Characterizing Voice Echo Response for Voice over IP Networks

This white paper defines echo and describes where it occurs in a voice network. It examines the basic aspects of echo analysis and describes the effects of various network elements on echo. This document also explains how echo is measured and provides assistance to the reader with methods to determine its impact on Quality of Service. Finally, it looks at customer and service provider expectations about echo, and explains how the Voice Echo Response (VER) measurement is a valuable tool in characterizing echo, especially under heavy call loading conditions.

askzeke@ameritec.com

Characterizing Voice Echo Response For Voice over IP Networks
Using GoldenVoice technology to determine the presence of echo and its impact on Quality of Service (QoS)

Contents

I. Introduction

Because of the fundamental delays associated with VoIP technologies, existing echos will be more annoying than with TDM, and even the normal operation of an echo canceler will be more apparent.

Echo problems are relatively rare in the PSTN, which has short delays; they are much more common over cellular and satellite long distance calls. Interestingly, they are also much more readily tolerated in cellular and long distance calls because customers have been educated to have lower expectations for such calls.

As long as VoIP calls continue to be terminated in analog tails, echo will be a problem. One major obstacle to widespread VoIP implementation is that many tail circuits have preexisting delays that will become noticeable only when service providers introduce digital segments to the networks.

These problems will gradually be solved as digital networks extend out toward homes and telephone endpoints. Until then, how much echo can be expected? One call in 50? 100? 1000? Even if customers are trained to complain only when an echo problem is persistent and repeatable, a service provider cannot hunt down and destroy every echo complaint. No one has sufficient resources to do this task, and hunting down an echo is a necessarily intrusive process

The challenge is to determine when an echo complaint is both solvable and worth solving. You know that the echo source is in the destination tail circuit. For an echo problem to be solved, the tail circuit needs to be accessible.

The goal of service providers in eliminating echos, therefore, is to identify clusters of echo complaints, look for common links, and then fix the echos. A PSTN has many *dirty tails*, and it is unrealistic to expect that every echo can be eliminated. The best that you can do is to make sure that your own network and tails are clean, which requires care in installation and provisioning, especially when connecting gateways to analog equipment.

II. Echo Analysis Basics

In a voice telephone call, an echo occurs when you hear your own voice repeated. An echo is the audible leak-through of your own voice into your own receive (return) path. Every voice conversation has at least two participants. From the perspective of each participant, there are two voice paths in every call:

- Transmit path—The transmit path is also called the send or Tx path. In a conversation, the transmit path is created when a person speaks. The sound is transmitted from the mouth of the speaker to the ear of the listener.
- Receive path—The receive path is also called the return or Rx path. In a conversation, the receive path is created when a person hears the conversation. The sound is received by the ear of the listener from the mouth of the speaker.

Figure 1 shows a diagram of a simple voice call between Bob and Alice. From Bob's perspective, the transmit path carries his voice to Alice's ear, and the receive path carries Alice's voice to his ear. Naturally, from Alice's side these paths have the opposite naming convention: the transmit path carries her voice to Bob's ear, and the receive path carries Bob's voice to her ear.

Figure 1 Simple Telephone Call

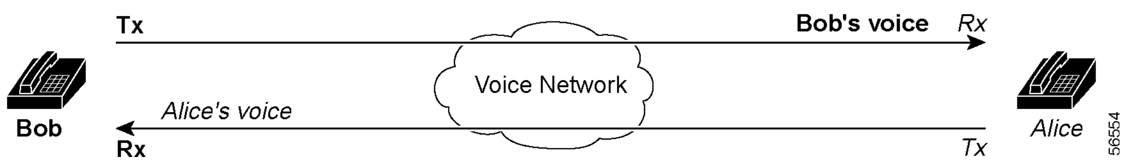
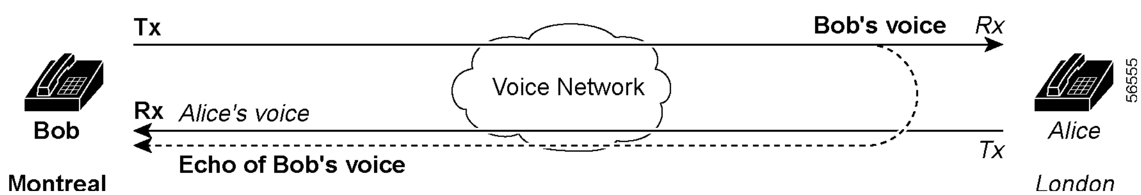


Figure 2 shows the same simple telephone call where Bob hears an echo.

Figure 2 Simple Telephone Call with an Echo



Bob hears a delayed and somewhat attenuated version of his own voice in the earpiece of his handset.

Initially, the source and mechanism of the leak are undefined.

One significant factor in echo analysis is the round-trip delay of the voice network. The round-trip delay of the network is the length of time required for an utterance to go from Bob's mouth, across the network on the transmit path to the source of the leak, and then back across the network again on the receive path to Bob's ear.

Two basic characteristics of echo are as follows:

- The louder the echo (echo amplitude), the more annoying it is.
- The longer the round-trip delay (the “later” the echo), the more annoying it is.

Table 1 shows how delay can affect the quality of a voice conversation.

Table 1 Effect of Delay on Voice Quality

One-Way Delay Range (ms)	Effect on Voice Quality
0–25	This is the expected range for national calls. There are no difficulties during conversation.
25–150	This is the expected range for international calls using a terrestrial transport link and IP telephony, which includes only one instance of IP voice. This range is acceptable for most users, assuming the use of echo control devices.
150–400	This is the expected range for a satellite link. Delays in this range can interrupt the normal flow of a conversation. A high-performance echo canceler must be used and careful network planning is necessary.
Greater than 400	*This is excessive delay and must be avoided by network planning.*

III. Locating an Echo

In Figure 2, Bob experiences the echo problem, which means that a signal is leaking from his transmit path into his receive path. The fact that Bob hears an echo illustrates one of the basic characteristics of echo: Perceived echo most likely indicates a problem at the other end of the call. The problem that is producing the echo that Bob hears, the leakage source, is somewhere on Alice's side of the network (London). If Alice was experiencing echo, the problem would be on Bob's side (Montreal).

The perceived echo leak is most likely in the terminating side of the network for the following reasons:

- Leak-through happens only in analog circuits. Voice traffic in the digital portions of the network does not leak from one path into another.

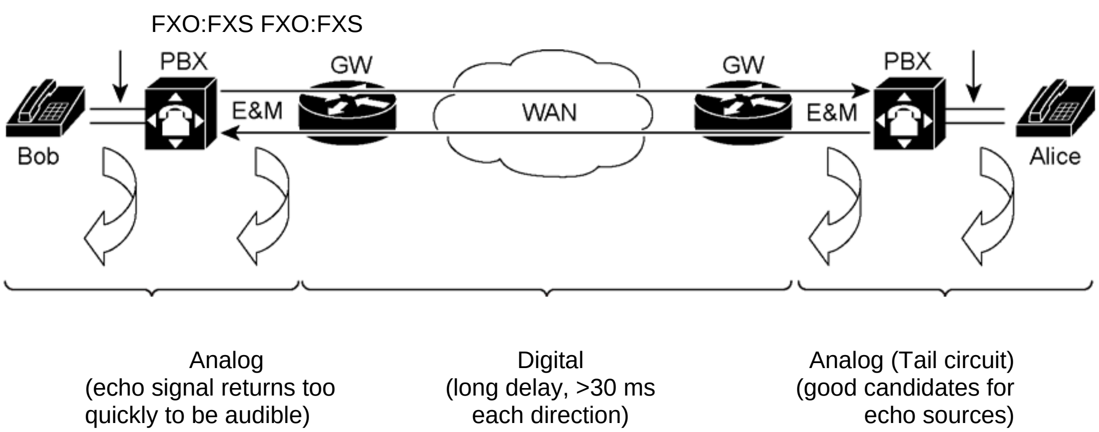
  Analog signals can leak from one path to another—electrically from one wire to another (called crosstalk), or acoustically through the air from a loudspeaker to a microphone. Also, analog

signals can be reflected in the hybrid transformer in the tail circuit. (See the section "Effect of Hybrid Transformers on Echo.") When these analog signals have been converted to digital bits, they do not leak.
All digital bits are represented by analog signals at the physical layer and these analog signals are subject to leakage. The analog signals that represent bits can tolerate a good deal of distortion before they become too distorted to be properly decoded. If such distortion occurred in the physical layer of the PSTN, the problem would not be echo. If you had connectivity at all, you would hear digital noise instead of a voice echo.

- This point is a corollary to the previous assertion that echos become increasingly annoying with increasing mouth-to-ear delay. A certain minimum delay is needed for an echo to become perceptible. In almost every telephone device, some of the Tx signal is fed back into the earpiece so that you can hear yourself speaking. This feedback is known as sidetone. The delay between the actual mouth signal and the sidetone signal is negligible, and sidetone is not perceived as an echo.
Also, your skull resonates during speech (an acoustic sidetone source) and the human auditory system has a certain integration period that determines the minimum time difference between events that will be perceived as separate events rather than a single one. Together, these phenomena create a minimum mouth-to-ear delay of about 25 ms before an echo signal can be perceived.

Given these two reasons—that echos must be delayed by at least 25 ms to be audible, and that leaks occur only in the analog portion of the network—you can deduce much about the location of the echo source. Figure 3 shows possible sources of echo in a simple Voice over IP (VoIP) network.

Figure 3 Potential Echo Paths in a Network with Both Analog and Digital Segments



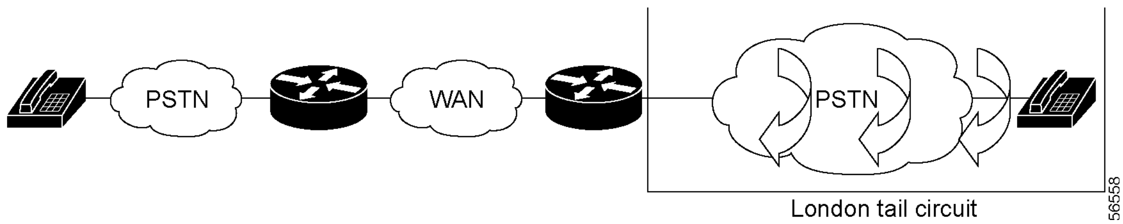
In this typical VoIP network, the digital packet portion of the network is located between two analog transmission segments. Bob in Montreal is connected by a 2-wire analog foreign exchange station (FXS) circuit to a local PBX, which is connected to a local VoIP gateway by a 4-wire analog recEive and transMit (E&M) circuit. The Montreal gateway communicates with the London gateway through an IP network. As we will discuss later, this packet transmission segment has an end-to-end latency greater than 30 ms. At the London end of the call, the gateway is connected in the same way to Alice's telephone (by E&M to the PBX and by FXS to the terminal).

Tail Circuits

The analog circuit in London is known as the *tail circuit*. It forms the tail or termination of the call from the perspective of the user experiencing the echo—Bob in this case.

A packet voice gateway is a gateway between a digital packet network and a public switched telephone network (PSTN). It can include both digital (TDM) and analog links. The tail circuit is everything connected to the PSTN side of a packet voice gateway—all the switches, multiplexers, cabling, and PBXs between the voice gateway and the telephone. Figure 4 shows that the PSTN can contain many components and links, all of which are potential echo sources.

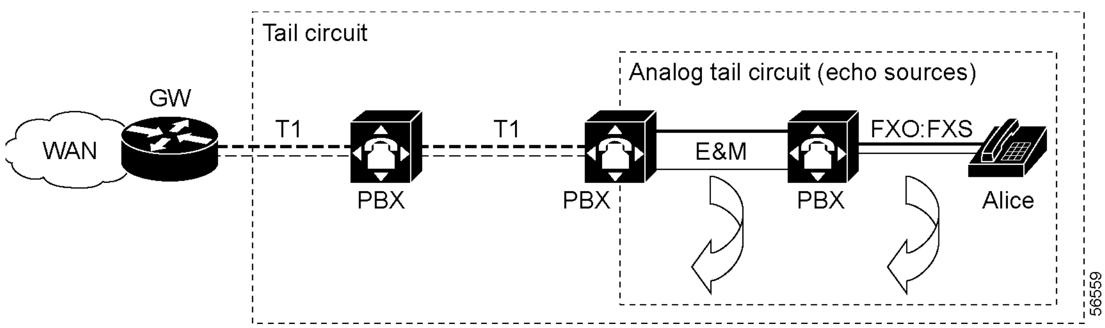
Figure 4 Echo Sources in the PSTN



London tail circuit

Packet voice gateways have two types of PSTN interfaces: digital (ISDN, BRI, T1/E1) or analog (E&M, FXO, FXS). Because bits do not leak, you can further refine your search for echo sources to the analog elements of the tail circuit. You can extend the echo-free digital zone out from the gateway to the point of digital-to-analog conversion in the PSTN, as shown in Figure 5.

Figure 5 Tail Circuit with Both Analog and Digital Links



Echo Sources

Suppose you want to locate potential sources of echo in the network in Figure 3. You know that bits do not leak, so you can disqualify the digital segment of the system. Therefore, the leak causing Bob's echo must be located in either the tail circuit in Montreal or the tail circuit in London. Any leak in the Montreal tail circuit would not have a long enough delay to be perceptible—echos there would be masked by Bob's sidetone. So the source of the echo must be the London tail circuit.

Remember that an echo problem has three ingredients, as follows:

- An analog leakage path between analog Tx and Rx paths.
- Sufficient delay in echo return for echo to be perceived as annoying.
- Sufficient echo amplitude to be perceived as annoying.

The digital packet link in Figure 3 requires a relatively long time for the analog signals entering this link to exit from the other side: the end-to-end delay of the link. This digital packet link delay comprises Tx, IP network, and Rx delays as follows:

- Tx delay includes primary processing delay (5 ms) and packetization delay (20 ms).
- IP network delays include transport propagation delay, packet-processing delay within the node, and serialization delay (depending on the interface choice).
- Rx delay includes jitter buffer delay (20 to 30 ms) and processing delay (5 ms).

IV. What Is Voice Echo Response (VER)?

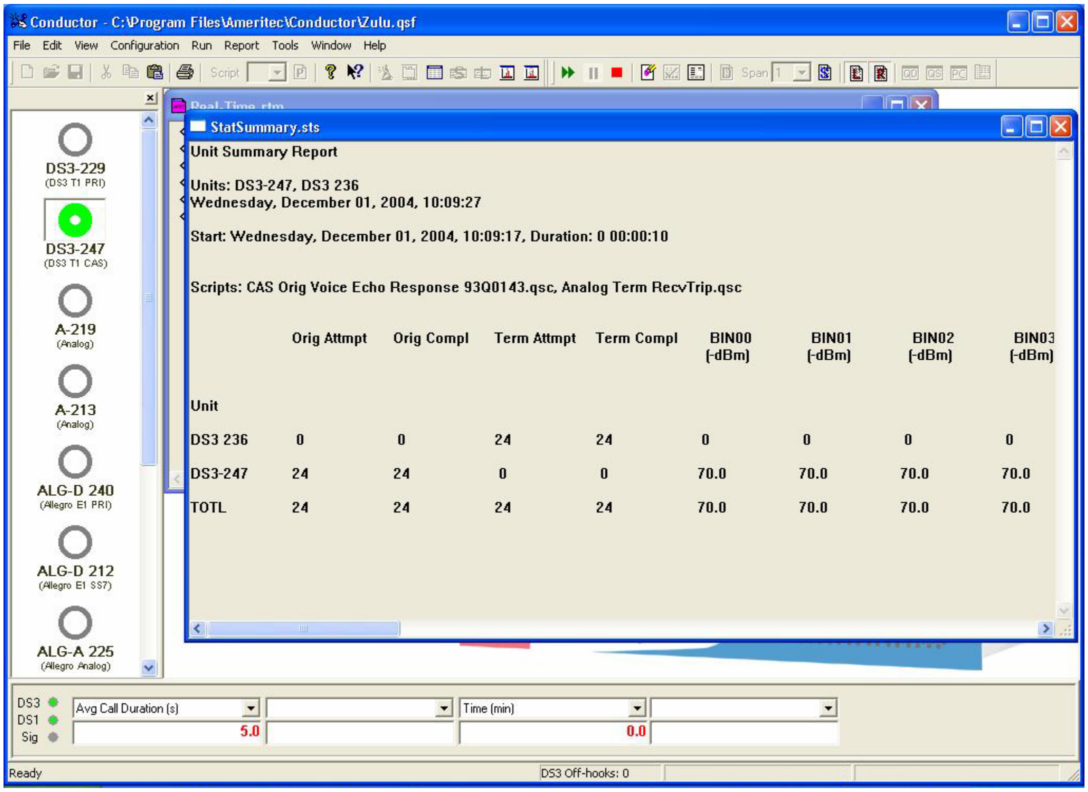
Voice Echo Response (VER) is Ameritec's method to detect the presence of echo during a call and to quantify the strength of the echoed signal as well as to record the time elapsed before the echo was detected.

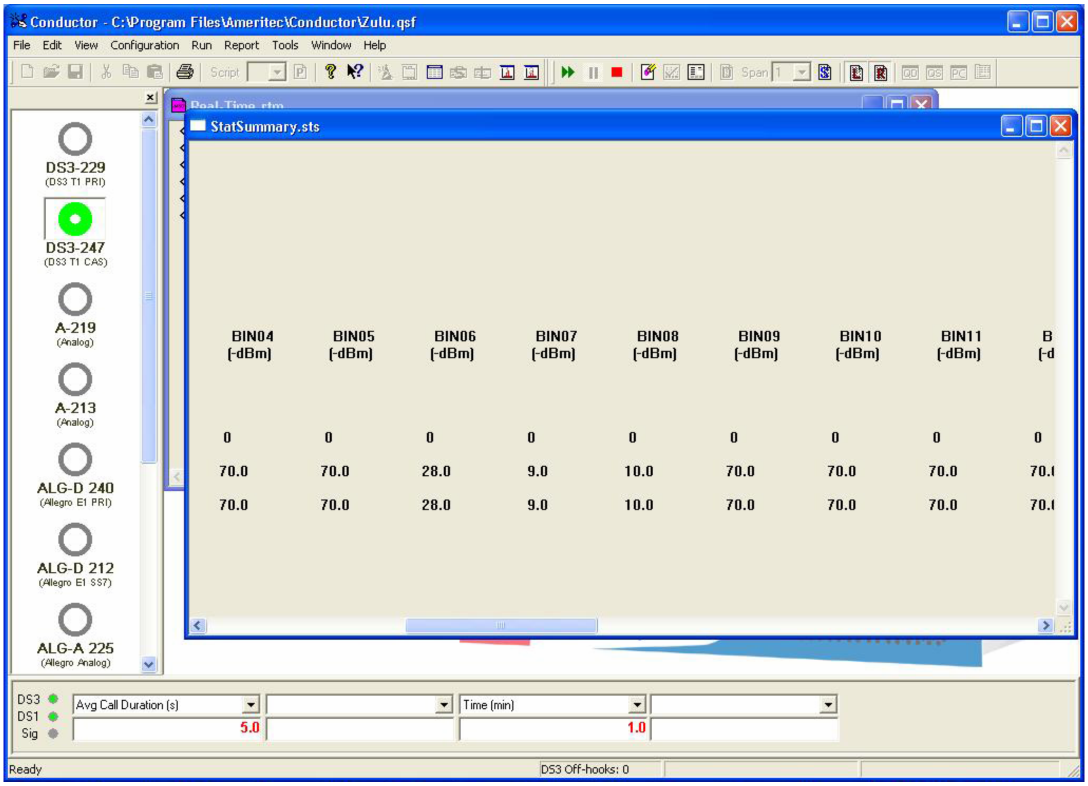
Each Fortissimo and Allegro call generator product utilizes a VER originate script that allows the user to send a burst of GoldenVoice tone, programmable in duration from 5ms to 4095ms. This tone can be controlled via the call script to vary the duration of the tone sent, the cadence of the tone or to accommodate other scenarios as deemed appropriate for echo detection. The terminate script will effectively remain in an off-hook, quiet state while this test proceeds. At the end of the tone burst, the call generator will detect and measure the reflected signal and log this signal strength into registers.

The user has the capability of selecting one of four different window sizes to look for the echo signal:

- 6.25ms listen time = 150ms (24 bins)
- 12.5ms listen time = 300ms (24 bins)
- 25ms listen time = 600ms (24 bins)
- 50ms listen time = 1200ms (24 bins)

Depending on the amount of delay anticipated for the circuit under test, the user should select the appropriate window size shown above. Selecting too short a window size could result in reflected energy being detected before an actual user would detect it or could result in delays that occur longer than 150ms to be missed in their entirety. Selecting too long a resolution may not provide the user with enough detailed information as to where in the circuit the delay occurs.


Conductor - C:\Program Files\Ameritec\Conductor\Zulu.qsf
File Edit View Configuration Run Report Tools Window Help
StatSummary.sts
Unit Summary Report
Units: DS3-247, DS3 236
Wednesday, December 01, 2004, 10:09:27
Start: Wednesday, December 01, 2004, 10:09:17, Duration: 0 00:00:10
Scripts: CAS Orig Voice Echo Response 93Q0143.qsc, Analog Term RecvTrip.qsc
Orig Attmpt
Orig Compl
Term Attmpt
Term Compl
BIN00 (-dBm)
BIN01 (-dBm)
BIN02 (-dBm)
BIN03 (-dBm)
Unit
DS3 236 0 0 24 24 0 0 0 0
DS3-247 24 24 0 0 70.0 70.0 70.0 70.0
TOTL 24 24 24 24 70.0 70.0 70.0 70.0
DS3-229
(DS3 T1 PRI)
DS3-247
(DS3 T1 CAS)
A-219
(Analog)
A-213
(Analog)
ALG-D 240
(Allegro E1 PRI)
ALG-D 212
(Allegro E1 SS7)
ALG-A 225
(Allegro Analog)
DS3
DS1
Sig
Avg Call Duration (s)
5.0
Time (min)
0.0
Ready
DS3 Off-hooks: 0



V. Summary

Echo is certainly a problem for any developer, engineer or user of a VoIP network. Most testing solutions for IP networks do not address the issue of echo and all QoS scoring techniques (R-Factor, PSQM, PESQ, etc.) also do not include echo as an affecting parameter of voice quality. Yet, any revenue-paying customer will insist there is nothing more annoying or quality affecting than echo.

Ameritec is now offering a means to test for echo along with its industry-leading QoS measurements as a method to provide a complete and comprehensive test suite for IP networks that connect to the traditional PSTN world. Tests can be configured to run on hundreds or thousands of lines simultaneously and it is possible to simulate a truly diverse traffic load with the Fortissimo and Allegro traffic generation products.

Ameritec Call Generators are available in a wide range of physical interfaces -- Analog, T1/E1 CAS, ISDN-PRI, SS7, DS3, SIP, OC3 and STM-1 – and provide interworking between the different interfaces. No longer is your testing confined to one piece of equipment in one location but with a GPS clocking source testing can also be done over a wide area duplicating the characteristics that are representative of actual deployed products.

You can count on Ameritec to provide the tools, the resources and the support to make you a success. If you have questions on testing your application: askzeke@ameritec.com

VI. Related Documents

• *Cisco IOS Voice, Video and Fax Configuration Guide,* Release 12.2
• Shenoi, K. *Digital Signal Processing in Telecommunications*. Prentice Hall PTR; 1995
• *Voice over IP Fundamentals,* Cisco Press, 2000
• *Echo Analysis for Voice over IP,* Cisco Press, 2002

