VoIP? A Question of Perspective

by Gary Audin
delphi-inc@att.net

Moving to VoIP raises fundamental questions for both data and voice sides of the ‘net. There is an inherent dilemma when it comes to building integrated voice/data networks—the problem of perspective. To the telephony world, the fact that voice is “mission critical” is so painfully obvious that it goes without mentioning.

However, in the data world—and particularly the IP world—applications are graded as “best effort,” “better-than-best effort” and “critical,” with “best effort” being the most common. Many applications have not moved to an IP network because of concerns about security, performance and reliability.

Given voice’s mission-critical status, any company planning an integrated voice/data network must consider a wide range of elements. While it has been proven that VoIP phone calls work, we must now examine all the other functions: VoIP’s features, administration, billing, performance and management issues.

Many business questions arise when confronting convergence.

- Should data or voice people manage VoIP?
- Can the voice management system modify the performance of the IP network in real time to improve voice quality?
- Will there be separate management systems for voice and IP network components?
- How is Quality of Service delivered for voice and who pays for it?
- Will the VoIP billing follow data billing concepts, for example, a flat fee?
- Is there a standard VoIP Call Detail Record (CDR) and how does it operate with existing billing and traffic analysis software?

Looking ahead, the most important task will be to make VoIP ubiquitous while at the same time making it interoperable with a variety of VoIP products, as well as with the legacy telephone environment.

VoIP Interoperability

Interoperability in VoIP transmission is a relatively easy process. Buy two IP gateways from the same vendor, hook them up and they should be interoperable. But use two different gateway vendors and who knows what will happen? Vendors do not usually carry identical H.323 software suites. Based on the work done by the *Network World* Test Alliance, only two or three vendors provide true H.323 support; most operate with modified or extended versions, which do not usually work together or interoperate at a basic function level.

Interoperability among IP phones and call servers (gatekeepers) is even more difficult to find. Vendors may talk about “open” VoIP systems, but they’re only “open” if you follow their approach. Open systems from two competing vendors cannot be expected to interoperate because an open system does not guarantee conformance with VoIP standards.

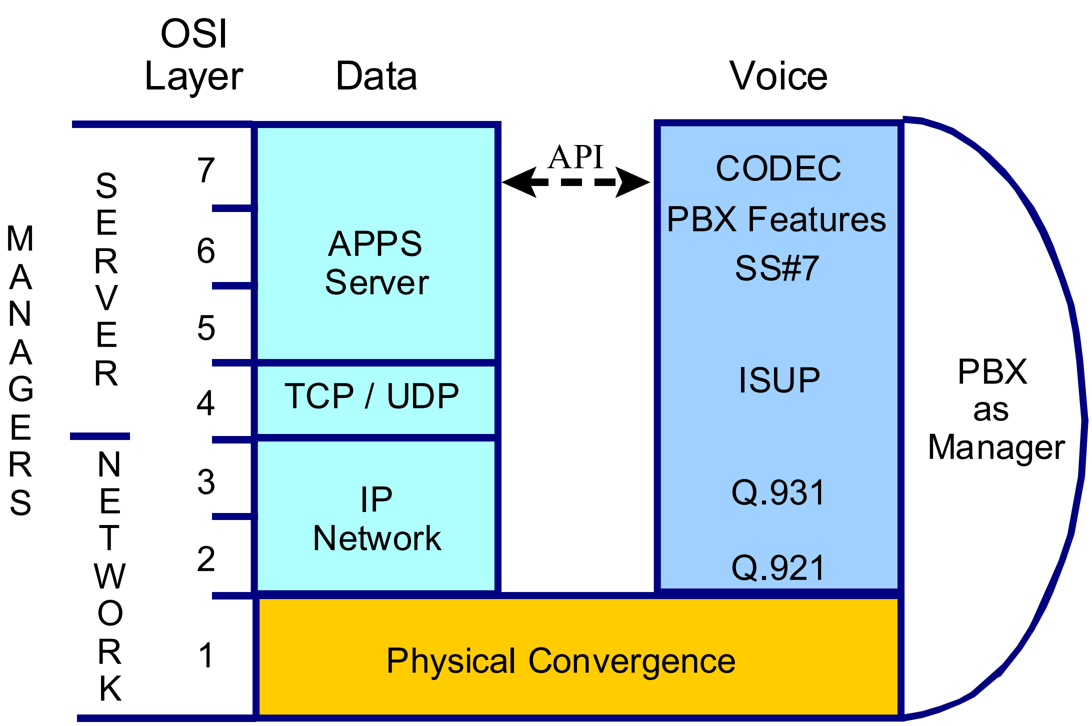
The good news is that some testing has been done. Mier Communications (www.miercom.com) and *Network World* have led the Alliance's testing on H.323 v2/v3/v4, Session Initiation Protocol (SIP), H.248/Megaco and Media Gateway Control Protocol (MGCP)-based products. This is a step in the right direction, but it also demonstrates that VoIP interoperability will emerge slowly, if at all, with some products.

VoIP's Place in the Network

So if VoIP is not yet ready to completely replace the world's legacy telephony environment, the next question is: Can VoIP work well within the enterprise? Probably, but only with sufficient planning and realistic expectations. And to gain that kind of perspective, it's first necessary to see where voice fits into existing data networks.

The place to begin is with the seven-layer OSI model (Figure 1). Today, the legacy telephony network and the IP network join only at the physical transmission Layer 1—for example T1/E1, etc. Above Layer 1, the IP and voice networks have their own worlds of protocols, applications, and network management. The management of IP applications is separate from IP network management systems, which assume no responsibility for any operation above Layer 3 except for Quality of Service (QoS) and Class of Service (CoS) functions.

Figure 1: Traditional OSI Layered Approach



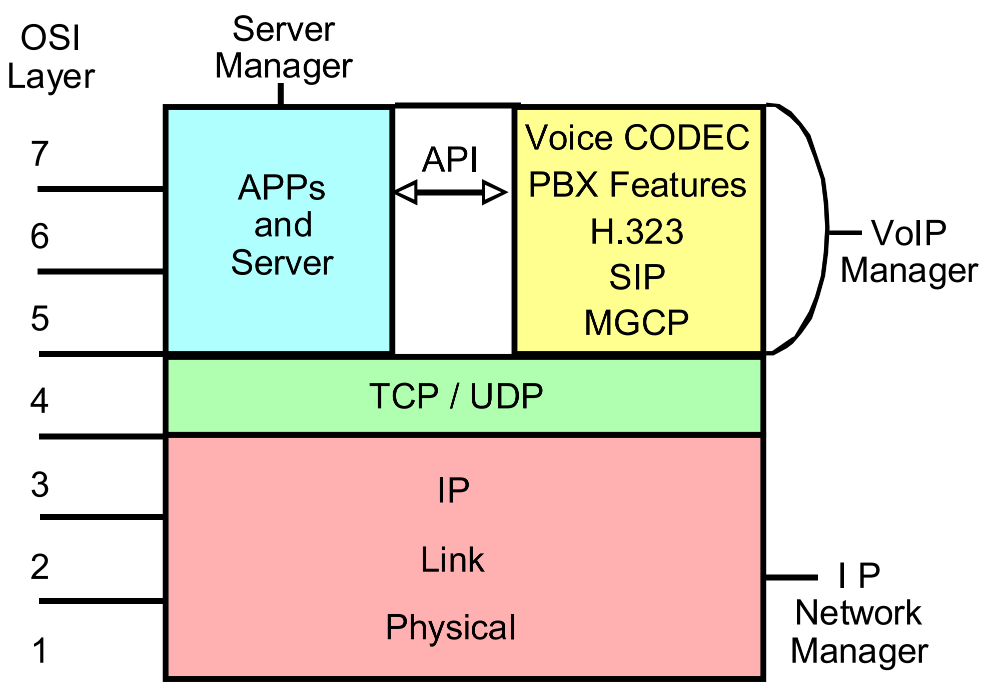
But with VoIP, voice shifts to OSI Layers 5 through 7: session, presentation and application (Figure 2). The voice applications *must* use the existing TCP/UDP/IP protocols without changing them. Therefore, all VoIP functions are really application programs; thus *voice management* becomes *application management*. The traditional voice physical transmission now belongs to IP managers, and all the voice signaling must also run as applications over TCP/UDP/IP.

Because the voice community will share the IP network along with all other data applications, job descriptions for voice managers and technicians will have to be rewritten, and server and software training will be required. The physical telephone moves, adds and changes (MACs) will shift outside the scope of the voice manager, although the logical changes will still be his or her responsibility. Voice personnel will no longer own the cable; they will not have separate wiring closets and they will not interface with the carriers supporting VoIP. In short, there will be a major shift in telephony management, knowledge and skill sets.

Everyday IP operations will also change. The current guidelines for IP network operations are not satisfactory for voice networks and must change to accommodate 99.99% network availability normally delivered by legacy voice networks. Moves, adds and changes will have to be better planned because, while IP PBXs are perceived to ease the MAC process, the methods to do so are non-standardized and must be a non-interruptible function.

Software upgrades will not be automatic simply because a vendor issues new software. While vendors often force IT managers to adopt upgrades through poor or non-existent support for older systems and equipment, the PBX world differs greatly. PBX releases last about five years, and businesses don't always adopt new upgrades just because

Figure 2 New Model from Data View
(Voice is an Application)



they're available. How VoIP equipment will fit between these two upgrade approaches will force a change in mindset for vendors, end-user businesses, or both. Similarly, the voice community expects the service level agreements (SLAs) that are delivered to meet their availability needs, so SLAs must be continuously monitored in near real time to ensure that the IP network's dynamic performance doesn't hurt call quality.

How to Think About VoIP

But understanding VoIP's place in the network is only one consideration. The ultimate concern is how this technology fits in a business. Should it offer:

- **Choice 1:** Fewer features, sometimes poorer voice quality, but reduced costs?
- **Choice 2:** A VoIP network equal to the present legacy world, but in a converged environment?
- **Choice 3:** A new network that supports legacy functions, performs better and brings new features to the desktop with little or no reduction in telephony costs?

The first choice fits the international arena, where the savings can be significant. Although international carriers are reducing their rates, they have not reached the nickel-or-less per-minute level.

However, the significant reduction in domestic rates makes Choice 1 harder to justify for a strictly national deployment. Further, any technology that *only* saves money usually has a short life span—perhaps two to three years— before it is usurped by a newer cost-cutting technology. For instance, when 100Mbps Ethernet first appeared, many businesses didn't buy it because of expense. However, as more businesses deployed the technology, it quickly became commoditized and the price differences between it and 10Mpbs weren't extreme. So reducing cost is only a short-term reason for choosing VoIP.

The second reason for data/voice integration—convergence—seems intuitively obvious to those in the IP networking community, but to voice professionals it may seem a little too familiar. The telecommunications industry has tried convergence, first with ISDN and then with ATM.

In ISDN, the voice designers took up the mantle and did it their way: voice first, data second, fitting into a 64Kbps bandwidth voice structure. ATM delivered a mixed design with greater emphasis on data. The industry , however, had to revise the use of ATM because ATM adaption layer designs did not fit the enterprise market; consider the demise of ATM's adaption layers 3 and 4 (AAL3/4) and the creation of AAL5. Now the IP data world thinks it can do a better job. Time will tell.

The third option—new features while preserving legacy features— is obviously the Holy Grail for integrated voice/data networks and should be the long-term reason for proceeding with VoIP. Adding functionality and features and possibly improving voice quality will make VoIP a long-term success. However, that success will not occur unless the data/voice network is designed from an approach that integrates legacy functionality—and this entails a host of issues.

Integration Issues: The Problems with IP Networks

Regardless of what side of the networking fence you sit on, voice/data convergence represents a *total* commitment. When certain elements are left out—such as converged transmission but separate management systems—only partial benefits can be expected. If convergence goes well beyond simply supporting phone calls on a data network then it requires much better planning and a wider vision of the problems that need to be solved.

Most of the new features will be supported by software rather than new hardware in order to more quickly and cheaply incorporate new features. Voice quality can be improved by using a better codec—such as G.722 instead of G.711—that will provide 7 KHz rather than the 3.4 KHz analog voice bandwidth used in the PSTN. New products support G.722 as an alternate codec option without requiring more digital capacity. G.711 and G.722 both operate at 64kbps.

However, VoIP must operate in the legacy telephony world. Legacy connections (T1, E1, PRI, analog) will exist for many years. This includes signaling protocols such as Q.931, ISDN user part (ISUP), and Signaling System 7 (SS7). Legacy interfaces will be necessary:

- When the VoIP WAN fails
- For overflow traffic from VoIP
- To handle local calls
- To provide 911 services
- To support non-VoIP off-net calls

If VoIP and the legacy telephony world are to co-exist, where does a designer begin? There are a number of elements, such as VoIP call servers, gateways, IP/Ethernet phones, etc., but the fundamental concern is the actual transmission of voice. If the IP network is to replace the current voice network, it must be ready to support both new and legacy capabilities, which is a difficult proposition given IP's bursty nature and voice's need for a constant bit rate.

Our focus here is in dealing with IP network design and operation for VoIP. Latency through the IP network is difficult to limit. A phone-to-phone latency of 100 ms would be ideal. A lower latency (less than 100 ms) is not worth pursuing because the average caller will not notice any difference. The individual delays that contribute to latency are caused by router processing and propagation times, but mostly by congestion buildup in the routers.

In contrast, the legacy telephone network is "gold plated"—it is better than it needs to be for voice. The quality variation (tolerance) a caller can sustain offers a chance for IP networks to provide satisfactory phone calls over an IP network (Table 1). When phone-to-phone latency reaches 150 ms, callers notice the decrease in quality, but usually find the call acceptable. Beyond 250 ms, callers must change the pace of their speech as they would on a satellite call.

TABLE 1 Voice vs. Data Networks				
Factor	**Legacy PSTN Delivery ***	**VoIP Tolerance**	**MIP or Intranet Delivery +**	**Internet Delivery +**
One-Way Latency	1 - 30ms	50 - 150ms	20 - 200ms	50 - 2000ms
Delay Variance (Jitter)	0 - 5ms	10 - 20ms	10 - 100ms	10 - 300ms
Loss	0%	1 - 2%	1 - 5%	1 - 30%
Out-of-Sequence Packets	Does not occur	Correction required but adds jitter to delay	Corrected	Corrected
Errors	Very low and ignored	Ignored; No retransmission	Low; Corrected by retransmission	Low; Corrected by retransmission

* Includes phones and gateways
+ Does not include phones or gateways
MIP = Managed IP Service

The next problem is jitter, or rapidly varying delay. The primary culprit here is varying network congestion. Processing delays and propagation times are nearly constant. Jitter introduces a further delay at the receiving end (gateway or VoIP phone). The early arriving voice packets must be held until the slower (delayed) packets arrive, so the words can be properly reconstructed. Jitter also complicates echo cancellation, which is corrected by the VoIP gateway or IP phone, not by the network.

Packet loss can be a major problem when it exceeds 2%, although that 2% figure is soft; some products can sustain losses of up to 5%, while others are very sensitive to loss. Packet loss is more tolerable with small voice packets, for example, 20 bytes of voice per packet. Loss is better tolerated when there is less voice compression, such as G.711 rather than G.729 or G.723.1.

The network can compensate for loss through techniques that substitute predicted voice bytes for the missing packet. The IP phone or gateway fills in the missing voice element(s) with an estimate of what it might have contained based on the contents of previous packets. However, when the loss increases significantly, or the loss is bursty rather than random, then speech will be distorted. All compensation techniques break down when packet loss is excessive—greater than 10%. Loss, like the other impairments, is due to congestion.

Out-of-sequence packets are caused by repeated changes in network routing tables. If the IP network changes the path every 30 seconds, as it can during Open Shortest Path First (OSPF) routing, then each change can create out-of-sequence problems every 30 seconds. This condition will only occur when there is heavy congestion on the network. Correcting the sequence does not affect voice clarity, but does introduce extra delay at the receiving end. Out-of-sequence problems are also related to congestion. The greater the congestion, the more likely the IP network will change the routing tables.

Another issue is that errors in voice transmission can be ignored. IP networks generally provide error rates well below those required for voice transmission. The signaling transmission must be error free. Retransmitting using TCP or error-correcting signaling protocols can correct these errors. Error correction for voice is unnecessary. However, error correction for signaling can increase the execution time for a feature, such as call setup.

Before voice can travel across IP, those issues must be addressed, and that is accomplished through properly preparing the network.

Readying IP Networks for VoIP

There are two primary solutions to improving IP network performance for voice: Allocate more bandwidth (reduce utilization) and QoS. How much bandwidth to allocate depends on:

- Packet size for voice (20 to 240 bytes)
- Compression technique (G.711, G.729, G.723.1, G.722, proprietary)
- Header compression (RTP + UDP + IP)
- Layer 2 protocols, such as point-to-point protocol (PPP), frame relay and Ethernet
- Silence suppression/voice activity detection.

The end result may range from 12Kbps to 26Kbps total bandwidth per voice call using 8Kbps G.729 compressed voice (see ”Budgeting One-Way Bandwidth,” Table 2 belowe). Tolly Research performed VoIP tests with G.711 and G.729 and found that G.711’s 64Kbps expanded to 110Kbps, and G.728’s 16Kbps expanded to 62Kbps with VoIP overhead. Tolly concluded that up to 80% of the needed bandwidth could be attributed to IP overhead with the rest of the bandwidth allocated to actual voice. (See *Network World*, “VoIP to Carry Bandwidth Tax” at http://nwfusion.com/news/2000/1030voip.html for more information on Tolly’s findings.) Most products are now designed to carry G.711 64Kbps voice within a total bandwidth of 80Kbps. Table 3 (Bandwidth Calculation Results) demonstrates the bandwidth requirements for different networks and codec compression types.

TABLE 2 Budgeting One-Way Bandwidth

Voice Digitization/Compression

G.711 = 64,000 bps or 8000 bytes/second
G.729 = 8000 bps or 1000 bytes/second

At 20 bytes/packet = 50 packets/second
At 40 bytes/packet = 25 packets/second when using G.729 compression

Protocol Overhead Using Frame Relay

	Protocol	Overhead in Bytes
	Frame Relay	6 (Header + Trailer)
Can be compressed 4 to 6 bytes	IP	20 (Header)
	UDP	8 (Header)
	RTP	12 (Header)
	Total	**46 bytes of overhead per packet**

For 20-byte packets for G.729
20 + 46 bytes/packet x 50 packets/sec = 3300 bytes/sec
x 8 bits/byte = 26.4Kbps voice overhead

For 40-byte packets for G.729
20 + 46 bytes/packet x 25 packets/sec = 2150 bytes/sec
x 8 bits/byte = 17.2Kbps voice overhead

The varying designs of packet size, voice compression choice and header compression make it difficult to determine the bandwidth to budget for a continuous speech voice call. A good rule of thumb is to reserve 18Kbps of IP network bandwidth per call for 8Kpbs (G.729-like) compressed voice.

If silence suppression/voice activity detection is used, the bandwidth consumption may drop 50%—to 8Kbps total per VoIP call. However, the assumption that everyone will alternate between voice and silence without conflicting with each other is not realistic. For instance, if four people are speaking one-quarter of the time during a four-way voice call, chances are that they will not take turns. Simultaneous bandwidth consumption is much more likely, so do not assume that there will be significant bandwidth savings for silence suppression. The value of silence suppression increases for trunking as the number of simultaneous calls on the trunk increases. Do not expect any bandwidth savings when there are 2 to 10 calls on an IP trunk.

Another aspect of this problem is determining how silence suppression would be useful when transmitting music on “hold” or for interactive voice response (IVR) announcements.

TABLE 3 Bandwidth Calculation Results*					
CODEC	Bandwidth	Packet Size (bytes)	Header Compression	Network	Approximate TOTAL Bandwidth
G.711	64Kpbs	160	No	Ethernet	85Kpbs
G.711	64Kpbs	160	Yes	Ethernet	70Kpbs
G.729	8Kpbs	10	No	Ethernet	30Kpbs
G.729	8Kpbs	10	Yes	Ethernet	15Kpbs
G.711	64Kpbs	160	No	Frame Relay	81Kpbs
G.711	64Kpbs	160	Yes	Frame Relay	67Kpbs
G.729	8Kpbs	10	No	Frame Relay	20Kpbs
G.729	8Kpbs	10	Yes	Frame Relay	10Kpbs
G.723	6.3Kpbs	30	No	Frame Relay	16Kpbs
G.723	6.3Kpbs	30	Yes	Frame Relay	8Kpbs

Reference: "IP Telephony/VoIP Design Guide"
April 1, 2003, Alcatel White Paper

Here is where QoS and CoS come into play. QoS allocates network capacity (for example bandwidth) by type of traffic, while CoS provides a preferred allocation of the network (such as less delay, less jitter, less loss, etc.), but does not guarantee any measurable amount.

QoS delivers guaranteed measurable, not-to-exceed numbers. If the CoS supporting voice is good enough and stable enough, no guarantees are required. Note, however, that with CoS and no bandwidth increases, the data traffic will experience poorer service because voice will be given preference and therefore will get bandwidth first. QoS requires very good network design, knowledge of all traffic, and performance tradeoffs.

A Question of Perspective

Converging networks will require converging mindsets between data and voice professionals. This means bridging a very wide chasm. To illustrate, BCR provides two seminars: "IP Telephony" and "IP Convergent Networks." The attendees' attitudes explain a lot about why moving to Choice 3 will be tough going.

The IP professionals that attend typically come from the mindset that voice is "just another application," and make comments such as:

- "I did not know you did that in a PBX/ACD/IVR/call center, etc."
- "Why do you want to do that?"
- "I can figure out how to do that in a few months if I have to."
- "That's going to change the way I run the IP network—and I finally have it running well now."
- "I have 10 features on my phone. My peers and I don't need any more."

Telephony and voice professionals that attend take for granted much of what has been incorporated in voice networks over the last 50 years. They do not expect to ask VoIP-related questions like:

- "How long do I wait to get a dial tone and where does it come from?"
- "What are the call setup and disconnect times?"
- "What happens when the telephone number directory doesn't have the IP address?"
- "Will I have trouble with viruses, worms and other security issues?"

The IP world is optimistic about VoIP. The legacy telephone community is skeptical. They are both correct. Each needs to learn more about the other's environment. Forethought, good planning and realistic expectations will make VoIP operate successfully.

Gary Audin is president of the network-consulting firm Delphi, Inc. and can be reached at delphi-inc@att.net. Audin is also the instructor for BCR's "IP Telephony" and "IP Convergent Networks" seminars.

This article was originally published in the April 2001 issue of Business Communications Review (BCR) and was revised in February 2004.