Summary

One of the most frequently asked questions from enterprises looking to implement a voice solution over a converged voice and data network is, "how will the telephone conversation sound when it is placed over my network?" And while the answer to this question is somewhat subjective due to the unique composition of every network, a set of internationally recognized Voice Quality Testing (VQT) standards is emerging to help objectively quantify this property.

This paper will describe VQT and address some of the general factors that affect (and can potentially impact) end-to-end voice quality over a network. MCK has rigorously tested its equipment in accordance with one of these standards (PAMS) and the results are included in this document.

MCK & Voice Quality Testing (VQT)

White Paper

Website: www.mck.com

September 2002

Corporate Headquarters – United States
117 Kendrick St ● Needham, MA 02494-0708
Toll Free 888-454-7979 ● Fax 617-454-6101

Research & Development Center
4040 Bowness Road NW ● Calgary, Alberta
T3B 3R7
Toll Free 800-661-2625 ● Fax 403-247-9078

Introduction

As telephone voice networks evolved to carry packet traffic, and new standards emerged to support the transport of voice over this medium, a unique set of problems arose with regard to voice quality.

Voice quality had hitherto been taken for granted, as it was typically not an issue with time division multiplexing (TDM) over the public-switched telephone network (PSTN). The standards that formed the backbone of the PSTN – known as SS7 – provided linear wavelengths, consistent bandwidth, manageable echo, and a predictably high level of quality that delivered real-time business quality voice anyplace in the world.

The emergence of voice over Internet Protocol (VoIP) has radically altered the world of telephony. VoIP has proven it can provide consistent high-quality voice, but it was necessary to create a tool that could cheaply and effectively measure voice quality over this type of network in order to reliably and objectively measure these results.

This led to more rigorous standards by the International Telecommunications Union (ITU) for the creation of voice quality testing (VQT) methods.

Voice Quality Defined

Before we begin our general discussion on how to test and measure voice quality, and review the factors that affect voice traffic over a packet network, a definition of what is meant by 'voice quality' is in order.

Voice Quality is Subjective

Voice quality is essentially a subjective measure of the clarity, inflection and tone of the conversation between a caller and the recipient. Although many characteristics influence the perception of quality (including environment and premise equipment, etc.), in general the human ear has come to define "acceptable" voice quality within a narrow range of values.

The challenge for VQT is to model this very human expectation of "quality" with mathematical equations that predict the user's experience in terms of objective and measurable criteria. The results are testing methodologies that produce a number that corresponds to how a vast majority of users will perceive the conversation.

Things that Impact Voice Quality

The things that impact voice quality can be roughly categorized into three areas:

1. **Delay** – the time it takes for a signal to travel from the caller to the recipient
2. **Echo** – the reflection of the user's voice over the network and back to the user
3. **Clarity** – the general signal strength, fidelity and clearness of the voice conversation

1. Delay

Delay, sometimes called latency, is caused by the overall network overhead introduced by switching equipment that buffer, switch and queue the voice stream as it travels over the network. Sometimes delay is related to the sheer distance of a PSTN call and the various loops over which it must transmit; but more often it is associated with IP network delay.

IP network delay is caused by the unique characteristics of IP that require the packets to travel through multiple "hops" and a variety of network equipment to reach its destination.

Delay is a characteristic inherent in the telephone network, and a little bit of delay is not even noticeable. However, the effect of too much delay is that voice queues can get misinterpreted. For example, the listener may experience abnormally long periods of silence at the end of a conversation, and may attempt to talk over his partner. The result is that conversation may become jumbled as two users attempt to converse.

How Much is Too Much Delay?

A general guideline is that below 200 ms in one-way delay is acceptable, with 100 ms and below being optimal.

MCK Delay Guidelines (One Way)	
Below 100 ms	Excellent. Delay generally not detectable
100 ms – 200 ms	Good quality. Slight delay or hesitation may become noticeable
Over 200 ms	Unacceptable delays; conversation becomes difficult

Jitter Delay

Another issue with delay, called jitter, is a variable change in the strength of the signal. Jitter can often result in the conversation of the two parties becoming jumbled together, making communication difficult.

MCK Jitter Delay Guidelines (One Way)	
Below 40 ms	Excellent. Jitter generally not detectable
40 ms – 75 ms	Good quality. Occasional delay or jumble noticeable
Over 75 ms	Unacceptable quality; conversation prone to jumble

Many network devices, including the MCK EXTender™ 6000 for Branch Offices, MCK EXTender 7000 for Branch Offices, EXTender PBXgateway™ and EXTender

PBXgateway II come with jitter buffer settings to help dampen the effect of jitter on the overall voice conversation. However, a jitter buffer - by default - adds additional delay that must be calculated into the overall network latency for performance measures.

2. Echo

Echo is caused when voice paths cross and the resulting energy is reflected back at the user. This generally happens when a voice stream transitions from one type of network to another, for instance, when a call path moves from a 4-wire E&M circuit to a 2-wire FXO line.

Like its name implies, echo sounds like the user's voice reflected back to him or her over the network. Some people describe the effect as carrying on a conversation in a large, hollow room. Echo and delay can go hand-in-hand, in that where delay is introduced, echo becomes more prevalent.

The solution to echo is having all of the voice traffic run through a carrier grade/toll grade circuit at the central office. This will eliminate many of the effects of the electrical crossover. Additionally, many gateway products, such as MCK's PBXgateway, come with industry-standard G.165 echo cancellation software, which lessens the impact of excessive echo.

3. Clarity

The final element that can affect voice quality is the clarity of the actual conversation.

There are four main things that can affect this:

Voice compression – voice compression codecs are used to optimize bandwidth by shrinking the size of the packet necessary to carry the conversation over the network. Voice compression, by its very nature, is a trade-off of quality versus bandwidth efficiency. Generally speaking, the greater the bandwidth, the higher the voice quality.

Packet loss – the nature of the IP network makes packet loss a reality. Packets that are lost are not retransmitted in a VoIP scenario, and a sizeable increase in the number of lost packets may result in a cessation of the voice stream and silence in the conversation. Different voice compression algorithms handle packet loss differently. For example, G.729A is a very efficient compression algorithm for handling packet loss; whereas, G.726 (24 Kbps ADPCM) is more susceptible to packet loss.

Service provider - the number of service providers through which the voice stream must travel on the network and the quality of the circuit through which voice travels will affect the clarity of the voice call. Limiting the number of service providers to a single provider for the entire WAN network is optimal.

Voice transcoding – transcoding refers to the use of multiple compression algorithms in a voice stream. An example of this is a cell phone call that uses one type of voice compression going through a gateway that compresses it into another type of compression. This can deteriorate the voice clarity of a conversation. Transcoding can also add delay to the overall voice stream.

Silence Compression and Comfort Noise Generation (CNG)

MCK equipment comes enabled with the ability to recognize silence in a voice conversation and to send this traffic as a single packet. This cuts down on the overall traffic sent over the network. A smaller, tighter stream of packets tends to arrive at its destination sooner.

At the same time that silence compression is transmitting, Comfort Noise Generation (CNG) adds background noise, some gentle hissing, and an occasional line pop to the conversation. This may sound counter-intuitive if the goal is to create excellent voice quality, but studies have shown that users prefer background noise as confirmation that the other party is still engaged in the conversation. If users perceive total silence, they may think the other party has dropped from the conversation and may hang up.

The Importance of Voice Quality Testing

Voice Quality Testing (VQT) is an objective and comprehensive methodology for testing end-to-end network voice quality. It is important, especially with the advent of Voice over Internet Protocol (VoIP) to quantify how the average human ear would perceive a conversation, especially before implementation of the solution.

Prior to modern testing methodologies, the measurement of voice quality over a TDM network was an inexact affair to say the least. Testing consisted of calculating the characteristics of various linear sound wavelengths, such as signal-to-noise ratio (SNR) and total harmonic distortion (THD). But testing was usually not necessary, as voice typically traveled over TDM without much distortion.

However, Internet Protocol (IP) and Voice over IP (VoIP) changed all that. Internet Protocol is a non-linear technology not measurable through sampling testing, and new low bit-rate voice compression codecs, such as G.729A, created packets instead of wavelengths.

No Quality of Service (QoS)

QoS can be thought of as the minimum parameters a network must maintain in order for acceptable quality voice to be attained. IP networks were designed with one purpose in mind: to carry "bursty" types of data over flexible paths, with varying amounts of bandwidth, and to connect a wide range of sites all sending and receiving data simultaneously. As such, IP networks have no built-in guaranteed quality of service (QoS). Voice is delay-sensitive and prone to quality deterioration if network parameters exceed limitations.

The complexity of the IP network also became a factor as software installed on gateways, routers and other equipment can manipulate characteristics of the packets that travel through them. What was needed was a new method of testing.

Two Methods for Quantifying Voice Quality

Two popular methods for measuring voice quality have emerged in recent years. They are PAMS and PSQM. Both standards have been embraced by the telecommunications and information technology communities, and are used continuously to measure end-to end-voice quality throughout the industry.

Note: These two methods may be used to measure VoIP, as well as traditional TDM networks.

PAMS – A Perceptual Analysis/Measurement System

PAMS is a speech quality metric that uses an auditory model to mathematically describe the way a human ear perceives voice, and performs an analysis of errors upon that model.

PAMS scores on a mean opinion score (MOS) of 1 - 5, where 5 is the best quality possible. A PAMS score of 4 or above is widely considered "business quality voice."

PAMS Voice Quality Scale Mean Opinion Scores (MOS)	
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

Note: PAMS MOS scores are usually expressed to two decimal places. (Ex. 4.84)

PAMS also splits its criteria into two different areas known as *listening effort* and *listening quality*:

Listening Effort (LE): The amount of effort a person must give to understand the conversation.

Listening Quality (LQ): The overall clarity and fidelity of the conversation.

PSQM – Perceptual Speech Quality Measurement (ITU-T) Rec. P.862

The PESQ is a recommendation from the International Telecommunications Union (Rec. P.862) for measuring distortion, noise and overall voice clarity. It was originally designed to measure voice codecs, but has since evolved to express voice quality in terms of an MOS.

Like PAMS, PESQ is measured on a MOS scale of 1-5, but the scale is reversed; on the PESQ scale, 1 is the best quality and 5 is the worst.

MCK Scores and Voice Quality Using PAMS

MCK has rigorously tested the voice quality from equipment that extends the features, functions and applications of the PBX to remote locations. The results of these tests and the PAMS score are included in this section.

How the Tests Were Conducted

MCK used its award-winning EXTender PBXgateway and MCK EXTender 6000 as the test products. The products were tested over a T1 line using a central office simulator with call generation/reception capabilities.

An audio file was generated, and a standard "non-extended" voice circuit provided the baseline for the test. The baseline tested at a perfect 5.0, the expected outcome for "non-extended" and uncompressed voice.

Next, this audio file was used to test an extended phone using each of the four compression algorithms that MCK supports and the results below were generated. Also the results include the use of non-linear processing (NLP) echo canceller.

MCK PAMS MOS Scores (Scale 1-5)

	G.711 (64K)	G.726 (32K)	G.726 (24K)	G.729A (8K)
PAMS LE	4.83	4.26	4.24	4.37
PAMS LQ	4.76	4.27	4.06	4.12

5.0 = perfect

These results are quantifiable and reproducible, so results can be objectively obtained with repeated trials.

MCK Voice Quality Disclaimer

It is important to note that QoS is a function of network, service provider and careful management of network delay, echo and jitter. MCK cannot guarantee voice quality over the network, but with careful management of the network, sustainable, high-quality voice is very achievable.

The PAMS scores included here depend on the audio file being used, and can vary from test to test depending on the quality of the audio file.

Conclusion

Although voice quality is subjective by nature, internationally recognized standards of PAMS and PESQ testing have led to very accurate, objective and measurable results in voice quality testing (VQT). This is helpful for enterprises looking to implement a Voice over Internet Protocol, such as the MCK EXTender 6000, MCK EXTender 7000, EXTender PBXgateway and EXTender PBXgateway II products.

As companies deploy VoIP solutions, it is important to manage elements such as delay, jitter and echo that could decrease the quality of voice over the network.

MCK's rigorous testing using the PAMS testing algorithms has yielded some objective, quantifiable and repeatable results for customers who are looking to implement a Voice over Internet Protocol.

Acronyms

Codec – compression/decompression algorithm

IP – Internet Protocol

ITU – International Telecommunications Union

MOS – mean opinion score

PSTN – public switched telephone network

QoS – quality of service

SS7 – Signaling System 7

TDM – time division multiplexing

VoIP – voice over Internet Protocol

VQT – voice quality testing

About MCK Communications

MCK Communications develops and markets distributed voice solutions that cost-effectively create a more open and interconnected business telephony environment which provides enterprises with extended business reach, lowered communications costs and options for flexible migration to next-generation environments while empowering individuals with portable business telephony. With more than 340,000 ports shipped worldwide, ISO-9001-certified MCK partners with such industry leaders as Alcatel, Anixter, Avaya, BellSouth, Catalyst Telecom, Ericsson, GBH Distributing, Iwatsu, Nortel, NEC, Infrontia (formerly Nitsuko), Panasonic, Toshiba, SBC, Sprint, Verizon and Voda One Corp. Headquartered in Needham, Massachusetts, the company has an R&D facility in Calgary, Alberta and field offices throughout the U.S., Canada and the U.K.