An innovative look at Microsoft's vision of voice communication technologies.

Microsoft® Voice and Unified Communications

Joe Schurman

"*Microsoft Voice and Unified Communications* is essential reading for anyone using—or considering—Microsoft's range of VoIP options, from consumers to small business owners to enterprise customers."

Xuedong Huang
General Manager, Microsoft Research Communications Innovation Center

"Joe Schurman has captured the essence of Microsoft's vision and implementation in the areas of Voice and Unified Communications. This is an important book for those interested in connecting the dots between the present and the future in human communications and understanding why things are evolving in that way."

Gurdeep Singh Pall
Microsoft Corporate Vice President, Unified Communications Group

Library of Congress Cataloging-in-Publication Data:

Schurman, Joe, 1977-
Microsoft voice and unified communications / Joe Schurman.
p. cm.
ISBN 978-0-321-57995-9
1. Internet telephony—Computer programs. 2. Multimedia communications—Computer programs. 3. Microsoft software. I. Title.

TK5105.8865.S368 2009
005.3—dc22

2008052525

ISBN-13: 978-0-321-57995-9
ISBN-10: 0-321-57995-X

Text printed in the United States on recycled paper at
R.R. Donnelley in Crawfordsville, Indiana.

First printing February 2009

Associate Publisher
Greg Wiegand

Acquisitions Editor
Loretta Yates

Development Editor
Kevin Howard

Managing Editor
Kristy Hart

Project Editor
Anne Goebel

Copy Editor
Geneil Breeze

Indexer
Erika Millen

Proofreader
Kathy Ruiz

Technical Reviewer
Tony Bradley

Publishing Coordinator
Cindy Teeters

Cover Designer
Gary Adair

Compositor
Gloria Schurick

Foreword

The evolution of communications is a fascinating subject. From the primordial event where two people at the same place exchanged a single piece of semantically significant information to Gene Roddenberry's resplendent "Transporter" vision you have to wonder what were the incremental and disruptive steps—what endured and what did not. Like historians then, we use that thesis to predict what will work in the journey into the future of communications, and what will not.

In our lifetime the world has evolved from analog phone calls and telegraphs to PC-based communications. We have watched the PC and the Internet grow to reach more than a billion people in a span of 15 years or so. We have witnessed how the union of software and the Internet has become the jet fuel for the evolution of communications—creating an amalgamated medium for innovation and adoption unprecedented in the history of communications.

What we call **Unified Communications (UC)** is really the Renaissance of communications, in the same way the Renaissance period was for culture from the fourteenth century to the seventeenth century. UC is transforming the fundamentals—user experience, programmability, accessibility, reuse of components, and infrastructure. Everything is being and will continue to be challenged. The only thing sacred is the principle of empowering the end user with new capabilities offered within a cognitive model that is intuitive and practical.

Joe Schurman captures the essence of Microsoft's vision and implementation in the areas of voice and unified communications. This is an important book for those interested in connecting the dots between the present and the future in human communications and understanding why things are evolving in that way. Above all, this is a thorough and practical book, useful for those evaluating and planning the next step with Microsoft's Unified Communications offerings.

—Gurdeep Singh Pall,
Vice President, Microsoft Unified Communications Group

Preface

The telecommunications industry is changing in ways not thought possible. We, as a human race, are seeing a complete transformation in how human communication can be used throughout applications and devices, truly connecting people and processes, regardless of geography, at a speed of innovation that is incomprehensible. The future of these successful innovations will lie within how human and machine-based communication will interact with tools that we use on a daily basis, liberating us from specific hardware and devices through the power of software. The purpose of this book is to introduce you to a company who is truly a leader in this strategy and vision and who is equipping the creative developers of tomorrow with the tools necessary to transform the world of voice and unified communications forever. This company is Microsoft Corporation, and the content within this book will help you understand what Microsoft is providing today and the vision for tomorrow throughout consumer, small business, and enterprise organizations on a global scale through the power of software.

CHAPTER 1

THE COMMUNICATIONS RENAISSANCE

Telephony Revolution

Fast forward to the near future—to the year 2010. I just finished yet another training seminar. Sweating a bit from all the excitement and talking, tired from standing all day, and a bit hoarse, I walk back to my hotel and at the same time, join a late afternoon conference call with one of my colleagues, Seth. Instead of just holding the phone up to my face generating even more heat, I am having a live video and voice conversation with him on my new mobile phone (see Figure 1.1).

FIGURE 1.1 Mobile video call

We wrap up the conference call and then I head up to the hotel to finish out the day with a cold one overlooking the European-esque architecture of beautiful Buenos Aires, Argentina, where I had been before.

Now, rewind to the last time I gave a speech in Argentina in the days before VoIP became prominently available. That is a much different story. After my speech, I wanted to call my wife, but I realized I did not have a phone connection—not even a roaming connection! There I was on my US, unstandardized CDMA phone in a GSM-supported location of Argentina with no phone access—none! I ended up walking through the pouring rain to the nearest mobile phone store to purchase a pay-as-you-go phone with a new SIM card running on the Personal network there in Buenos Aires. After slogging through water up to my ankles back to my hotel, I tried to make the call. After several failed dialing attempts and

several calls to a Spanish-speaking only network service line, I learned that that the pay-as-you-go phone did not have the ability to make international calls, so the entire adventure was all for naught. In the end, I used the hotel phone to make the hundred-dollar call to my wife (see Figure 1.2).

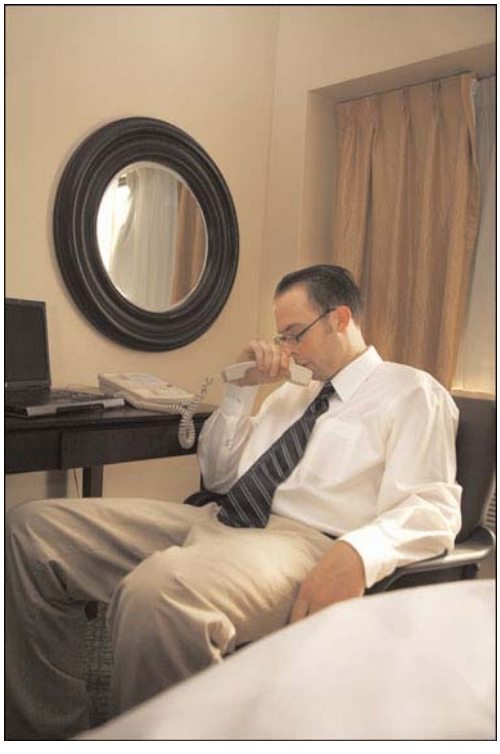


FIGURE 1.2 Perfect depiction of how I felt

Now, back to the near future in my mobile video conference with Seth! I'm not sweating at all thinking about the past situation because now, the video and voice call I just made, unlike before, was covered in my $29 USD per month plan. No wireless PSTN surcharge, no $100 USD per month wireless plan, no international roaming, no GSM, CDMA, 3G, and so on. My new wireless broadband phone that has HD-video and crystal clear HD-audio with no interruption is leveraging a technology made available for consumer use in the 1990s called **Voice over Internet Protocol (VoIP)**.

That's right! VoIP has been around for more than a decade! What's most important to note is that somewhere, some business executive who works for a wireless provider or telecommunications (Telco) provider is starting to sweat. Why? Because this executive is beginning to realize what is happening—the end of wireless and wireline communications services as they exist today. No more SIM cards, no more international roaming, no more dictatorship! For more than three decades telecommunications providers have, similarly to how energy companies have, controlled the cost and service to their consumer and commercial customers by charging whatever rates and fees they want at will.

No more!

This time marks the beginning of the end of traditional telecommunication services and the provider's ability to enforce unjust fees and limitations to its customers. This is the birth of a new world of telecommunications leveraging the power of voice and unified communications software! Your charter is to understand what this change means and prepare for what is to come.

For the past two decades many organizations have designed and developed technologies that leverage voice and unified communications, but have failed miserably, or they have introduced these services before thoroughly testing them for mass use. This has in some cases given VoIP a bad name, which is why my book is not entitled, Microsoft VoIP, but Microsoft Voice and Unified Communications. The purpose of this book is to identify the most innovative voice and unified communications provider on a global basis and its products and services for consumers, small businesses, and enterprise organizations.

You may know already that providers are developing new voice and unified communications technology to change the way we communicate forever, and for the better. Some typical names are Nortel, Cisco, Avaya, and other leading providers in the telecommunications industry. What you may not realize is that a company based out of Redmond, Washington that is popular in the area of software development and manufacturing will forever shape the future of voice-based communication and collaboration technology. This company's name is Microsoft, and the purpose of this book is to provide insight into what these technologies are and to prepare you for the next wave of communications innovation.

Telephony Evolution

To understand how Microsoft's Voice and Unified Communications vision will change the telecommunications industry, it's first important to understand how the telecommunications industry has evolved. I love how some movies or books start with a "In the beginning…." In keeping with tradition, I will also say, "In the beginning, there was **Plain Old Telephone Service (POTS)**." POTS was the first communications layer to enable one person to talk to another without having to ride a horse, fly a plane, or drive a car to see someone in person. Some say that this was the beginning of the end in human communication in that we are now seeing more and more individuals hiding behind the phone and spending less time in person, which I completely agree with, but hope to live in a time where we can evolve communications to enforce visual presence. However, POTS enabled the first wave of communications. End users of a POTS line would use a first edition phone device, designed and manufactured by Alexander Graham Bell himself, to communicate to the same type of device held by another end user with a POTS phone (see Figure 1.3).

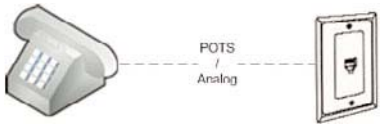


FIGURE 1.3 POTS (analog) line

Lines were terminated by switches, which soon led to a release of a new technology launched in the vibrant 1970s, the **Public Switch Telephone Network (PSTN)**, which enabled a company, remember Ma Bell, to terminate calls to enable long-distance calling internationally. Back then, we all paid, well, you paid, because I was too young to have a phone in my own name, exorbitant fees at each termination resulting in an expensive phone bill if long distance was used. Basically the PSTN connected POTS phones across cities, states, countries, and ultimately oceans to enable voice packets to be sent and received across PSTN networks in each region of the world (see Figure 1.4).

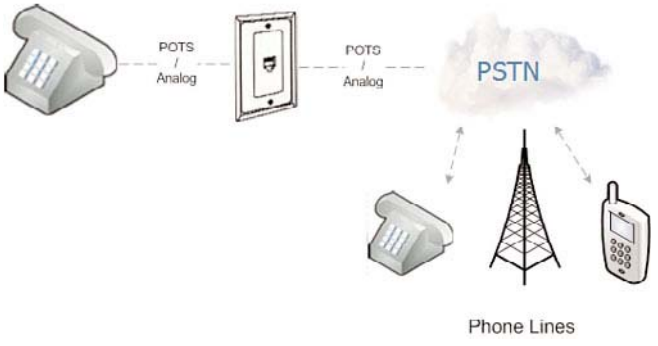


FIGURE 1.4 POTS/PSTN integration

Within the same decade, **Private Branch eXchange (PBX)** systems, illustrated in Figure 1.5, were released to enable corporations to host their own telephone network without having to pay for individual POTS lines for each office worker. PBX telephone users would call each other using a four- or seven- digit extension, and if they had to dial one of the enabled outside POTS lines, they would normally dial a 9 and then the number. Remember dialing 9 at school to call home? Yep, that's where this comes from. Anyway, POTS lines are shared within the PBX network so that 16 to 20 employees may use eight POTS lines in a given company, and that's how most PBX systems are still sold today in an 8x16 model. PBX systems connect to PSTN systems as well as connect PBX phones to long-distance callers.

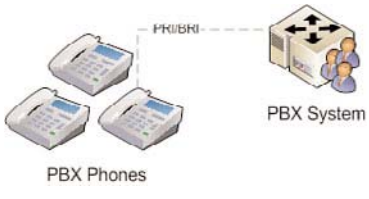


FIGURE 1.5 Private Branch eXchange

Around the 1990s, PBX and PSTN systems and networks started to accept digital voice packets. By accepting digital packets, PBX systems were able to advance using **Internet Protocol (IP)** communications connecting voice over the Internet and creating a new way to communicate using what we now know as VoIP. VoIP is a protocol or a vessel by which communications including voice, video, and data pass over an IP line. VoIP uses another protocol called **Session Initiation Protocol (SIP**; see Figure 1.6) to pass this data. (This will be important for you to remember later.)

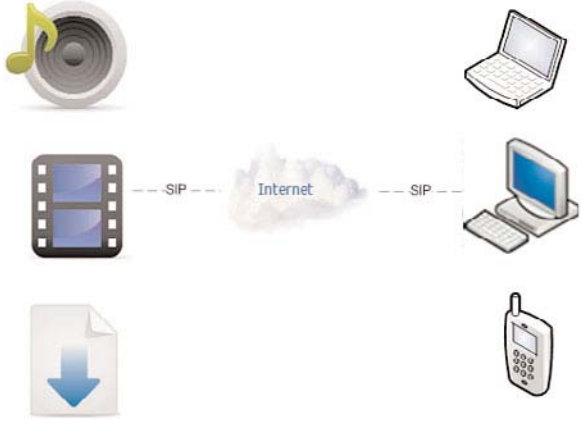


FIGURE 1.6 Session Initiation Protocol

What is needed for this to work? Simple, an Internet connection. With new Wi-Fi and WiMax technology, obtaining an Internet connection is easier and more broad reaching than ever before, and in the future, Internet connections will blanket the earth. PSTN networks are also crucial in this development as voice calls are carried from a VoIP service provider, known as an **Internet Telephony Service Provider (ITSP)** to the PSTN using digital packets to terminate a connection to a POTS line as well as cellular/mobile lines.

As depicted in Figure 1.7, PBX systems can take advantage of this because they can now utilize SIP as the primary communications layer for external communication instead of having to rely on POTS or analog lines. Internally, PBX systems also upgraded their service protocols by moving

away from **Primary Rate Interface (PRI)** and **Binary Rate Interface (BRI)** lines to IP-based Ethernet-enabled cabling, creating the IP-PBX as we know today.

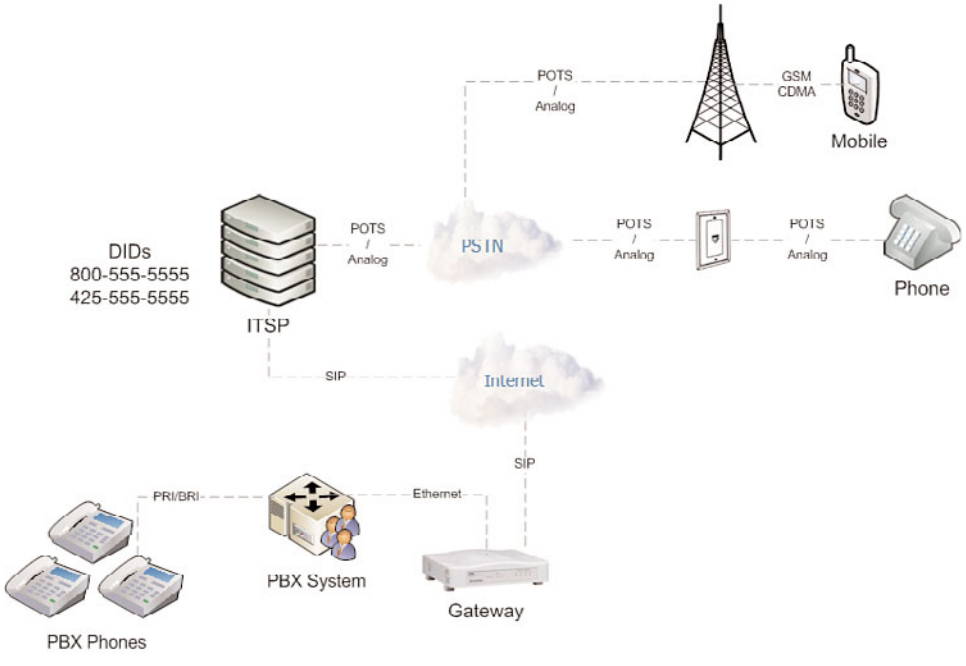


FIGURE 1.7 PBX/SIP integrated architecture

SIP'ing VoIP

VoIP services and IP-based communications services, such as the IP-PBX, leverage protocols to carry audio, video, and data. The two most popular VoIP protocols are H.323 and SIP. For detail on these protocols, visit the **Internet Engineering Task Force (IETF)** via its Web site at http://www.ietf.org searching for RFC 2543 for SIP, and the **International Telecommunications Union (ITU-T)** Web site via http://www.itu.int for details on H.323. As depicted in the figures earlier in this chapter, SIP is the standard VoIP protocol used to carry audio, video, and data communications and is the protocol that Microsoft chose to build its communication

platform of products on due to the ability to provide more flexibility and customization in how audio, video, and data are handled between applications (see Figure 1.8). H.323 was the previous industry standard, but based on its rigidness and lack of flexibility, was not the protocol of choice. Another major reason why H.323 was not chosen has to do with bandwidth. The goal of VoIP is to work within ubiquitous networks and lower the overall threshold of the communication pipe to lower bandwidth requirements enabling lightweight applications and devices that leverage VoIP.

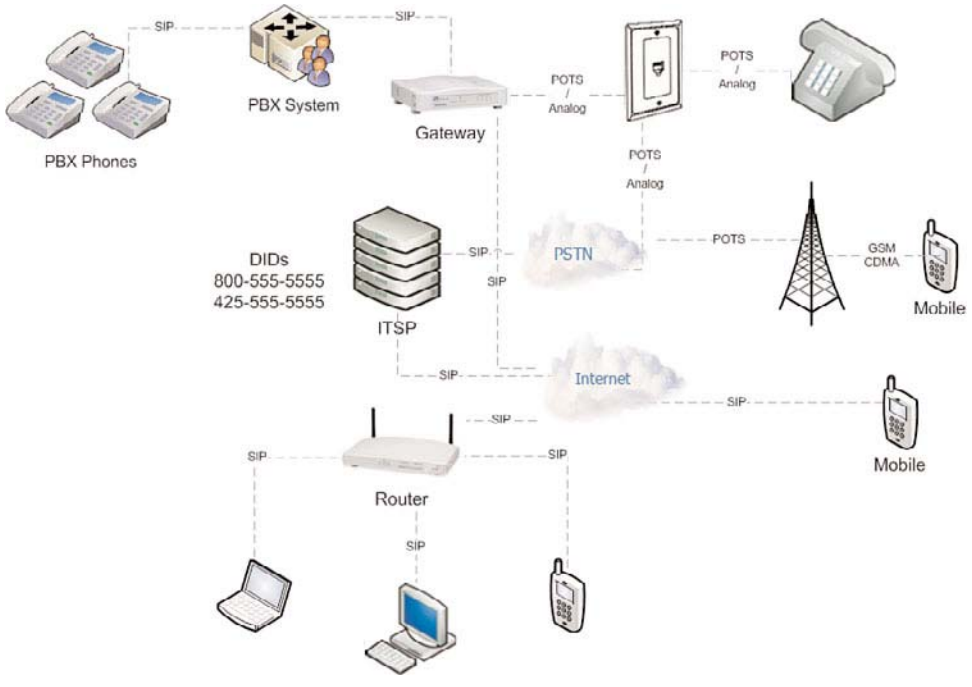


FIGURE 1.8 SIP architecture

Leveraging SIP as the foundation protocol for VoIP enables Microsoft to build communications products that carry the richness of audio, video, and data as well as provide the capability of custom application development to enhance these SIP-based products using third-party integrated solutions from Microsoft's worldwide partner community. In the future, as expressed in my hypothetical 2010 scenario described earlier in this chapter, these flexible and feature-rich solutions will be extended to mobile

users leveraging SIP over wireless broadband networks such as WiMax towers that spread a broadband Internet connection across entire regions. Like you see in Japan and in spy movies, you will be able to communicate, via mobile, audio, video, and data without using your mobile wireless carrier network, saving you time and at a fraction of the cost.

The Presence of SIP

Leveraging SIP further, Microsoft, unlike its competitors including Cisco and IBM, has taken a different approach in respect to providing voice communications services to its customers. Microsoft has built its services around the core human element, presence (see Figure 1.9). Understanding Microsoft's focus on presence, its vision of human-based presence services, will unlock the understanding of Microsoft's vision of future communications and collaboration services. Think about this for a minute. Telco and networking providers have built communications technologies based on the technology itself, not considering the human element. They have confined you as an individual and companies to the actual service itself and then have built enhancements on top of this service. Microsoft took another approach. Microsoft thought about you, the individual. What good is the technology if you are not there to use it? What good is the technology if it cannot adapt to you as a person? The Discovery channel has a series called *How It's Made*, which focuses on the manufacturing of consumer goods from crayons to baseball bats to even popsicles. These machines have advanced cameras that can sort items based on color and size. These cameras identify motion and color, and are sensitive to the touch of the item as well. Why can't technology like this be applied to the way we communicate? When I wake up in the morning and rise out of bed, why can't a camera or sensor be educated/programmed enough to sense my presence, my schedule, or even my health and determine what services I need in respect to contacts I can communicate with or meet with, or even to see whether I'm feeling blue? No other manufacturer of communications-based technologies is focusing on the human element. One company in particular is advertising this focus, but it is simply advertising. Microsoft is actually walking the walk without the hype and advertisement.

FIGURE 1.9 SIP/Presence

They have focused their communications and collaboration products based on the human element, based on human presence; and out of this foundation, they have built additional services on top. The focus on this core, critical human element, will far surpass any of its competitors as technology changes in that the core will remain the same, providing the ability to add better performing and functioning applications and hardware surrounding presence. By focusing the development of these products on the foundation of presence-based communications, Microsoft will outwit its competitors "hand over fist."

Leveraging this presence-based model as identified in the RFC 3856 documentation, which can be found via http://www.ietf.org/, enables Microsoft applications to integrate with telephony equipment such as phone devices and PBX systems. The presence-based model also works with applications to provide better insight to users on a contacts list to identify each other's true availability as mentioned earlier in the form of

presence status indicators such as Away, Busy, Online, or Offline, which are the standard, but also extend to On the Phone, In a Conference, and In a Meeting as well as customized presence functionality so you can configure your own, such as "Gone Fishing." Regardless, this focus on presence is the central core of Microsoft's portfolio of voice and unified communications products and services for the future, benefiting the actual user of the technology to increase productivity and enhance collaboration capabilities.

The Battle for Voice Quality

Even though IP-based communications have matured since the 1990s, many organizations and individuals are still skeptical in relying on VoIP for reasons such as dependence on power and the quality of the service compared to traditional telephony services, put simply, the analog dial tone. Since Microsoft's entry into the Unified Communications marketplace, the question of voice quality was the primary target of Microsoft's competitors, specifically Cisco. Telecommunication providers such as Cisco, Avaya, Nortel, and others sell **Quality of Service (QoS)** along with their devices and application platforms. QoS networking routers and switches provide compression on the codecs that are used to transmit the audio that initiates from one phone to the other or one PBX to the other. These existing telephony players use legacy codecs as well, which force them to use QoS devices and services to ensure that the quality of the call is clear. For example, the audio codec used for SIP calling is G.711. For more detailed information about this codec, visit the ITU Web site via http://itu.int. The PBX uses this codec at a rate of 64KB per second to transmit the audio. The telephony provider will usually provide a QoS router that keeps this compression low to make sure that disturbances such as jitter, lossless audio, and echoing do not occur during the transmission. Bottom line, telephony providers are obsessed with this feature, and for good reason, but it seems to be the only thing they are obsessed with outside of raping the customer with overpriced, unnecessary hardware.

Microsoft took a different approach to QoS by adding **Quality of Experience (QoE)** to the equation. To explain QoE, you need to know that all Unified Communications providers, including Microsoft, use **Mean of Opinion (MOS)** scores to determine the measurement of quality of the communications infrastructure. MOS can be used for video and voice as well. Table 1.1 outlines an example of a MOS score report covering many different codecs within a company's telephony environment.

Table 1.1 MOS Scores

Codec	Kilobytes/Second	MOS Score
G.711 (ISDN)	64	4.3
iLBC	15.2	4.14
AMR	12.2	4.14
G.729	8	3.92
G.723.1r63	6.3	3.9
GSM EFR	12.2	3.8
G.726 ADPCM	32	3.8
G.729a	8	3.7
G.723.1r53	5.3	3.65
GSM FR	12.2	3.5

Source: http://en.wikipedia.org/wiki/Mean_opinion_score

MOS scores are rated between a range of 1 and 5 based on categories of rating in Table 1.2.

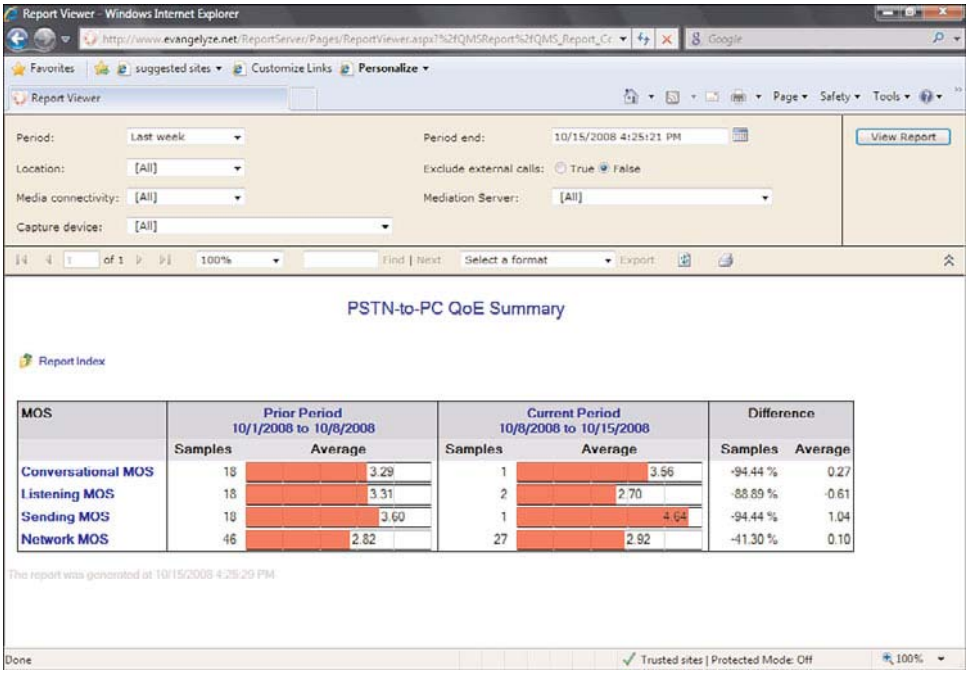
Table 1.2 MOS Score Ratings

MOS	Quality	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

Source: http://en.wikipedia.org/wiki/Mean_opinion_score

Microsoft took this a step further by adding QoE to not only provide quality of the communications service from a networking perspective, but also from a user experience perspective. While QoS looks only at the hard evidence of the system, QoE factors in the actual user experience. This means that even though the QoS reports may be perfect, some of the users within the same reported environment may still experience static or some other kind of line trouble, so the QoS may result overall in a good score, but bottom line the service is still not operational, especially if the CEO is

the user experiencing the line interference. Microsoft created its own QoE Monitoring Server that diagnoses and collects reports on the experience of each communication endpoint (see Figure 1.10). As soon as a call or session is completed, reports and statistics are sent to the server and are available for review. Metrics are taken in real time during each user session so that the true user experience is captured to ensure the quality of the overall service.



MOS	Prior Period 10/1/2008 to 10/8/2008		Current Period 10/8/2008 to 10/15/2008		Difference	
	Samples	Average	Samples	Average	Samples	Average
Conversational MOS	18	3.29	1	3.56	-94.44 %	0.27
Listening MOS	18	3.31	2	2.70	-88.89 %	-0.61
Sending MOS	18	3.60	1	4.64	-94.44 %	1.04
Network MOS	46	2.82	27	2.92	-41.30 %	0.10

FIGURE 1.10 An example of a QoE report generated from the Microsoft QoE Monitoring Server

Voice Protocols and Codecs

Rounding out this introduction of voice and unified communications concepts, we now take a look at which protocols are used that provide the audio, video, and data experience. So far, I have mentioned SIP, which is used as the protocol for audio, video, and data. Another key protocol used by VoIP that is important to understand is **Real-time Transport Protocol (RTP)**. RTP works with both H.323 and SIP, explained earlier in this chapter, and was originally created as the standard protocol used for VoIP. RTP is the delivery mechanism that carries audio and video over an IP network as well as the Internet.

In addition to VoIP protocols, as mentioned earlier in the QoS/QoE discussion, codecs are used to provide the actual band of voice communication. The most popular of which, and Microsoft designed, **Real-time Audio (RTAudio)** is a wide-band speech codec used by Microsoft's voice communications products to compress the speech/audio used in a multiperson or two-way conversation (see Figure 1.11).

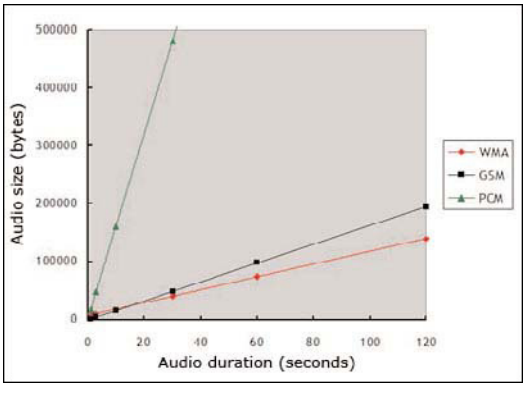


FIGURE 1.11 An example of RTAudio sample rates

This level of audio compression gives Microsoft the competitive advantage over traditional Telco providers in that the audio compression provided by the RTAudio codec ensures nonpacket loss, which ultimately means less or no jitter on the line. RTAudio is susceptible to delays, though, which is why QoS and QoE are used to overcompensate for any flaws that may be present within this voice communication infrastructure.

Securing Voice Communications

There are many areas where voice and unified communications can be secured, depending on the level of security needed by specific organizations. Chapter 8, "Securing Voice," explains this security information in more detail. There is the filtering of bad traffic in the form of viruses and buffer overflows as well as the identification and encryption of SIP communication between servers and clients. Microsoft offers security for both models, using its Microsoft Forefront security product line for filtering attacks, viruses, and SPIM, which is SPAM for Instant Messaging and if it finds its way into your network, say hello to the most annoying experience of your IT administrative life. To encrypt the communications, SIP uses **Transport Layer Security (TLS)**. More detailed information on TLS can be found via the IETF Web site at http://www.ietf.org under the RFC 2246. TLS is an evolution of **Secured Sockets Layer (SSL)**, which is heavily used in the configuration of Web site and e-commerce applications online. Both require certificates that are generated from a **Certificate Authority (CA)**. A CA can be private or public, meaning you can use the CA service provided within an internal Microsoft Windows Server environment to generate certificates for your internal SIP servers. An example of a public CA would be providers such as VeriSign, Entrust, and even GoDaddy that generate certificates that are widely trusted and are already installed as preconfigured root certificates on every new PC and Mac as well as on mobile devices such as Windows Mobile and Blackberry devices.

TLS uses an architecture that includes a CA that generates a root certificate trust and also generates certificates for servers and clients that trust against the root CA. This brokering of certificates as depicted in Figure 1.12 enables a VoIP/SIP environment to establish secure sessions between servers and between clients.

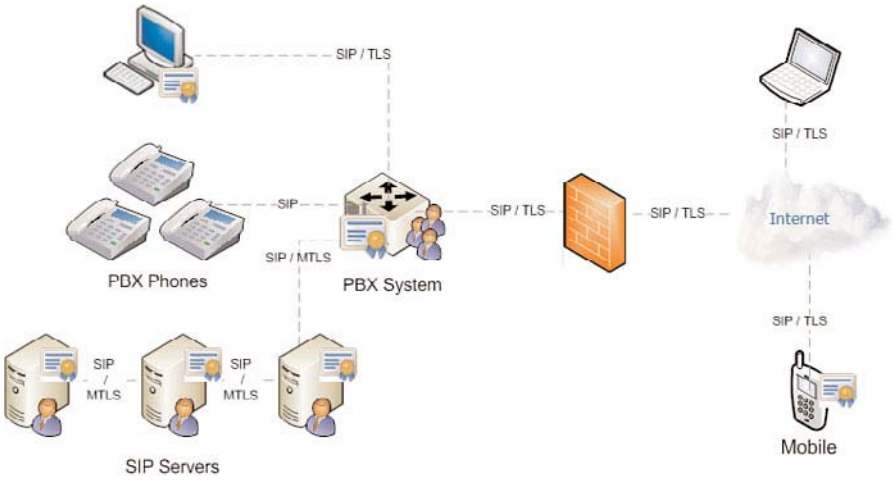


FIGURE 1.12 TLS/MTLS architecture

To provide security between servers exclusively, the use of **Mutual Transport Layer Security (MTLS)** certificates configure a trust between VoIP/SIP servers. Within the Microsoft Unified Communications infrastructure itself, Microsoft enforces the use of TLS and MTLS to ensure that the UC servers that communicate with one another and clients that communicate with these servers, between each other, and between federated partners or public ITSP networks, are trusted and secure.

The Microsoft Vision of Software-Powered Voice

As mentioned earlier, Microsoft has taken the software approach to providing voice and unified communications products instead of relying on hardware devices such as PBX systems and phone devices to make these innovative end-user communication features usable.

Building off Microsoft-designed and industry standards-based protocols and codecs, combined with learning from the mistakes of traditional Telco manufacturers and injecting some of the industry's top voice and unified communications product specialists, Microsoft has the opportunity to take voice services, applications, and devices to an entirely new level.

What inspires me the most about Microsoft's vision for voice and unified communications solutions has to do with the determination to build an open platform, but also to design and develop new technologies, including applications and devices. But most inspiring are the underlying voice protocols, codecs, and the obsessiveness to ensure the quality of these services to provide customers a high-fidelity, price conscious, and adaptable voice solution despite the size of the customer. My vision of where Microsoft will clearly succeed in the areas of voice and unified communications technologies is within the focus of software plus services. Many of Microsoft's competitors are busy designing the latest and greatest switches, conferencing devices, phones, and headsets. To be completely honest, this technology is archaic and rudimentary. What will change the way people truly communicate over the next 5 to 10 years lies within the ability to transform voice through applications. My vision is that we will be able to develop line-of-business and vertical industry applications that will completely change the way business processes are executed and will provide organizations with the ability to provide services that make a difference in the way humans communicate and collaborate in a way that is not yet understandable. The only true leader in this space, the only company focused on providing visionary developers and strategic organizations with the tools to build these innovative voice solutions, is clearly and always has been Microsoft Corporation. The focus on software, the ability to integrate voice and unified communications into applications, and the future of these technologies will enable an experience that crosses devices, applications, Web sites, mobility, and virtualization like we have never seen before, and this all comes down to one fundamental aspect, human presence. We will now explore these technological breakthroughs in more detail to give you a better understanding of what is ahead and how to prepare for the coming Communications Renaissance!

INDEX

C

D

E

L

M

N

O

P

Q

T

U

V

W-X-Y-Z