

ARTISAN®
TECHNOLOGY GROUP

Your **definitive** source for quality pre-owned equipment.

Artisan Technology Group
(217) 352-9330 | sales@artisantg.com | artisantg.com

Full-service, independent repair center
with experienced engineers and technicians on staff.

We buy your excess, underutilized, and idle equipment
along with credit for buybacks and trade-ins.

Custom engineering
so your equipment works exactly as you specify.

- Critical and expedited services
- In stock / Ready-to-ship
- Leasing / Rentals / Demos
- ITAR-certified secure asset solutions

Expert team | Trust guarantee | 100% satisfaction

All trademarks, brand names, and brands appearing herein are the property of their respective owners.

Find the *Keysight / Agilent J1981C* at our website: *Click HERE*



Agilent Technologies

Testing Voice Quality on Next Generation Networks Using the Agilent VQT

Application Note 1376

Agilent Technologies

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Introduction 4

What is Voice Quality? 4
Why Test Voice Quality? 4
Network Impairments 5

Testing Voice Quality with the VQT - An Overview 6

VQT System and Its Deployment 6
Common Testing Tasks 8
Benchmarking 8
Baselining 9
Detecting Impairments 10
Isolating and Troubleshooting Impairments 10
Optimizing Networks 10

Testing Voice Clarity 11

Clarity Metrics 12
Measuring Clarity with PSQM+ 13
How the Measurement Works / Setup Tips 13
Interpreting PSQM+ Measurement Results 14
Other PSQM Test Techniques 22
Measuring Clarity with PAMS 23
How the VQT Measures Clarity with PAMS 23
Interpreting PAMS Scores 24
Interpreting PAMS Graphs 25
Detecting Impairments and Optimizing Networks for Clarity 27
Using the Clarity Trend and Clarity File Measurements 28
Testing Voicemail and Other Non-realtime Systems 29
Other Clarity Test Strategies 30
Traffic Loading 30
One-way and Round-trip Clarity 30
Analog vs. Digital Interfaces 31
Use of Training Cycles 32

Testing Voice Delay 32

Network Delay Budget 32
How the VQT Measures Delay 33
Interpreting the Delay Graph 34
Benchmarking your delay measurements 36
One-way and Round-trip Delay Measurements 37
Other Delay Measurement Techniques 38
Measure Delay in Both Directions 38
Delay Trending 38
Use of Pre-average Cycles 38
Detecting Impairments and Optimizing Network Performance 39

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Testing Echo 39

How the VQT Performs Echo-PACE Measurements 40
Test Strategies for Echo 41
Choice of Interfaces Play Significant Role 41
Echo Measurement's Network Simulator 43
Benchmark Echo Measurements 43
Determining Acceptable Echo 44
Determining Presence and Location of Existing Network Echo 45
Determine If and Where Echo Cancellers Are Needed 45
Testing Echo Canceller Performance 46
Use Network Simulator to Dominate Immediate Echo on FXO Source 46
Testing Echo Cancellation During Doubletalk Conditions 47
How the VQT Performs DTalk Measurements 47
DTalk Testing Strategies 47

Testing Silence Suppression 48

What is Silence Suppression? 48
Testing Silence Suppression with the VQT 49

Testing DTMF 49

DTMF and VoIP 49
Testing DTMF with the VQT 50

Other IP Telephony Testing Strategies with the VQT 51

Measuring Call Setup Time 51
Tasklists and Autotasker 52
Network Simulator 52
Testing Networks with Multi-stage Dialing 53

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Introduction

Over the past century, traditional circuit-switched networks have been designed and optimized for the time-sensitive delivery of voice traffic. Public Switched Telephone Networks (PSTN) provide a highly predictable quality of service for voice, and people are accustomed to holding the PSTN as the standard-bearer for voice quality.

Next generation networks, particularly voice-over-packet networks, are now being deployed to exploit the benefits of reduced network capital and operating costs, shared bandwidth efficiencies, and simplified network management. With these benefits, however, come significant challenges in delivering the quality services that have long been taken for granted. Private voice networks, which were designed to emulate the PSTN with regard to voice transport and quality, are some of the first networks to take advantage of voice-over-packet technology. At the same time, many new carriers are building global IP telephony networks to compete head-on with traditional PSTN carriers in the public market.

Voice quality, at one time very predictable via the PSTN, has now become a key differentiator for network equipment and services and a critical factor to the success of both public and private networks. Consequently, the field of voice quality testing is developing very rapidly, and sound strategies and methods are being deployed effectively. Next generation carriers and enterprise network operators must pay close attention to these strategies in order to successfully deploy voice services on data networks. Likewise, network equipment manufacturers are being scrutinized very closely by their service provider customers when it comes to delivering voice quality.

Agilent Technologies' Voice Quality Tester (VQT) is the industry's premiere tool for testing voice quality over communications networks. The VQT provides the most comprehensive and robust set of features for measuring voice quality objectively. This application note offers some techniques and strategies for testing voice quality using the VQT.

What is Voice Quality?

Voice quality is one of the most important aspects of overall Quality of Service (QoS) for voice telecommunications. Overall QoS is comprised of:

- Voice quality (clarity, delay, echo, clipping, level).
- Call completion (network reliability and availability as measured by call completion rates, routing confirmation, and call setup delay).
- Support Services (billing, customer care, assistance services).

Voice quality is, in essence, in the ears of the beholder. Although there are dozens of metrics that characterize the electrical signals used on the PSTN to carry voice traffic, none accurately reflect voice quality on next generation networks.

While there are many factors that affect voice quality, there are three key parameters that characterize it: clarity, delay, and echo. These parameters, as well as several other aspects of voice quality and VoIP network performance are addressed in this Application Note.

Please refer to the VQT's *Getting Started Guide*, online Help, or the Agilent Technologies *Voice Quality in Converging IP and Telephony Networks* white paper for more information on general voice quality definitions and issues.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Why Test Voice Quality?

Circuit-switched networks (e.g. the PSTN) have always provided a predictable level of quality by allocating dedicated bandwidth and using non-compression analog-to-digital encoding techniques. However, with the deployments of converged networks including voice-over-IP, ATM, frame relay, and the PSTN, voice quality is neither guaranteed nor predictable. Voice quality has become a key differentiator for networks, equipment, and services. In fact, low-cost service is no longer regarded as the primary business driver for VoIP. End-users, who are used to paying for the predictably high quality provided on the PSTN, will not tolerate substandard voice quality. Therefore, testing voice quality on new networks is imperative to ensure that they are competitive to each other and to the PSTN. Some applications for testing voice quality include:

- Compare different equipment (gateways, PBX's, phones, routers, IADs, MTAs)
- Compare different networks (PSTN, VoIP, VoATM, different router or switch backbones)
- Compare different technologies (codecs, VADs, echo cancellers)
- Determine impact of different configurations (jitter buffers, router networks, QoS settings, echo cancellers)
- Test new hardware and software releases
- Regression testing of services and networks
- Certify network deployments
- Compare voice quality with competitors' networks, products, or services
- Monitor network for impairments to voice quality

Voice quality is subjective. However, for the sake of consistency, reliability, repeatability, and accuracy, it is vital to measure it with objective means.. It is also important to measure voice quality from the perspective of the end user since it is here that voice quality matters the most. The VQT is the right system to use to objectively and accurately test voice quality from end user perspectives as well as from different access points in the network.

Network Impairments

There are various ways to deploy voice services over an IP, ATM, or Frame Relay network. However, most deployments share the same basic processes and, therefore, the same impair ments to voice quality. These impairments are often the result of the trade-offs between voice quality, delay, bandwidth usage, and CPU utilization.

Figure 1 illustrates many of the sources of voice quality impairments in a typical IP telephony network. Together, these impairments affect the clarity, delay, echo reception, and level of voice in the network. Knowing where these impairment sources are in a network, how these sources and their configurations impact voice quality, and how different network resources can be used to balance voice quality with bandwidth efficiency, are crucial to successful deployments of voice services on next generation networks.

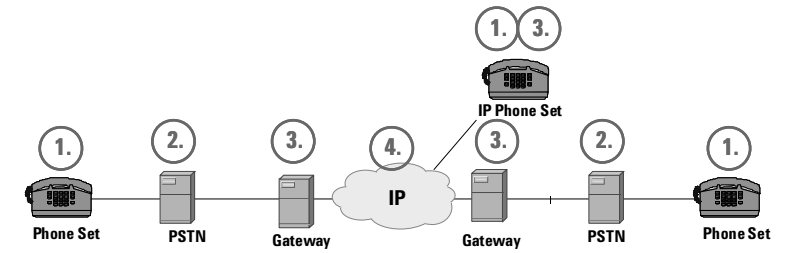


Figure 1: Sources of VoIP Network Impairments

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Testing Voice Quality with the VQT – An Overview

The VQT provides objective measurements of voice quality by employing robust mathematical techniques that provide consistent, reliable, and repeatable results. The VQT can be used in a variety of situations and in a number of modes. This section provides an overview of VQT deployments and test situations.

Note: this application note discusses generally how to measure voice quality with the VQT. You can find detailed procedures, connection diagrams, and feature descriptions in the VQT's *Getting Started Guides* and online Help.

VQT System and Its Deployment

The VQT system can be used in many different environments to accomplish different objectives. VQT servers provide analog, T1, and E1 interfaces that allow the user to test at various access points and across various types of networks. Figure 2 shows several examples of where you might connect VQT servers to access and test voice quality.

The VQT system tests voice quality by establishing a connection (placing and/or answering one or more telephone calls) across a voice network. The user can establish connections between ports/channels on a single VQT server or between servers. Once a connection has been set up, the VQT system transmits and receives test signals across the connection to evaluate various aspects of voice quality.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

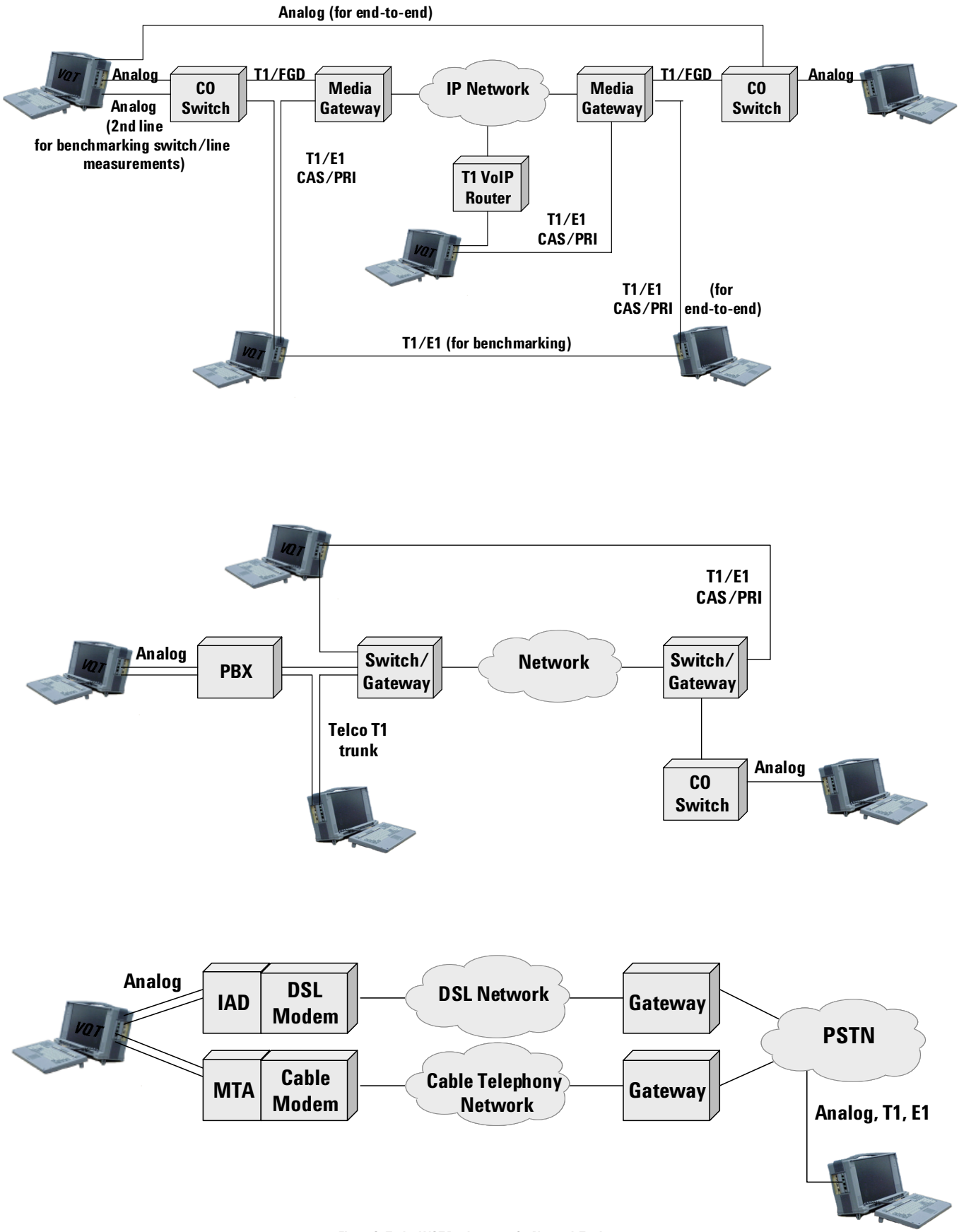


Figure 2: Typical VQT Deployments for Network Testing

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Other aspects of how the VQT system is used are:

- **Local vs. Remote Control**
  Depending on the type of server, each of the testers shown in Figure 2 can be locally controlled via its own monitor, keyboard, and mouse, or remotely controlled across an IP management network. This allows users to run voice quality tests from centralized locations.

- **Single vs. Multiple VQT Servers**
  As implied above, many test scenarios can be performed with a single, locally or remotely controlled VQT server. However, there are numerous test techniques that rely on test signals being transmitted between two VQT servers (for example, one-way Clarity as described later in this document).

- **Call Mode vs. Audio Path**
  Another important characteristic of the VQT system is the distinction between the call that is placed between ports/channels and the path taken by the audio test signal once that call is set up. Some types of voice quality measurements only make sense when the audio path is one-way. Other measurements can take advantage of round-trip audio paths.

  Please refer to the VQT's online Help and *Getting Started Guide* for more detailed information about how to connect and configure VQT equipment, how to place/answer calls, and how to run measurements. Also use these other sources of information to learn about the VQT application and its features.

Common Testing Tasks

Regardless of the tool, there are certain voice quality testing tasks that can be performed. It is a good idea to understand the overall testing process before learning about detailed descriptions of using the VQT.

Benchmarking

To implement an effective overall voice quality testing program for next generation voice-over-packet networks, it is crucial to establish the appropriate benchmarks to which subsequent measurements can be compared. Many different sources for benchmarks are available:

- Private or public circuit-switched networks that represent the voice quality most people are used to. In particular, the PSTN provides the most valuable source.

- Ideal networks that impose no processing. One example is time-domain multiplexing (TDM) hairpin circuits within switches, gateways, and other edge devices. Another example is a line simulator.

- Established standards such as ITU recommendation G.114 for delay.

- Subjective test results such as MOS (Mean Opinion Scores).
  Figure 3 shows an effective method that can be used to benchmark a VoIP network against the PSTN. The VQT connects to a Class 5 switch to access both the VoIP network and the PSTN. Since the VQT uses the same interfaces and access lines for measurements across both networks, a valid comparison measurement results can be made. In other words, differences between measurements can be attributed to the different networks and not changed connections.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

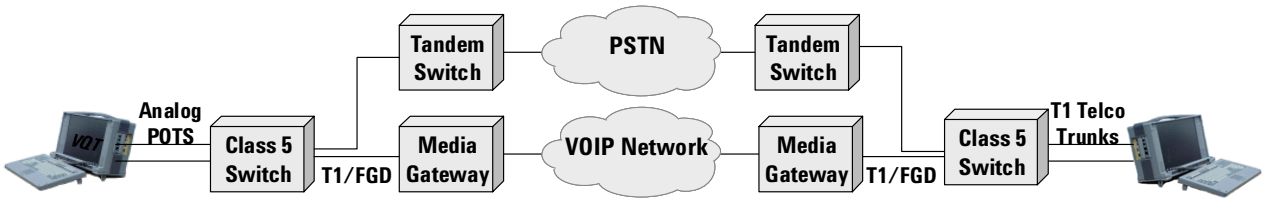


Figure 3: PSTN Benchmarking for VoIP Network Testing

Some important considerations and strategies for benchmarking are:

- Benchmark each measurement type. Determine the values for clarity, delay, front-end clipping, and other measurements that will serve as the standard to be met.

- Benchmark the impact of individual resources and their impairments. For example, determine how different codecs affect clarity, and how different levels of packet loss affect a particular codec with respect to clarity.

- For live networks, benchmark against the PSTN using the same endpoints for the call (as shown in Figure 3).

- For labs, benchmark against line emulators and TDM hairpin connections.

- Benchmark via subjective testing:
  - Determine what is acceptable to end users, in terms of delay, clarity, echo, and front-end clipping.
  - Determine minimum quality values (the lower end of acceptability) and optimum quality values (the lower end of "impressive" quality) for each metric.

Baselining

Another important aspect of a comprehensive voice quality test program is baselining. Baselining is measuring a network to determine nominal performance for metrics such as delay and clarity. Typically, baselining requires that measurements be made regularly over some extended period of time (i.e. measurement trending). Baseline measurement results can be valuable when detecting and isolating impairments, and can help expose undesirable changes in network performance before it becomes critical.

Baselining is different than benchmarking. For example, your benchmark for delay may be 150 milliseconds (based on ITU G.114). However, your network may consistently deliver delays of 100 milliseconds (that is, your network's delay baseline is 100 ms). While a measured delay of 140 ms meets your benchmark, it exceeds your baseline and may indicate an impairment such as IP network congestion.

In a lab environment, you can use the VQT's traffic loading feature on the T1/E1 interface to characterize network performance against varying traffic conditions.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Detecting Impairments

Once network performance is characterized and/or benchmarks established, the VQT is used to detect impairments by setting thresholds to established baseline values or benchmarks, and running trending measurements.

For clarity, an average score is not enough. A transient impairment, such as bursty packet loss, can cause a noticeable effect, but for not enough time to negatively impact the clarity score average. You need detailed time-domain scoring, such as provided by the VQT, to detect transient impairments.

It is important to know when a detected impairment really is an impairment. This is enabled with trending measurements. For example, when running hundreds of PSQM+ measurements repetitively, a single test may produce a result above a threshold. However, this may be an anomaly, especially if all other results are below the threshold. Trending is used to highlight when a series of measurements indicates an impairment.

You can use of T1/E1 traffic loading can detect clarity impairments and excessive delay inflicted by VoIP DSP resources due to increased traffic loads.

Isolating and Troubleshooting Impairments

While the VQT can detect voice quality impairments on an end-to-end network, isolating the impairment to determine exactly where it is requires more effort. Typically, locating the cause of a particular impairment requires either segmenting the network or replacing individual elements/resources in an incremental manner, and performing measurements each time to determine at which point the measurements indicate quality is severely affected.

The VQT provides detailed analysis techniques that can expose sources of impairments. Some of these techniques are described in this paper and are related to each measurement. The analysis of graphs and numerical measurements, along with techniques for isolating impairments, can enable the determination of the source of voice quality problems.

Optimizing Networks

The VQT can be used to optimize networks for voice quality. Network design often involves many trade-offs between voice quality and network efficiency. With each design change or even simple re-configuration, you can use the VQT to obtain a quantitative and accurate measurement of the impact on voice quality. By performing numerous measurements with the VQT, the design or configuration that delivers the optimal voice quality can easily be determined.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Testing Voice Clarity

Clarity refers to the clearness or fidelity of voice as it is reproduced by a network and perceived by a listener. It is affected by different types of distortion introduced by various network elements. For the most part, it is independent of delay in that the clearness or intelligibility of a voice signal can be maintained when transported with zero, little, or excessive delay across a network (however, delay *variance* can influence clarity). Clarity is also independent of echo since echo is perceived by the speaker and clarity is perceived by the listener.

Factors that influence the clarity of voice on a communications networks are:

- Waveform speech encoding which introduces quantization distortion. This includes traditional analog-to-digital conversion using pulse code modulation (PCM) as defined in G.711 and Adaptive Delta PCM (ADPCM) using G.726.
- Non-waveform speech compression which introduces non-linear distortion. Examples are CS-ACELP (G.729A), LD-CELP (G.728), and LPAS (G.723.1).
- Packet loss
- Packet jitter (a.k.a. delay variance)
- Front-end clipping (time clipping) due to Voice Activity Detectors
- Echo cancellation
- Muting and attenuation
- Background noise and EMI (electro-magnetic interference)

Most impairments to voice clarity occur at the edge of a network where most VoIP processing is located. This is another reason why it is important to measure clarity from end users' perspectives; that is, at points of call origination using analog or T1/E1 interfaces. Other impairments, including packet loss and packet jitter occur in the core network.

Measuring clarity can be challenging because impairments to voice clarity are not additive. In other words, the presence of multiple impairments can have a non-linear effect on clarity. A certain impairment can also influence the effect of another impairment in different ways. These characteristics prohibit the use of linear models for network planning such as the QDU method outlined in G.113. Designing a non-linear network (such as a VoIP network) to meet voice clarity objectives requires careful measurement of the actual network in addition to modeling and analysis on paper. For example:

- Transmitting voice through a compression encoder/decoder twice will introduce more than twice the distortion of a single encoder/decoder.

- Excessive packet jitter can be mitigated with the use of jitter buffers on terminating gateways. However, jitter buffers introduce more delay, particularly when they must adapt to greater values of packet jitter. Limiting jitter buffer delay, however, can result in higher packet loss (more late packets will be dropped) and, thus, more impairments to voice clarity.

- A single value of packet loss (e.g., 5% random) will affect voice encoded with G.723.1 differently than voice encoded with G.729A.
  For these reasons, it is important to measure the effects of individual impairments on clarity by isolating the impact of each.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Clarity Metrics

Operators of traditional circuit-switched networks have employed metrics such as Signal-to-Noise Ratio and Total Harmonic Distortion to measure the quality of network transmissions. However, these metrics are insufficient when faced with the non-linear characteristics of next generation networks because they do not accurately measure voice clarity as it is perceived by actual end users (i.e. they do not measure 'perceptual voice clarity'). Next generation networks of today require measurement techniques that take human perception into account. One such technique is to perform subjective testing in which real people judge and score the clarity of a particularly voice signal. This technique is known as Mean Opinion Scoring (MOS) because the mean value of a large number of human opinion scores is calculated.

There are several limitations to MOS testing. It is inconsistent and unrepeatable due to its subjective nature. It is inefficient and very expensive. Finally, it is impractical in an environment in which the impact of frequent design changes and re-configurations must be measured, or in situations in which ongoing service monitoring is needed.

To address this problem, many mathematical techniques have been developed and implemented in software to measure perceptual voice clarity. The two most valuable and popular techniques are the Perceptual Speech Quality Measurement (PSQM) and the Perceptual Analysis Measurement System (PAMS). PSQM was originally developed to measure clarity across speech codecs, and was approved by the ITU as Recommendation P.861. A later improvement called PSQM+ accounts for packet network transmission impairments such as packet loss. PSQM and PSQM+ produce numerical scores that represent the 'perceptual distance' between an input and output signal, in which the distance is an indication of the perceptible distortion added to the output signal.. PSQM and PSQM+ scores can be converted to estimated MOS scores.

PAMS is a similar technique that uses different signal processing methods, and produces scores that correlate directly with MOS scores.

It is important to note that PSQM+ and PAMS are two different models for measuring perceptual quality. Although PAMS provides scores that correlate directly to MOS scores and PSQM+ provides estimated MOS equivalent scores converted from average PSQM+ scores, the estimated MOS equivalent scores from PSQM+ will not necessarily match the PAMS LQS or LES scores. PSQM+ and PAMS predict the results of subjective testing in different ways, and are objective methods for predicting a very subjective test result.

The VQT system uses both the PSQM+ and PAMS algorithms to measure clarity. Please refer to the VQT's online Help or *Getting Started Guide* to learn more about these measurements.

For more information on perceptual clarity metrics, refer to Agilent's whitepaper "Methods for Measuring Perceptual Speech Quality".

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Measuring Clarity with PSQM+

This section discusses how to test clarity using the VQT's Clarity (PSQM) measurement.

How the Measurement Works / Setup Tips

To measure voice clarity using PSQM+, the VQT places a telephone call over the network-under-test, and then transmits a user-selectable voice signal from one end of the call, and records that signal at the other end of the call after it has passed through the network. The VQT then time-synchronizes the transmitted and received signals (which are saved as wave files) to remove effects of network delay, and inputs the time-synchronized signals into the PSQM+ algorithm. The PSQM+ algorithm compares the two signals and produces a time-correlated voice signal / PSQM+ score graph, and a statistics spreadsheet. These measurement results indicate, among other things, the amount of perceivable distortion present in the received signal. A brief mention of other aspects of this analysis is useful:

- The VQT calculates and plots PSQM+ scores every 16 milliseconds. This is not PSQM+ "sampling", but rather frame-by-frame scoring.

- While the PSQM+ algorithm is a consistent and repeatable one, it does not account for time synchronization; it assumes the test and reference signal are already time-synchronized. Thus, synchronizing the test and reference signals is a crucial function of any PSQM test tool. The VQT synchronizes the test and reference signals by correlating them on a piecemeal basis. That is, the VQT examines each time segment of the test and reference signals, determines a mathematically best fit of the two signals within the time segment, and performs a time shift of the test signal segment to synchronize it with the reference signal. This is the most robust method available of any tool on the market and will produce the most reliable results.

- As a result of the above described time correlation, an average time-shift is measured and reported as an estimated delay. This delay result should be close in value to the result of the VQT's proper Delay measurement. However, the VQT's proper Delay applies a more robust method and will produce more accurate results. In most cases, the difference in values between the VQT's proper Delay measurement and the PSQM+ delay result is less than 5%.

- The VQT may fail to correlate the test and reference signals, resulting in a "correlation timeout". This may be due to any of a few reasons:
  - The test signal is severely distorted to the point that it cannot be recognized.
  - There is excessive noise in the transmission.
  - There is excessive delay variances in the test signal.

  Very often, it is valuable to first perform a proper Delay measurement (discussed later in this document), and then perform a PSQM+ measurement. If the measured delay in the PSQM+ measurement is close to the result of the proper Delay measurement, then the PSQM+ signal correlation can be trusted. If the measured delay in the PSQM+ measurement is not at all close to the result of the proper Delay measurement, but correlation did not timeout, then a mis-correlation may have occurred resulting in unacceptable PSQM+ scores.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Finally, it is important to note that consistency is key to effective clarity measurements. For example:

- Clarity metrics are best used as relative measurements and should not be thought of as absolute.
- Comparing PAMS or PSQM+ scores made by different test tools is not effective and potentially misleading. However, comparing results obtained by using two VQTs is effective because each has the same implementation of the PSQM+ or PAMS algorithms and voice acquisition hardware and will measure clarity the same way. The VQT provides the most robust implementation of the clarity measurement. But comparing a PSQM+ score made a VQT with a PSQM+ score made by another test tool will not produce the same results.
- Comparing PSQM scores made with different voice wave files is not as effective as using the same wave file.

Interpreting PSQM+ Measurement Results

As mentioned briefly before, the VQT's Clarity (PSQM+) measurement provides statistical and graphical results. This section discusses how to understand and interpret both of these data types. For more details on setting up and running the VQT's Clarity measurement, refer to the online Help or the *Getting Started Guide*.

PSQM+ Numerical Scores

PSQM+ is intended to be used as a relative measurement tool. That is, rather than provide absolute values such as those that can be obtained using the Delay measurement, PSQM+ scores are used to compare the clarity of voice between different networks, different configurations of a network, the impacts of different elements within a network, and the impacts of different conditions (e.g., traffic loads) on a network. PSQM+ scores are essentially a measure of distortion. A 0 indicates no distortion, or perfect clarity, while higher scores represent higher degrees of distortion, or degraded clarity.

Although PSQM+ is best used as a relative measurement, it is nonetheless useful to know what a "good" PSQM+ score is. PSQM+ scores can be roughly assigned to an absolute scale based on real-life experience.

While it is best for to develop one's own scale based on benchmarking and end user needs, a nominal scale for average PSQM+ scores can be used as follows:

0 – 2.5	PSTN voice quality
2.5 – 4.5	Acceptable but noticeably degraded
4.5 and above	Unacceptable

Average PSQM+ scores above 10 are uncommon. That is, once the signal becomes this degraded, it is very difficult to perform test/reference signal correlation.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

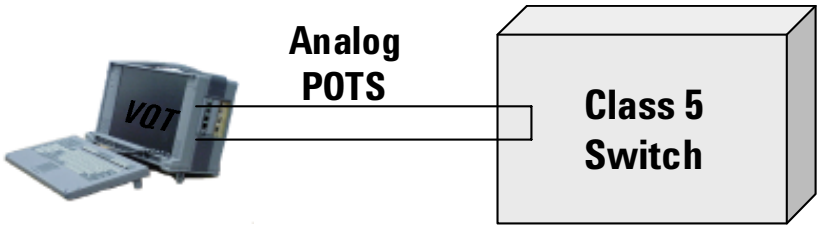
This scale reflects numerous PSQM+ tests performed on the actual PSTN, in which the highest scores (poorest quality) were 2.5 or below. It is the objective of commercial carriers using VoIP to meet or exceed PSTN quality. As such, they strive to design networks that can deliver PSQM+ scores of 2.5 or below.

Prior to executing a PSQM+ test plan, it is important to establish benchmarks. Two effective techniques for benchmarking are:

- **Using the PSTN**. This is useful for testing overall network performance to ensure that PSTN quality can be met. For example:

  (1) Use the VQT to place a local Class 5 call over the PSTN. Measure PSQM+ for this call, and use it as an access network benchmark. See Figure 4.
  (2) Use the VQT to place a call across the PSTN with the same endpoints as a call placed across the network under test. Measure PSQM+ for the PSTN call, and use that measurement result as a benchmark for a complete network call. See Figure 3.

The first method is used to benchmark the PSTN access-only portion of a hybrid PSTN/VoIP network call. This is needed to understand that the entirety of a PSQM+ score is not attributed to just the IP telephony portion. This first method is also used to benchmark a local call. The second method is used to benchmark a long-distance or non-local call on the PSTN.



Figue 4: Using Class 5 switch for PSQM+ Benchmarking

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

- **Using Time-Domain Multiplexing (TDM) hairpins**. This is useful for testing the impact of individual elements on clarity. ('Hairpin' refers to a specific type of connection in which the voice signal is not packetized or processed in any way. This eliminates the possibility of distortion due to compression or other VoIP processing.) For example:

  (1) Use the VQT analog interface to place a call from Port A or C to Port B or D, using a line simulator, router, switch, gateway, or any element that can switch the call directly back to the VQT without any processing.

  (2) Use the VQT T1/E1 interface to place a call from Port 1 to Port 2 directly, using a T1 crossover cable. To do this, simply connect the crossover cable from one port to the other, select a CAS signaling variant in the Port Setup tool, and use the Call Control tool to place the call (any telephone number will work).

  (3) Use the VQT T1/E1 interface to place a call from Port 1 to Port 2 using a line simulator, router, switch, gateway, or any element that can hairpin the call.

  Using either (or all) of the methods above, you can measure PSQM+ to use as an initial benchmark. Then add processing and network elements to the call path, one by one, and measure PSQM+ on each one to determine the incremental impact.

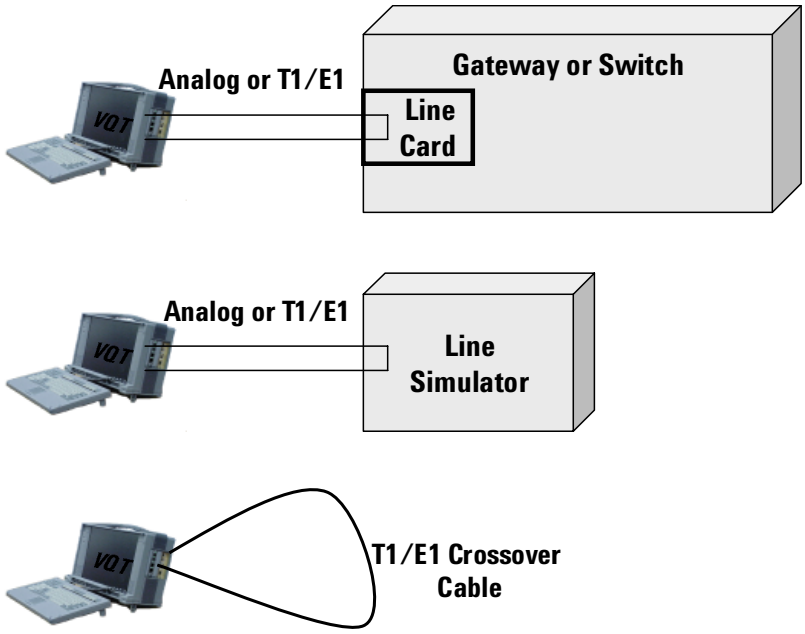


Figure 5. Simple TDM Calls for Establishing Benchmarks.

When comparing PSQM+ results, it is important to use the same speech sample from test-to-test. Scores on the same call will vary slightly with different voice samples. This is due to the unique audio dynamics of each voice sample, such as the loudness bursts of certain words, and the spectral differences (especially between male and female speakers).

It is likewise important, when performing an off-line PSQM+ test using the Clarity File measurement, to use the exact same file that was selected for transmission. These requirements dictate a well controlled and consistent file-naming scheme for PSQM+ wave files, particularly when performing PSQM+ measurements between two different VQT servers.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

An average PSQM+ score is just the beginning…

The key metric in a PSQM+ test is the average PSQM+ score. This is the time-averaged result of all the individual PSQM+ scores calculated across the voice sample. However, as will be shown later, the VQT provides more detailed analysis that is important for assessing the quality of a network.

A poor PSQM+ score may be the result of many different impairments, and it is valuable to know which factors influenced the score. Likewise, a good PSQM+ score may hide the fact that transient impairments exist. For example, a burst of packet loss can cause a noticeable dropout in the audio voice, and result in a PSQM spike. But because of its short duration, this dropout may not significantly affect the time-averaged PSQM + score. If given only an average PSQM+ score, one would not be aware of this noticeable impairment affecting the users' experience.

Front-end clipping due to Voice Activity Detectors also represents a short transient impairment that, while noticeable to end users, may not significantly influence an average PSQM+ score. Further analysis of the result provided by the VQT is valuable in assessing the complete picture of voice clarity.

Other Types of Statistical Analysis

Other statistical data are provided numerically in the VQT's spreadsheet to allow further analysis. In particular, these are the PSQM+ Maximum, Standard Deviation, and Outliers Percentage. These, in combination with the Average score, enable a more complete assessment of clarity.

As mentioned, an average PSQM+ score can hide a transient but noticeable impairment. The maximum PSQM+ score provides easy visibility to these impairments. For example, a burst of packet loss may cause a PSQM+ spike of 10, but be short enough to not affect the average score, which may come in at 1.0.

The Standard Deviation will reflect the consistency of PSQM+ scoring during a single test across a single voice sample. A single spike will not drastically affect standard deviation, but many spikes will.

Outliers Percentage is a powerful measurement result because it provides a more direct method than Standard Deviation for detecting PSQM+ spikes (or outliers) and for detecting transient impairments. To make good use of the Outliers statistic, set the Maximum PSQM Threshold (in the Clarity measurement's Advanced Configuration dialog box) to a value that is considered the high-end of acceptable quality (4.5 is a good nominal value). An outliers percentage more than 2 or 3 percent should alert you to possible transient impairments such as a burst of packet loss, in which the time duration of the impairment may not be great enough to cause the average PSQM+ score to be high.

As an example of the above discussion, consider the following results:

Average PSQM	1.34	Maximum PSQM	6.67
Outliers percentage	2.5	Std Deviation	0.9136

The average score is quite good. However, there was at least one spike at 6.67, and with the Maximum PSQM threshold set at 4.5, the Outliers Percentage was relatively high. This indicates that there were enough PSQM+ spikes to warrant investigation, and to listen to the recorded file to see if errors were noticeable. The Standard Deviation is relatively low and indicates overall quality was consistently good, despite transient impairments.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Estimated MOS Equivalent Scores

Another important statistic provided by the VQT is the Estimated MOS Equivalent Score. Because PSQM+ is designed to predict the results of subjective (e.g., MOS) tests, but is on a different scale than MOS, the VQT calculates this estimated MOS equivalent score to assist in interpreting PSQM+ scores. This calculation is based on a correlation curve of average PSQM+ scores to MOS scores, with the curve itself being based on results of subjective vs. objective testing. The estimated MOS score corresponds to the Listening Quality Score opinion scale for the Absolute Category Rating testing method specified in ITU Recommendation P.800.

It is important to note that the relationship between the MOS and PSQM values is not necessarily the same for different languages or even for different subjective tests within a language. P.861 itself specifies that "It is difficult to determine a unique function which transforms the PSQM value to the estimated MOS value. In practice, therefore, it is necessary to derive such transformation functions for individual languages and individual subjective tests in advance".

MOS scores are results of subjective testing, with all the dependencies on human nature and perceptions built-in. On the other hand, PSQM+ scores are results of objective testing, using a consistent and repeatable mathematical algorithm. While a PSQM+ score will yield a certain estimated MOS equivalent, this does not mean that any subjective test on a particular call will result in the same MOS score. The MOS equivalent is used to translate a PSQM+ score for those more familiar with the MOS 1-5 scale. It also enables more direct comparisons of PSQM+ scores with other voice clarity measurement results such as MOS and PAMS. PSQM+ and estimated MOS scores are most valuable when comparing tests using the same speech samples.

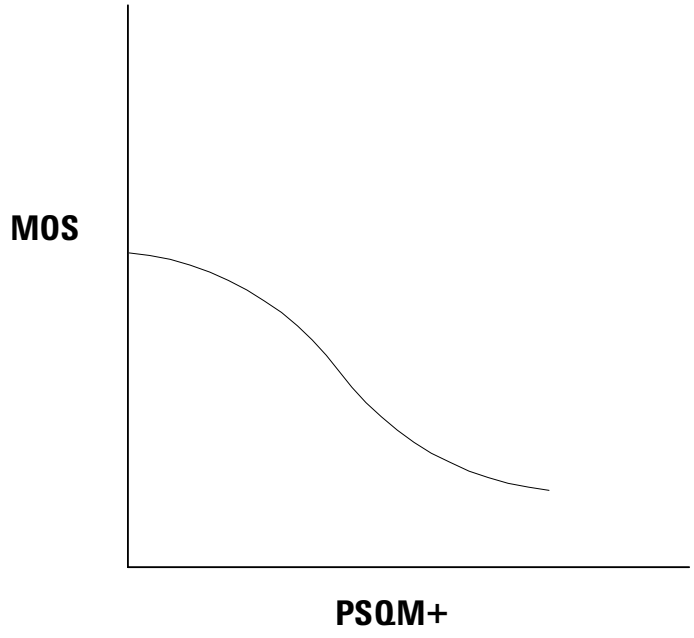


Figure 6. Sample PSQM / MOS Correlation Curve

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Interpreting the PSQM+ Graph

The VQT correlates individual PSQM+ calculations with the test and reference signal amplitude/power vs. time graphs. While this provides some useful analysis techniques, it is important to remember that PSQM+ is not necessarily a direct measurement of power deviations nor waveform deviations between the transmitted test signal and received test signal (PSQM+ measures perceptual distortion regardless of how it manifests itself in graphed representations of voice signals). Care must be taken when interpreting the graphs. For example:

- The transmitted and received test signals may appear to be identical in the amplitude/power vs. time graph, but high PSQM+ scores may still be observed. This might indicate that PSQM+ is measuring a distortion in a limited frequency band that is audible but not reflected in the graph, or may be due to a limitation in the time resolution of the amplitude graph itself.

- The graph might show amplitude/power distortion between the test and reference signals but show low PSQM+ scores. This might represent distortion that is not perceptible to a human ear.

  However, there are some impairments measured by PSQM+ that can be detected in the amplitude/power vs. time graph, making this graph a useful analysis tool. For example:

- **Dynamic Delay (Packet Jitter)**
  A variation in network delay during a PSQM+ measurement will cause some segments of the test signal not to be time-synchronized with the reference signal. The out-of-sync portions of the test signal are often visible in the graph and can be visually correlated with PSQM+ spikes. Delay variance can sometimes be observed at the beginning of the test signal, where it may be worse due the initial adaptation of jitter buffers. You can turn off the Clarity measurement's training cycles (in the Advanced Configuration dialog box) and measure PSQM+, looking for evidence of jitter at the beginning of the test. Regardless of where it occurs during the measurement, if delay variance is too great, the Clarity measurement could fail to correlate the test signals.. If correlation timeouts occur, or if you suspect delay variance due to PSQM+ spikes and/or visible out-of-sync conditions in the test signal graph, perform a proper Delay measurement. If delay correlation succeeds (and there is delay variance on the network), the Delay sawtooth graph will show a large spread. Please read the section on delay for more information.

  Figure 7 illustrates the results of a PSQM+ measurement showing evidence of packet jitter at the beginning of a speech sample. Large PSQM+ spikes above 10 can be observed, correlated with where the test signal can be observed as being out of synch with the reference signal. The remaining test signal is in synch, and PSQM+ scores are very low. Note the average PSQM+ score of 0.60, which hides the fact that a large impairment occurred. This is apparent in both the graph and the Maximum PSQM+ score.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

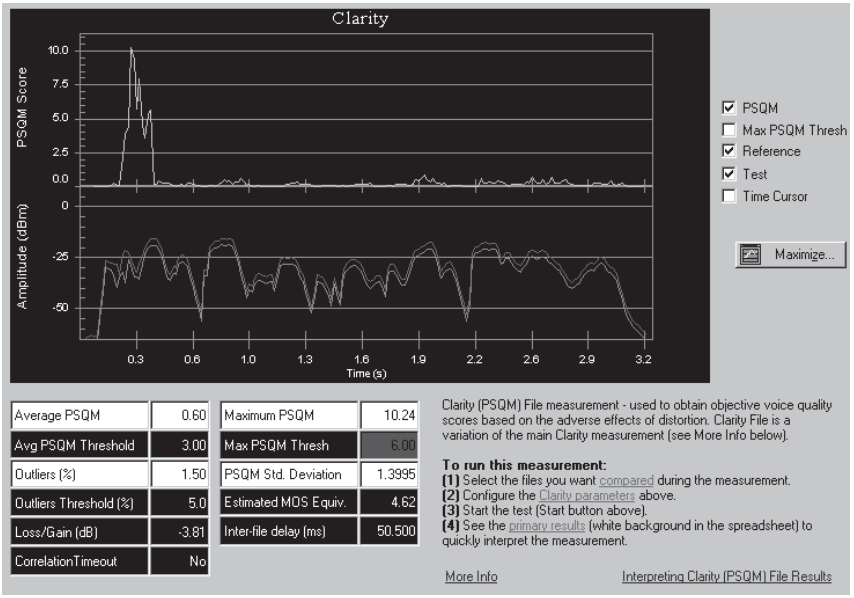


Figure 7: Delay Jitter in a Clarity (PSQM) Measurement

- **Front-end Clipping**

  If there is a large PSQM+ spike at the beginning of the transmission, look for front-end clipping in the waveform graph. If suspected, perform a VAD measurement to confirm clipping and quantify it.

  Figure 8 illustrates the results of a PSQM+ measurement showing about 40 milliseconds of front-end clipping. A large PSQM+ spike above 4 can be observed, correlated with where the test signal can be observed as being clipped when compared with the reference signal. Note the average PSQM+ score of 0.25, which hides the fact that an impairment occurred. Based on this observation, a VAD measurement should be performed.



Figure 8: Front End Clipping in Clarity (PSQM) Measurement

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

- **Dropouts and Loss**
  If there is a large, but narrow PSQM+ spike anywhere in the transmission, look for a narrow signal dropout in the test signal graph. This might be evidence of packet loss. It is possible that packet loss will not be shown in the test signal graph if loss correction mechanisms are used in the network. Such mechanisms, however, might not accurately reproduce the lost signal, and PSQM+ spikes can still occur.

  Figure 9 shows the results of a PSQM+ measurement during which two dropouts of about 40 milliseconds each occurred. Large PSQM+ spikes above 10 can be see to be correlated with dropouts in the test signal. Note the average PSQM+ score of 0.60, which hides the fact that an impairment occurred. If you suspect packet loss on your network, the Agilent Advisor provides powerful IP telephony test tools.

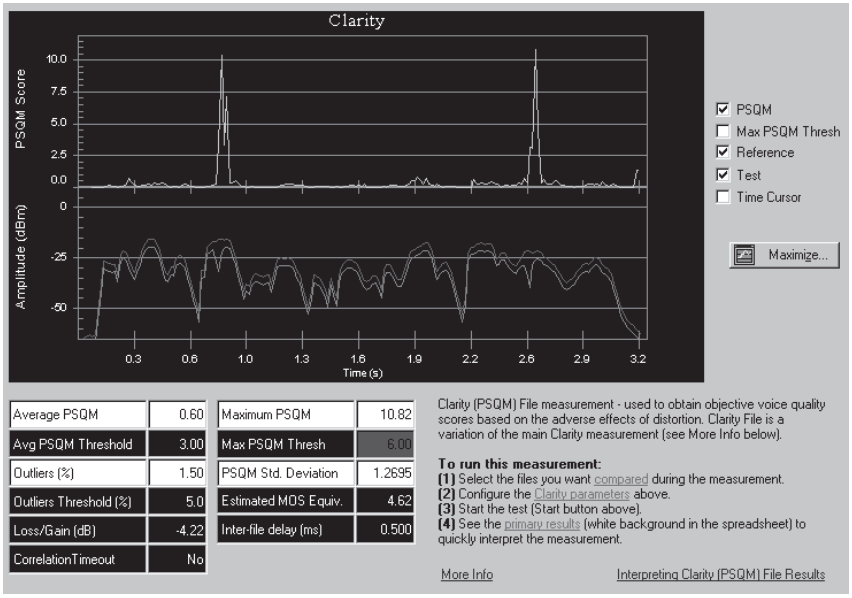


Figure 9. Drop-outs in Clarity (PSQM) Measurement

- **Coding Distortion**
  A consistently high PSQM+ score (as opposed to spikes) can indicate coding distortion that is consistent throughout the transmission. To further understand this problem, you can perform another PSQM+ measurement using G.711 coding, with all other conditions identical, and compare graphs. PSQM+ spikes can also reflect coding distortion. A short burst of loudness, for example, will experience greater coding distortion and will result in a PSQM+ spike. Look for increased volume in the reference signal correlated with a PSQM+ spike. It is useful to benchmark G.711 quantization distortion to provide a "ground-zero" PSQM+ trace. To do this, establish a connection between one T1 or E1 port and the other using just a T1/E1 crossover cable connected between ports. Perform a PSQM+ measurement. The results will be completely attributable to 8-bit G.711 encoding.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

- **Attenuation**
  PSQM+ compensates for overall system gain/attenuation by performing global scaling before comparing the input and output voice signals. Thus, for the most part, gain or attenuation imposed by the network under test will not affect PSQM scores. However, in the case where the VQT scales up an attenuated output signal (presumably attenuated by the network), any noise introduced by the network will be scaled upward as well, resulting in higher PSQM scores due to the amplified noise. This effect is typically present when output signals have been attenuated more than 10dB. This may be visible in the test signal graph, in which an attenuated test signal is shown with a high PSQM+ score. Severe attenuation may interfere with correlation, causing high PSQM+ scores due to mis-correlation, or correlation timeouts.

  Why is there still variance of PSQM+ scores across a "clean" transmission, such as a T1 cable loop or a TDM hairpin? Note the effects of large signals through m-law and μ-law PCM: greater quantization errors at higher levels. Also, across a network with vocoders, under severe conditions, non-linear vocoders can produce level clipping. These impairments will result in distortion, either added components or lost components in coding, and will be measured by PSQM+. Thus, a single voice sample played across a codec may produce PSQM+ scores that vary in time and that are dependent on utterances within the sample (e.g., an accented emphasis on a word may produce a higher PSQM+ score than non-emphasized words).

Other PSQM Test Techniques

Finally, there are a few other PSQM test techniques worth mentioning.

Using Different Languages

Voice codecs are known to exhibit different levels of performance (for example, produce different levels of clarity) when exposed to different languages. This fact is validated by the ITU in their recommendation for converting PSQM scores to MOS scores. ITU Recommendation P.861 for PSQM states: "Since the relationship between the MOS and PSQM values is not necessarily the same for different languages or even for different subjective tests within a language, it is difficult to determine a unique function which transforms the PSQM value to the estimated MOS value. In practice, therefore, it is necessary to derive such transformation functions for individual languages…"

To test the clarity performance of a network for different languages, the VQT provides over 150 speech samples in nine different languages. It is a good idea to perform clarity measurements using a variety of languages.

Measuring PSQM+ on Networks with Large Delays

Normally, the VQT cannot time-synchronize the received test signal with the transmitted test signal if the received test signal experiences network delay of more than two seconds. For networks that exhibit delays greater than 2 seconds, there are two techniques that can increase the delay the VQT will tolerate:

- Eliminating the training cycles that are configured in the Clarity measurement's Advanced Configuration dialog box will enable PSQM+ correlations to occur for delays up to about 3.6 seconds.

- Using an audio file editor to pad the trailing-end of the reference wave file with a few seconds of silence will enable PSQM+ correlations to occur up to about 5 seconds.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

The following sections describe how to evaluate clarity with the VQT's Clarity (PAMS) measurement.

How the VQT Measures Clarity with PAMS

The VQT uses the PAMS method to measure clarity much the same way it uses PSQM: it places a telephone call across the network, transmits and receives a test signal, and then inputs the sent and received test signals into the PAMS algorithm for analysis. Likewise, PAMS is similar to PSQM+ in that it compares the two test signals using a model based on human perceptual factors, and produces a clarity score that predicts the results of subjective testing. However, PAMS differs from PSQM+ in many ways. It uses different signal processing techniques prior to performing the comparison, uses different techniques for mapping perceptual audible errors to scores, and produces a pair of scores (in the 1 to 5 range – 5 being the best quality) which are a direct prediction of mean opinion scores (MOS). More specifically, the PAMS algorithm performs the following:

- The PAMS model accounts for network delay by performing time correlation. The test signal is synchronized with the reference signal in time segments, making PAMS less susceptible to correlation timeouts due to delay variance.
- PAMS also performs level alignment and frequency equalization to lessen the effects of attenuation and analog line filtering, respectively.
- Once these pre-processes have occurred, PAMS transforms both signals into the time-frequency domain and calculates the sensation surface of each signal. The sensation surface is a time-frequency representation of the perceived loudness of each signal. It is analogous to the spectral power density, but is based on how a human would perceive the signal in each time-frequency cell. PAMS then determines the audible difference in sensation surfaces of the test and reference signals, the result being the error surface. The error surface represents the audible errors, in the time-frequency domain, found in the test signal as compared to the reference signal.
- Finally, the audible errors are evaluated and mapped to predictive scores based on correlation with the results of subjective testing.

By performing time alignment, level alignment, and equalization up-front, PAMS attempts to remove the effects of delay, analog line attenuation, and analog line filtering. It, therefore, focuses on distortion introduced by signal processing and packet transmission.

Please refer to Agilent's *Methods for Measuring Perceptual Speech Quality* white paper for more detailed information on the PAMS algorithm.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Interpreting PAMS Scores

PAMS produces two scores, the Listening Quality Score (LQS) and the Listening Effort Score (LES). These scores, based on two different opinion scales, predict the results of subjective testing. These scales are standard in the industry and are defined in ITU recommendation P.800. These scales are shown below, along with the parameter (for example, "quality of speech" or "effort require") that subjects are asked to evaluate. Both scores lie between 1 and 5 and are usually quoted to two decimal places.

Quality of the speech
5 Excellent
4 Good
3 Fair
2 Poor
1 Bad

Listening quality scale

Effort required to understand the meaning of sentences
5 Complete relaxation possible; no effort required
4 Attention necessary; no appreciable effort required
3 Moderate effort required
2 Considerable effort required
1 No meaning understood with any feasible effort

Listening effort scale

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

LQS and LES scores give a measure of users' perception of quality. A PAMS score of 5 means that no distortion is measured. Because they relate to different aspects of subjectivity, LQS and LES scores are normally different if there is perceived distortion; LES is usually higher than LQS.

The same methods and conditions for using PSQM+ apply to using PAMS. PAMS is best used as a relative measurement, although it does provide the benefit of subjective qualifiers (for example, "excellent" for LQS=5 or "poor" for LQS=2) assigned to the LQS and LES scales. You should practice the same guidelines as described for PSQM+:

- Establish benchmarks by using a quality network standard; for example, the PSTN or a short TDM network.
- Establish benchmarks for impairments by performing PAMS measurements on individual elements, isolated in a controlled environment. Begin with an ideal network (for example, a TDM loop or hairpin) and add or change elements and configurations, performing the measurement each time.
- Perform PAMS measurements on different networks, different configurations of a network, the impacts of different elements within a network, and the impacts of different conditions (for example, traffic loads) on a network. Ensure all other conditions are identical. Compare scores on a relative basis.

  Use results of network baselining, or use established benchmarks, to set thresholds and detect impairments. Perform PAMS measurements repeatedly over time, comparing results to these benchmarks and thresholds, and to expose trends.

Interpreting PAMS Graphs

The Error Surface graph presented in the VQT's Clarity (PAMS) measurement provides a valuable tool for analyzing clarity and detecting impairments. The Error Surface represents the difference in the Sensation Surfaces of the test and reference signals, showing the audible differences introduced by the system under test.

To obtain the Error Surface, the PAMS algorithm subtracts the reference signal's sensation surface from the received test signal's sensation surface. Additive distortion (for example, noise) are color-coded as dark blue for small values and go up the color scale to aqua-blue for large values. Subtrac tive distortion (for example, parts of the signal that have been lost or were not encoded by a vocoder) have negative values and are color-coded as maroon for small values, and go down the color scale to yellow for large values. Zero distortion, or time-frequency segments in which no audible difference in test and reference signals were detected, are represented as black. The amplitude of distortion as shown by the color coding is related to how audible and annoying the PAMS algorithm predicts they will be.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

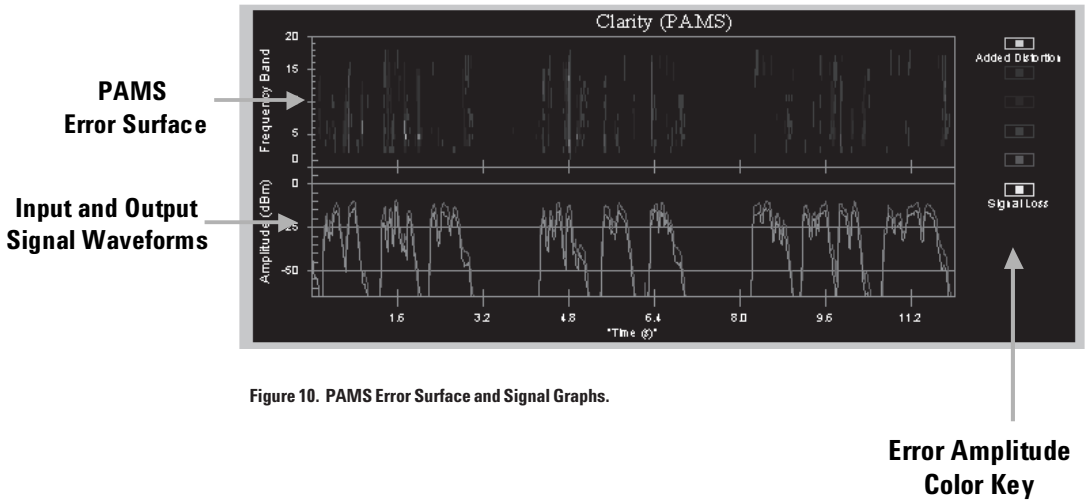


Figure 10. PAMS Error Surface and Signal Graphs.

The Error Surface graph is shown in two dimensions. Frequency is represented on the vertical axis and is scaled into 19 frequency bands with the labeling referring to numbered bands and not Hertz. Time is represented on the horizontal axis, and is correlated with the test and reference amplitude/power graphs.

To evaluate the amplitude of errors, first determine "ground-zero" by noticing the color-coding between speech utterances (that is, visually correlate the Error Surface graph with amplitude/power vs. time graphs). The color-coding of the Error Surface that corresponds to time-segments **between speech utterances** represents low distortion levels. It will be black if the call is clean and true silence is received. It will be dark blue if background noise or comfort noise is present, which represents small added distortion to the original reference signal. From this reference point, you can identify time-frequency segments within the Error Surface that are color-coded differently.

As is true with the VQT's Clarity (PSQM) graph, you can see evidence of various types of clarity impairments:

- **Front-end Clipping**
  Causes large amplitude but narrow time duration subtractive distortion at the start of speech bursts. Look for narrow vertical lines of red and yellow. Vertical lines indicate the loss occurred across all frequencies.

- **Packet Loss**
  Causes short bursts of loss, similar to front-end clipping, but can happen at anytime during the transmission. Look for narrow vertical lines of red and yellow. Figure 11 shows an example of packet loss.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

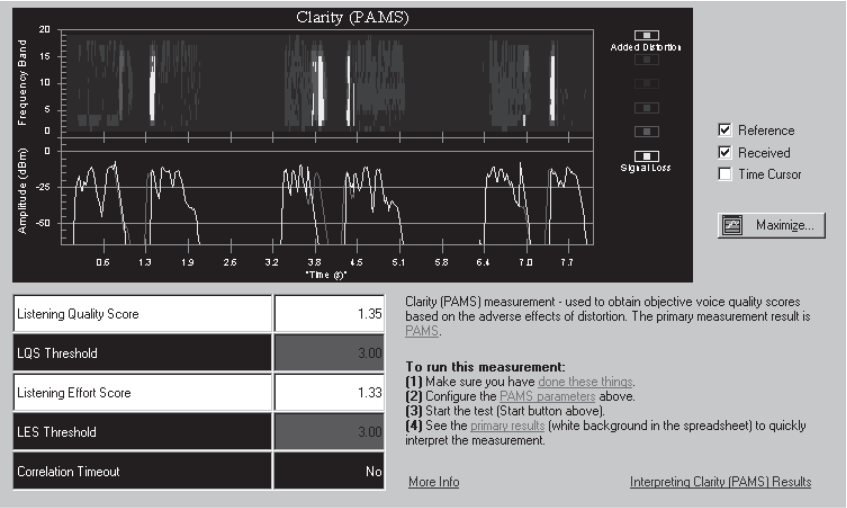


Figure 11. Clarity (PAMS) Measurement Showing Loss/Clipping.

- **Muting**
  Can be seen as prolonged subtractive distortion during speech, where the degraded sensation surface falls close to zero. Look for shadings of red and yellow across multiple frequency bands.
- **Addition of Background Noise**
  Can be seen as additive distortion, and is most apparent in silent periods. Look for shades of blue and aqua-blue between speech utterances.
- **Coding Distortion**
  Generally causes low-level distortion throughout speech bursts, and is very codec-dependent. Look for shades or dark blue and maroon during speech utterances.
- **Bit or Frame Errors**
  Causes localized distortion which may be additive or subtractive. This effect is dependent on the codec and any error concealment algorithm used in the network under test.
- **Dynamic Delay**
  Even though PAMS performs time-alignment within speech utterances to properly correlate the input and output test signals under conditions of dynamic delay, audible effects of delay variance are preserved and will be shown in the PAMS graph (representing how a human would perceive these effects). Therefore, audible distortion due to delay changes (as opposed to codec distortion) will appear in the PAMS graph as 'severe added distortion' or 'severe signal loss'.

Detecting Impairments and Optimizing Networks for Clarity

You can use the VQT's Clarity measurements to detect impairments in the network, and to optimize network performance for voice quality. By performing clarity measurements on individual network segments, or on a complete network, and changing individual elements, measurement results become relative benchmarks used to compare different designs, configurations, technologies, networks, and conditions, and how each impact clarity.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

The VQT's different interfaces can also be used to segment a network for measurements. For example, you can use the analog interface to test from an end-user POTS line to another end-user POTS line. Then you can use the T1 interface to test from the originating VoIP gateway to the terminating VoIP gateway for the same network. Differences between these two tests can be attributed to the IP telephony network from gateway to gateway. You can segment the network further by re-configuring the network itself. Refer to Figure 2 for some examples. Other examples include:

- Measure clarity end-to-end across a VoIP network using a G.711 codec. Then enable a G.729A codec and measure clarity under exact same conditions to determine the impact of G.729A compression.

- Measure clarity end-to-end across a VoIP network with a certain level of QoS implemented in the IP network. Adjust that level to result in greater or lesser voice packet loss, and determine the impact on clarity.

- Measure clarity end-to-end across a VoIP network using the same codec, but under different network conditions. For example, use the T1/E1 VQT interface to vary traffic loading from zero to full. Determine the impact of varying traffic loads on performance of a codec resources.

Remember that PSQM+ and PAMS scores are not additive; the scores obtained from individual segments cannot be added to predict the end-to-end score. Segments or elements that significantly impair clarity will have a much poorer clarity score than other elements, even though they might be a small part of the end-to-end score.

Once a network has been installed, configured, certified, and made operational, the VQT's Clarity Trending measurements can be used to monitor it and detect impairments based on established thresholds.

Using the Clarity Trend and Clarity File Measurements

The VQT also provides special versions of the Clarity measurement for both PSQM+ and PAMS. For example, it is important to perform multiple repetitions of clarity using the Clarity Trend measurement because it provides essential data to understand network performance variations over the course of an hour, a day, or more. You can use results of network baselining or established benchmarks to set thresholds in the Clarity measurement itself. Then, during Trend measurements, you can see evidence of network/voice impairments when those thresholds are exceeded.

The VQT's Clarity File measurement is a valuable tool for measuring clarity off-line. For example, you can test voice quality on non-realtime systems such as messaging and storage systems as described in the next section.

The Clarity File measurement is also valuable for performing "What If" Analysis. You can create a distorted voice audio file with actual network-induced impairments using File Play and Record; or you can create 'customized' distortion using an audio file editor (e.g., CoolEdit) by Syntrillium Software Corporation). You can then perform Clarity File measurements to determine how clarity is influenced by these induced distortions.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Testing Voicemail and Other Non-realtime Systems

You can measure the clarity of speech as it is reproduced by voicemail and other non-realtime systems. Leaving and retrieving a voicemail message using the VQT can be accomplished by using a Phone Line Splitter and an Analog Telephone to generate DTMF tones as follows:

Leave a message:

1. Connect the VQT and an Analog Telephone to a Phone Line Splitter as shown in Figure 12.
2. Connect the Phone Line Splitter to the FXO network.
3. Configure the Call Control tool to Assume Connection after a 10-second timeout using Call Control's Advanced Configuration dialog box.
4. Using Call Control, place a call to the voicemail system.
5. When the voicemail system answers, use the Analog Telephone's DTMF pad to navigate the voicemail system. You may not need to navigate if the voicemail system will eventually prompt you to leave a message.
6. Switch to File Play/Record and prepare to "Play" the voicemail message.
7. When prompted by the system, start playing the selected audio file.
8. When File Play/Record is finished, use Call Control to terminate the call.

Retrieve the message:

9. Configure Call Control to Assume Connection after a 10-second timeout using Call Control's Advanced Configuration dialog box.
10. Using Call Control, place a call to the voicemail system.
11. When the voicemail system answers, use the Analog Telephone's DTMF pad to navigate the voicemail system.
12. Switch to File Play/Record and prepare to "Record" the voicemail message.
13. Using the Analog Telephone, enter the final digit to play the voicemail message.
14. Simultaneously, start File Record to record the incoming message.
15. When File Record is finished, use Call Control to terminate the call.
16. Perform a Clarity File measurement on the recorded file.

Note: It is often possible to remove the Analog Telephone's handset to avoid unwanted background noise.

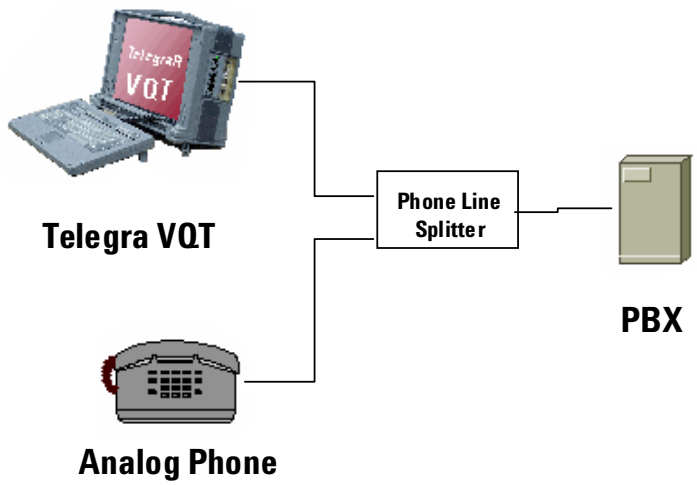


Figure 12. Setup for testing voicemail and other non-realtime systems with the VQT

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Other Clarity Test Strategies

The previous sections have discussed the core use of the VQT's Clarity measurements. The following sections discuss various other clarity test strategies that can enhance your overall test plans.

Traffic Loading

Using the T1 or E1 interface, you can measure clarity under varying traffic conditions. To do this, place calls on all channels of a T1 (or E1) port. Measure clarity across one or more channels (trending is recommended). Then add loading to a few channels at a time while keeping all calls in progress (this is important to ensure that the same network resources are being used throughout this test). Perform clarity measurements with each incremental increase in traffic load until all channels are loaded.

One-way and Round-trip Clarity

The VQT provides both one-way and round-trip Clarity measurements (refer to VQT System Deployment earlier in this document) for both the PSQM+ and PAMS algorithms. However, it is strongly recommended that you perform one-way Clarity measurements only for two reasons:

- First, the purpose of measuring clarity is to determine how end users experience the network service. An end user does not speak and then hear her voice returned to her to judge its quality. Rather, an end user hears what is spoken by the other party in a one-way audio path. For this reason, one-way clarity testing is really the only valuable method.

  It is important to measure one-way clarity in both directions, and to do so one right after the other to capture the same network condition. Any significant difference in scores can indicate an impairment or defect in a single element, such as the gateway that encodes one speaker and decodes the other.

- Second, it is very difficult to deduct any meaningful information from a round-trip clarity measurement. Clarity measurements are actually measurements of distortion, and unlike delay, distortion on most next generation networks is neither linear nor additive. For example, if delay in one direction is 70 milliseconds and delay in the other is 30 milliseconds, round-trip delay is 100 milliseconds. However, if PSQM+ in one direction is 2 and PSQM+ in the other is 3, you cannot assume that round-trip PSQM+ is 5. In a round-trip audio path, a voice signal will be encoded/decoded at least twice. This will have a non-additive effect on the final distortion measurement. While echo is not an issue when performing one-way clarity measurements, when performing round-trip clarity measurements, echo can play a significant role. Any echo on the call path will be recorded along with the returned test signal. This will be detected by the clarity algorithms as distortion and will impact the clarity result.

  You can perform one-way Clarity measurements with a single VQT server, but the end-points of the call on which this measurement is run need to be in close physical proximity. To address this concern, the VQT system enables you to perform one-way Clarity measurements between two VQT servers. In this way, you can measure clarity across a real network the same way that end users experience it.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Round-trip clarity measurements do have a value in network deployments which use enhanced service plaforms or application platforms on a circuit network spur to the VoIP network. In these situations, the voice signal will actually pass through an encoding/decoding and packetization process twice. This will have a detrimental effect on clarity that can be characterized using a single-leg VoIP network and the VQT with round-trip clarity measurements.

As shown in Figure 13, an Enhanced Service Platform (ESP), such as an Interactive Voice Response node, is a T1 spur off of the VoIP network. The voice signal from the telephone passes through an encoding/decoding process on Media Gateways twice before reaching its destination. Performing round-trip clarity measurements emulates this environment.

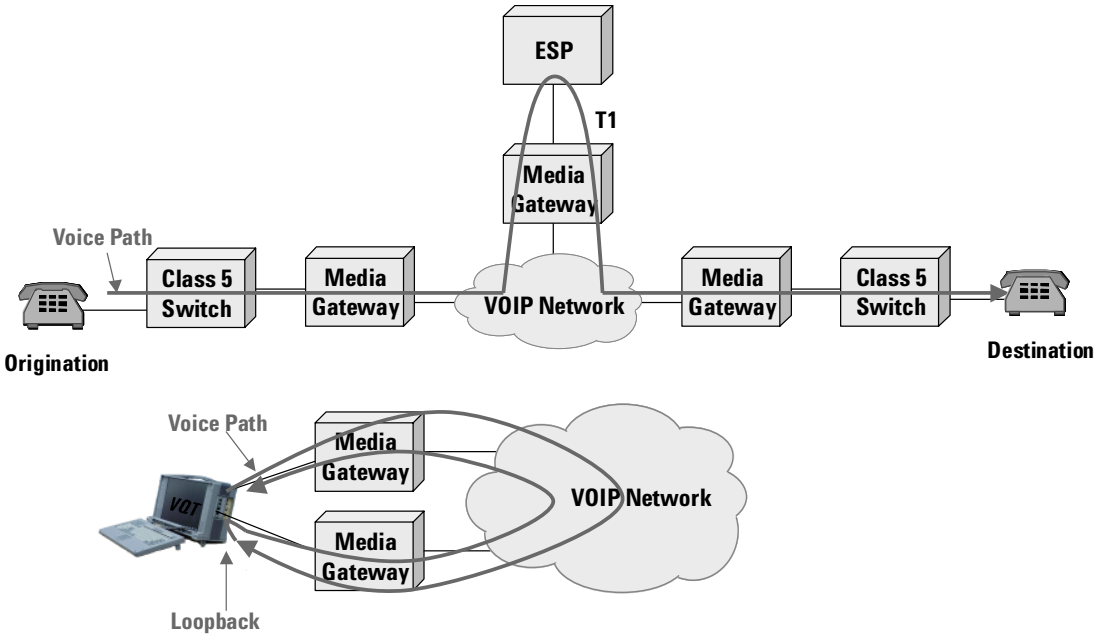


Figure 13. Round-trip Clarity Measurements.

Analog vs. Digital Interfaces

You might get different clarity results across the same call path using an analog VQT server vs. a T1 or E1 VQT server. This is due to the effects of encoding and quantization distortion for the different interfaces, as well as the effects of analog line transmission.

To prepare test signals to be transmitted onto analog interfaces, the VQT decodes clarity speech samples from a 16-bit encoded, 8kHz sampled file to an analog signal. To prepare test signals to be transmitted onto T1 or E1 lines, the VQT transcodes the voice file using G.711 8-bit non-linear encoding. These different encoding/decoding schemes will result in slightly different clarity results due to slight differences in quantization distortion.

To benchmark the impact of T1 encoding on clarity, place a call from Port 1 to Port 2 using a T1 crossover cable that directly connects Ports 1 and 2. This limits the focus of the test to distortion introduced by T1 encoding. Perform a clarity measurement, and use the results as a general benchmark of distortion introduced by G.711 encoding.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Use of Training Cycles

Training Cycles (set in the Clarity measurement's Advanced Configuration dialog box) are used to enable measurement of clarity under steady-state conditions; that is, they enable the Clarity measurement to operate on a network that is not adapting to varying traffic load or packet jitter. However, human conversation is bursty by nature; training cycles are not part of human conversation. It is effective to measure clarity with and without training cycles. Any differences will be noticeable at the very beginning of the transmission as a short but significant decrease in clarity.

Testing Voice Delay

Delay is the time it takes a voice signal to travel end-to-end between talker and listener and often manifests itself as an apparent time lag between when the talker speaks and when the listener responds. When perceptible delay is present, conversations can seem 'cold' or uncomfortable. When too much delay is present, voice conversations can become impossible.

With the emergence of voice-voice-over-IP (VoIP), delay has perhaps become more of an issue than with traditional PSTNs. But delay continues to be an important performance parameter both for individual VoP devices and integrated PSTN/VoP systems because of the various factors that contribute to it and because it can exacerbate other existing system conditions that adversely affect voice quality (such as echo).

The majority of delay on a PSTN is related to transmission distance and is usually quite low. A VoIP network will incur the same transmission-related delay, since it also must transmit over miles of copper or fiber. However, the VoIP network will also incur large processing delays. Delay in a VoP network is caused primarily by the time it takes codecs to digitize and packetize voice signals. Also, because packets can exhibit varying arrival times (jitter), compensating network processes called jitter buffers add even more delay. Routine router or switch processing add small amounts of delay as well. This means that it is nearly impossible to reach PSTN standards of delay with a VoIP network.

Network Delay Budget

How much delay is too much? Setting a maximum threshold for one-way voice delay is a subjective determination, and is based on the purpose of the network and the needs of the end users and network administrators. There are some valuable benchmarks, however, that can be followed:

- ITU Recommendation G.114 recommends a maximum of 150 milliseconds for one-way delay for most user applications.
- A PSTN call consisting of a 3000 mile call path typically exhibits delays of less than 50 milliseconds.
- In subjective testing, users typically do not perceive delays of less than 125 milliseconds. Annoyance is usually not experienced for delays less than 200 milliseconds. As a result, many VoIP carriers are reasonably striving to keep one-way delays under 125 milliseconds.

Remember, perceiving one-way delay refers to the perception of long pauses between speaker exchanges in a two-way conversation. Because a user is only on one end of a call, they do not have "visibility" (or "audibility") to the one-way delay experienced by their voice.

Setting a delay budget is useful for optimizing a network for competing objectives: delay, speech clarity, and bandwidth efficiency. Allocating delay values to different elements allows one to carefully weigh a reduction in delay with the added benefit of a resource such as bandwidth reduction.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

How the VQT Measures Delay

The VQT uses a powerful and unique technique for measuring delay. This technique, referred to as normalized signal cross-correlation, measures the impulse response of the system or circuit under test by transmitting a burst of pseudo-random, pattern-repeating noise referred to as Multiple Length Sequence (MLS). The MLS signal covers a broad spectrum of frequencies within the telephony band, and provides a highly predictable signal pattern to enable very accurate signal correlation. (Note: the impulse response function for a system may be estimated when the input is not an impulse and the system is non-LTI. In this case, sampling of the input and output signals is used for correlation (as opposed to using the entire input and output signal in a single function). Measuring delay by cross-correlating waveforms is not feasible in non-LTI systems such as many VoIP networks. Therefore, a cross-correlation of signal energy via a piecemeal impulse response measurement is performed.

To perform a normalized cross correlation, the VQT transmits the MLS signal from its source port or channel and, at the same time, records the received MLS signal on the destination port or channel. The VQT normalizes and correlates in time the test signal with the reference signal, and determines where in time the distribution of received signal energy correlates the best with the transmitted signal. The result of this correlation is a measured time-shift equal to the delay of the received test signal – the value is displayed in a spreadsheet. In addition to the reported delay value, the distribution of the signal energy that is detected and correlated is graphed with respect to the measured values of time-shift or delay. The VQT determines the highest concentration of energy within the distribution, and identifies the corresponding value of delay as the overall delay of the signal. Figure 14 illustrates the above process.

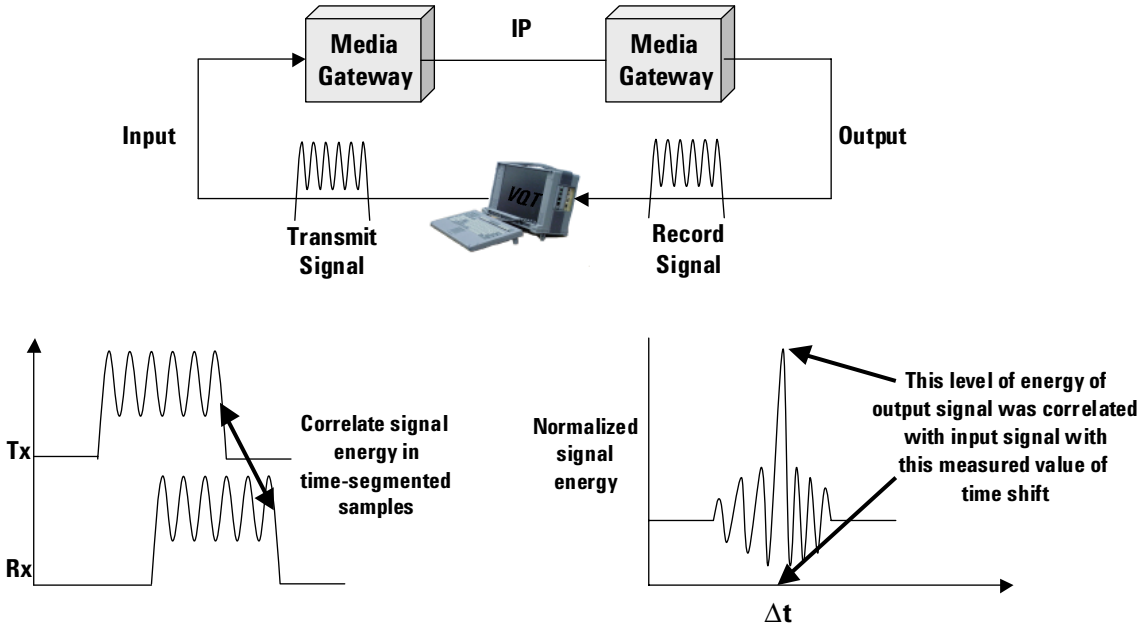


Figure 14. VQT Delay Measurement.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

This method offers several advantages over other methods that measure delay via an acoustic ping or tone:

- It provides a more robust method of time synchronization between the test and reference signals. Time synchronization relies heavily on signal detection, and signal detection within the VQT's Delay measurement is not dependent on detecting a single short tone or ping.

- It is less susceptible to front-end clipping, noise, attenuation, and loss, any of which can inadvertently cause the test device to have problems accurately detecting the returning test signal (false positive detection, failed detection, etc.). Acoustic ping or tone methods are more susceptible to these types of problems. For example, imagine trying to time-align signals based on the initial sample points under conditions of VAD front-end clipping. The front-end of the received signal may be missing due to VAD clipping, and a later sample point may be mistaken for the first sample point.

- It measures delay across all frequencies.

- It provides visibility into dynamic delay.

Interpreting the Delay Graph

The Delay measurement graph shows the results of the sent and received test signal cross-correlation as well as the measured value of time-shift (that is, the delay) for the distribution of correlated test signal energy. The horizontal axis measures time shifts or delays in milliseconds, and the vertical axis measures the amplitude of sent and received signal correlation – the point in time when the highest correlation amplitude is detected is reported as the delay value. The result shown in the graph is generally a sawtooth pattern with the spread of the pattern due to two phenomenon: a time-spread distribution (that is, a non-perfect transmission) of output signal energy that was correlated with the input signal, and possible delay variance (jitter) of the network. A network that exhibits no delay variance will still show a spread in the delay graph pattern, but will show a dominant spike with minimal spread. If the results show dominant spike(s) and large spread, this is likely due to delay jitter.

Figure 15 shows a graph of a narrow set of spikes, and with a single spike that is very narrow and tall. This indicates an excellent correlation of test/reference signals and no delay variance in the network.

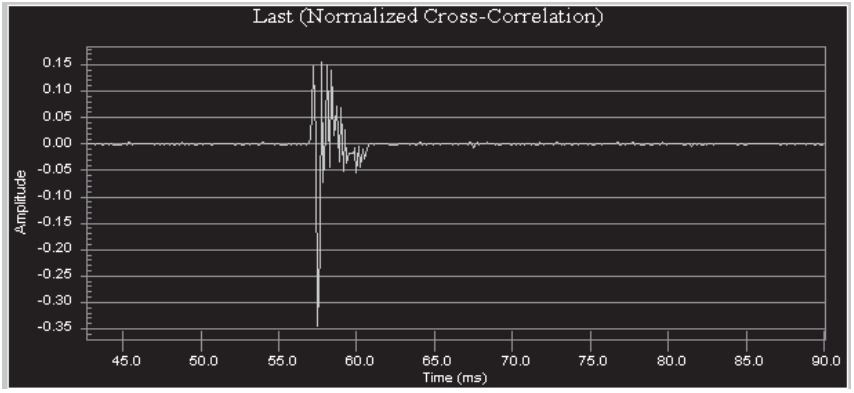


Figure 15. Delay graph with good correlation and good results, minimal spread

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Figure 16 shows a graph of a sawtooth pattern that is has more spread and with no dominant spike or with a relatively short dominant spike. This could be the result of good correlation between the test and reference signals but with delay variance occurring on the network during the Delay measurement transmission. This pattern could also indicate that the signal correlation was disrupted by severe test signal degradation.

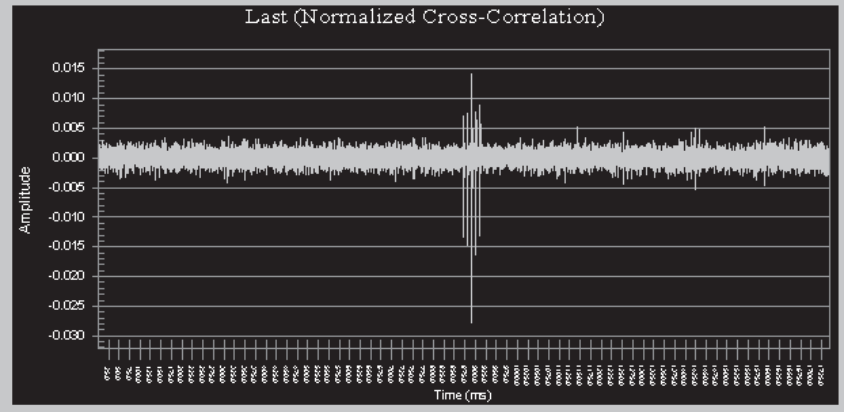


Figure 16. Delay graph with good correlation but showing some delay jitter.

Figure 17 shows a graph that indicates the received test signal has been corrupted with distortion such that it cannot be correlated with the reference signal. In this case, delay results cannot be trusted, and you should run the measurement again.

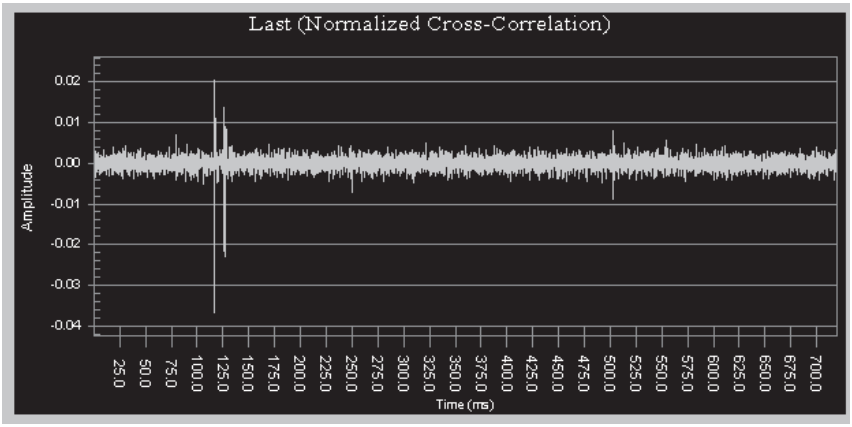


Figure 17. Delay graph with bad correlation results.

It is often a good idea to perform a delay measurement on a simple, deterministic circuit network to observe a reference graph (to which you can compare later) for the "shape" of a good sawtooth pattern.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Benchmarking Delay Measurements

Prior to executing a delay test plan, it is important to obtain delay benchmarks. By providing a reference point for subsequent delay measurements, benchmarks make it easier to detect network failures and trends. Two effective techniques for benchmarking are as follows:

- **Using the PSTN.** This is useful for testing overall network performance to ensure that PSTN quality can be met. Use the VQT to place a call on the PSTN with the same endpoints as the call you want to test on your network. Measure delay across this PSTN call to use as your total network benchmark.

- **Using Time-Domain Multiplexing (TDM) hairpins.** This is useful for testing the impact that individual elements have on delay. TDM hairpins are also very useful for setting a reference for the ideal "shape" of the delay graph. When interpreting the delay graphs of measurements made on actual networks, you can compare their spread to the spread of this reference graph. To perform TDM hairpins:

  (1) Use the VQT analog interface to place a call from Port A or C to Port B or D through a line simulator, router, switch, gateway, or any element that can switch the call directly back to the VQT without any processing. This is known as "TDM hairpinning".

  (2) Or use the VQT T1/E1 interface to place a call from Port 1 to Port 2 directly, using a T1 crossover cable. Remember to use a CAS signaling variant.

  (3) Or use the VQT T1/E1 interface to place a call from Port 1 to Port 2 using a line simulator, router, switch, gateway, or any element that can hairpin the call.

  Once you have setup the circuit under test as described above, measure delay to use as your initial benchmark. Then add processing and network elements to your call path, one by one, and measure delay on each one to determine the incremental impact. Unlike clarity, delay values are additive. Thus, you can breakdown an end-to-end delay of a complete network path to delays of individual elements.

  It can be very important to measure the delay of a complete network design to determine if the design meets standards. Possible standards that can be used are:

- ITU G.114 (150 ms)
- Other network benchmarks like the PSTN.
- Results of subjective testing to determine what is acceptable.
- Other carriers' objectives (e.g., 125 ms).

  Another effective technique is to place a call across a non-VoIP network (e.g. single switch) and measure delay. Then place a call across a VoIP network with all other factors constant. Measure delay and compare results to determine the delay added by VoIP processing.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

One-way and Round-trip Delay Measurements

Both one-way and round-trip delay measurements are useful when testing overall voice quality. One-way measurements characterize the time needed for a network to transmit a voice signal from a speaker's mouth to a listener's ear, and are most useful for measuring and optimizing the performance of a network. When network performance metrics are mentioned, it is the one-way delay value that is most often referenced. One-way delay measurements are also valuable in troubleshooting network delay problems. For example, delay in one signal path direction may be more than delay in another, if the terminating gateway's jitter buffer used in the first signal path is set to a greater value than the terminating gateway's jitter buffer in the other. A round-trip delay measurement will not uncover this phenomenon.

As implied above, round-trip delay measurements are equally valuable because they more accurately characterize a user's experience with regard to delay; that is, round trip delay is the primary cause of disruptive conversational pauses. For example, if you add round-trip network delay to the delay of the other speaker's verbal response, you can assess the "pauses" in conversational exchanges.

For local measurements, in which a single VQT provides both the source and destination ports of a call, one-way measurements are most meaningful and are very accurate.

For distributed measurements, in which a call is placed between two VQT servers, round-trip delay measurements are made. Dividing a round-trip delay measurement by two is a close approximation of one-way delay. This is because:

- The same processes (voice signal detection, encoding, packetization, re-assembly, decoding, etc.) are used in both directions.
- Different packets often will take the same router path except for only those brief periods under maximum congestion.
- Although different routes might be taken, routes vary on a packet-by-packet basis in addition to a directional basis. Hence, differences average out in both directions.
- The physical transmission distances are nearly the same in most cases. In few instances will these distances vary enough to have a significant impact on delay values in each direction.

Figure 18 illustrates a setup for measuring round-trip delay. Each VQT has one T1 link to a VoIP network (under test) and one T1 link to the PSTN. Simultaneous connections to the PSTN via the same endpoints allow the PSTN to provide a valuable benchmark.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

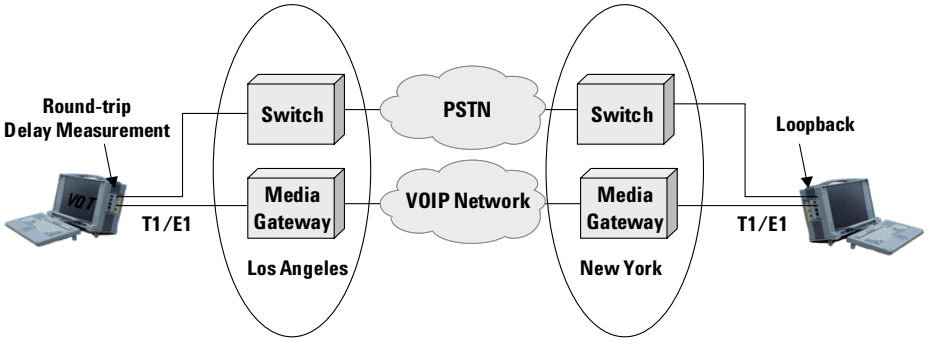


Figure 18. Setup for measuring round-trip delay using PSTN for benchmark.

It is important to note that for round-trip delay measurements, the physical interface at both the source and destination ports is important. If FXO is used for the source port, immediate echo will be present and will be mistaken for the round-trip test signal, resulting in measured round-trip delay values less than 3 milliseconds. To address this characteristic of analog networks, the VQT provides the Loopback Delay / Blind Window parameter in the Delay measurement's Advanced Configuration dialog box. The same problem exists if the destination port is FXO – the best solution in that case is to make sure it is an E&M or T1/E1 interface.

Other Delay Measurement Techniques

Regardless of the techniques discussed so far, there are other delay measurement techniques that can enhance your voice quality testing.

Measure Delay in Both Directions

Be sure to measure delay in both directions, under the same conditions, to detect any differences in delay in one direction. Perform the measurement in one direction immediately after performing it in the opposite direction in order to capture the same network conditions. Since certain delay-impacting processes (such as jitter buffers) are located on different network elements for various audio paths, delay may vary in different directions.

Delay Trending

It is also important to perform multiple repetitions of delay measurements. This provides essential data to understand network performance variations over the course of an hour, a day, or more, and under varying conditions such as traffic load. A common use of trending is to detect impairments in a network. Use results of network baselining, or use established benchmarks, to set thresholds. When these thresholds are exceeded during trending, this is often evidence of network impairments.

Use of Pre-average Cycles

Pre-average Cycles (set in the Delay measurement's Advanced Configuration dialog box) are used to cause dynamic network operations, such as jitter buffers, to reach a steady state prior to performing the actual delay measurement. One useful testing technique is to set the pre-average cycles to 1, then perform multiple repetitions of the delay test, and view the "summary" graph. This will show how delay can vary (typically will decrease) as the network adapts.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Detecting Impairments and Optimizing Network Performance

Many of the same techniques described in "Detecting Impairments and Optimizing Networks for Clarity" can also be practiced for delay. Refer to that section for these techniques. These techniques are useful for:

- Segmenting networks to isolate a delay value to individual segments and elements.
- Measuring end-to-end delays of networks when changing individual elements or configurations to determine the change's impact on delay.

Testing Echo

Echo in a telephony environment is usually caused by hybrid wire junctions in a network in which a portion of the energy from a person's voice signal is reflected back to the them on the same line. Echo diminishes the person's perception of call quality because it makes it more difficult to have a conversation. The two characteristics of echo that affect a person's experience are echo signal delay and echo signal strength. The greater an echo signal is delayed and/or the greater an echo signal is in strength, the more perceptible and annoying it becomes.

When using a two-wire analog access line, a person will receive an echo from the hybrid junction of their access at the local central office. This echo has minimal delay (less than 3 milliseconds) and, because it cannot be perceived, is not annoying. However, the speaker may also receive an echo from the hybrid junction of the tail circuit at the call destination, if this echo is not cancelled. This "far end" echo will potentially have a larger delay (greater than 25 milliseconds), thus becoming perceptible and possibly disruptive. Figure 19 illustrates these conditions. Typically, echo delays larger than 25 milliseconds become perceptible and annoying.

The PSTN deals with echo by deploying echo cancellers. Traditionally, the PSTN deployed echo suppressers, which sought to only attenuate the echo. The modern PSTN has deployed echo cancellers, which detect and subtract echo signals, then employ a non-linear processor to further attenuate echo.

There is a problem with PSTN echo canceller policies, however, in that, because they are expensive resources, echo cancellers are deployed only where needed. Typically, echo cancellers are deployed on long-distance circuit paths (for example, inter-machine trunks on physical transmission distances greater than 700 miles), because it is only on long call paths that echo can be delayed enough to be perceptible.

When connecting a VoIP network to a PSTN, the VoIP network introduces non-distance related delays to the call path. As a result, a call path may be too short (distance-wise) for the PSTN to have deployed echo cancellation, but the VoIP processing introduces enough processing delay to make echo delays large enough to be perceptible. Hence, echo has become a problem for VoIP networks. This is shown in Figure 19.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

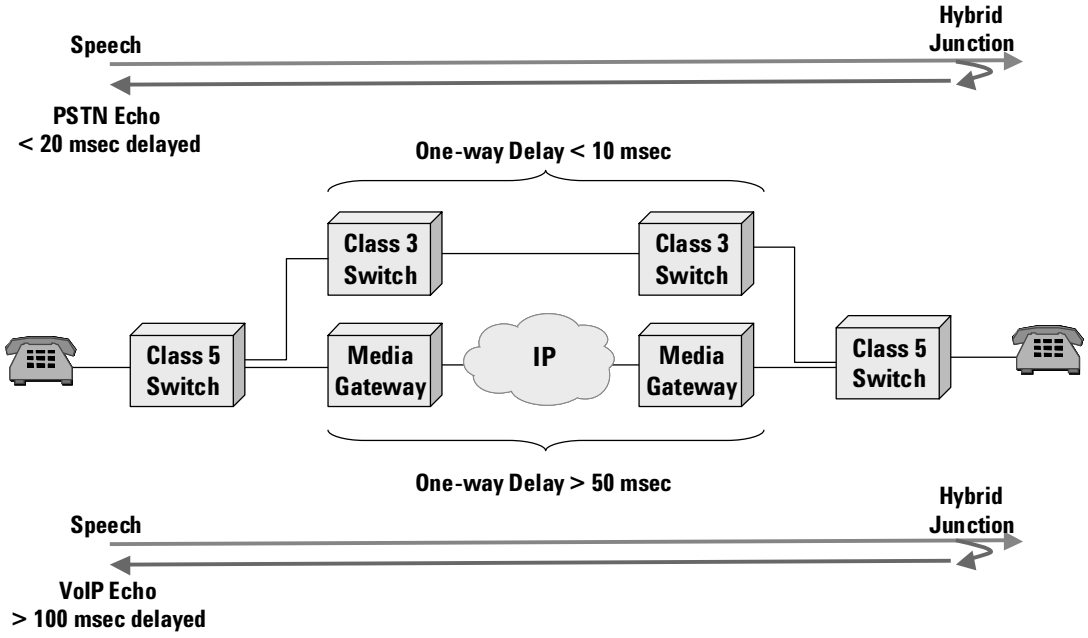


Figure 19. PSTN and VoIP networks, with same transmission distance and no echo cancellers.

The VQT provides powerful tools for not only detecting and measuring echo, but for determining its impact on call quality. These tools, called Echo-PACE and Echo-DTalk, enable you to:

- Detect any echo on a network and measure its impact on quality.
- Test the performance and operation of echo cancellers in the network.
- Determine where in the network echo cancellers are needed.
- Optimize the performance of echo cancellers.
- Assess an echo canceller's impact on 'interrupting' voice signals.

Please refer to the VQT's online Help system to read more about the causes of perceptible echo in a VoIP environment.

How the VQT Performs Echo-PACE Measurements

The VQT's primary echo measurement tool is Echo-PACE. To perform an Echo-PACE measurement, the VQT places a call over the network-under-test and transmits (from the source port or channel) a user-selectable signal of actual human voice. Simultaneous to this transmission, the VQT records any signal received on the *source* port (presumably, this received signal is returning echo). The VQT graphs the transmitted reference signal along with the received (echo) signal, and displays information about the received signal's delay and strength (again, presumably this is echo delay and echo strength).

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Along with this direct measure of the returning echo's characteristics, the VQT also performs a "Perceived Annoyance Caused by Echo" (PACE) measurement. Because conversational quality is impacted by echo, and because conversational quality is subjective in nature, measuring echo level and delay is not enough. The VQT uses a perceptual algorithm to represent the combined effects of echo delay and strength on perceived voice quality. The VQT does this by adding the received echo signal to a copy of the transmitted reference signal to create a 'combined signal', and compares the combined signal with the reference signal using the PSQM+ algorithm. Any "distortion" to the original reference signal that is created by adding the echo signal to it, is measured by PSQM+. Both the delay and the strength of the echo signal, when added to the reference signal, can distort the reference signal. No time synchronization is performed, so that the effects of echo delay are preserved and measured. The greater the echo delay and strength, the more distortion is added. PSQM+ measures this distortion, and is used to determine the impact of an echo signal on call quality.

Test Strategies for Echo

The following sub-sections provide information that will help you make the best use of the VQT's Echo measurements.

Choice of Interfaces Play Significant Role

Because of the causes of echo, and because it is usually only the far-end echo that is of interest, the interfaces to which you connect play an important role with regard to getting useful echo measurement results. Specifically, it is generally best to perform Echo measurements using the VQT's four-wire E&M, T1, or E1 interfaces as the source port. When using the two-wire FXO interface as the source port, a nearly immediate echo will be received that will dominate Echo measurement results. Although the VQT will detect and measure all echoes, this strong immediate echo will have virtually no delay and will, therefore be synchronized with the reference signal rather well (typical Echo measurement results for this case are shown in Figure 20). Thus, the immediate echo will have very little impact on perceived quality (and PSQM+ scores) and will obscure the evaluation of the far-end echo. However, the four-wire interface will experience no immediate echo and will provide a much cleaner measurement of the far end echo. Typical results for this case are shown in Figure 21.

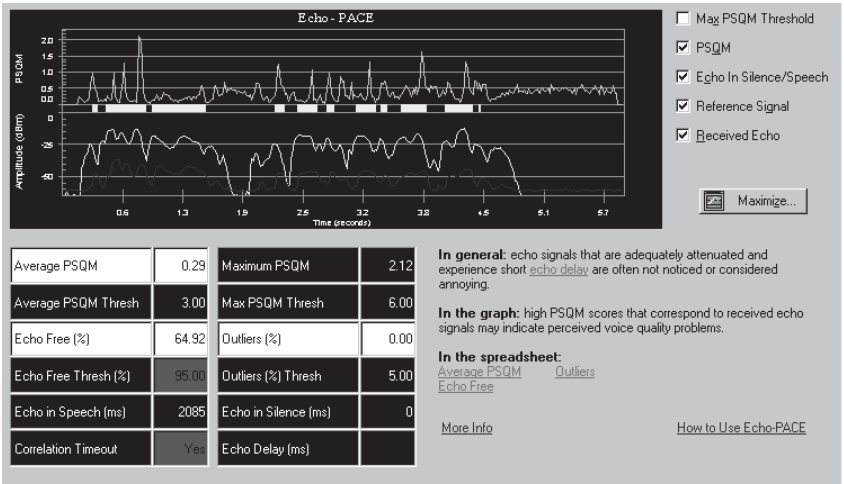


Figure 20. Echo PACE using FXO source port, with access circuit echo visible.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

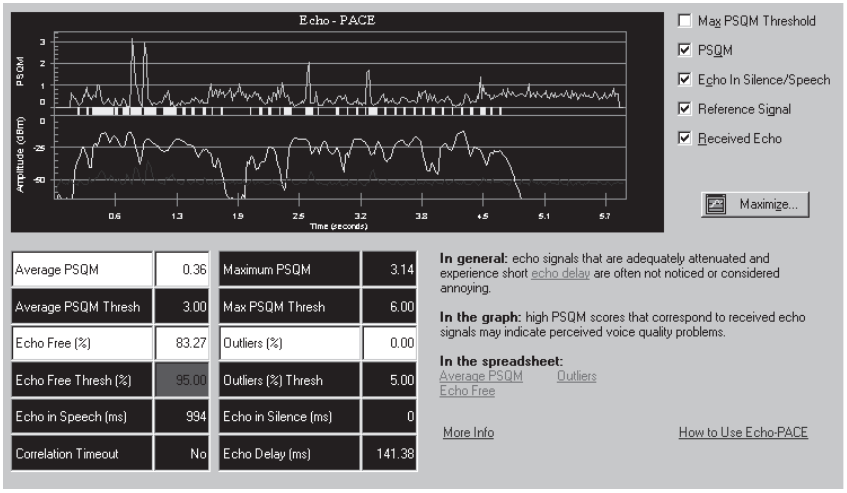


Figure 21. Echo PACE using E&M source port, with only tail circuit echo visible.

Effectively, there are four combinations of interfaces for performing Echo measurements (remember that FXO is always 2-wire, E&M can be two-wire or four-wire depending on the signaling, and T1/E1 is four-wire in the context of the VQT):

	Source Port	Destination Port
1	Two-wire	Two-wire
2	Two-wire	Four-wire
3	Four-wire	Two-wire
4	Four-wire	Four-wire

While each combination has its use, 3 and 4 in the table are most effective. These are used to measure the impact of far-end tail-circuit echo, which is of most interest since it is the most annoying.

When using a two-wire destination port, real tail-circuit echo will be generated on the network. This is useful for measuring the performance of echo cancellers and for determining if echo cancellers are needed at a particular point.

Note, as shown in Figure 22, that there can be multiple echoes on a network; for example:

- If using FXO, hybrid junction on access circuit
- If using FXO destination, hybrid on tail circuit
- VQT Network Simulator on destination port

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

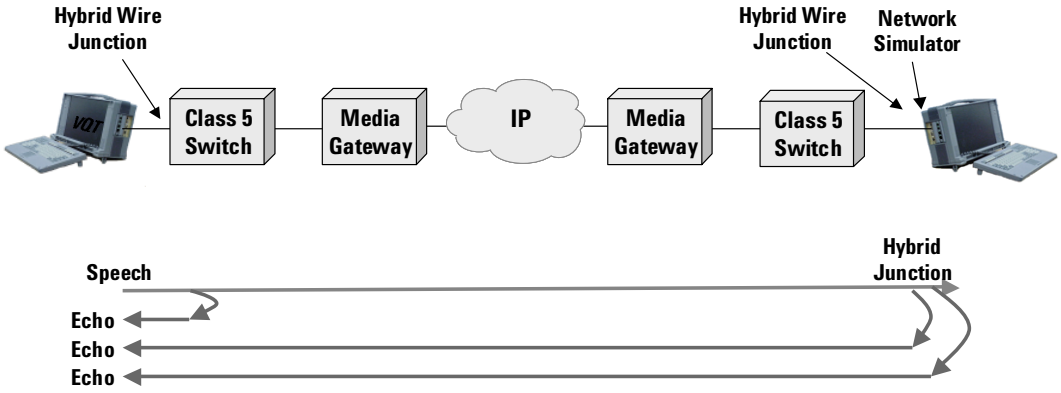


Figure 22. Network Echo From Multiple Sources.

Echo Measurement's Network Simulator

Within the Echo measurements themselves, the VQT provides a Network Simulator function that can be used to generate simulated echo from the VQT's destination port or channel. This is especially effective when using a four-wire destination port, since no hybrid junction exists that produces 'real' echo for testing. The VQT's echo simulation can also be used on two-wire destination ports, but in this case, the VQT's simulated echo will be mixed with any real echo generated by the network. Typically, the Network Simulator is used to generate the kind of echo that would originate from an actual tail circuit, thus simulating the tail circuit. The Network Simulator enables you to specify both return loss and delay parameters:

- Echo return loss should be set to meet minimum values of standard loss plans. ITU G.168 specifies tests using a minimum echo return loss of 6 dB, indicating that 6 dB is a typical worst case value for most networks. An echo return loss of 10 dB is a good average and is the VQT's default. When entering values for Return Loss, use a positive value to indicate attenuation.

- Delay can be varied to emulate a range of possible tail-circuit delays. Typical delays may range from 1 millisecond to 50 milliseconds, although the VQT allows much higher values to be entered. Echo cancellers typically have windows for detecting delays up to 64 milliseconds.

Benchmark Echo Measurements

When benchmarking echo, it is not enough to measure only echo strength and echo delay. You also need to establish a benchmark for echo's perceived annoyance. To establish these benchmarks, you use the Echo-PACE measurement.

Begin by placing a call with a telephone across a minimal network (for example, through just a line simulator) into the VQT, and disable any echo cancellation. Using the VQT's standalone Network Simulator tool, loop the audio signal back to the source telephone with some value of signal loss and delay, thus simulating tail-circuit echo. While listening with the handset of the telephone, vary the loss and delay parameters to determine the threshold for echo perception and annoyance (these thresholds may or may not be the same). Then place the same call using one VQT port as the source (the point from which the telephone placed the call) and with the other port at the same location as was used previously (the destination port). Configure the Network Simulator in the Echo-PACE measurement with the loss and delay values you obtained previously and run the Echo-PACE measurement.

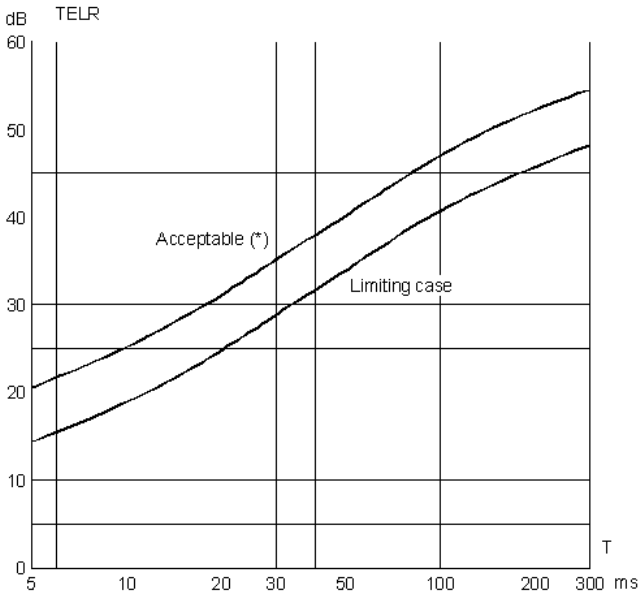
Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

The PSQM+ scores obtained in this case are representative of the threshold of perception and annoyance, and can be used as benchmarks. You can also perform the Echo-PACE measurement with zero tail-circuit echo (that is, with no loss and delay), and use the PSQM+ graphs and scores as benchmarks. The previous example shows how you can obtain loss, delay, and PSQM+ scores, based on your own perception, as benchmarks. You can also use PSQM+ scoring as benchmarks for actual existing network echo. Echo will be present on most telephone calls; it is a matter of delay and signal level that makes echo a problem, and PSQM+ applied as PACE is an effective measurement for determining this.

Determining Acceptable Echo

ITU Recommendation G.131 describes the relationship between the loudness of a speaker's echo and its delay, and how these two parameters combine to negatively impact conversational quality. An echo's loudness is measured as a parameter called Talker Echo Loudness Rating (TELR). TELR is the difference in dB in signal level between the speaker's voice signal and the received echo signal. A large value of TELR represents a large difference between the speaker's voice level and the received echo, and thus represents a relatively low echo. The smaller the TELR and the longer the delay, the more annoying the echo becomes. Thus, a large TELR and a small one-way transmission delay is desirable.

Figure 23 is based on a figure from ITU G.131 which shows an acceptable range of values of TELR versus one-way transmission delays.



TELR	Talker Echo Loudness Rating
T	Mean one-way transmission time
(*)	The "Acceptable" curve is equivalent to the curve with "1%" probability of encountering objectionable echo.

Figure 23. Acceptable TELR vs. one-way transmission delay.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

By determining a network's one-way transmission delay, you can determine what value of TELR is needed to fall within the acceptable range. To do this, first establish a call and perform a series of Delay measurements to determine the one-way delay across the circuit under test. Then refer to the above chart to determine what minimum value of TELR is needed; this value will be the TELR value that falls on the "acceptable" curve for the corrsponding value of delay that was measured. Then set the "Echo in Speech Detect Level" in the Echo-PACE measurement's Advanced Configuration dialog box to the minimum TELR value just determined, and perform the Echo-PACE measurement. The VQT will show if any received echo was detected during speech that caused the TELR to fall below the acceptable value.

Determining Presence and Location of Existing Network Echo

One of the most fundamental testing tasks you can perform with the Echo-PACE measurement is to determine the presence and location of existing network echo. Because you do not want to introduce echo into the circuit under test simply by virtue of the VQT's connection point, you need to use a four-wire E&M, T1, or E1 interface for the source and destination ports. In addition, do not enable the Echo-PACE measurement's Network Simulator.

Simply perform the Echo-PACE measurement across the target circuit and compare the results with benchmarks of acceptable levels of echo. PSQM+ results and the Echo in Speech / Echo in Silence values are the most effective in determining if the detected echo is a problem. You can set echo detection thresholds in the Advanced Configuration dialog box to values obtained during benchmarking and see if echo durations during speech and silence total more than an acceptable percentage (for example, > 3% duration). Remember that because the delay on a network can vary, so also can the "perceived annoyance caused by echo".

Perform the VQT's Echo-PACE measurement on a periodic basis and to different terminations, using automated testing, to determine if echo is present on a network. Refer to the "VQT Automated Testing Guide", which can be found as a PDF document on the VQT Web page: www.agilent.com/comms/onenetworks.

While it is difficult to determine the actual location of an echo source, the VQT can be used to determine a location relative to certain elements. You can determine, for example, if the echo originates at the destination port. To do this, use a four-wire E&M, T1, or E1 interface as the source port. Perform an Echo-PACE measurement and note the echo delay. Then perform a round-trip delay measurement on the same call. If the echo delay is about the same as the round-trip delay, the echo originates at the destination port.

Determine If and Where Echo Cancellers Are Needed

Echo cancellers are expensive resources, and while it may be practical in some networks to deploy echo cancellers on every circuit, it is often cost-effective to deploy them only where needed. But where they are needed can change as VoIP networks are deployed and interact with existing PSTNs. There are many calls in which echo is present and measurable but is not a problem because it is not perceivable. However, and as discussed earlier, if network delay is introduced into a call path due to VoIP processing, this echo becomes perceivable and perhaps disruptive. You can use the Echo-PACE measurement to not only detect and measure echo, but also to determine if it is perceivable (based on comparisons with benchmarks). PSQM+ scores and echo duration in speech/silence are effective metrics for this determination. It is often a good idea to perform these measurements both before and after VoIP elements are introduced to determine where on the existing network echo cancellers will need to be deployed.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Testing Echo Canceller Performance

Testing the performance of echo cancellers can take one of two forms: you can measure actual echo on the network across a known echo canceller or you can create echo and impose it on a known echo canceller.

To determine whether a given echo canceller is canceling actual sources of echo in the network, you will need to connect the VQT and place the call so that the destination port is connected to the tail circuit that can potentially produce the echo and that has an echo canceller provisioned to it. You will need to use a four-wire interface for the source port, and an FXO port for the destination port. Do not enable the Network Simulator. Perform the Echo-PACE measurement. Any measured echo will be the result of uncanceled tail-circuit echo from a hybrid junction on the network.

To test the performance of echo cancellation by inducing echo, place a call across the system under test using a four-wire E&M or T1/E1 interface for both the source and destination ports. This will focus the measurement on emulated far-end (destination) tail-circuit echo that you will generate using the Echo-PACE measurement's Network Simulator function. You can apply varying degrees of delay and loss to see how the echo canceller will react. You can also use a zero-echo measurement as a benchmark, then compare echo levels, delays, durations, and PSQM+ scores of measurements with emulated echo.

Other thresholds that can be set to further stress an echo canceller are in the Echo-PACE measurement's Advanced Configuration dialog box. There are different thresholds for echo detection in speech and echo detection in silence. Thresholds for detection in speech are specified in dB relative to speech. Thresholds for detection in silence are specified in absolute dBm. Both of these thresholds should be benchmarked, but using the default values of 20 dB and –30 dBm, respectively, is very effective.

Use Network Simulator to Dominate Immediate Echo on FXO Source

When using an FXO port on the VQT as the source port for an Echo-PACE measurement, an immediate echo from the near-side hybrid junction will be received. This echo will dominate the "received echo" measurement. Tail-circuit echo may also be received and measured, but it will lose visibility to the immediate echo. To deal with this problem so that you can see if the tail-circuit echo is passing through the network and to the speaker's ear, use the Network Simulator to generate tail-circuit echo on the destination port. By setting a strong return signal and delay, you can make sure the tail-circuit echo is visible. Then set the Return Loss to 0 and delay to at least 50 milliseconds. Then perform the measurement and notice the difference in results. The tail-circuit echo will begin to gain visibility in addition to the immediate echo. As the delay is increased, the tail-circuit echo becomes more apparent. You can also specify negative Return Loss values for a stronger signal.

Be careful in using this technique, as you may be creating tail-circuit echo signals that would never exist on a real network. Echo cancellers are set to converge on a tail-circuit echo based on expected parameters and they may not cancel out an echo signal that is too strong or delayed too much.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Testing Echo Cancellation During Doubletalk Conditions

Doubletalk refers to the condition in which both parties on a telephone call are speaking simultaneously. This poses a challenge for echo cancellers because they must distinguish between legitimate speech and unwanted echo on the same telephone line. Figure 24 shows a common condition in which the echo canceller facing tail circuit B must cancel the echo from Speaker A while transmitting the 'interrupting' signal from Speaker B unimpaired.

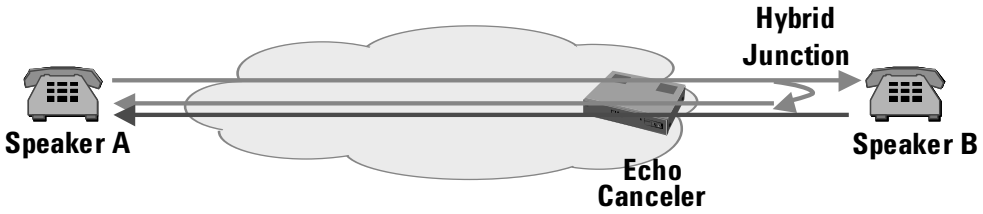


Figure 24: Doubletalk

How the VQT Performs DTalk Measurements

In order to measure how distorted the 'interrupting' voice signal is due to echo canceller performance, a single VQT is connected to both ends of a telephone call as shown in Figure 25. The VQT transmits a reference signal from port C to D (speaker A's voice) and, at the same time, transmits an 'interrupting' signal (speaker B's voice - also known as the DTalk signal) from port D to C. When the DTalk signal is received at port C, the VQT compares it to a 'clean' original using the PSQM+ algorithm. Any distortion due to uncanceled echo, or due to echo canceller performance problems imposed by the interrupting signal itself, will be reflected in the PSQM+ score.

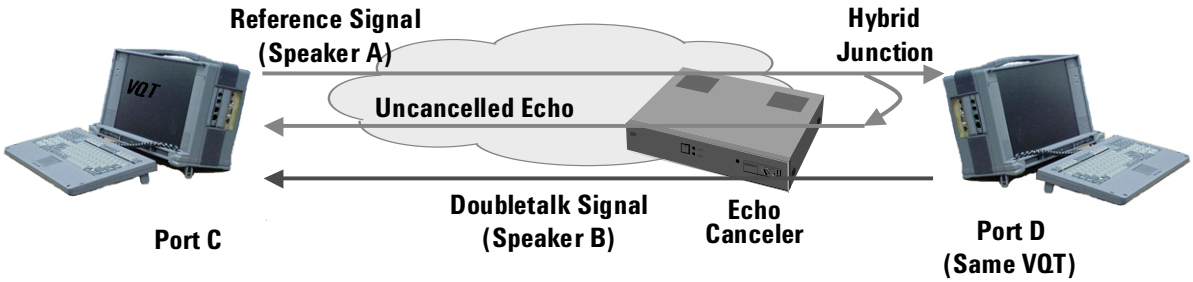


Figure 25: Echo Doubletalk Measurement

As in Echo-PACE, it is best to use the four-wire E&M, T1, or E1 interface as the source port for this measurement to avoid the interference caused by immediate echo.

Also as in Echo-PACE, the VQT provides a Network Simulator tool in the Echo-DTalk measurement to generate echo exhibiting varying degrees of return loss and delay, thus simulating various tail circuits. This is especially useful when using a four-wire destination port where no real network echo will likely be generated.

DTalk Testing Strategies

First, benchmark your network performance by running the Echo-DTalk measurement using a four-wire destination port and zero echo simulation (that is, do not use the Network Simulator). If possible, perform the measurement with echo cancellers disabled. The measurement results you obtain under these conditions (both numerical and graphical) provide a benchmark for later measurements.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Then, enable the echo cancellers on your network and repeat the measurement, again, without using the VQT's Network Simulator with zero echo. These measurement results should match your benchmark.

Finally, add echo with the Network Simulator. Echo cancellers in the network can be configured for optimal performance by running the Echo DTalk measurement several times, with the Network Simulator generating echo with various delay and return loss settings. Ensure PSQM+ scores, which reflect the performances of echo cancellers, are within acceptable range of your benchmark measurements. You can do this type of testing periodically to ensure acceptable performance.

Testing Silence Suppression

As the amount of voice signals that are packetized and transmitted across VoIP and other packet networks increases, the need to preserve network bandwidth also increases. Since, during a conversation, each speaker is usually silent about 50% of the time, it makes sense to prevent the packetization and transmission of silence. One process used to reduce unnecessary packet transmission (thus preserving bandwidth) is silence suppression performed by voice activity detectors or VADs.

What is Silence Suppression?

Silence suppression is enabled with Voice Activity Detectors (VAD) on the origination side of a VoIP network. Typically, VADs are implemented with voice codecs in a gateway or access device. For example, G.729A compression uses G.729B VAD.

One significant impairment introduce by VADs is front-end clipping. Front-end clipping (FEC) is the portion of speech that is clipped due to the time it takes a VAD to detect *and then react* (that is, *open* to let speech through) to a speech signal. During that time, the speaker's voice is not transmitted, and is perceived as front-end clipping to the listener.

Although not an impairment, hold-over time (HOT) is an important parameter of a VAD. Hold-over time is the amount of time it takes the VAD to close (that is, stop transmitting) once a speaker stops talking. VADs are designed to continue transmitting for a time after voice signal detection ends. This ensures that the VAD is not opening and closing too frequently, which can increase the occurrence of front-end clipping.

Along with silence suppression, many networks enable Comfort Noise Generation (CNG). The object of CNG is to match the natural background noise as much as possible. Natural background noise will be very audible to a listener between utterances of a speaker and during the hold-over time. If the CNG does not match the natural background noise well, when CNG starts, it will be very noticeable and annoying to the listener.

Some specifications, like G729B, provide a method for natural background noise to be sampled and input to a CNG on the terminating process end. In this way, CNG attempts to match the audible characteristics of the natural background noise.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Testing Silence Suppression with the VQT

The VQT's VAD measurement provides both spectrum vs. time and amplitude vs. time graphs to indicate the amount of front-end clipping and hold-over time imposed by a VAD on a particular voice circuit. The VAD measurement also provides a method for evaluating the spectral match between CNG and 'natural' background noise. Typically, you use the spectrum graph to determine if a VAD is enabled on the voice circuit under test and then the amplitude graph to determine the actual FEC and HOT values.

As with other measurements, it makes sense to first establish benchmarks for front-end clipping. You can determine acceptable levels for front-end clipping by using subjective testing. On a given phone call with a VAD active, determine by listening when front-end clipping becomes noticeable as VAD parameters are changed incrementally. When it becomes noticeable, measure its exact value with the VAD measurement and use this value as a benchmark. Typically, front-end clipping becomes noticeable at about 20 milliseconds.

Other uses for the VAD measurement include:

- Use the VAD measurement on an ongoing basis to detect front-end clipping impairments above your benchmark thresholds.
- Front-end clipping affects clarity. If front-end clipping is observed in the Clarity measurement's PSQM+ graph, use the VAD measurement to confirm and quantify it.
- Determine the maximum noise floor for the VAD under test. Vary the noise level to see how much of a noise floor will keep the VAD open. If a consistent noise floor is enough to keep the VAD open, silence suppression has been rendered ineffective.

Testing DTMF

Dual tone multi-frequency (DTMF) is a signaling system that uses pairs of sinusoidal tones to represent the digits on a telephone dial pad. DTMF tones must be transmitted and received within certain tolerances for both their amplitudes and frequencies. Voice-over-packet systems, particularly those using first generation gateways, can distort the amplitudes and frequencies of DTMF tones. The VQT provides valuable techniques for testing methods of DTMF transmission.

DTMF and VoIP

DTMF signals can be transported on a VoIP network in two ways: in-band (that is, on the same channel as voice signals) and via carriage techniques (also known as DTMF relay). When DTMF signals are transmitted in-band, they are encoded, packetized, reassembled, and decoded in much the same way as voice and other audio signals. When DTMF signals are transmitted via carriage techniques, the originating gateway or access device detects a DTMF signal on the circuit network, and instead of encoding the DTMF signal as voice, the gateway or access device relays a message to the terminating gateway or access device, indicating the DTMF signal received and its duration. The terminating gateway or access device then reproduces the DTMF signal.

It is important to note that when DTMF signals are transmitted in-band, they are subject to the same impairment conditions as voice, but often with more serious results. For example, low bit-rate vocoders are designed to encode/decode complex voice signals. This, however, may have an unpredictable effect on a pair of sinusoidal tones. In particular, the high frequency tone and the low frequency tone can be attenuated differently (as shown in Figure 26). This effect is known as DTMF twist, and is a common distortion incurred on a VoIP network.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

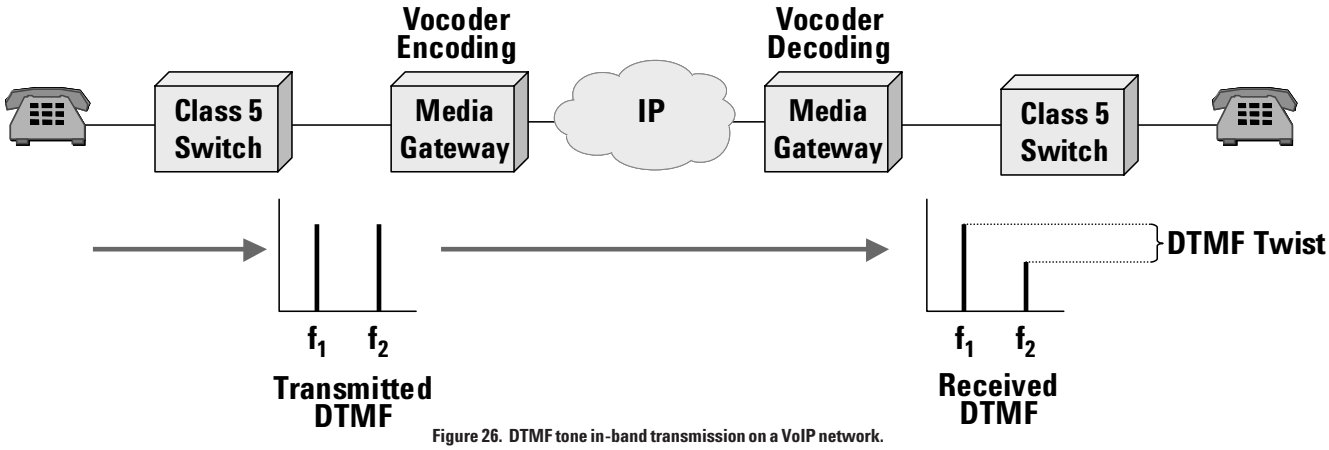


Figure 26. DTMF tone in-band transmission on a VoIP network.

The result of excessive DTMF twist is that a system designed to detect DTMF signals may not recognize them. The DTMF twist effect of a network can vary with each of the 16 DTMF signals, due to the different frequency spreads of each signal. For example, 3 is typically more problematic due to its wide frequency spread.

To test DTMF, the VQT transmits tone pairs representing the digits on a telephone's keypad, records the received signals, and performs the analysis to determine whether the tones or their respective levels have been altered by transmission through the network.

Testing DTMF with the VQT

Often, the first step is to establish benchmarks for DTMF measurement results using the PSTN or other circuit-switched network. Essentially, it is low-bit rate compression that will distort DTMF signals, so you can establish benchmarks using any G.711 or analog-only network. Be sure to test all 16 DTMF digits.

When you run the VQT's DTMF Tone measurement, the VQT presents the extent of twist and attenuation that occurred for each individual signal. You can perform this test for different codecs to determine its impact on DTMF quality.

As shown in Figure 27, for networks that employ DTMF carriage or relay, there are three tests you can perform:

- **Test the tolerance for DTMF signal recognition.**
  You can introduce twist in the VQT's source port transmission to test tolerances of DTMF detection in the gateway or access device. Begin with zero twist, then gradually increase the amount, until the gateway or access device can no longer recognize the signal. Once DTMF twist becomes so severe that the gateway can no longer recognize it, the DTMF tone will be carried in-band, and codec processing will further distort them.

- **Test the quality of DTMF signal reproduction.**
  Transmit DTMF signals with both zero and small amounts of twist, and determine if the received signal (which was reproduced by the terminating gateway or edge device) is free of twist.

- **Determine if DTMF relay is enabled.**
  Introduce small amounts of twist in the VQT's source port transmission. If the received signals are better (less twist) than the transmitted signal, this is an indication that DTMF relay is enabled.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

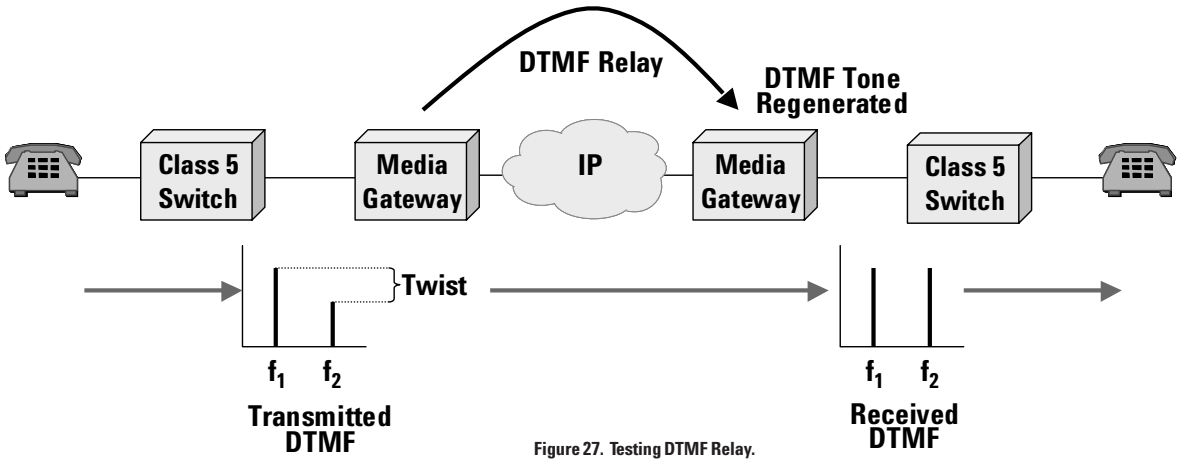


Figure 27. Testing DTMF Relay.

Other IP Telephony Testing Strategies with the VQT

More or less independent of all the testing techniques and strategies discussed so far, the following items can, nonetheless, enhance your voice quality testing activities.

Measuring Call Setup Time

Using the VQT Call Control tool, you can measure a network's call setup delay in seconds. What you will be measuring is the post-dial delay plus the amount of time needed to answer the call, which is usually less than one second. To measure call setup time, open the Call Control tool's Advanced Configuration dialog box, set Answer Mode to Auto Answer 100 milliseconds (minimum value) after ringing is detected. This minimizes the time added to line seizure. Place a call and allow the destination port to answer. Note the running clocks on Port A Duration and Port B duration. A short period of time after the call connection, hang-up the call to stop the clocks and note their values. The difference in clock times is the call setup time. The example in Figure 28 shows a call setup time of 4 seconds.

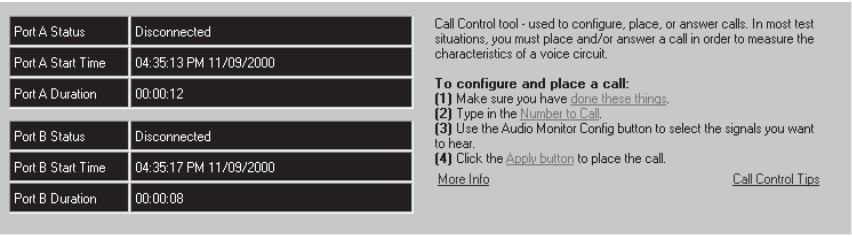


Figure 28: Call setup time shown in the Call Control tool

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Tasklists and Autotasker

As mentioned briefly in an earlier section, the VQT software provides customizable TaskLists that make it easy to navigate the user interface and lead you through common voice quality testing tasks. You can execute TaskLists using the VQT's AutoTasker to perform repetitive tests, mass calling, and other complex functions. You can also use external mechanisms applied to simple TaskLists. For example:

- In the first example, you can build a TaskList that places many repetitive calls, and to perform various tests. These TaskLists can get very long, but should pose no problems.

- In the second example, you can build simple TaskLists designed to perform simple functions, such as place a single call and perform measurements. Then you can use an external program to execute TaskLists in a certain sequence or according to a specific schedule, and to use certain dialed numbers. This method is more efficient and easier to perform.

  For more information on how to perform complex scheduled and automated testing using the Autotasker, refer to the "VQT Automated Testing Guide", which can be found as a PDF document on the VQT Web page at: www.agilent.com/comms/onenetworks.

Network Simulator

The VQT's Network Simulator (mentioned earlier in the context of echo testing) can be used to simulate an analog audio channel by applying gain/attenuation, delay, and limited frequency response behavior to a signal that is passed through the VQT. Once common use of this tool is to simulate actual tail-circuit networks.

To simulate a tail circuit, you can first perform an Impulse Response measurement on a linear time-invariant (LTI) network. (In a PSTN-VoIP hybrid network, the tail-circuit usually exhibits LTI behavior.) The VQT captures the response and saves it to an IR file which contains the coefficients of a polynomial characterizing the impulse response of the tail circuit. You can then use the Network Simulator to apply the IR file applies the IR file to a single-port loopback or a port-to-port loopback as shown in Figures 29 and 30.

Now, using one VQT, place a call to another VQT running Network Simulator. You can place a one-call-leg call and perform a single-port loopback, or you can place a two-call-leg call and perform a port-to-port loopback, as shown.

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

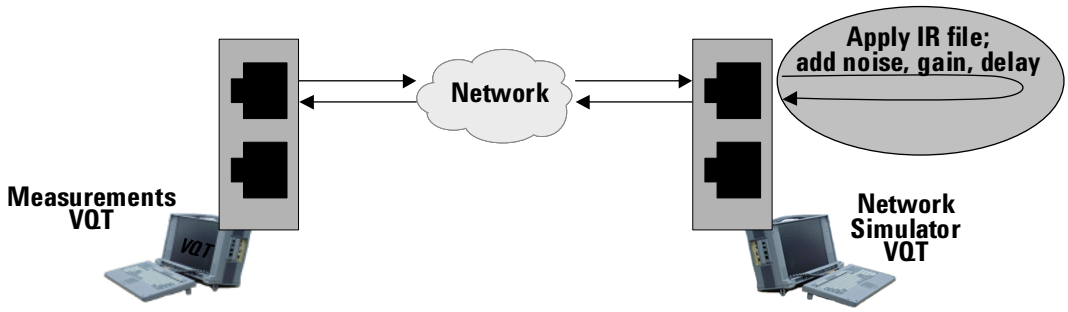


Figure 29. Using Network Simulator with Two VQTs and Single Port Loopback.

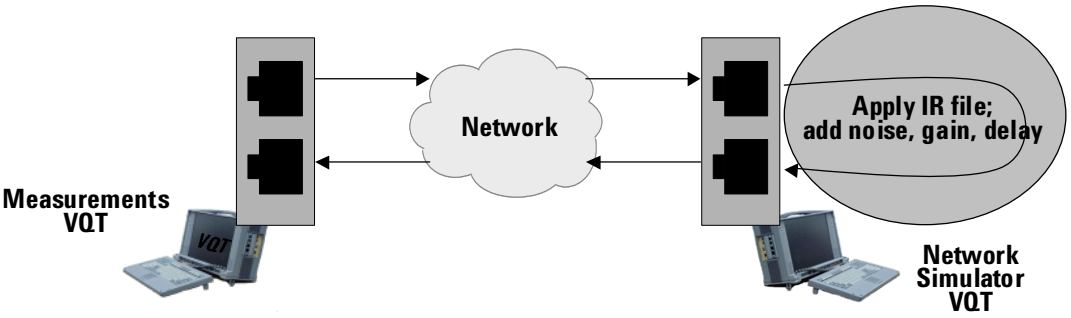


Figure30. Using Network Simulator with Two VQTs and Port-to-port Loopback..

When the Network Simulator is enabled and running, use the first VQT to perform measurements across the network. This enables you to characterize the entire end-to-end network including the behavior of the simulated tail-circuit. You can vary the simulated tail-circuit characteristics by applying gain, noise, and delay to the looped-back signal.

These techniques are effective for performing "what-if" analysis and for optimizing network designs in a lab or for real networks that will connect to various tail-circuit networks. The Network Simulator can also be used to simulate networks for other test devices.

Testing Networks with Multi-stage Dialing

Many IP Telephony networks utilize multi-stage dialing techniques in which a VoIP call is setup on a Media Gateway by dialing multiple numbers. For example, an access number to a Gateway network is first dialed, then a customer account authorization code is dialed, and finally a destination telephone number is dialed. The VQT can be used to test across multi-stage dialing network by setting up the call using one of two methods.

The first method involves using a telephone splitter and analog telephone to outpulse DTMF digits after the first stage of the call completes. This method is described previously in the section "Testing Voicemail and Other Non-realtime Systems".

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

The second method is described below:

Type in each stage and outpulsed DTMF string in a single line in the Call Control tool using the "?" character to separate stages. The "?" will introduce an indefinite pause in the outpulsing, causing the system to wait for input from the user before outpulsing the next dial string. The Audio Monitor should be configured to play "Receive" or "Transmit and Receive" on the selected channel, so the user can hear the interactive system's prompts and responses.

For example, in a 3-stage dialing network described above, the network access number is "800-555-1234", the authorization code is "999-0000", and the destination phone number is "222-555-1212". The user will enter the following string in the Call Control tool's Number to Call box for a selected channel or port:

8005551234?9990000?2255512l2

The VQT will outpulse "800551234", and then it will prompt the user to continue dialing. When the user hears the network's prompt for an authorization code, the user presses the Enter key. The VQT then outpulses "9990000", and again prompts the user to continue dialing based on the "?" character placement. The third and final stage is then executed.

Before placing the call, configure the Call Control tool's Advanced Configuration dialog box as follows:

Answer Mode: Manual Answer

Connection Control: Timeout after 10 seconds; assume connected

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Notes

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

www.agilent.com

Agilent Technologies' Test and Measurement Support, Services, and Assistance

Agilent Technologies aims to maximize the value you receive, while minimizing your risk and problems. We strive to ensure that you get the test and measurement capabilities you paid for and obtain the support you need. Our extensive support resources and services can help you choose the right Agilent products for your applications and apply them successfully. Every instrument and system we sell has a global warranty. Support is available for at least five years beyond the production life of the product. Two concepts underlie Agilent's overall support policy: "Our Promise" and "Your Advantage."

Our Promise

Our Promise means your Agilent test and measurement equipment will meet its advertised performance and functionality. When you are choosing new equipment, we will help you with product information, including realistic performance specifications and practical recommendations from experienced test engineers. When you use Agilent equipment, we can verify that it works properly, help with product operation, and provide basic measurement assistance for the use of specified capabilities, at no extra cost upon request. Many self-help tools are available.

Your Advantage

Your Advantage means that Agilent offers a wide range of additional expert test and measurement services, which you can purchase according to your unique technical and business needs. Solve problems efficiently and gain a competitive edge by contracting with us for calibration, extra-cost upgrades, out-of-warranty repairs, and on-site education and training, as well as design, system integration, project management, and other professional engineering services. Experienced Agilent engineers and technicians worldwide can help you maximize your productivity, optimize the return on investment of your Agilent instruments and systems, and obtain dependable measurement accuracy for the life of those products.

By internet, phone or fax, get assistance with all your Test and Measurement needs.

Online assistance:
http://www.agilent.com/find/assist

United States:
(Tel) 1 800 452 4844

Canada:
(Tel) 1 877 894 4414
(Fax) (905) 282 6495

China:
(Tel) 800-810-0189
(Fax) 1-0800-650-0121

Europe:
(Tel) (31 20) 547 2323
(Fax) (31 20) 547 2390

Japan:
(Tel) (81) 426 56 7832
(Fax) (81) 426 56 7840

Korea:
(Tel) (82-2) 2004-5004
(Fax) (82-2) 2004-5115

Latin America:
(Tel) (305) 269 7500
(Fax) (305) 269 7599

Taiwan:
(Tel) 080-004-7866
(Fax) (886-2) 2545-6723

Other Asia Pacific Countries:
(Tel) (65) 375-8100
(Fax) (65) 836-0252

Product specifications and descriptions in this document subject to change without notice.

©Agilent Technologies, Inc. 2000-2001

Printed in U.S.A. October 8, 2001

5988-3036EN

Use this link to go directly to our network troubleshooting solutions:
http://www.agilent.com/comms/onenetworks

Agilent Technologies

Artisan Technology Group - Quality Instrumentation ... Guaranteed | (888) 88-SOURCE | www.artisantg.com

Artisan Technology Group is an independent supplier of quality pre-owned equipment

Gold-standard solutions

Extend the life of your critical industrial, commercial, and military systems with our superior service and support.

We buy equipment

Planning to upgrade your current equipment? Have surplus equipment taking up shelf space? We'll give it a new home.

Learn more!

Visit us at **artisantg.com** for more info on price quotes, drivers, technical specifications, manuals, and documentation.

Artisan Scientific Corporation dba Artisan Technology Group is not an affiliate, representative, or authorized distributor for any manufacturer listed herein.

We're here to make your life easier. How can we help you today?

(217) 352-9330 | sales@artisantg.com | artisantg.com

ARTISAN® TECHNOLOGY GROUP