ECHO CONTROL IN VOIP

Samy El-Hennawey, Ph.D., P.Eng., SM-IEEE
Nortel - Enterprise Solutions
Belleville, Ontario, Canada, K8N 5M1
Email: hennawey@nortel.com

1. INTRODUCTION

Echo control in public-switched telephone networks (PSTN) using echo cancellers have been exercised for nearly the past 40 years. While a great deal of success has been achieved, the recent deployment of voice over the Internet protocol (VoIP) equipment has faced additional challenges. As it is well known, perception of echo and its annoyance is a function of the level of echo as well as its associated round-trip delay. While the percentage of all-PSTN calls that requires echo control is relatively very small, all VoIP calls need echo control due to the inherent latency of all VoIP calls accessing any 2-wire connection including PSTN. As such, echo cancellation is an inevitable functionality of any voice media gateway.

2. SOURCES OF ECHO

In telephone conversation, there are primarily two classes of echo. The first class is due to the use of a 2-to-4 wire conversion using hybrid circuits. Echo control of this class is achieved by a line echo canceller (LEC). The other class is due to the use of handsfree mechanisms where there is significant acoustical feedback coupling between the receiver speaker and the receiver microphone. Echo control of this class is achieved using an acoustic echo canceller (AEC). This paper is primarily concerned with the first class. Additional sources of echo need to be dealt with the LEC as shown in Figure 1 below.

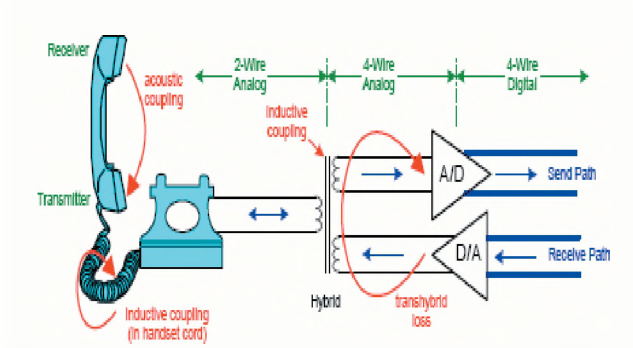

Figure 1 Echo Sources

3. ROUND-TRIP DELAYS IN VoIP

In any VoIP call, voice is packetized either at the IP-phone itself or in the voice gateway. While voice packets can vary in duration, a typical value is 20 milliseconds. At each side of the IP network of the call, a jitter buffer is part of the receiver to account for the asynchronous reception of the packets due to passing through the IP network. A typical inserted delay in the jitter buffer is twice the packet size, i.e. about 40 milliseconds. The path of the speech packets through the IP network involve many switches and routers and as such will encounter additional delays. Noting that echo encounters both ways, it is the round-trip delay and its typical value is in the neighborhood of 200 milliseconds. In reality, one should expect longer round-trip delays. It is well known that the maximum round-trip delay NOT needing any echo control is approximately 50 milliseconds. Therefore, the use of echo cancellation is inevitable when there is echo.

4. PERCEPTION OF ECHO

As mentioned earlier, the perceived annoyance of echo is a function of the echo level in dB as well as the round-trip delay. Figure 2 below depicts the level of annoyance with these two parameters.

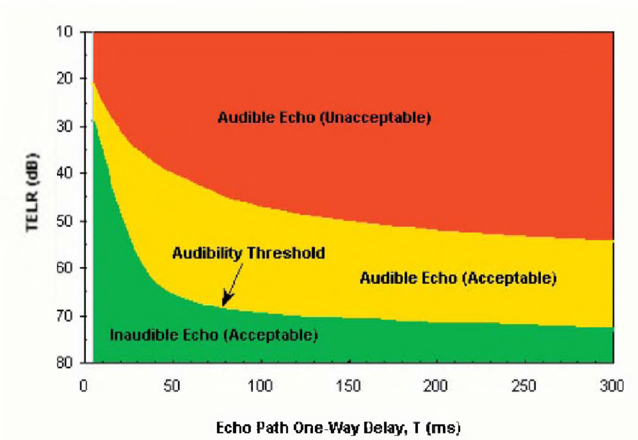

Figure 2 Echo Perception

With the minimum round-trip delay of 200 milliseconds, and with the help of Figure 2, the "total echo loss rating (TELR) has to exceed 55 to 70 dB, with the upper range meant for high-profile users. This definitely represents a major challenge facing LEC designers. Achieving this level of cancellation using fixed-point DSP techniques in an adaptive filter is rather impossible. Therefore, echo suppression mechanisms are used in addition to the adaptive filter operation for better cancellation. This is done by using a non-linear processor (NLP) that is activated only when the signal coming is only echo, i.e. the person at the other end is not talking. This mandates the existence of a double-talk detector and a sophisticated NLP mechanism. NLP has become a standard module of an LEC as shown in Figure 3 drawn from the G.168 Recommendation of the ITU-T.

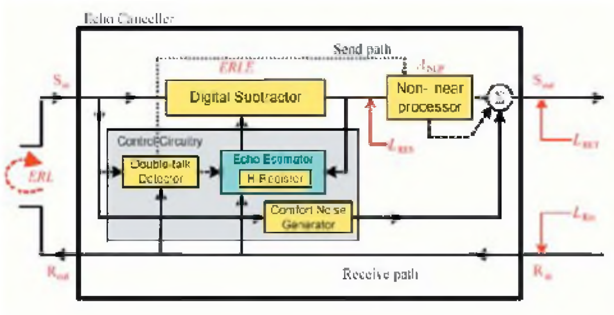

Figure 3 The Line Echo Canceller Modules

The design of double-talk detectors is another challenge for VoIP gateway design. On one hand an aggressive operation of the NLP will lead to cut or clip portions of real speech from the other end. This leads to double-talk choppiness distortion that is very annoying. On the other hand, a very lenient NLP operation leads to echo leakage in single-talk, which is also very annoying. Most double-talk detectors are energy-based. With the non-homogeneous loss plan with a wide variety of call scenarios in addition to the interruptability caused by the long delay add to the challenge.

5. ADDITIONAL CHALLENGES

Conferencing also represents a major challenge to the LEC designers. Some conference participants are IP phone users and others are connected to the bridge from the PSTN through a voice gateway. The number of PSTN participants could be large and each of them usually joins the conference bridge at a different time. Each PSTN participant then may have a different echo path with different loss and round-trip delay. This means the LEC initially adapts to the first PSTN participant echo path. Each time a new PSTN participant joins the conference bridge the overall echo path changes with possible dispersion different from the previous ones. This mandates the use of either a very long tail delay full filter LEC or sparse LEC. In both cases, a re-training procedure has to take place that may cause echo to occur for some duration of time. In the case of full filter LEC, the number of filter coefficient becomes too large which will slow the speed of convergence, hence possibility of echo leakage for longer duration of time. In the case of sparse LEC, a new allocation of window is needed for one or more participants depending on the new dispersion location(s). This adds to the complexity of the canceller hence limiting the channel capacity.

Another challenge seen in deploying VoIP with voice gateways to interface with PSTN is that some PSTN local loops have peculiar characteristics. Some local loops are nonlinear in nature and if this nonlinearity is significant the LEC will be unable to cancel echo efficiently unless using an aggressive NLP with its associated choppiness possibility. Some other local loops are seen to use pair gain technique where some very primitive cancellers are used. This results in canceller tandeming. The voice gateway will track the other primitive canceller and thus seeing a time-varying echo path. This will eventually lead to longer convergence time and as such more leakage of echo or choppiness if an aggressive NLP operation is used.

6. ECHO MEASUREMENT

With multi-vendor networks, it is expected to have the DSP module of the voice gateway manufactured by a different vendor than the network management and QoS monitoring module. Therefore, it is highly desirable to measure echo at the end user of an IP phone. One way to achieve this is through the use of the RTCP-XR protocol of the IETF. However, this relies on the LEC itself for estimating the echo leakage. Another more efficient process is developed at Nortel to provide a complete suite of impairments measurement located at the end user of an IP phone. An important impairment, that is to be measured, is echo with its echo path loss and its associated round-trip delay. The effect of each impairment on the voice quality is reported along with the overall quality measure based on the mean opinion score (MOS). More details can be found in the references provided. The work is also being discussed for the new P.CQO Recommendation of the ITU-T that is in progress.

7. DEMONSTRATION

During the presentation of the paper, it is intended to provide demonstrations of some calls recorded in real life at customer sites showing some of the challenges faced the deploying team. As well, a simple demonstration will be given to illustrate the effect of echo amplitude and delay in the perception of echo.

REFERENCES

M.S. El-Hennawey, R.A. Goubran, L. Ding, A. Radwan, "Non-intrusive single-ended voice quality assessment in VoIP based on voice payload", United States patent application filed by Nortel Networks on June 15, 2005.

M.S. El-Hennawey, R.A. Goubran, L. Ding, A. Radwan, "Method and Apparatus for Non-Intrusive Single-Ended Voice Quality Assessment in VoIP", European Patent Application: WO2006136900 , Publication Date: December 28, 2006.

L. Ding, M.S. El-Hennawey and R.A. Goubran, “Nonintrusive Measurement of Echo-Path Parameters in VoIP Environment”, IEEE Trans. on Instrumentation and Measurement, Vol. 55, No. 6, December 2006.