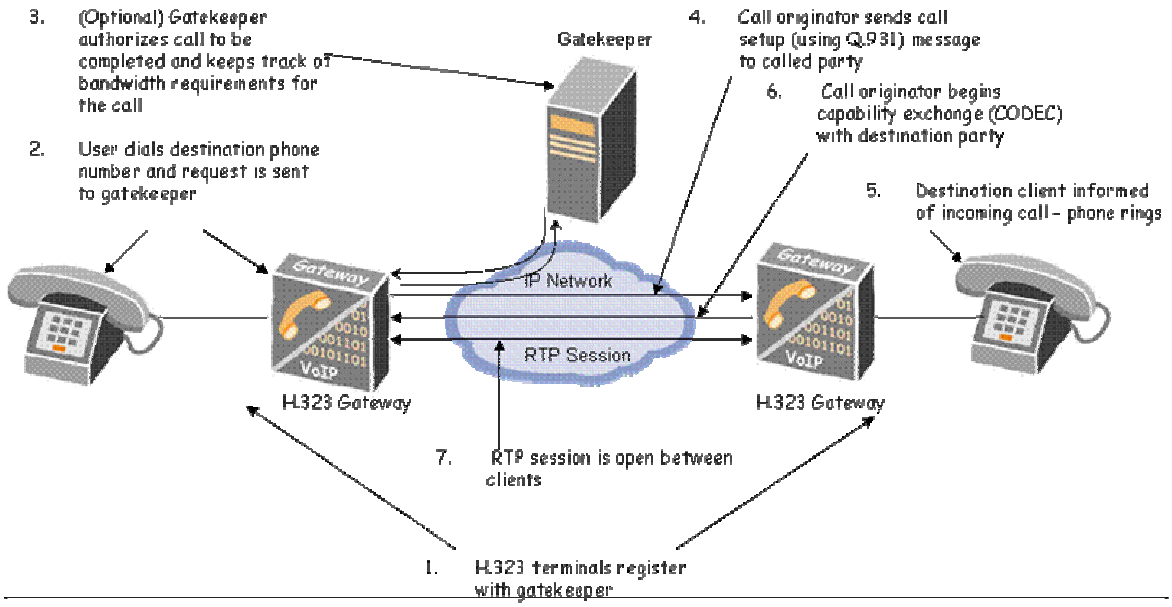

Figure 2: H.323 Call Setup Process

Establishing VoIP Connections with SIP

Many VoIP networks use the IETF's signaling protocol, SIP, to handle the setup and tear down of multimedia sessions between endpoints. This lightweight, text-based signaling protocol is transported over either Transmission Control Protocol (TCP) or UDP. SIP uses *invitations* to create Session Description Protocol (SDP) messages to carry out capability exchange and to setup call control channel use. These invitations allow participants to agree on a set of compatible media types.

The powerful SIP client-server application supports user mobility with two operating modes: *proxy* and *redirect.* In *proxy* mode (shown in Figure 3), SIP clients send requests to the proxy server. The proxy server either handles the requests or forwards them to other SIP servers. Proxy servers can insulate and hide SIP users by proxying the signaling messages. To the other users on the VoIP network, the signaling invitations look as if they are coming from the proxy SIP server.

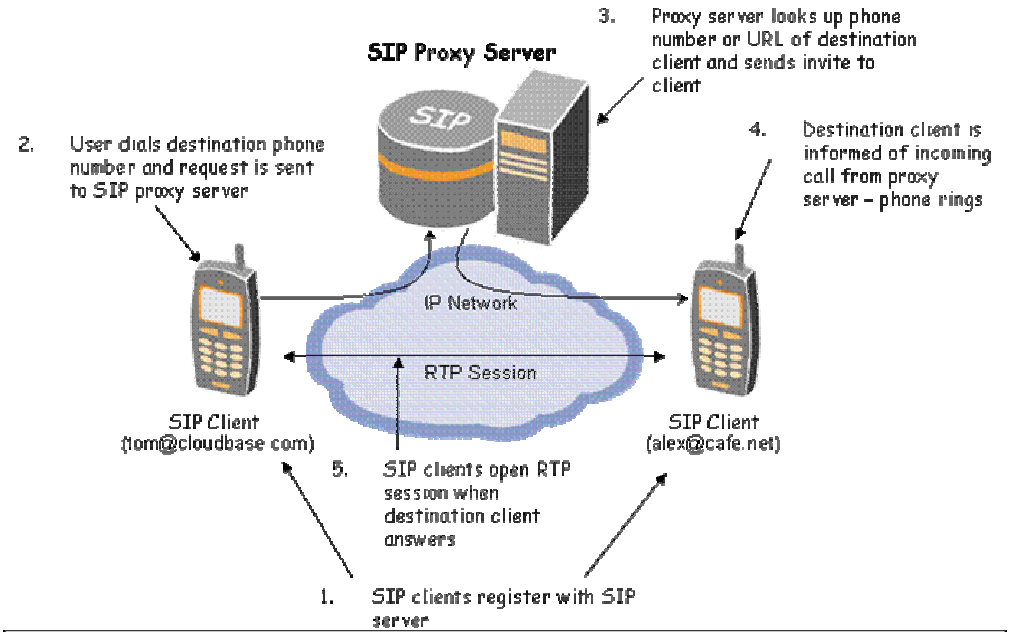

Figure 3: SIP Proxy Operation

When the SIP server is operating in redirect mode (shown in Figure 4), the SIP client sends its signaling request to a SIP server, which then looks up the destination address. The SIP server returns the destination address to the originator of the call, who uses it to signal the destination SIP client.

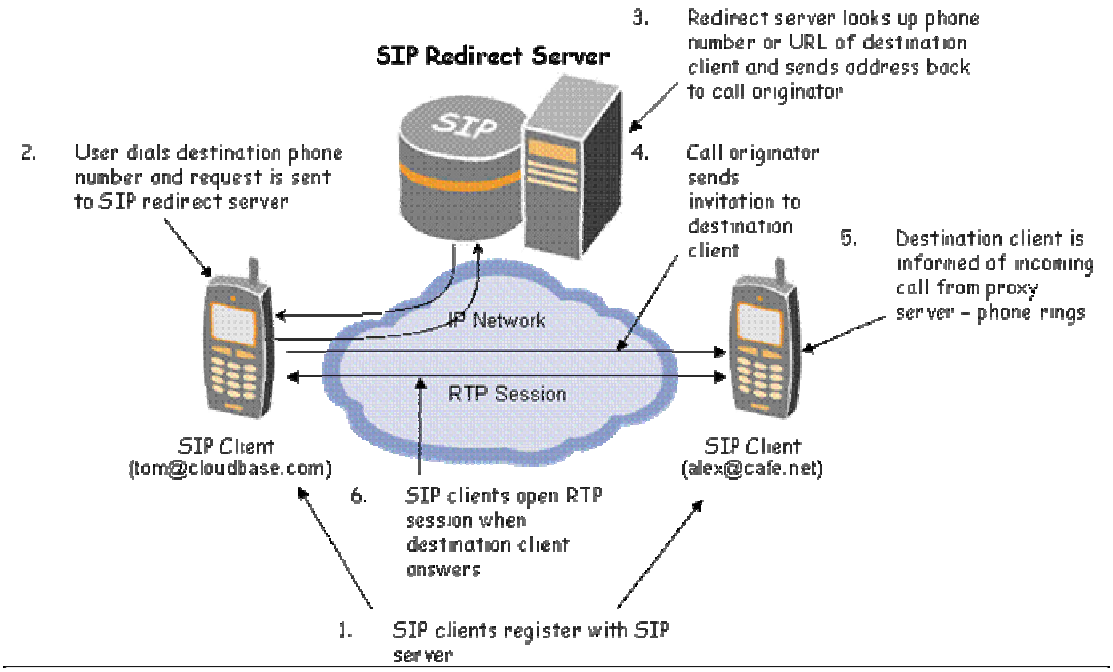

Figure 4: SIP Redirector Server

The ability to proxy and redirect requests to the end user's current location is critical to supporting a highly mobile voice user base. SIP enables users to inform the SIP server of their current location (IP address or URL) by sending a registration message to a *registrar*. As a result, although early VoIP

deployments were based on H.323, SIP has become the protocol of choice.

In fact, SIP (as defined in RFC 2543) is the basis for the IP Multimedia Subsystem multimedia data and control protocol framework that the IETF is developing in conjunction with the Third Generation Partnership Project (3GPP). IMS uses SIP and other standard interfaces between applications, network layers, and back-office systems to create a flexible framework that can deliver any kind of traffic—voice, data, video, or multimedia—over any wireless or wireline access network.

Signaling Control Information Between VoIP Network Elements

The ITU and IETF also define VoIP control protocols oriented at signaling control information between VoIP network elements. The earlier of these protocols, MGCP, is currently supported by the IETF as an informational standard. MGCP is an incomplete predecessor to H.248, which both the IETF and ITU use to signal control information between VoIP network elements. This section briefly examines both protocols.

An Early Approach: MGCP

As described in RFC 2705, MGCP mimics a softswitch architecture by dividing the role of traditional voice switches into three functional units: the media gateway, media gateway controller, and signaling gateway. This enables operators to manage each VoIP gateway independently as a separate entity.

MGCP is a master-slave control protocol that coordinates the actions of media gateways. The media gateway controller, also known as a call agent, manages the call-related signaling control intelligence, while the media gateway informs the media gateway controller of service events. As Figure 5 shows, the call agent instructs the media gateway to create and tear down connections when the calls are generated. In most cases, the call agent informs the media gateways to start an RTP session between two endpoints.

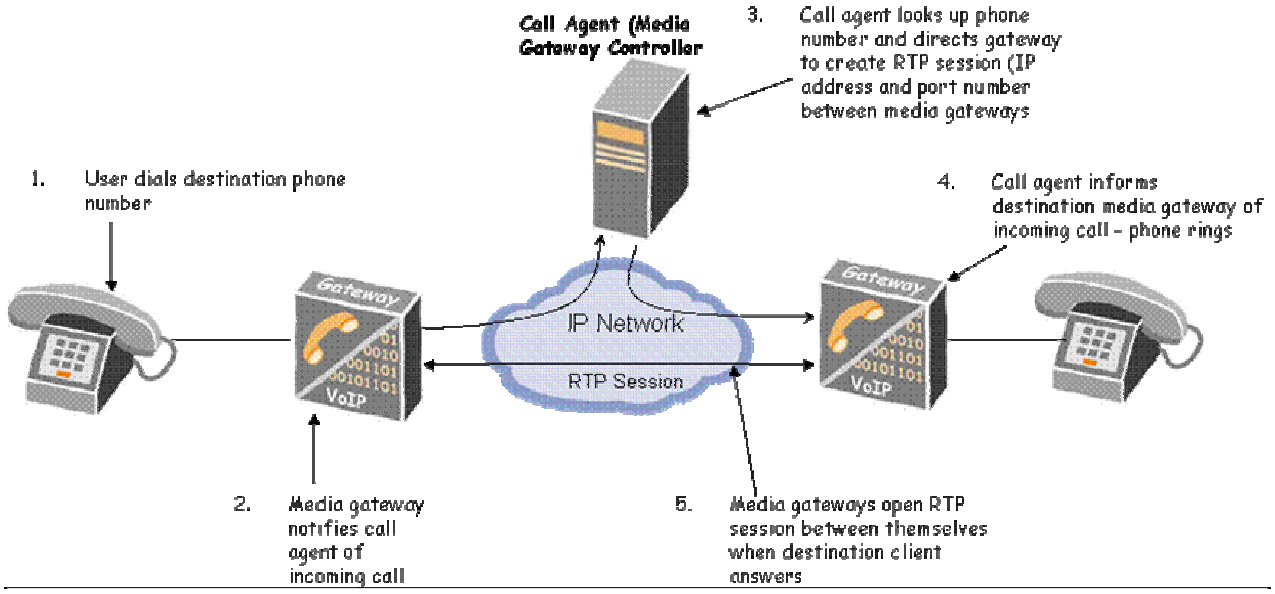

Figure 5: How MGCP Coordinates the Media Gateways

The Next Generation: Megaco/H.248

A newer standard, Megaco/H.248, is the result of collaboration between the IETF and ITU standards bodies. The Megaco architecture defines media gateways that provide media conversion and sources of calls, while media gateway controllers provide call control. The architecture's many similarities to MGCP are not coincidental, as MGCP is an interim version of the eventually completed Megaco/H.248. In fact, because Megaco addresses the same requirements as MGCP, there is an effort underway to merge the protocols.

Megaco defines a series of transactions coordinated by a media gateway controller to establish call sessions. Its primary focus is the promotion of standardized IP telephony equipment. Megaco design goals include:

- An IP phone that meets the basic needs of the business user from day one
- A path for rapid expansion to support sophisticated business telephony features
- A wide range of telephones and similar devices to be defined from very simple to very feature rich
- A simple, minimal design
- Device costs that are appropriate to the capabilities provided and package and termination types with characteristics that enable reliability
- An IP phone that meets the appropriate Megaco/H.248 protocol requirements, as provided in the Megaco requirements document, and that uses a straightforward application of the Megaco/H.248 protocol.

Handling VoIP Service Requirements

As is the case with most real-time services, VoIP demands that the network provide predictable performance within a constrained boundary of transport parameters. This section surveys the key networking issues that an organization or service provider must carefully consider when deploying a VoIP solution.

Minimizing Latency

Latency (also referred to as delay) is the time that it takes a packet to make its way through the network to the destination or terminating device. In other words, latency is the time it takes the speaker's voice to reach the listener's ear. Large latency values do not necessarily degrade the sound quality of a phone call, but large latency values can result in a lack of synchronization between the speakers. This can cause hesitations during the voice conversation and make it difficult to interact.

Several factors contribute to creating this latency in a multiservice network, including:

- The time it takes for the endpoints to create the packets used in voice services, known as *packet creation latency*

- The time it takes to serialize the digital data onto the physical links of the interconnecting equipment
- The time it takes an electrical (or photonic) signal to travel the length of a conductor, known as the *propagation delay*
- The time that a packet remains buffered in a network element while it awaits transmission, referred to as the *queuing delay*
- The time it takes a network device (router, switch, firewall, etc.) to buffer a packet and make the forwarding decision, known as *packet forwarding delay*

When designing a multiservice network, the total delay that a signal or packet exhibits is the sum of all the latency contributors. Generally, it is accepted that the end-to-end latency should be less than 150 ms for toll quality phone calls. The remainder of this section describes some basic steps operators can take to assess and mitigate the impact of each latency contributor in a multiservice network.

Assessing packet creation latency

Packet creation concerns exist on both the originating and destination endpoints of a given voice connection. At the originating endpoint, the packet creation delays vary according to the amount of time it takes the endpoint to fill a packet with data. Generally, voice packets tend to be smaller, which helps to minimize the amount of latency added by the packet creation process.

On the receiving side, the media gateway must remove and further process the packet data. All considerations being equal, nominal operation of any media gateway unit should not exceed 30 ms.

Mitigating serialization delays

The second source of latency, the serialization delay, is inversely proportional to the link speed. The faster the media, the less time it takes to serialize the digital data onto the physical links, and the lower the overall latency. The impact on latency depends somewhat on the link technology used and its access method. For example, it takes 125 microseconds to place one byte on a 64Kb circuit. Placing the same byte on an OC-3/STM-1 circuit takes 0.05 microseconds.

Although some delay is unavoidable regardless of the bandwidth used, keeping the number of intervening links small and using high bandwidth interfaces reduces the overall latency.

Calculating propagation delays

Because the speed of electrical or photonic signals through a conductor is always slower than the speed of light, there will always be some propagation delay. However, propagation delay only becomes an issue for signals or packets that must travel a great distance.

Operators use the following formula to calculate the propagation delay:

Propagation delay = Circuit km / (299,300 km x .6)

Example: To calculate the one-way propagation delay of a 6,000 km fiber run (discounting any signal repeaters in between), solve the equation:

Propagation delay = 6000 km / (299,300 km x .6) = 0.0334 sec

By this calculation, the latency contributed by propagation delay alone is 33.4 ms.

Mitigating queuing delays

The fourth latency contributor, the queuing delay, varies according to the network traffic load. Operators can usually configure the length of time a packet waits in a given network element's buffer before being transmitted—with a smaller number being better for latency values. However, queuing delays also depend on the amount of traffic the network element is trying to pass through a given link, so will increase with network load. Consequently, operators must set aside adequate bandwidth and resources for voice traffic. Queues that are not serviced fast enough and that are allowed to grow too large will result in greater latency.

Assessing packet forwarding delays

The final latency contributor, the packet forwarding delay, is determined by the time it takes a router, switch, firewall or other network device to buffer a packet and make the forwarding decision. Among the forwarding considerations are which interface to forward the packet to and whether to drop or forward the packet against an Access Control List (ACL) or security policy. Packet forwarding delay varies depending on the function and architecture of the networking device. If a packet must be further buffered as a part of its processing, greater latency is incurred.

Assessing the Impact of Jitter on Voice Quality

Jitter is the time difference between when a packet is expected to arrive to when it actually arrives. In other words, given a constant packet transmission rate of every 20 ms, new packets would be expected to arrive at the destination exactly every 20 ms. Unfortunately, as Figure 6 shows, this is not always the case. In Figure 6, packet one (P1) and packet three (P3) arriving when expected, but packet two (P2) arrives 12 ms later than expected and packet four (P4) arrives 5 ms late.

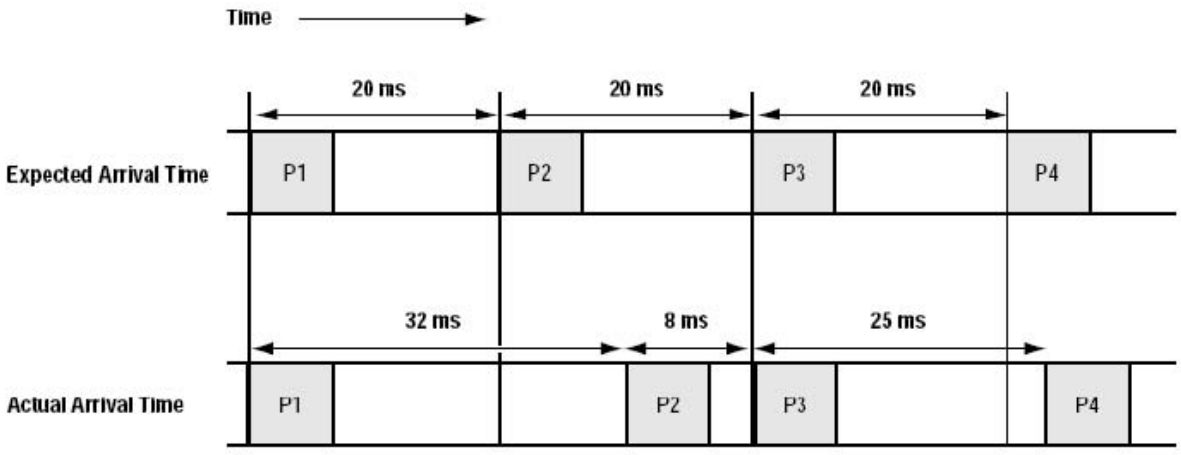

Figure 6: Jitter Example

Jitter in a multiservice network is most frequently caused by queuing variations due to dynamic changes in network traffic loads. Jitter also results when one or more packets take a different equal-cost link that is not physically (or electrically) the same length as the links other voice packets are using.

Most media gateways have *play-out buffers* that buffer a packet stream, so that the reconstructed voice waveform is not affected by packet jitter. Play-out buffers can minimize the effects of jitter, but cannot eliminate severe jitter.

Although some amount of jitter is to be expected, severe jitter can cause voice quality issues because the media gateway might discard packets arriving out of order. In this condition, the media gateway could starve its play-out buffer and cause gaps in the reconstructed waveform.

Calculating Bandwidth Requirements

Operators can calculate the amount of bandwidth required to support voice traffic with fairly simple math. But deciding how much bandwidth to allocate to each service in a converged voice and data network requires careful consideration of the organization's priorities and the available bandwidth that organization can afford. Allocating too little bandwidth to voice services could result in unacceptable quality. Among the factors to consider when making bandwidth calculations are:

- *The impact of bandwidth priority:* Voice traffic typically receives higher priority than other data traffic due to the sensitive nature of VoIP packets.
- *The trade-off between compression and voice quality:* Operators can use CODEC technology to compress VoIP packets and reduce the amount of bandwidth. However, compression decreases the overall quality of the voice call, forcing operators to balance the bandwidth saved against quality considerations.
- *The projected peak use:* Operators allocate network bandwidth based on the projected number of calls during peak hours. Oversubscribing voice bandwidth can decrease voice quality. Operators must also set aside adequate bandwidth for signaling to ensure that calls are complete and to reduce service interruptions.

The remainder of this section illustrates how these factors come into play in a typical bandwidth calculation.

A Sample Bandwidth Calculation

The formula for calculating required voice traffic bandwidth is relatively straightforward. Operators calculate RTP bearer voice bandwidth usage for a given number of phone calls as follows:

bits per sec = samples per sec x packet size x number of calls x 8 bits per sec

where samples per sec = 1,000 ms / packet creation rate

Example: To calculate the bandwidth required to support 2,000 full-duplex G.711 encoded voice channels that have a packet creation rate of 20 ms and a packet size of 200 bytes (40 byte IP header + 160 byte payload), perform the following calculations:

Samples per second = 1,000 ms / 20 ms = 50

Bandwidth (bits per sec) = 50 samples/sec x 200 bytes x 2,000 calls x 8 bits/sec = 160 Mbps

Note that this number is a raw measure of IP traffic and represents only the bearer (voice) content. It does not take into account the overhead used by the transport media (the links between the routers) and data-link layer protocols. To determine the link speed needed to support this number of calls, operators must add this raw IP value to that of the overhead.

Signaling bandwidth requirements vary depending on the rate at which the calls are generated and the signaling protocol used. If a large number of calls are initiated in a relatively short period, the peak bandwidth needs for the signaling could be quite high. In general, the maximum amount of bandwidth required by an IP signaling protocol is roughly three percent of all bearer traffic. Using the previous example, signaling bandwidth requirements if all 2,000 calls were initiated in one second, would be approximately 4.8 Mbps (3 percent of 160-megabits).

Therefore, the total bandwidth needed to support the bearer and signaling traffic generated by 2000 G.711 encoded calls would be approximately 164.8 MB. This bandwidth requirement is a theoretical maximum for this specific case. Changes in the call initiation rate, voice encoding method, packet creation rate, compression and silence suppression or other parameters would result in changes to the bandwidth requirements.

For large VoIP implementations with sizable bandwidth requirements, it is imperative that the IP network delivers the needed service with predictably high performance.

Compensating for Packet Loss

Packet loss can occur for many reasons, and in some cases, is unavoidable. During network congestion, routers and switches can overflow their queue buffers and be forced to discard packets. Packet loss for non-real-time applications, such as Web browsers and file transfers, is undesirable, but not critical. The protocols used by non-real-time applications, usually TCP, have retransmission capabilities that enable them to tolerate some amount of packet loss.

Real-time applications based on UDP are significantly less tolerant of packet loss. UDP does not have retransmission facilities, however, retransmissions would almost never help. In an RTP session, by the time a media gateway could receive a retransmission, it would no longer be relative to the reconstructed voice waveform; that part of the waveform in the retransmitted packet would arrive too late.

Although packet loss of any kind is undesirable, some voice packet loss can be tolerated as long as the loss is spread out over a large amount of users. Voice quality is not generally affected if the amount of packet loss is less than five percent for the total number of calls.

Ensuring Reliability

Although network failures are rare, it is essential to plan for them. Operators need failover strategies for situations when network devices malfunction or links are broken. To address these situations, operators must establish redundant links between network devices and/or deploy redundant

equipment. Planning redundant schemes for media gateways and media gateway controllers can also increase reliability.

IP networks use routing protocols to exchange routing information. As part of their operation, routing protocols monitor the status of interconnecting links. Routing protocols typically detect and reroute packets around a failure if an alternate path exists. The time required to detect and recalculate an alternate path can vary depending on the interconnecting media used for these links.

Having media gateways and media gateway controllers that can actively detect the status of their next-hop address (default gateway) as part of their failover mechanism decreases the likelihood of a large service disruption. Operators could also directly connect the media gateway and media gateway controller to the router. Depending on the nature of the failure, this setup could enable the network devices to immediately detect a link failure and take appropriate action.

Setting Security

VoIP networks are vulnerable to many of the same security risks that data networks are, including Denial of Service (DoS) attacks, service theft, tampering, and fraud. Many conventional firewalls cannot combat VoIP attacks because VoIP is implemented at both the signaling and media layers. To secure VoIP solutions, the security device must be able to support VoIP protocols like SIP, MGCP, and H.323, and to associate state at the signaling layer with packet flows at the media layer.

VoIP presents its own set of security vulnerabilities, mainly related to the VoIP signaling path and media exchange path. The good news is that SBCs offer a number of security features. For instance, SBCs filter VoIP sessions based on multiple criteria. Only trusted traffic that meets service provider-defined conditions is allowed to traverse the SBC to pre-determined destinations. They also screen traffic classification and packet processing engines to protect internal VoIP and non-VoIP equipment in the network against DoS attacks.

Firewalls can perform security functionality. But to be able to dynamically open and close ports for the VoIP traffic only for the duration of a call, firewalls need to understand the VoIP signaling protocols used in the network. Otherwise, VoIP calls cannot go through the firewall unless a range of ports are opened – which exposes the network for unauthorized access.

Operators should establish firewall policies to protect communications between servers and VoIP end-devices. These policies should restrict VoIP communication based on authorized end-devices or traffic sourced or destined for a particular IP address or interface. Firewalls can also be used to segment the VoIP network, separating the voice traffic from other traffic to ensure appropriate priority and policies are applied. Firewalls may also be placed to mitigate DoS attacks and to create logs for forensics. Finally, operators can deploy intrusion prevention systems to help detect and prevent certain attacks, such as manipulated Dynamic Host Configuration Protocol (DHCP) messages or flooded forwarding tables.

Supporting Commonly Available VoIP Network Solutions

VoIP and other real time applications bring new challenges for service providers and enterprises. Networks need to be more intelligent, secure and

have a higher level of performance. Fortunately, new technologies are available that service providers and enterprises can leverage to help meet these new requirements.

When designing a network to support VoIP and real time applications, some of the factors service providers need to consider include application requirements, available budget, quality of service requirements, and downtime ramifications. This section describes three common approaches to designing a network to support VoIP services:

- Best-effort network design
- Differentiated services approach
- Traffic Engineered Multiprotocol Label Switching (MPLS-TE) network

Each approach presents a trade-off between benefits and cost. Best-effort networks are less complex and more cost-efficient but do not address most of the requirements of VoIP. The differentiated services approach adds some network intelligence but still falls short compared to MPLS-TE networks. As this section will show, MPLS-TE networks are the best equipped to address the VoIP requirements that are so critical to providing the level of service voice users expect.

Supporting VoIP Services with a Best-Effort Network

A best-effort network provides just what its name describes—a network that does its best to deliver packets in a timely manner. It is the least complex and costly network approach and is the design most networks use today. Best-effort networks work well for non-sensitive traffic types such as web browsing and e-mail since delays in these services generally do not significantly impact the user experience.

Best-effort networks leverage Interior Gateway Protocol (IGP) technologies such as Open Shortest Path First (OSPF) and Intermediate System to Intermediate System (IS-IS) to provide connectivity for routing packets between hosts. IGP protocols use a Shortest Path First (SPF) algorithm to build routing tables. The routing engine references these routing tables at each router hop traversed by packets.

Operators provide QoS by over-provisioning their best-effort network so that network congestion does not introduce unwanted levels of latency, jitter, and packet loss. Over-provisioning a network would appear to be a simple way to provide the necessary level of QoS. But though this approach requires a significant amount of added bandwidth, it still cannot guarantee service levels in an operational environment. Historically, to provide toll-grade services during predictable periods of network usage, operators must over-provision bandwidth resources by a factor between 10:1 and 20:1.

Best-effort networks are not ideal in dealing with temporary or permanent outages. IGP will advertise the outage and initiate a route table reconstruction based on the modified topology. This process is called route re-convergence, and if not designed carefully or optimized, can take seconds to stabilize. Web browsing or Email users do not typically notice these seconds. However, this can be detrimental to VoIP users in the middle of a conversation since both latency and jitter can be impacted.

Supporting VoIP with a Differentiated Services Network

Some operators enhance the best-effort approach by adding differentiated service capabilities to their networks. Differentiated service adds state information to each packet—allowing the router to identify different service flows and direct queuing and forwarding treatment appropriate to the service requirements. This enables routers to identify voice packets and mark them for higher priority treatment over less sensitive packets. The IETF created this technique, known as Differentiated Services (DiffServ), as part of its Differentiated Service Code Points (DSCP) specification and the associated QoS framework.

To implement a Differentiated Services approach, the network administrator defines a number of service classes that corresponds to the number of QoS-dependant services that will be transported within a network domain. Each service class is assigned a 32-bit code (value) and a corresponding per-hop treatment. This treatment, often referred to as the *Forwarding Equivalency Class (FEC) policy*, infers handling priority and queue depth, as well as the rate policing or drop policy. The typical router interface can support eight different queues. Each queue can support a provisioned service class policy or DSCP. At the ingress interface, the router classifies each packet for queuing based on its DSCP value (or any single or combination of layer 2 and/or layer 3 parameters). The router services its queues according to a provisioned algorithm, such as weighted round-robin (WRR) and strict-priority queuing (SPQ).

Operators commonly define two service classes to handle bearer (audio) and signaling traffic in their VoIP solutions. Bearer traffic benefits from a high priority FEC with a shallow depth queue and no Random Early Discard (RED) rate policing drop profile. If voice is deployed in a network transporting only voice and one class of data, the FEC may be specified as a SPQ. But in converged multiservice networks that also carry video, no service should receive SPQ FEC treatment, as it will cause unwanted service delays in non-SPQ services that require high priority treatment.

The differentiated services approach enables operators to manage QoS schemes by assigning voice traffic priority. However, because this approach provides QoS awareness independently for each node during packet transport, it does not provide any fault tolerance when the service needs it most. Additionally, the differentiated services framework alone does not support any proactive traffic engineering design capabilities.

Supporting VoIP Solutions with Traffic-Engineered MPLS

MPLS is widely recognized as the most modern approach to providing the network functionality required by VoIP and other real-time applications. MPLS combines layer 2 and layer 3 technologies to support a mix of services. From the lower-layer ATM and PSTN technologies, MPLS has borrowed the concepts of simple label lookup-based packet forwarding, a-priori out-of-band forwarding path allocation, and service-time resource commitment. From the layer 3 IGP and Differentiated Services QoS framework, MPLS borrows the concept of advertising network topology and link attributes and providing multiservice forwarding paths to mitigate network complexity and enhance scale.

Traffic Engineered MPLS technologies have been designed from the start to address the complexities and high availability requirements of carrier-grade VoIP and other premium services. Fast re-route, auto-bandwidth

provisioning, and virtual path TE tunnels are some of the MPLS capabilities that guarantee the high level of quality and reliability that we expect from telephony services.

MPLS defines Label-Switched Paths (LSP) which are simple uni-directional forwarding paths constructed by wrapping ATM, IP, and other transport protocols packets in MPLS frames. MPLS identifies each frame with a label. The ingress Label Edge Router (LER) provisions the labels and distributes them to Label-Switching Routers (LSRs) using a signaling protocol such as Label Distribution Protocol (LDP) or Resource Reservation Protocol-traffic engineering (RSVP-TE) prior to enabling transport across the path. The label distribution process involves an automated sequence of resource requests and acknowledgements that create a path between two points in the network. When using RSVP-TE, QoS parameters may be specified as a requirement to each LSR. When acknowledged, these QoS parameters represent an agreement to provide that level of QoS, or FEC, to packets forwarded along the path.

Service providers can construct customized LSPs that support specific application requirements. Operators can design LSPs to minimize the number of hops, meet certain bandwidth requirements, support precise performance requirements, bypass potential points of congestion, direct traffic away from the default path selected by the IGP, or simply force traffic across certain links or nodes in the network.

An important benefit of the label-swapping forwarding algorithm is its ability to take any type of user traffic, associate it with an FEC, and map the FEC to an LSP that has been specifically designed to satisfy the FEC's requirements. Adding DSCP support to the MPLS network allows the network to populate a single LSP with multiple FECs.

The MPLS-TE approach also enables network administrator to provision Fast Re-Route (FRR) paths for LSPs and associated backup paths while minimizing physical LSR overlap between primary and backup paths. FRR limits path outage times to milliseconds by pre-negotiating resource borrowing from LSR neighbors and localizing the event signaling that implements the FRR operation.

Deploying technologies based on label-swapping forwarding techniques offers network administrators precise control over traffic flow in their networks. This unprecedented level of control results in a network that operates more efficiently and provides more predictable service.

Choosing a VoIP Equipment Vendor

Just as data networks require different equipment and protocols based on an operator's business and technical requirements, individual VoIP solutions can use different equipment and protocols, depending on the business and technical requirements at hand. Although the variety of VoIP protocols has caused some confusion in the marketplace, it is precisely this protocol flexibility that makes VoIP-based voice systems so much more useful than legacy voice systems. Companies should assess VoIP equipment vendors according to their:

- Ability to support different network transport service models, such as Traffic Engineered MPLS.
  Care must be taken in addressing VoIP's extremely rigid QoS and reliability requirements. Choosing an experienced vendor with

expertise in deploying large, complex multiservice IP networks will pay significant dividends.

- Commitment to supporting open standards within their products. Any vendor should be actively developing voice strategies that consider interoperability with all VoIP protocols. Without this commitment, VoIP systems are in danger of becoming as proprietary as legacy voice systems.
- Ability to support multiple protocols.
  Support for multiple protocols creates a VoIP solution that is better positioned to handle future system migrations and incorporate products that support a different protocol without requiring wholesale infrastructure redeployments or significant network upgrades.
- End-to-end support for all VoIP protocols.
  Vendors must provide solutions that work in both single and multi-protocol environments.

Conclusion

Although VoIP is an attractive alternative to traditional PSTN voice services, deploying VoIP is not a simple process. Before choosing a VoIP solution, organizations should consider both the required functionality and the potential issues associated deploying a VoIP network. These service considerations drive the protocol and equipment choices organizations when designing their VoIP solution. Although the wide range of VoIP protocols has caused some confusion in the marketplace, it is precisely this protocol flexibility that makes VoIP-based voice systems so much more useful than legacy voice systems.

In designing their VoIP solution, organizations also need to consider how their chosen solutions will address the latency, jitter, bandwidth, packet loss, reliability, and security issues raised in this paper. By working with vendors that can provide this VoIP flexibility, companies can take advantage of the efficiencies of VoIP while building scalable and reliable networks that can meet the needs of the next generation of services.