VoIP Development in China

Several factors have spurred the explosive growth of VoIP phone use in China, including customer incentives such as improved voice quality and lower cost per call, and provider incentives such as higher profits and upgrade paths to next-generation technologies.

Runsheng Wang
Beijing University of Posts and Telecom

Xiaorui Hu
Saint Louis University

The *Voice over Internet Protocol*, also called IP telephony, offers a new type of service that uses the Internet Protocol, intranets, and extranets to deliver voice information. In contrast to traditional telephone services, which operate through a circuit-switched network, VoIP uses a packet-switched network. This distinction results in differences in implementation, quality of service (QoS), and operating costs.

During its long history, traditional telephone service has expanded across most of the world. In contrast, VoIP offers a relatively new type of telephone service that has been available for fewer than 10 years. At this point, the development of VoIP telephone services cannot be totally separated from traditional telephone networks. Although VoIP services have partially supplanted traditional toll telephone services, when users make a VoIP telephone call, they must still go through a local telephone network.

Since the service was introduced to the public in China in April 1999, VoIP toll telephone traffic has increased with astonishing speed. By the end of 2002, VoIP toll telephone traffic had surpassed traditional toll telephone traffic in China in both domestic long-distance and international call areas, including phone calls to and from Hong Kong, Macao, and Taiwan. Four factors have contributed to this phenomenon:

- *Price advantage*. Given current price regulations on toll telephone services in China, VoIP toll charges are only about one-third to one-half the cost of traditional telephone charges. Although VoIP telephone providers cannot guarantee that voice quality will match that of traditional telephone service, the price advantage alone remains attractive enough to draw customers to VoIP phone services.
- *New and profitable area for ISPs*. VoIP telephone services provide an opportunity for Internet service providers to earn higher profits, especially those ISPs new to the telecommunications market. This new opportunity has fostered a new class of ISPs in China: Internet telephone service providers.
- *Benefits for traditional telephone service providers*. Although VoIP has made inroads into the toll telephone service market, total demand for traditional toll telephone services has actually increased. For example, from 2000 to 2002, domestic and international toll telephone traffic in China increased 107 percent and 110 percent, respectively. Both domestic and international calls more than doubled within this three-year period. Because most VoIP toll telephone traffic must use the local telephone networks, the increase in VoIP services brought an increase in local telephone network usage as well. Thus, the growth of VoIP services has actually increased profits for traditional telephone service providers.
- *Potential value in the transition to next-generation networks*. Development of VoIP telephone services can facilitate the transition from current telecommunications networks to IP-based next-generation networks. This tech-

Published by the IEEE Computer Society 0018-9162/04/$20.00

Table 1. Comparison of traditional and IP telephone services.

Traditional telephone service	IP telephone service
Circuit-switching technology	Packet-switching technology
Uses synchronous time-division multiplexing in transmission, resulting in lower channel utilization	Uses asynchronous time-division multiplexing in transmission, resulting in higher channel utilization
When network congestion occurs, calls will be blocked, but once call connection is established, the voice signal will not be lost	When network congestion occurs, calls can be blocked or IP packets can be lost, resulting in reduced voice quality
Uses the G.711 Pulse Code Modulation voice-encoding scheme without compression and achieves a transmission speed of 64 Kbps	Usually uses voice-compression encoding, with the bit rate of encoded voice data ranging from as high as 16 Kbps to as low as 5.3 Kbps
Short end-to-end transfer delay except in satellite communications and limited delay variation	Relatively long end-to-end transfer delay and significant delay variation
Guaranteed good voice quality	Voice quality affected significantly by the IP network's quality of service; absent specific measures, voice quality cannot be guaranteed
Given the separate network built to provide telephone services, reducing operational costs is difficult	Sharing network resources by combining with data and other multimedia services on the same IP network helps reduce operational costs

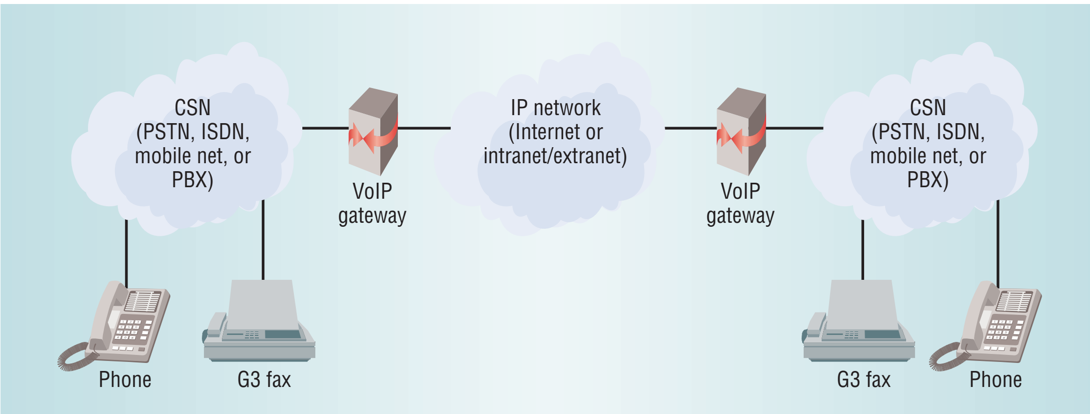


Figure 1. VoIP phone-to-phone implementation. A gateway must be set up to facilitate internetworking between the circuit-switched and IP networks.

nology, and the benefits it will provide, make its implementation crucial to telecommunications network and service providers.

In China, VoIP achieved rapid growth in both public and private networks. Recently, some of China's large-scale intranets have begun integrating phone and data services into the same network through VoIP, reducing telecommunications costs by eliminating the use of a separate phone network. Table 1 provides a point-by-point comparison of VoIP and traditional telephone services.

VOIP AND RELATED PROTOCOLS

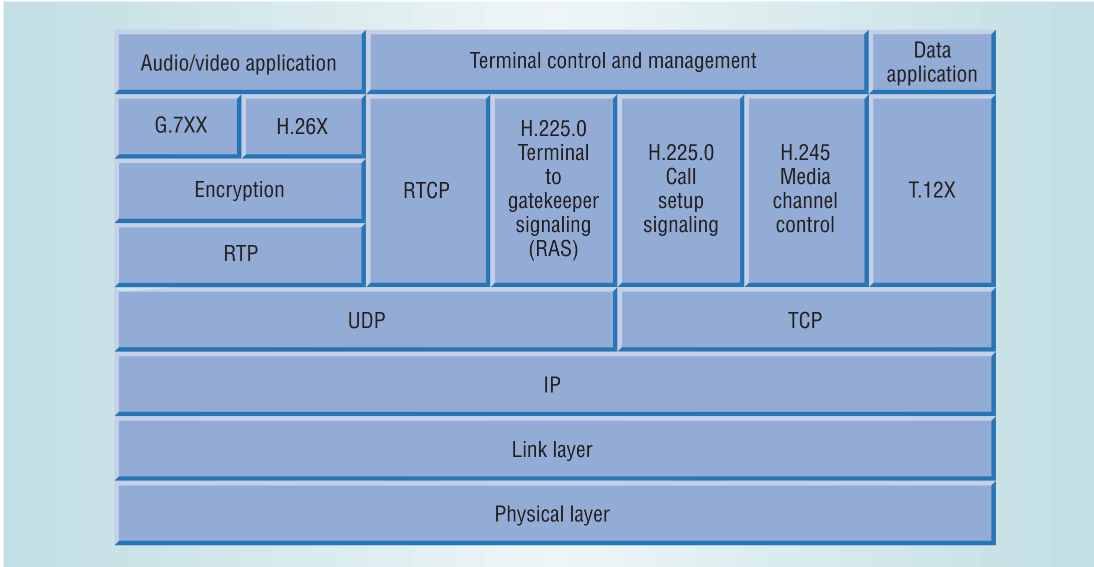
Currently, VoIP implementations can choose from three terminal types: PC with a microphone, traditional phone, and an IP phone. Each of these terminals connects to a different network, and they should be able to internetwork with each other. A new phone type, the IP phone, can connect to IP networks through a local area network (LAN).

Practically, a phone-to-phone implementation still predominates in VoIP communications. The phones connect to the IP network through a local circuit-switched network—either PSTN, ISDN, mobile net, or PBX—at both users' sites, with the VoIP serving as a new type of trunk that connects two local circuit-switched networks to provide toll telephone services. For the purpose of internetworking between the circuit-switched network and the IP network, a special device—the gateway—must be set up between the two different networks, as Figure 1 shows.

When a PC functions as a VoIP user terminal, it can directly connect to the IP network without a gateway. However, in this case, special-purpose software, such as Microsoft NetMeeting, should be installed on the PC. The IP phone is commonly used in the enterprise LAN environment, where a specific server on the LAN acts as a signaling gateway that performs signaling-conversion and -control functions. Specifically, the gateways perform the following tasks:

- establish and release the call connection;
- encode and decode the voice signal and packetize and depacketize the encoded voice data;
- implement the voice-activity-detection function at the sending end and insert the comfortable background noise signal at the receiving end;
- eliminate the effect of echo, then insert a buffer at the receiving end to reduce the effect of delay variation;
- support various voice-encoding standards,

Figure 2. H.323 protocol stack architecture applied to an IP network. The G.7XX represents a series of voice-encoding standards, H.26X represents a series of image-encoding standards, and T.12X represents a series of data-encoding standards.



including G.711, G.728, G.729, and G.723.1;

- implement the conversion between different voice-encoding schemes;
- implement signaling protocols and signaling conversion;
- differentiate between voice and fax traffic automatically;
- support T.30 and T.38 facsimile communications procedures;
- fulfill communications and internetworking tasks between the gatekeeper and other gateways;
- provide original charge information;
- communicate with and respond to the network management center;
- provide physical interfaces with communications links and perform clock synchronization; and
- perform necessary network testing functions, including QoS testing.

These functions mainly conform to the ITU-T Recommendation H.323,[1] currently the principal international standard for VoIP implementation. Rec. H.323, which is actually a protocol stack, is also suitable for other IP-based multimedia communications, including videoconferencing and distance learning. The standards most closely related to VoIP include the following:

- setup signaling, voice data packetizing, and media control—H.225.0,[2] Q.931,[3] and H.245;[4]
- voice encoding—G.711,[5] G.728,[6] G.729,[7] and G.723.1;[8] and
- real-time IP network applications—RTP and RTCP.[9]

Figure 2 shows the H.323 protocol stack's architecture applied to IP networks. G.7XX represents a series of voice-encoding standards, H.26X represents a series of image-encoding standards, and T.12X represents a series of data-encoding standards. H.225.0 consists of two parts:

- a registration, admission, and status protocol used between the H.323 terminal equipment, which can be a PC, an IP phone, a gateway, or a multipoint control unit (MCU), and the gatekeeper; and
- the signaling for call setup, based on the ITU-T Q.931 signaling protocol; H.245 is a protocol for media channel control.

Several other protocols relate to VoIP, such as the T.30[10] and T.38[11] facsimile communications protocols, the RADIUS[12] protocol for user authentication and accounting, and the SNMP[13] protocol for network management.

According to H.323, to provide VoIP services in public networks or large-scale private networks, in addition to gateways, one or more gatekeepers should be in place. Each gatekeeper should communicate with all the gateways within its controlled zone and perform the following main functions:

- handle address resolution by translating the telephone office number into the corresponding VoIP gateway's IP address;
- provide information to the authentication and accounting center and cooperate with it to perform user authentication and accounting functions;
- provide routing information to the gateways within its controlled zone;
- when used as a VoIP terminal, participate in the call setup and release process between gateways or between the gateway and PC when required;
- administer the bandwidth—an optional function; and

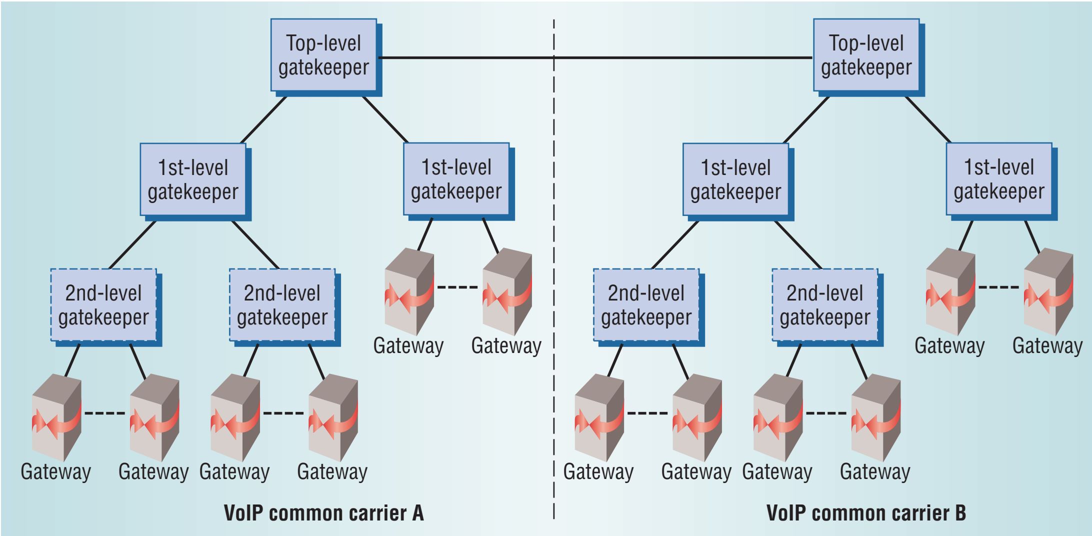


Figure 3. VoIP networks in China. These common-carrier-provided VoIP networks divide gatekeeping into a top level and a first level—a necessary strategy for managing large-scale networks.

- communicate with and respond to the network management center.

A large-scale VoIP network will be divided into several zones, and each zone needs at least one gatekeeper. When gateways within different zones communicate with one another, the gatekeepers in corresponding zones should perform address resolution. The system also must exchange address resolution information among gatekeepers.

In large-scale VoIP networks, developers usually organize the gatekeepers into a hierarchical structure. As Figure 3 shows, according to the VoIP specifications laid down by the Telecommunications Administration Department in China, any large-scale public network that a VoIP common carrier provides needs to implement at least two levels of gatekeepers: a top level and a first level.

The first-level gatekeeper performs the VoIP functions within the zone it controls, while the top-level gatekeeper manages address resolution between different first-level gatekeepers that belong to the same VoIP common carrier. It is also responsible for communications with other common carriers' VoIP networks. In addition, top-level gatekeepers handle the setup and release of all international calls. A large VoIP network requires second-level gatekeepers. In this case, the first-level gatekeeper will be responsible for address resolution among all the second-level gatekeepers below it. Similarly, in a small VoIP network, the top-level gatekeeper can be omitted, and one of the first-level gatekeepers can carry out its functions.

The public VoIP network also needs an authentication and accounting center that connects to the gatekeeper. This center accepts the original user and accounting information from the gatekeeper and uses it to perform the user authentication and accounting functions.

Figure 4 shows how to set up a VoIP call and establish voice media communication channels between two VoIP gateways within the same zone, managed by a single gatekeeper. This process, which conforms to H.323, unfolds as follows: Phone Subscriber 1 inputs the VoIP number and password to Gateway 1 by dialing through the local telephone exchange office. The authentication process between Gateway 1 and the gatekeeper then executes using the registration, admission, and status protocol. Next, the local exchange forwards the called subscriber's phone number to Gateway 1 when the authentication process completes. At this point, the address process resolves.

When Gateway 1 knows Gateway 2's IP address and TCP port number, it sends the Call Setup signaling message to Gateway 2. In response, Gateway 2 sends the Call Proceeding message back to Gateway 1. Meanwhile, Gateway 2 sends an admission request to the gatekeeper. When Gateway 2 receives the admission confirmation from the gatekeeper, it sends the Call Setup Request message to the called phone Subscriber 2. Next, when Gateway 2 receives the ringing indication signal from the local telephone exchange office that connects to Subscriber 2, it sends an alerting signal to Gateway 1. Phone Subscriber 1 hears the ringing-back signal sent from the local telephone exchange office, and phone Subscriber 2 performs off-hooking, then Gateway 2 sends a connect signal to Gateway 1 and finally to the phone Subscriber 1. In response, the signaling process for opening the logical channel between Gateway 1 and Gateway 2 occurs in conformance with H.245.

The new version of H.323 defines the *fast call* mechanism, which combines the process for establishing the logical channels with the call setup procedures, reducing connection setup time.

Unlike file or e-mail transfers, TCP cannot be used to transport encoded voice data because TCP's automatic-retransmission-request mechanism cannot satisfy VoIP's real-time requirement. VoIP uses the User Datagram Protocol, a general-purpose

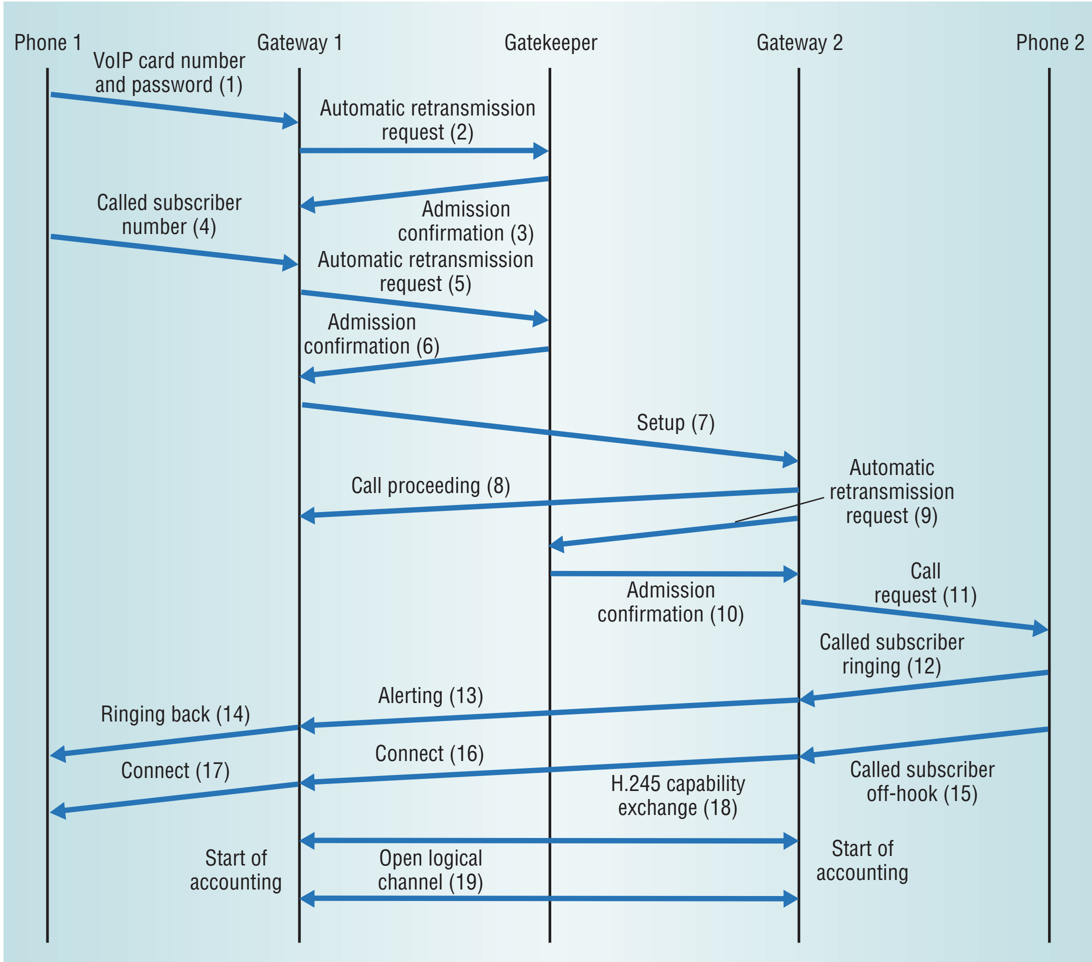


Figure 4. Establishing a VoIP call. This multiple-step process establishes voice media communication channels between two VoIP gateways within the same zone. A single gatekeeper manages the process.

transport layer protocol that also cannot fulfill VoIP's real-time service requirement.

In H.323, two other protocols—RTP and RTCP—are recommended for use with UDP. In this case, RTP mainly supports recovery of the real-time voice signal at the receiving end by adding a time stamp and serial number to each packet before sending it. RTP also provides information indicating the type of payload the packet contains and identifying the data source. All these measures are used to reassemble the received voice packets and then recover the original voice signal at the receiving end. RTP and H.225.0 also jointly define the encoded voice data's packetization method.

RTCP mainly provides feedback information concerning the quality of transmission to the sending end or, in the case of a conference, to all participants. RTCP also can provide other information, such as the IDs of a conference's participants.

RTP and RTCP package transmissions are all based on UDP, but RTCP transmits packages periodically. The transmission period should be reasonably set so that it will not cause the IP network to experience an inappropriate load.

Currently, H.323 is the most important and widely used international standard for implementing VoIP. In consideration of the country's practical situation and the prevalence of H.323, China's Telecommunications Administration Department established a series of professional standards related to the implementation and application of IP telephony, including

- general technical requirements of IP telephony and IP fax services;
- technical requirements and testing methods for VoIP gateway and gatekeeper equipment;
- technical specifications for the interoperability of VoIP gateways; and
- rules for numbering IP telephony.

These standards greatly contributed to the development of VoIP networks and services in China, especially in the area of public networks. They also fostered the development and production of household VoIP equipment.

Currently, two standards systems for IP-based multimedia communications exist: first and foremost, the serial recommendations established by ITU-T, including Rec. H.323; and, second, the multiple RFCs established by the Internet Engineering Task Force, especially the Session Initiation Protocol (SIP)[14] and its associate, the Session Description Protocol (SDP).[15]

Table 2. Comparison of H.323 and SIP standards.

Parameter	H.323 standard	SIP standard
Organization that established the protocol	ITU-T	IETF
Protocol complexity	Complex	Simple
Control method	Peer to peer	Client-server
Message presentation format	ASN.1 (binary)	Text
Media capabilities description	H.245	SDP
Core control equipment	H.323 gatekeeper	SIP server (proxy/redirect server)
Functional expandability	Poor	Good

SIP, a session layer control protocol, can be applied to VoIP for signaling use. SIP's functions resemble those of H.225.0, Q.931, and H.245 included in the Rec. H.323. SIP can be used to create, modify, and terminate a VoIP session, which may take the form of a conference and have multiple participants. SIP's designers referred to and absorbed the valuable design lessons learned from developing other Internet protocols such as HTTP, which makes SIP simple, open, compatible, and functionally expandable.

SIP uses the client-server architecture and represents messages in the text format commonly used in other Internet applications. Compared to H.323, SIP has many advantages. Table 2 shows a detailed comparison of these two protocol types. SIP does not concern itself with the voice encoding and transfer of RTP and RTCP data packages in IP networks, and thus SIP and H.323 do not differ in this respect.

Currently, in large-scale VoIP networks and especially in public VoIP networks, H.323 remains the preferred standard. Developed by ITU-T, an international standardization organization that includes each country's official telecommunications administration department and some authorized telecommunications operating companies, H.323 and similar standards make compatibility with existing telecommunications networks and related standards a priority.

H.323 specifies signaling protocols based on the existing ISDN standard Q.931 and other related signaling standards. Thus, while H.323 is currently the main standard for VoIP implementation, new standards may someday replace it. Several such standards have already been proposed.

VOIP PERFORMANCE

Since the debut of VoIP phones, both users and service providers have valued the technology's performance above all other factors. We evaluate this parameter the same way we would evaluate traditional toll telephone service: by assessing how well VoIP phones establish a call and maintain good voice quality during that call.

Establishing a call

This measure can be gauged using two measures: the time required for call establishment and the ratio of successful call connections. The end-to-end call establishment time derives from the time spent in the circuit-switched network and the time necessary for setting up the two-way media channels between the VoIP gateways.

The allowable end-to-end call establishment time should be fixed in accordance with the specific situation. Usually, the call establishment time for international calls is longer than that for domestic toll telephone calls, but both should be fixed at less than 10 seconds.

A successful call connection means that a call reaches the called subscriber and sends a ringing signal to the called subscriber, and the called subscriber performs off-hooking. Multiple factors associated with the network design affect the ratio of successful call connections, such as the trunk capacity between the circuit-switched network and VoIP gateways, the IP networks' bandwidth, and the processing capability of VoIP gateways and gatekeepers. Generally, the ratio of successful call connections should not drop below 80 percent.

Maintaining voice quality

A variety of factors influence VoIP phone voice quality, but three are especially important: voice packet transfer delay, delay variation, and packet loss ratio.

The end-to-end transfer delay of voice packets includes the time delay that circuit-switched networks, VoIP gateways, and the IP network cause. The IP network and VoIP gateways play a major role in these delays. The time delay that the IP network causes varies depending on the network's design, configuration, and operation and also on its variable load condition. Usually, the time delay from the IP network should be less than 200 to 250 ms, and the overall end-to-end transfer delay should not exceed 400 ms.

IP networks are the main source of variations in end-to-end transfer delay. This variation significantly affects the voice quality of IP telephone communications and generally should be restricted to 80 ms or less.

Loss of voice packets can occur in the transmission path in IP networks or in the VoIP gateways when traffic becomes congested. This packet loss significantly affects the quality of voice recovered at the receiving end, especially in the case of compressed voice data.

VoIP usually uses compression algorithms such as G.729 or G.723.1 to encode the digitized voice signal. The corresponding bit rates of the encoded voice data in these two standards are 8 Kbps and 6.3 to 5.3 Kbps, respectively. Generally, a higher compression ratio has a more adverse effect on voice quality when packet loss occurs. For VoIP applications, packet loss should be limited to 5 percent or less.

The voice-compression algorithm that VoIP gateways implement also plays a crucial role in assuring good voice quality. Although H.323 lists several voice-compression algorithms, including the ITU-T Rec. G.7XX series, G.729 is currently the most popular choice. This algorithm can achieve voice quality comparable to traditional toll telephones under ideal transfer conditions. Other than that, some VoIP gateway equipment characteristics, such as echo cancellation and VAD functions, also can affect the quality of the recovered voice.

Quality of service

Many other factors can affect VoIP performance, but most important is the IP network's own QoS. Because the IP network is a connectionless best-effort network, it alone cannot guarantee adequate QoS. In this case, the QoS parameters primarily include the transfer delay, the delay variation, and the packet loss ratio.

To provide real-time services such as VoIP, the IP network QoS issue becomes crucial. Along with the deployment of VoIP and other IP-based real-time multimedia services, this problem has become increasingly important with regard to further IP networking development.

Using RTP and RTCP does not solve this problem. Within the scope of IETF, several possible technical solutions have been proposed, such as the int-serv,[16] diff-serv,[17] and MPLS.[18] However, applying these potential solutions to the existing Internet will not be easy because it would require upgrading all existing routers. Using the int-serv or MPLS will make the router more complicated than the best-effort case and consume extra router processing capabilities and network resources.

In reality, most VoIP service providers prefer to build a dedicated IP network to reduce the degradation of voice quality that QoS tradeoffs cause. However, doing so risks losing the original benefits of VoIP phone deployment because it is contrary to the intention of combining data, voice, and other multimedia traffic into one network to reduce operational costs, share network resources, and promote development of new IP-based services. Therefore, at present, using a dedicated IP network to provide VoIP service is only a temporary solution.

Solving QoS issues in intranets will be relatively easy because they carry a more predictable traffic load, thus the solutions we propose are more plausible. In China, where some intranets combine data and voice traffic into one network, the diff-serv scheme has already been implemented and has resulted in significantly improved voice quality.

Although VoIP remains a work in progress, developers have identified its potential development paths. First, we can expect more value-added services associated with VoIP, such as call centers, Web services, and videophones. Second, in regard to VoIP's technical development, removing the signaling conversion and control functions from gateways and making the gateway a simple-function, purified-media gateway will be an important trend.

Because the media gateway fulfills only the tasks of media flow conversion and matching to the networks connected with it, a functionally independent media gateway controller will govern it. In addition, a separate signaling gateway will exclusively perform the signaling conversion functions. This will increase VoIP network expandability and flexibility and will accelerate the development of new services.

Based on this thinking, the ITU-T and IETF have jointly proposed the Media Gateway Control Protocol, ITU-T Rec. H.248 or IETF RFC 3015.[19] Meanwhile, researchers have developed a new technology and new device, the soft-switch,[20] which is viewed as a next-generation network core technology. Recently, the telecommunications world has begun paying more attention to the H.248 protocol and the associated soft-switch technology, with new soft-switch products just now coming to market. In China, field trials in this area are ongoing. ■

Acknowledgments

The authors thank Gary S. Robinson and Richard Mathieu for their valuable suggestions, which substantially improved this article.

References

1. ITU-T Rec. H.323, *Packet-Based Multimedia Communications Systems*, Int'l Telecommunication Union, 1998.

2. ITU-T Rec. H.225.0, *Media Stream Packetization and Synchronization for Visual Telephone Systems on Non-Guaranteed Quality of Service LANs*, Int'l Telecommunication Union, 1998.
3. ITU-T Rec. Q.931, *ISDN User-Network Interface Layer 3 Specification for Basic Call Control*, Int'l Telecommunication Union, 1998.
4. ITU-T Rec. H.245, *Control Protocol for Multimedia Communications*, Int'l Telecommunication Union, 1998.
5. ITU-T Rec. G.711, *Pulse Code Modulation (PCM) of Voice Frequencies*, Int'l Telecommunication Union, 1988.
6. ITU-T Rec. G.728, *Coding of Speech at 16 Kbit/s Using Low-Delay Code Excited Linear Prediction*, Int'l Telecommunication Union, 1992.
7. ITU-T Rec. G.729, *Coding of Speech at 8 Kbit/s using Conjugate Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP)*, Int'l Telecommunication Union, 1996.
8. ITU-T Rec. G.723.1, *Speech Coders: Dual-Rate Speech Coder for Multimedia Communications Transmitting at 5.3 and 6.3 Kbit/s*, Int'l Telecommunication Union, 1996.
9. RFC 1889, *RTP: A Transport Protocol for Real-Time Applications*, IETF, 1996.
10. ITU-T Rec. T.30, *Procedures for Document Facsimile Transmission in the General Switched Telephone Network*, Int'l Telecommunication Union, 1996.
11. ITU-T Rec. T.38, *Procedures for Real-Time Group 3 Facsimile Communication Over IP Networks*, Int'l Telecommunication Union, 1998.
12. RFC 2138, *RADIUS: Remote Access Dial-In User Service*, IETF, 1996.
13. RFCs 1441 to 1448, *SNMPv2: Simple Network Management Protocol*, v.2, IETF, 1994.
14. RFC 2543, *SIP: Session Initiation Protocol*, IETF, 1998.
15. RFC 2327, *SDP: Session Description Protocol*, IETF, 1998.
16. RFC 2210, *The Use of RSVP with IETF Integrated Services*, IETF, 1997.
17. RFC 2475, *An Architecture for Differentiated Services*, IETF, 1998.
18. B. Davie and Y. Rekhter, *MPLS—Technology and Applications*, Morgan Kaufmann, 2000.
19. ITU-T Rec. H.248, *Media Gateway Control Protocol*, Common text with IETF RFC 3015, Megaco Protocol, Int'l Telecommunication Union, 2000.
20. International Soft-Switch Consortium, *Softswitch Consortium Reference Architecture*, 1999.

***Runsheng Wang** is a professor of computer science at the Beijing University of Posts and Telecom. His research focuses on data communications and computer networking. He received an MS in telecommunication engineering from the Beijing University of Posts and Telecom. Contact him at wangrs@public.bta.net.cn.*

***Xiaorui Hu** is an assistant professor of management information systems at the John Cook School of Business, Saint Louis University. Her research focuses on trust-related issues in electronic commerce, B2B markets, the telecommunications market, and the cultural impact on international business. She received a PhD in economics from the University of Texas at Austin. Contact her at hux2@slu.edu.*