



# Understanding Packet Voice Protocols

## Introduction

Over the past decade, the telecommunications industry has witnessed rapid changes in the way people and organizations communicate. Many of these changes spring from the explosive growth of the Internet and from applications based on the Internet Protocol (IP). The Internet has become a ubiquitous means of communication, and the total amount of packet-based network traffic has quickly surpassed traditional voice (circuit-switched) network traffic (DataQuest, 1998).

In the wake of these technology advancements, it has become clear to telecommunications carriers, companies, and vendors that voice traffic and services will be one of the next major applications to take full advantage of IP. This expectation is based on the impact of a new set of technologies generally referred to as voice over IP (VoIP) or IP telephony.

VoIP supplies many unique capabilities to the carriers and customers who depend on IP or other packet-based networks. The most important benefits include the following:

- *Cost Savings:* By moving voice traffic to IP networks, companies can reduce or eliminate the toll charges associated with transporting calls over the public switched telephone network (PSTN). Service providers and end users can also conserve bandwidth by investing in additional capacity only when it is needed. This is made possible by the distributed nature of VoIP and by reduced operations costs as companies combine voice and data traffic onto one network.
- *Open Standards and Multivendor Interoperability:* By adopting open standards, both businesses and service providers can purchase equipment from multiple vendors and eliminate their dependency on proprietary solutions.
- *Integrated Voice and Data Networks:* By making voice "just another IP application," companies can build truly integrated networks for voice and data. These integrated networks not only provide the quality and reliability of today's PSTN, but they also enable companies to quickly and flexibly take advantage of new opportunities within the changing world of communications.

In 1995, the first commercial VoIP products began to hit the market. These products were targeted at companies looking to reduce telecommunications expenses by moving voice traffic to packet networks. Early adopters of VoIP networks built toll-bypass solutions to take advantage of favorable regulatory treatment of IP traffic. Without any established standards, most early implementations were based on proprietary technology.

As these packet telephony networks grew and interconnection dependencies emerged, it became clear that the industry needed standard VoIP protocols. Several groups took up the challenge, resulting in independent standards, each with its own unique characteristics. In particular, network equipment suppliers and their customers were left to sort out the similarities and differences between four different signaling and call-control protocols for VoIP:

- H.323
- Media Gateway Control Protocol (MGCP)
- Session Initiation Protocol (SIP)
- H.248/Media Gateway Control (MEGACO)

In the process of implementing workable VoIP solutions, network engineers had to determine how each of these protocols worked and which ones were best for particular networks and applications.

This paper provides some guidance and understanding of these VoIP protocols and tries to clarify some of the confusion in the marketplace.

## Topics

1. The Great Voice Myth
2. Making Sense of the VoIP Standards
3. Understanding Centralized and Distributed Architectures
4. H.323
5. MGCP/H.248/MEGACO
6. SIP
7. Interconnecting VoIP Protocols
8. Conclusion: Packet Voice and VoIP Is About Services, Not Protocols

Self-Test

Correct Answers

Glossary

# 1. The Great Voice Myth

"The Great Voice Myth" states that there is only one way to build voice networks and that there should be only one voice protocol for each function in a packet voice network. Although this vision of network nirvana has been discussed in many academic circles, the reality is that multiple VoIP protocols and architectures have already been deployed, they will exist for the foreseeable future, and many networks will continue to be built using multiple VoIP protocols.

Just as today's data networks were built over time using multiple protocols and applications, the VoIP networks of today and tomorrow will be constructed using the protocols and applications that best fit the associated technology and business requirements.

The question that companies must ask is not "Which protocol is best?" but "Which services do we want to deploy, and which VoIP protocols best support those services?" The answer to the first question may reflect the bias of a particular vendor or standards body. The answer to the second question depends entirely on the unique requirements of each network implementation. Because no two businesses or networks are exactly the same, each company will answer the second question in its own way.

## 2. Making Sense of the VoIP Standards

VoIP comprises many standards and protocols. Basic terminology must be understood in order to comprehend the applications and usage of VoIP. The following definitions serve as a useful starting point (the protocols are listed in alphabetical order):

- *H.248* is an ITU Recommendation that defines "Gateway Control Protocol." H.248 is the result of a joint collaboration between the ITU and the IETF. It is also referred to as IETF RFC 2885 (MEGACO), which defines a centralized architecture for creating multimedia applications, including VoIP. In many ways, H.248 builds on and extends MGCP.
- *H.323* is an ITU Recommendation that defines "packet-based multimedia communications systems." In other words, H.323 defines a distributed architecture for creating multimedia applications, including VoIP.
- *IETF* refers to the Internet Engineering Task Force (<http://www.ietf.org>), a community of engineers that seeks to determine how the Internet and Internet protocols work, as well as to define the prominent standards.

- *ITU* is the International Telecommunication Union (<http://www.itu.int/home>), an international organization within the United Nations System (<http://www.unsystem.org>) where governments and the private sector coordinate global telecom networks and services.
- *MEGACO*, also known as IETF RFC 2885 and ITU Recommendation H.248, defines a centralized architecture for creating multimedia applications, including VoIP.
- *MGCP*, also known as IETF RFC 2705, defines a centralized architecture for creating multimedia applications, including VoIP.
- *Real-Time Transport Protocol (RTP)*, also known as IETF RFC 1889, defines a transport protocol for real-time applications. Specifically, RTP provides the transport to carry the audio/media portion of VoIP communication. RTP is used by all the VoIP signaling protocols.
- *SIP*, also known as IETF RFC 2543, defines a distributed architecture for creating multimedia applications, including VoIP.

### 3. Understanding Centralized and Distributed Architectures

In the past, all voice networks were built using a centralized architecture in which dumb endpoints (telephones) were controlled by centralized switches. Although this model worked well for basic telephony services, it mandated a trade-off between simplified management and endpoint and service innovation.

One of the benefits of VoIP technology is that it allows networks to be built using either a centralized or a distributed architecture. This flexibility allows companies to build networks characterized by both simplified management and endpoint innovation, depending on the protocol used.

In general, centralized architectures are associated with MGCP and H.248/MEGACO protocols. These protocols were designed for a centralized device—called a media gateway controller or call agent—that handles switching logic and call control. The centralized device talks to media gateways, which route and transmit the audio/media portion of the calls (the actual voice information).

In centralized architectures, the network intelligence is centralized and endpoints are relatively dumb (with limited or no native features). Although most centralized VoIP architectures use MGCP or H.248/MEGACO protocols, it is also possible to build SIP or H.323 networks in a centralized fashion using back-to-back user agents (B2BUAs) or gatekeeper routed call signaling (GKRCS), respectively.

Advocates of centralized VoIP architectures favor this model because it centralizes management, provisioning, and call control. It simplifies call flows for replicating legacy voice features. And it is easy for legacy voice engineers to understand. Critics of centralized architectures claim that it stifles innovation of endpoint features and that it will become a hindrance when building VoIP services that move beyond legacy voice features.

Distributed architectures are associated with H.323 and SIP protocols. These protocols allow network intelligence to be distributed between endpoints and call-control devices. *Intelligence* in this instance refers to call state, calling features, call routing, provisioning, billing, or any other aspect of call handling. The endpoints can be VoIP gateways, IP phones, media servers, or any device that can initiate and terminate a VoIP call. The call-control devices are called gatekeepers in an H.323 network, and proxy or redirect servers in a SIP network.

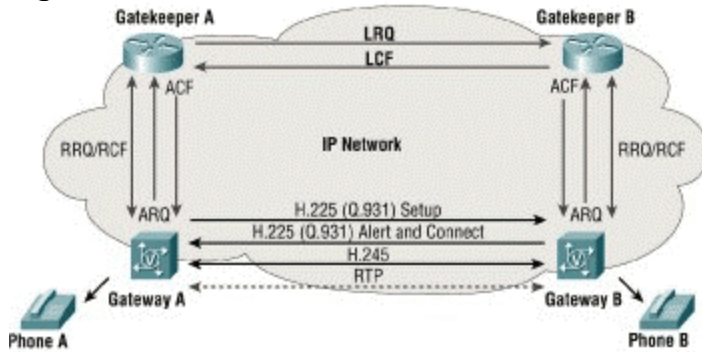
Advocates of distributed architectures favor this model because of its flexibility. It allows VoIP applications to be treated like any other distributed IP application, and it allows the flexibility to add intelligence to either endpoints or call-control devices, depending on the business and technology requirements of the network. Distributed architectures are usually well understood by engineers who run IP data networks. Critics of distributed architectures commonly point to the existing PSTN infrastructure as the only reference model that should be used when trying to replicate legacy voice services. They also note that distributed networks tend to be more complex.

## 4. H.323

H.323 was originally created to provide a mechanism for transporting multimedia applications over local-area networks (LANs). Although H.323 is still used by numerous vendors for videoconferencing applications, it has rapidly evolved to address the growing needs of VoIP networks. Because of its early availability and these advancements, H.323 is currently the most widely used VoIP signaling and call-control protocol, with international and domestic carriers relying on it to handle billions of minutes of use each year.

H.323 is considered an "umbrella protocol" because it defines all aspects of call transmission, from call establishment to capabilities exchange to network resource availability. H.323 defines Registration, Admission, and Status (RAS) protocol for call routing, H.225 protocols for call set-up, and H.245 protocols for capabilities exchange.

**Figure 1. H.323 Networks**



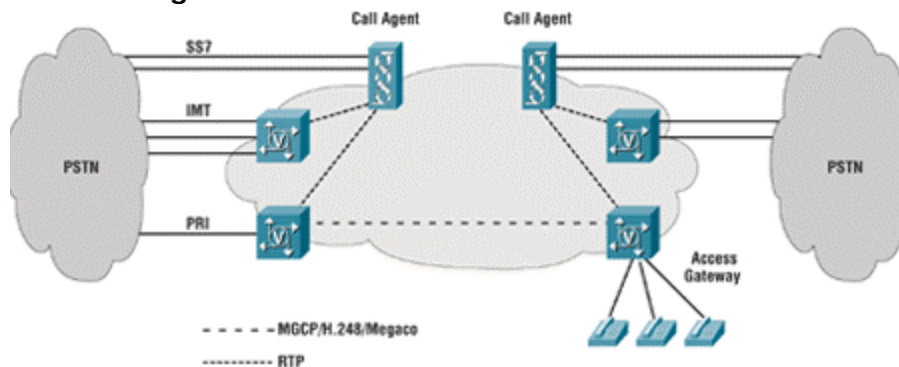
H.323 is based on the Integrated Services Digital Network (ISDN) Q.931 protocol, which allows it to easily interoperate with legacy voice networks such as the PSTN or Signaling System 7 (SS7). As a protocol used in a distributed architecture, H.323 allows companies to build large-scale networks that are scalable, resilient, and redundant. It provides mechanisms for interconnecting with other VoIP networks and supports network intelligence on either the endpoints or the gatekeepers.

## 5. MGCP/H.248/MEGACO

MGCP and H.248/MEGACO were designed to provide an architecture where call control and services could be centrally added to a VoIP network. In that sense, an architecture using these protocols closely resembles the existing PSTN architecture and services.

MGCP and H.248/MEGACO define most aspects of signaling using a model called packages. These packages define commonly used functionality, such as PSTN signaling, line-side device connectivity, and features such as transfer and hold. In addition, session definition protocol (SDP) is used to convey capabilities exchange.

**Figure 2. MGCP/H.248/MEGACO Networks**



In a centralized architecture, MGCP and H.248/MEGACO allow companies to build large-scale networks that are scalable, resilient, and redundant. It provides

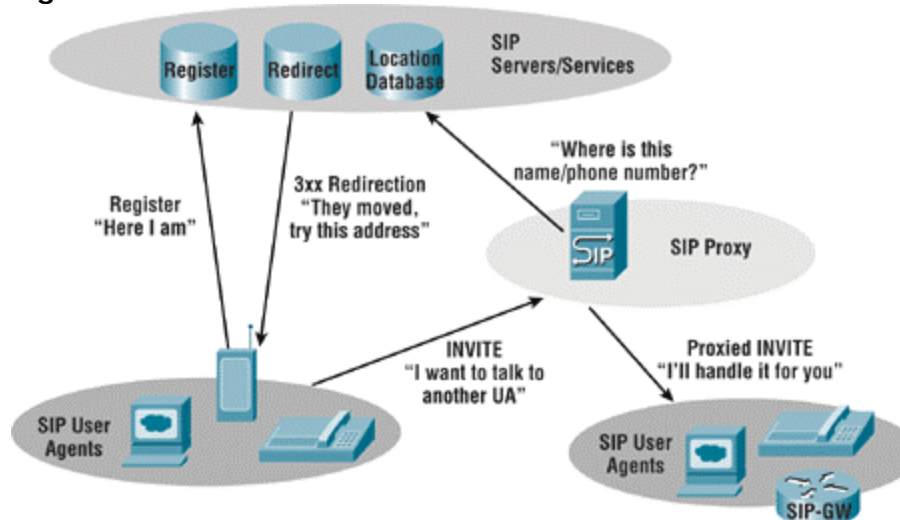
mechanisms for interconnecting with other VoIP networks and for adding intelligence and features to the call agent.

## 6. SIP

SIP was designed as a multimedia protocol that could take advantage of the architecture and messages found in popular Internet applications. By using a distributed architecture—with universal resource locators (URLs) for naming and text-based messaging—SIP attempts to take advantage of the Internet model for building VoIP networks and applications. In addition to VoIP, SIP is used for videoconferencing and instant messaging.

As a protocol, SIP only defines how sessions are to be set up and torn down. It utilizes other IETF protocols to define other aspects of VoIP and multimedia sessions, such as SDP for capabilities exchange, URLs for addressing, Domain Name Systems (DNSs) for service location, and Telephony Routing Over IP (TRIP) for call routing.

Figure 3. SIP Networks



Although the IETF has made great progress defining extensions that allow SIP to work with legacy voice networks, the primary motivation behind the protocol is to create an environment that supports next-generation communications models that utilize the Internet and Internet applications.

As a protocol used in a distributed architecture, SIP allows companies to build large-scale networks that are scalable, resilient, and redundant. It provides mechanisms for interconnecting with other VoIP networks and for adding intelligence and new features on either the endpoints or the SIP proxy or redirect servers.

## 7. Interconnecting VoIP Protocols

VoIP networks continue to be deployed at a rapid pace, and VoIP vendors and service providers continue to add new functionality. Because vendor support for each protocol differs and companies have varying business requirements, it is very likely that VoIP networks will continue to be made up of multiple protocols.

Having various protocols gives customers the flexibility that they need to connect services from multiple carriers. Using standards, even multiple standards, still simplifies deployment of multivendor endpoints and increases options for network management and provisioning.

As companies expand their networks, they are faced with choices about how to interconnect segments using differing VoIP protocols. These choices often fall into one of three categories:

- *Translation through Time Division Multiplexing (TDM):* In this model, a company uses either TDM equipment or VoIP gateways to translate from one protocol domain to another. The benefits of this model are that it can be used today. The downside is that it introduces latency into the VoIP network and involves yet another protocol translation (VoIP no. 1 <—> TDM <—> VoIP no. 2). This model is usually considered as a short-term solution until IP-based protocol translators are available.
- *Single Protocol Architecture:* In this model, a company moves all its VoIP devices and services to a single protocol, simplifying the network as a whole. The downside to this approach is that it might not be possible to migrate existing equipment to support the new protocol, a situation that can limit the company's ability to take advantage of some existing services. In addition, it limits the potential connectivity to other networks that are using other VoIP signaling protocols.
- *Protocol Translation:* In this model, a company uses IP-based protocol translators to interconnect two or more VoIP protocol domains. IP translators allow a company to retain the flexibility of using multiple VoIP protocols, do not introduce the delay problems that additional TDM interconnections do, and do not require a wholesale replacement or swap of existing equipment.

The downside to this approach is that there is no standard for protocol translation, so not all VoIP protocol translators are exactly the same. Although the IETF has attempted to define a model for translating H.323 to SIP, it involves more than just building a protocol-translation box.

As shown in *Table 1*, although protocols are somewhat similar, they do have some differences. Vendors of protocol translators need in-depth knowledge of all the

protocols being used in the VoIP network, and they must be aware of how various VoIP components utilize different aspects of the protocol.

For example, H.323 and SIP can send dual-tone multifrequency (DTMF) digits in either the signaling path or the media path (via RTP). But H.323 mandates only that the H.245 signaling path be used, and SIP does not specify how DTMF should be carried. This means that SIP devices could be sending DTMF in the media path (RFC 2833), and H.323 devices could be sending DTMF in the signaling path (H.245). If the VoIP protocol translator cannot properly recognize both the signaling path and the media path, then it might not function properly.

**Table 1. Details of VoIP Products**

	<b>H.323</b>	<b>SIP</b>	<b>MGCP/H.248/MEGACO</b>
Standards Body	ITU	IETF	MGCP/MEGACO— IETF;H.248—ITU
Architecture	Distributed	Distributed	Centralized
Current Version	H.323v4	RFC2543-bis07	MGCP 1.0, Megaco, H.248
Call control	Gatekeeper	Proxy/Redirect Server	Call agent/media gateway controller
Endpoints	Gateway, terminal	User agent	Media gateway
Signaling Transport	Transmission Control Protocol (TCP) or User Datagram Protocol (UDP)	TCP or UDP	MGCP— UDP;Megaco/H.248—both
Multimedia capable	Yes	Yes	Yes
DTMF—relay transport	H.245 (signaling) or RFC 2833 (media)	RFC 2833 (media) or INFO (signaling)	Signaling or RFC 2833 (media)
Fax—relay transport	T.38	T.38	T.38
Supplemental services	Provided by endpoints or call control	Provided by endpoints or call control	Provided by call agent

## 8. Conclusion: Packet Voice and VoIP Is About Services, Not Protocols

Just as companies choose various protocols for their data networks, they will choose various protocols for their VoIP requirements, depending on the business and technical requirements at hand. Although the variety in 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 choose vendors based on three very important requirements:

- Customers need vendors that are committed to supporting open standards within their products and are 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.
- Customers need products that support multiple protocols. This way, if a company finds that it needs to migrate its systems or add products that support a different protocol, it will not be required to perform upgrades to the network.
- Customers need voice solutions with end-to-end support for all VoIP protocols, meaning vendors must provide solutions that work in both single-protocol and multiprotocol environments.

By working with vendors that can provide this VoIP flexibility, companies can focus on building scalable and reliable networks that support the requirements of next-generation networks.

### Self-Test

1. The total amount of traditional voice (circuit-switched) traffic is greater than the total amount of packet-based network traffic on public networks.
  - a. true
  - b. false
2. The IETF and ITU have agreed on a single VoIP protocol for all networks.
  - a. true
  - b. false

3. The IETF defines which VoIP protocols?
  - a. SIP and MGCP
  - b. H.323 and H.248
  - c. MGCP only
  - d. SIP and H.323
  
4. What is the standard protocol for Fax-Relay, used by SIP, H.323, and MGCP?
  - a. T.39
  - b. T.38
  
5. SIP is the most widely used protocol in VoIP networks today.
  - a. true
  - b. false
  
6. Service providers should look for VoIP vendors that provide which of the following?
  - a. Support for standard VoIP protocols
  - b. Support for multiprotocol VoIP networks
  - c. Support for end-to-end interoperability of all VoIP protocols
  - d. None of the above
  - e. All of the above
  
7. Distributed architectures are associated with the H.323 and SIP protocols.
  - a. true
  - b. false
  
8. H.248 and MEGACO are the same protocol.

- a. true
  - b. false
9. Which are characteristics of centralized architectures? (select all correct answers)
- a. Intelligent endpoints
  - b. Features deployed on a call agent
  - c. Easily replicates TDM features
  - d. Works best when using the H.323 protocol
10. Which are characteristics of distributed architectures? (select all correct answers)
- a. Uses H.248 or MEGACO
  - b. Uses SIP or H.323
  - c. Uses gatekeepers for call routing
  - d. Dumb endpoints

## Correct Answers

1. The total amount of traditional voice (circuit-switched) traffic is greater than the total amount of packet-based network traffic on public networks.
- a. true
  - b. false**
2. The IETF and ITU have agreed on a single VoIP protocol for all networks.
- a. true
  - b. false**
3. The IETF defines which VoIP protocols?
- a. SIP and MGCP**

- b. H.323 and H.248
  - c. MGCP only
  - d. SIP and H.323
4. What is the standard protocol for Fax-Relay, used by SIP, H.323, and MGCP?
- a. T.39
  - b. T.38**
5. SIP is the most widely used protocol in VoIP networks today.
- a. true
  - b. false**
6. Service providers should look for VoIP vendors that provide which of the following?
- a. Support for standard VoIP protocols
  - b. Support for multiprotocol VoIP networks
  - c. Support for end-to-end interoperability of all VoIP protocols
  - d. None of the above
  - e. All of the above**
7. Distributed architectures are associated with the H.323 and SIP protocols.
- a. true**
  - b. false
8. H.248 and MEGACO are the same protocol.
- a. true**
  - b. false

9. Which are characteristics of centralized architectures? (select all correct answers)

- a. Intelligent endpoints
- b. Features deployed on a call agent**
- c. Easily replicates TDM features**
- d. Works best when using the H.323 protocol

10. Which are characteristics of distributed architectures? (select all correct answers)

- a. Uses H.248 or MEGACO
- b. Uses SIP or H.323**
- c. Uses gatekeepers for call routing**
- d. Dumb endpoints

## Glossary

### Definitions

#### **ACF (Admission Confirmation)**

The confirmation message in response to an H.323 ARQ.

#### **ARQ (Admission Request)**

An H.323 message from an H.323 endpoint to an H.323 gatekeeper, requesting the IP address of a remote device.

#### **DNS (Domain Name System)**

System used in the Internet for translating names of network nodes into addresses.

#### **DTMF (Dual Tone Multifrequency)**

Use of two simultaneous voice-band tones for dialing (such as touch tone.)

#### **H.248**

An ITU recommendation that defines “gateway control protocol.” H.248 is the result of a joint collaboration between the ITU and the IETF. It is also referred to as IETF RFC 2885 (MEGACO), which defines a centralized architecture for

creating multimedia applications, including VoIP. In many ways, H.248 builds on and extends MGCP.

### **H.323**

H.323: An ITU recommendation that defines “packet-based multimedia communications systems.” In other words, H.323 defines a distributed architecture for creating multimedia applications, including VoIP. H.323 is an extension of the ITU–T standard H.320 that enables videoconferencing over LANs and other packet-switched networks, as well as video over the Internet.

### **IETF (Internet Engineering Task Force)**

A community of engineers that seeks to determine how the Internet and Internet protocols work, as well as to define the prominent standards. This task force consists of more than 80 work groups and operates under the auspices of ISOC (Internet Society, an international nonprofit organization, founded in 1992, that coordinates the evolution and use of the Internet. They are headquartered in Reston, Virginia, United States).

### **IMT (Inter-Machine Trunk)**

The bearer channels associated with SS7 links.

### **IP (Internet Protocol)**

Network layer protocol in the TCP/IP stack offering a connectionless internetworking service. IP provides features for addressing, type-of-service specification, fragmentation and reassembly, and security.

### **ISDN (Integrated Services Digital Network)**

Communication protocol, offered by telephone companies, that permits telephone networks to carry data, voice, and other source traffic.

### **ITU (International Telecommunication Union)**

An international organization within the United Nations System (<http://www.unsystem.org>) where governments and the private sector coordinate global telecom networks and services (a function formerly performed by the CCITT, Consultative Committee for International Telegraph and Telephone).

### **LCF (Location Confirmation)**

The confirmation message in response to an H.323 LRQ.

### **LRQ (Location Request)**

A request message between two H.323 gatekeepers to find the address of an H.323 endpoint.

### **MEGACO**

Also known as IETF RFC 2885 and ITU Recommendation H.248, defines a centralized architecture for creating multimedia applications, including VoIP.

### **MGCP (Media Gateway Control Protocol)**

Also known as IETF RFC 2705, defines a centralized architecture for creating multimedia applications, including VoIP.

### **PRI (Primary Rate Interface)**

ISDN interface to primary rate access. Primary rate access consists of a single 64-Kbps D channel plus 23 (T1) or 30 (E1) B channels for voice or data. Compare with BRI (Basic Rate Interface), ISDN interface composed of two B channels and one D channel for circuit-switched communication of voice, video, and data.

### **Protocol translation**

In this model, a company uses IP-based protocol translators to interconnect two or more VoIP protocol domains. IP translators allow a company to retain the flexibility of using multiple VoIP protocols, do not introduce the delay problems that additional TDM interconnections do, and do not require a wholesale replacement or swap of existing equipment.

### **PSTN (Public Switched Telephone Network)**

General term referring to the variety of telephone networks and services in place worldwide. Sometimes called POTS.

### **RAS (Registration, Admission, and Status Protocol)**

The suite of messages between H.323 endpoints and H.323 gatekeepers. The messages are made up of Requests (xRQ), Confirmations (xCF), and Rejections (xRJ). The “x” letter denotes the type of messages, including Admission, Registration, Bandwidth, and Disconnect.

### **RCF (Registration Confirmation)**

The confirmation response from the H.323 gatekeeper to an H.323 endpoint.

### **RRQ (Registration Request)**

The Registration request from an H.323 endpoint to an H.323 gatekeeper.

### **RTP (Real-Time Transport Protocol)**

Also known as IETF RFC 1889, defines a transport protocol for real-time applications. RTP is designed to provide end-to-end network transport functions for applications transmitting real-time data—such as audio, video, or simulation data—over multicast or unicast network services. RTP provides services such as payload type identification, sequence numbering, timestamping, and delivery monitoring to real-time applications. RTP is used by all the VoIP signaling protocols.

### **SDP (Session Definition Protocol)**

RFC 2327. The media negotiation signaling used by SIP and MGCP.

### **Single Protocol Architecture**

In this model, a company moves all of its VoIP devices and services to a single protocol, simplifying the network as a whole. The downside to this approach is that it might not be possible to migrate existing equipment to support the new protocol, a situation that can limit the company's ability to take advantage of some existing services. In addition, it limits the potential connectivity to other networks that are using other VoIP signaling protocols.

### **SIP (Session Initiation Protocol)**

Also known as IETF RFC 2543, defines a distributed architecture for creating multimedia applications, including VoIP.

### **SS7 (Signaling System 7)**

Standard CCS system used with B-ISDN and ISDN. Developed by Bellcore. CCS (Common Channel Signaling) is a signaling system used in telephone networks that separates signaling information from user data. A specified channel is exclusively designated to carry signaling information for all other channels in the system.

### **TDM (Translation through Time Division Multiplexing)**

Technique in which information from multiple channels can be allocated bandwidth on a single wire based on pre-assigned time slots. Bandwidth is allocated to each channel regardless of whether the station has data to transmit.

### **TRIP (Telephony Routing over IP)**

RFC 3219. A protocol for distributing E.164 numbering plans between VoIP devices, similar to how BGP or OSPF distribute IP routes between routers.

### **VoIP (Voice over IP)**

The ability to carry telephony-style voice over an IP-based network with POTS-like functionality, reliability, and voice quality. VoIP enables an IP network to carry voice traffic (for example, telephone calls and faxes). VoIP occurs because the TDM voice is packetized by a DSP, which resides in a VoIP gateway, of the VoIP phone. Once the voice is packetized, it can then be routed over an IP network. VoIP separates the signaling over voice calls from the bearer portion of voice calls (the audio). The signaling is handled by H.248, H.323, MGCP, or SIP. The bearer is carried by RTP.