Troubleshooting Audio Quality Problems in the Field

Issue 1.0
February 2006

Abstract

This document provides troubleshooting procedures for resolving audio quality issues in the field.
Table of Contents

1. Introduction ................................................. 3
2. System Status and Environment .................... 4
3. Problem Definition ...................................... 5
4. Data Gathering ........................................... 5
5. Troubleshooting .......................................... 6
6. Audio Quality Anomalies ............................ 7
7. Other Information ...................................... 11
8. References ................................................. 12
APPENDIX A: Audio Processing Components & Audio Quality Terminology 13
APPENDIX B: Basic IP Telephony Solution Diagram 26
1. INTRODUCTION

Audio quality issues experienced by customers have many and varied causes. The converged voice and data solutions being deployed today require additional consideration when troubleshooting these issues as compared to the legacy voice solutions of the past. Some problems may be caused by equipment configuration or administration, others by network or environmental conditions, some by user misuse, and some are simply a customer expectation of the solution’s operation as opposed to the solution’s inherent design intent.

This document provides a guide to troubleshooting a customer’s audio quality issue(s) and determining what component(s) in the call topology is causing the problem. This is accomplished by first analyzing the customer system and environment, and then defining the customer’s audio quality issue using an industry accepted definition of the anomaly and the typical causes of that anomaly to thereby identify the area of the customer’s solution (or network) in which to concentrate the troubleshooting effort.

Appendix A provides a glossary of audio processing components and audio quality terminology. Appendix B provides a sample basic IP Telephony solution diagram as a reference.
2. SYSTEM STATUS AND ENVIRONMENT

As part of the normal Services troubleshooting routine, it is necessary that system components are up to date with the correct software/firmware releases, hardware versions, etc. Once the system components are established as having the latest updates, it is highly recommended that Best Practices documentation be consulted. Areas not meeting the initiatives of Best Practices documentation should be corrected and the audio problem retested before proceeding in this document.

Environmental conditions may cause or contribute to audio quality problems. Local noise sources such as computer fans, PC speakers, noise from air conditioning (HVAC), ambient speech or noisy office environments (as in call centers); can have an affect on audio quality being perceived by the user when using handset, headset or speakerphone.

For optimum speakerphone performance, make sure the phone is sufficiently clear of any shelves or clutter near the phone that could provide an echo path during speakerphone use. Note that environments with hard and reflective surfaces (such as metal walls and glass windows) can promote reverberation when speakerphones are used and result in echo complaints (from the far-end).

1Best Practices documents are intended for Avaya Services Installation teams and Business Partners, and can be found on the Avaya Enterprise Portal Support page via the Anatomy of a Successful Cut link: https://enterpriseportal.avaya.com/ptlWeb/appmanager/ptl/internal/?nfpb=true&_pageLabel=support. The link provides documents such as the Echo Prevention Checklist which can be obtained by clicking on Checklists/FAQs and then CSD - CM.
3. PROBLEM DEFINITION

Assuming the system components are up to date (latest version HW, SW, FW) and the solution design meets Avaya’s IP Telephony deployment requirements\(^2\), the audio quality complaint from the customer should be translated into an industry accepted definition of audio or voice quality related terminology as appropriate. Appendix A provides a glossary of these terms.

4. DATA GATHERING

Beyond the standard set of component model and release information, gather additional meaningful information such as call type (station to station or PSTN trunk call), call traversal and solution topology (within port network, across port network, LAN, WAN, etc.). Information gathering on whether the user experiences the anomaly on handset, headset, or speakerphone; or whether any operation is being performed (transfer, conference, etc.) when the anomaly occurs can help in determining the cause of the anomaly.

\(^2\)Avaya solution deployment guidelines and network requirements are contained in document ID 555-245-600 or can be accessed at the following URL: http://support.avaya.com/elmodocs2/comm_mgr/r3/pdfs/245600_3_4_1.pdf
5. TROUBLESHOOTING

With the information gathered (as described in section 4), it is important that you outline the exact call flow or the end-to-end route the voice signal takes from the service provider network (if applicable) through the system and across the customer network to the end network device. Appendix B provides a network diagram with measurement points identified (A – D) for troubleshooting an incoming PSTN call to an enterprise user. The following paragraphs are meant to assist in determining the call flow.

If the audio quality problem involving Legacy Telephones is intermittent, the support engineer will be required to be available onsite or remotely at the time the anomaly occurs (this may require a pager or other notification process be established between customer and support engineer so that the trouble can be addressed in real time).

If the audio quality problem involving IP Telephones is intermittent, a network analyzer (such as Agilent) should be set up for 24/7 operation such that the call information is captured (and buffered) at the time the anomaly occurs. The key to ensuring the anomaly is captured is by stopping the trace within a short time after the problem occurs. This will require the user to notify the customer’s IT department when the anomaly occurs so that trace can be stopped ensuring pertinent trace information is not overwritten. A pager or other notification should be sent to the Avaya Services Engineer to alert them that the capture has been obtained.

If it has been established that the audio quality problem occurs on trunk calls, it is recommended that a trunk ID button be programmed on the user’s phone so that when the audio anomaly occurs, the specific trunk accessed can be obtained.
If the **audio quality problem is reproducible**, establish a call with the phone experiencing the audio issue and note the call traversal (the path the call takes through the system). Once the call flow has been identified, all the network components involved are now known. At this point, try to determine whether the problem is limited to specific system components by initiating new calls that traverse a fewer number of system components. For instance, test if the audio issue can be reproduced on a station to station call between two DCP Telephones on the same port network. If the audio anomaly occurs when IP Telephones are used, determine whether the anomaly exists with IP to IP direct (shuffling) or whether the anomaly only exists when calls traverse a Media Gateway VoIP resource or only over a trunk call.

### 6. AUDIO QUALITY ANOMALIES

Having established the path the call traverses when the customer audio anomaly occurs, combine this knowledge with the definitions in Appendix A to determine what system component is causing the problem or if further measurements are required. The following sections describe the three major categories of audio quality anomalies (distortion, noise, and echo) and the typical causes.

**Distorted Speech**

Speech distortion(s) are typically caused by low-rate coders, improperly matched hybrid trunk circuits, and stutter or speech-level pumping. Audio distortion (amplitude clipping) coming from outside the customer’s network is typically cause by the far-end phone.

---

3 Note that taking a trace to an IP Phone while an anomaly is occurring is a quick test to identify if the problem is in the phone or not by listening to the RTP stream and comparing that to what is heard in the phone.
Distortion from inside the customer network could be caused by, for example, a headset user with the microphone too close to their mouth and overdriving the microphone circuitry.

1. **Distorted Speech** can be caused by a **damaged microphone** in the handset, headset, or speakerphone circuit. Swap out the equipment to determine if this is the cause of the distortion. Use of **Low-rate coders** can cause **muffled speech, synthetic, mechanical** or **robotic speech** when network impairments are present. Take network QoS measurements and note the values of the typical characteristics (packet loss, delay and jitter).

   - If the user experiences any of these types of speech distortions while on a call using the 4610/20/21/22 IP Phones, the user (or Services engineer) may press the Network Audio Quality menu item (hit the Options button and select Network Audio Quality) to get an instant measurement of packet loss, delay and jitter being experienced by that phone.

   - If the problem is intermittent, working with the customer to set up a network analyzer to monitor the network 24/7 such that when the user experiences the problem, the trace can be stopped and RTP analyzed. Use this data in conjunction with the network topology and call flow to determine what network element is causing the anomaly.

2. **Amplitude clipping** can be determined by running an Ethereal trace and converting the RTP stream to a .wav file and analyzing that file with a tool such as Audacity. The clipped speech signal can be observed graphically with that tool.
If amplitude clipping is present, identify solution elements that add gain and verify each element is applying the expected amount of gain.

3. **Clipping** can be caused by packet loss which can be determined using a network analyzer. Also, e-can settings 1 thru 4 on the TN464GP boards are very aggressive and can cause clipping. Administering the DS1 (GP) e-can setting to 8 or 9 can reduce/eliminate the clipping. On DS1 HP boards, e-can setting 3 is a less aggressive echo canceller and can reduce/eliminate clipping.

4. **Clipping during double-talk** can be caused by the suppressor function of an echo canceller in the network. DS1, MedPro/Cruiser, Crossfire, and ATM all use e-cans with suppressor functions.

5. **Stutter** can be traced to it’s origin by setting up a network analyzer to analyze the call topology and where the bottleneck is occurring.

6. **Speech-level pumping** can be caused by the DS1 e-can setting 4 (on GP cards) or by the telephone’s AGC. Adjusting the GP type DS1 e-can setting to 7 or 8 or turning off AGC are fixes for this problem. Also, high transmit gain levels can cause pumping. Reducing the TX gain in the 46xx Settings file use by the IP Telephones can also reduce/eliminate pumping.

---

4. Changing the e-can setting affects active calls. For this reason, it’s beneficial to change he e-can setting while the audio anomaly is occurring in order to get immediate feedback on the change.
5. This should only be performed by Tier IV or under Tier IV supervision.
Noise

Noise occurs in most telephones at low levels. It’s only when the levels are perceptible that complaints arise. The following provides noise types and ways in which to capture information in order to analyze and pinpoint the cause of the noise.

1. **Hiss** or **white noise** is common at low levels. Requirements allow for up to 20dBrnC of idle-channel noise to appear in the receiver of a handset or headset. Typical idle-channel noise should measure approximately 15dBrnC. This can be found using a pair of Sage 925VSTs. Or a Sage 925VST with a Sage Responder.

2. **Static** causes can be identified by using a network analyzer to determine if there is packet loss or bit errors.

3. **Motor-boating** or **Hum** can be caused by power-line interference. Checking grounds and any nearby high voltage sources can uncover the cause for this problem.

4. **Distorted Music-on-Hold** is typically caused by a low-rate coder. If so, changing the coder will fix this problem.
Echo

Echo, either acoustic or electric, is a function of delay in the call path\(^6\). The steps below address how to determine the cause of the delay.

1. Speech back to the far-end. If the far-end no longer complains of echo after muting the near-end phone, then this confirms the near-end is the source of the echo. If the near-end is speakerphone, verify the environment setup. If near-end headset user, verify proper cords used. If headset with M12 amplifier, verify the amplifier settings and proper cords are used. Taking a trace on an IP Phone while echo is heard can identify if the echo is in the phone or from the far-end (if no echo is heard in the trace, then it’s in the phone).

2. Electrical echo (constant, intermittent, residual, distorted) is either at the analog trunk or in the PSTN. Terminating the trunk locally and testing for echo will identify if the source of echo is in local or in the PSTN. Verify DS1 e-can is being utilized and not the MedPro e-can resources.

7. OTHER INFORMATION

It is highly recommended that you do not adjust the Loss Plan or Terminal Parameter gain settings. This is not a recommended option for addressing audio impairments.

For optimum headset performance, use only Avaya headsets best suited for the environment\(^7\).

\(^6\) For more in-depth echo discussion, see Echo and Echo Control document - COMPAS ID 106541.

\(^7\) [http://www.avayaheadsets.com](http://www.avayaheadsets.com)
8. REFERENCES


[2] Echo and Echo Control, CID # 106541, Issue 0.4, June, 2005.

APPENDIX A: Audio Processing Components & Audio Quality Terminology
In a typical telephony call, speech from talker to listener often passes through the following processing components and in the order identified in Figure 1.

![Diagram of speech processing components](image)

**Figure 1.**

The talker’s voice enters at the microphone on the left side of Figure 1, then to the microphone expander, voice coder, network transport, voice decoder, packet-loss concealment, echo controller, automatic gain control and, finally, the listener’s ear. Brief definitions of these components are given below.

**Echo controller.** This is a broad term meaning an echo canceller, echo suppressor, or a combination of the two. Speakerphone algorithms are also included. An echo controller prevents a talker from hearing distant reflections (echoes) of his/her own voice, reflections caused by acoustic or electrical reflection points within the telephone network and end-user equipment. Echo controllers are often only partially successful, and this is why echo is sometimes heard even though the call path is known to include echo controllers. Often, people use the “echo canceller” when in fact what is being referred to is an echo controller.

**Echo canceller.** Software or hardware implementations of a digital signal processing algorithm designed to model and subtract-out – or cancel – the reflection, or echo, of a speech signal. Strictly speaking, an
echo canceller does not introduce attenuation or suppression into the speech paths to reduce the loudness of echo. The term canceller refers to an adaptive digital filter that models the physical echo path and subtracts that (excited) model from the return speech path.

**Echo suppressor.** Like echo canceller, above, except the echo level is reduced or eliminated by applying suppression or attenuation to the return speech channel. The use of attenuation causes other audio artifacts, including chopping or clipping of speech utterances and/or pumping of the loudness level of a caller’s speech.

**Microphone expander (and or noise reduction).** A microphone expander is a traditional and relatively simple method of improving the speech-signal-to-background-noise ratio emanating from the microphone path. An expander attenuates weak room background noises while passing unaltered the relatively loud speech of the talker addressing the handset (or headset, or speakerphone).

**Speech coder (encoder and decoder).** The raw speech signal, once digitized, is often digitally encoded for transmission into the telephone network. Encoding has one purpose, namely, to reduce the bits-per-second rate of transmission required to communicate voice from one end to the other. Highly compressive codecs, such as a G.729 codec, reduce speech to a low transmission rate (8000 bits-per-second), but sacrifice voice quality in doing so. Higher voice quality is experienced in systems using the traditional G.711 codec (mu-law codec), since G.711’s higher transmission rate of 64,000 bits-per-second better captures the nuances of speech. Regardless of coding scheme, at the receiving side, the speech decoder reconstructs (an approximation to) the original speech for playback.
Packet-loss concealment. Often combined with speech decoders. When the network path includes packet-speech transmission links, like VoIP, speech packets can be lost because of network failures. In such cases, a concealment algorithm attempts to fill-in missing speech samples. Concealment can work well when the rate of lost speech is very low, say, less than 2% of transmissions.

Automatic gain control. Automatic gain control devices apply signal gain or loss automatically in an attempt to keep the speech sound level at the listener’s ear relatively constant. Therefore, AGC boosts low-level speech while reducing speech levels that are too loud. Such devices have been used for decades in audio broadcasting and recording applications.

The following sections outline the troubleshooting steps in order to identify audio issues.
Speech Distortions

Distorted Speech
Speech accompanied by an unnatural buzzing or raspy sound. A classic example of distortion occurs in the case of a far party who is speaking too loudly or too close to the handset or headset microphone. The far party’s speech saturates either the mechanical or electrical capabilities of the handset, causing overload distortion or amplitude clipping. Other sources of speech distortion include: defective or improperly (impedance) matched hybrid trunk circuits, damaged microphones, loudspeakers or earpieces (a rarity), and low-bit-rate speech coders (VoIP and wireless).

Muffled Speech
Speech that has an unnatural loss of high-frequency content. Muffled speech may be caused by, for example, poorly designed microphone assemblies in handsets (in particular, wireless handsets) and low-bit-rate speech coders (VoIP or wireless).

Reverberant speech (also hollowness or speaking-in-a-tunnel effect)
Sounds like the person you are listening to is speaking in a barrel or large empty room. This can be the case when the talker is using a speakerphone, but it can also be the case when there is network echo, e.g., in a teleconference without echo control.

Also see ECHO section anomalies: hot sidetone, short-path acoustic echo, and electrical echo.
**Synthetic, Mechanical, or Robotic Speech**
This can be very subtle or very severe, or very consistent or intermittent. In the most severe case, the pitch information has been lost making the speech sound monotonic and robotic. Recognizing who is speaking is often difficult. Faulty or bad acting hardware (e.g., a DSP chip in a cell phone, base station, or VoIP gateway) can cause this. This may be port/resource dependent but consistent on the faulty port or resource. If this is occurring intermittently, sudden bursts of packet loss (or frame erasure in a wireless network) may be the cause as the packet-loss concealment scheme struggles to interpolate the speech signal. A less extreme case in which you may recognize the speaker, but yet the speech sounds somewhat artificial or synthetic has been known to occur as well. Sometimes the effect includes a swirling sound accompanying or in place of certain frequencies of the voice signal. It may be very subtle and is often linked to the usage of a low-bit-rate speech codec, such as G.729 in the presence of background noise, or possibly the occurrence of packet loss or frame erasures, as occurs in VoIP and wireless networks.

**Amplitude Clipping**
See definition for distortion above.

**Clipping**
Portions of the speech signal are not heard. This can occur in packet-switched networks when, for example, large numbers of successive speech packets are not received because of excessive network congestion. Common in wireless phones, where the RF-signal strength fades as the user moves within the environment.
Clipping during Double-talk
Clipping, as defined above, but heard only when both parties of a telephone call talk at the same time. When it occurs, this effect is almost always caused by the excessive use of echo suppression (see definition) at some point within the network. In this case, clipping of speech utterances is not caused by lost speech packets or, in the case of wireless phones, RF fades, though those artifacts may also be present in the same call.

Stutter
This is often used to describe an effect caused by repetition of short bursts of noise or speech, such as “da-da-da-da...” or “fa-fa-fa-fa...” Stutter distortion can occur in packet-speech networks when one or more network elements (e.g., router or switch) become a bottleneck to the timely transmission of speech packets.

Speech-level pumping
Pumping is often used to describe a varying speech-loudness level, that is, were the speech gets louder, softer, then louder again, etc., over the course of a call, often over a period of just several seconds. Automatic gain control devices can cause audible and distracting pumping.

NOISE

Hiss or White Noise
Relatively natural-sounding noise containing energy at all frequencies. Low-level, idle-channel hiss noise can be perceived on nearly every telephone call when no person is speaking.
**Static**
Impulsive, ticking noise, similar to the sound of an AM radio when tuned to a very weak or nonexistent radio station. In a packet-speech network, can be caused by lost speech packets and/or bit errors. May also be used to describe power-line hum (see definition below).

**Motor-boating**
Repetitive noise that is separate and distinct from the talker’s voice. Motor-boat noise differs from static in that it is repetitive or non-random. Motor-boat noise can be caused by periodic disruptions (errors) to digital speech circuits and power-line interference at telephone endpoints.

**Hum**
Sounds like humming, as in “Hmmmmm...” Hum noise often occurs when a source of 50 Hz or 60 Hz electrical power is located near a telephone. The power source emits an RF (radio frequency) field that induces a hum-like noise that is heard through the phone’s handset/headset earpiece or speakerphone loudspeaker.

**Distorted Music-on-Hold**
Low-bit-rate codecs such as G.729, and G.723, were created to efficiently encode and transport speech but not music. Thus the usage of these and other codecs may distort and ruin the music signal. This can be subtle or severe depending on the music source. As a remedy you can try to experiment with various music sources, but it would be best to use a more appropriate codec such as G.711, G.726, or G.722.
There are only two physical sources of echo in telephony: electrical echo, or network echo, and acoustic echo. Electrical echo is caused by a reflection of the speech signal at 2-to-4-wire hybrid circuitry. This circuitry is present in our own analog trunk cards, and it also exists deep in within the PSTN (at customer premises, for example). Acoustic echo is caused by the physical coupling (air path, appliance-body path) between a loudspeaker and a microphone, for example, in a speakerphone, a handset and a headset. Whether or not a talker actually perceives electrical or acoustic echo depends on the loudness of his/her reflected voice signal and the roundtrip delay that that reflection suffers. The loudness of the reflection at the point of reflection depends upon the electrical impedance mismatch, for electrical echoes, and the acoustic gain of the loudspeaker-to-microphone path, for acoustic echoes. The roundtrip delay is a function of the path the reflected signal traverses, which in turn is a function of the call topology.

**Electrical echo, also called network echo**

Reflection of a talker's speech signal at a point of 2-to-4-wire conversion caused by an impedance mismatch at the point of analog-to-digital conversion.

**Acoustic echo**

Reflection of a talker's speech signal at an acoustic endpoint caused by the acoustic coupling between the loudspeaker and microphone.
**Constant echo**
When talking, the perception of echo with every utterance. Such cases occur when there is a physical electrical or acoustic echo path but no echo controller in the call topology to control echo. Additionally, constant echo may result even though an echo controller is known to be in the call path; this indicates a complete failure of the echo controller, usually because the capabilities of the echo controller are exceeded (e.g., the echo tail length exceeds the specifications of the echo controller).

**Intermittent echo**
When talking, the occasional perception of echo. Intermittent echo often caused by the intermittent failure of an echo controller in the call path. The echo suppressor within the echo controller may fail to engage (to apply echo attenuation) when necessary, with the result that short bursts of echo become audible. In acoustic echo control applications (speakerphone) in which people or objects close to the speakerphone are moving, the change to the physical echo path often results in audible intermittent acoustic echo to listeners at the other end of the call.

**Residual echo**
When talking, the perception of very low-level (quiet) echo. The echo could be either constant or intermittent. Residual echo can be caused by PSTN electrical echo that is not entirely removed by the echo controller in the call path.

**Distorted or buzz-like echo**
When talking the perception of a distorted echo or buzz-like sound. This can be caused by a non-linear echo source. An example of this is saturation distortion at an analog trunk interface. In this case, signals low in amplitude are reflected cleanly, but signals high in amplitude are returned with significant distortion making it difficult for an echo canceller to control echo.
Such distorted echo can be perceived constantly or intermittently, depending on the degree of distortion and the echo canceller(s) involved.

**Slapback or kickback acoustic echo**

This is strictly a phenomenon of acoustic echo. With speakerphones, slapback or kickback echo is the intermittent echo perceived at the ends of one's utterances. This can occur with both older-model half-duplex speakerphones and newer-model acoustic-echo canceling speakerphones. For example, a talker speaking into a handset utters the phrase “Please send me the check” and perceives echo primarily at the end of his/her sentence. This echo is described as hearing just the sound “eck” or “k” of the word “check,” or as a slapping sound such as that made by slapping one’s palm against a desktop. Commonly, slapback/kickback echo is caused by acoustically reverberant rooms. Large offices and conference rooms can have long reverberation times. In such rooms, the speakerphone senses at its microphone a reverberated version of the word “check” (our prior example) several tens or even hundreds of milliseconds after the far talker has finished saying the word “check.” The speakerphone algorithm detects this reverberated speech at its microphone, detects no speech at its receive-path driving the loudspeaker, and decides to transition to transmit mode. The reverberated version of "check" is transmitted back to the far talker, where it is perceived as echo.

Whether a slapback reflection occurs depends on the speakerphone acoustics, speakerphone algorithm and room acoustics. Whether slapback echo is perceived as echo by the far talker depends on the loudness of the echo and the roundtrip delay. Some have described slapback echo as speakerphone false switching, where it is meant that the speakerphone has switched too quickly from the receive mode to the transmit mode. This is not actually a false switch, however; the
speakerphone merely rapidly switches between states because it cannot distinguish between microphone energy that is due to a talker who is local to the speakerphone room and energy due to reverberation of the received speech. A speakerphone that enforces a longer hold-time governing the smallest amount of time for a switch between the receive and transmit states will in general suffer less from slapback echo. A common drawback, however, of a longer receive-to-transmit hold-time is a higher perception of clipped utterances during doubletalk; in other words, a greater perception at both ends of the speakerphone call that the speech path is half duplex.

Sidetone
In handsets and headsets, a portion of the microphone energy is fed back to the earpiece so that the user of the handset/headset does has a psychoacoustic experience that simulates the case in which the user's ear is not occluded by an object (the handset earpiece). Without sidetone injection, the user experiences the psychoacoustically bothersome condition that can be demonstrated to oneself by pressing a finger into one ear while speaking. With one ear occluded, the sound of one's own voice is dominated by the path through the interior of the head (skull, etc.) instead of around the head, an effect that most people find objectionable.

Hot sidetone
In a handset or headset, microphone-to-earpiece sidetone injection is not normally noticed. Some digital phones, in particular, IP phones in which the internal audio processing frame rate is 5 ms or greater, inject sidetone with an appreciable delay (e.g., 5 ms) in the microphone-to-earpiece signal path. This delay causes the sidetone to sound reverberant and/or louder than normal, or hot. Though hot sidetone is a type of echo source – because some people may use the term “echo” to describe hot sidetone – it is generated local to the telephone, not at some point within the telephone network.
**Short-path acoustic echo, short-path electrical echo**

Acoustic or electrical echo that occurs in a very short roundtrip call topology. This type of echo is commonly described as a hollow sound or sound of speaking in a barrel (see 2.1.3). In a digital-to-digital phone call (think DCP-to-DCP), station-to-station, the roundtrip delay is usually very small, less than 10 ms. Some digital speakerphones produce significant acoustic echo, which is not canceled, suppressed, or otherwise controlled in this simple call topology. In these cases, and depending on the volume setting of the far-party’s speakerphone and near-party’s listening handset, the near party may perceive echo and refer to this as hot sidetone. Again, this is truly acoustic echo from the speakerphone but is returned to the talker with such a short roundtrip delay that it is perceived as hollowness or reverberance rather than as classic echo. Because of the short roundtrip delay in this case, it can be difficult to distinguish between hot sidetone (see definition) and short-path echo.

**Short-path electrical echo** refers to an electrical echo source in which the roundtrip echo path is also short, say less than 10 ms. Again, in such cases, it may be very difficult to distinguish between short-path echo and hot sidetone (see definition).
APPENDIX B: Basic IP Telephony
Solution Diagram
Figure 2.