XVII IMEKO World Congress
Metrology in the 3rd Millennium
June 22–27, 2003, Dubrovnik, Croatia

ANALYSIS OF ECHO IN PSTN-TO-VOIP TRANSMIGRATION

Vedran Srsen[1], *Sandor Dembitz*[2], *Petar Knezevic*[2]

Coting d.o.o.[1], Faculty of Electronical Engineering and Computing[2], Zagreb, Croatia

Abstract - The growth of the Internet and its impact on other communication technologies is nothing new today. One of the issues is the migration of voice traffic from Public Switched Telephone Network (PSTN) to packet networks. In this paper we will show the impact of that migration on the occurrence of echo, and the use of solutions that are implemented in the VoIP network.

Keywords: echo, VoIP, ELR

1. INTRODUCTION

Every conversation in Public Switched Telephone Network (PSTN) consists of two paths: transmit and receive path. Transmit path (TP) is created when a person speaks and the sound is transmitted from the speaker's mouth to the listener's ear, and receive path (RP) is created when a person hears the sound and the sound is received by the listener's ear. An echo occurs when a part of the signal from the TP leaks to the RP. When this happens the speaker hears his voice delayed, which can be annoying depending on the amplitude and the delay of the echo. The leakage occurs only on the analog circuits, electrically from one wire to the other. If these analog signals are converted to digital signals, the leakage doesn't occur [1].

Human hearing system has a minimum time interval between sound events that determines whether those events will be perceived as a single event or as two separate events. This minimum delay that separates two events is about 20 ms. In telephone devices some of the signal from the mouthpiece is directly sent to the earpiece (sidetone), so it uses the above mentioned human hearing phenomena to mask all echo that is delayed less than 20 ms. Because analog transmission is very fast, PSTN connections have a very short round-trip time delay generally lower than 20 ms, so even very loud echoes are imperceptible cause they are masked by the sidetone.

Voice over Internet Protocol (VoIP) was introduced as an application of the packet networks for transporting voice. The main issue is satisfying the end users concerning the quality of transported speech on which they were used to when using PSTN. Our main quality consideration in this paper is echo. The question is: How will the migration from the PSTN to the packet networks (e.g. VoIP) affect the issue of echo, if we said that echo is only the problem in the analog networks? The answer is the fact that every packet network inserts extra delay and the amount of the delay

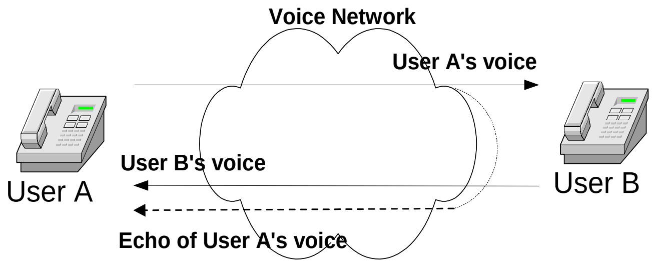

Fig.1. Echo occurs on the terminating side of the call

depends on the network size (the number of the network elements that process the packets, thus inserting new delay). The use of the packet network transmission link imposes an extra delay [2], which makes the existing echoes that were imperceptible in "pure" PSTN network, perceptible. With our measurements we will show that migration to VoIP networks does not cause echo, but exacerbates existing unnoticeable echo problems by increasing network delay.

2. ECHO CANCELLER

In our measurement we tried to eliminate echo from our test network environment. We did that by using echo cancelers that are implemented in the voice gateways (gateways between PSTN and IP network, GW1 and GW2 in Fig.1). An echo canceler is a part of a voice gateway that is used for reducing the level of echo that leaked from the TP to the RP (if observing from the User A side). The echo canceler removes echo from the receiving signal by learning the electrical characteristics of the remaining analog part of the network between gateway and the end user (tail-circuit). It then creates an estimation of echo signal based ob the current and passed transmitted signal. Then it subtracts the

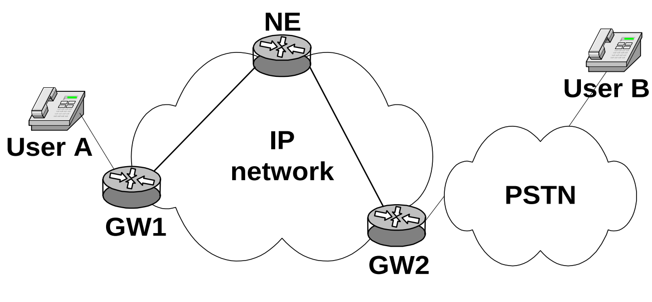

Fig.2. Test network configuration

estimated echo from the received signal and if the estimation was accurate only the voice from the User B is transmitted to the packet network.

2.1. Echo Canceler Operation

The echo canceler operation is based on the fact that the tail circuit can be represented by a mathematical formula, or, in other words, there is a formula that can describe the relationship between the input signal (User A's voice) and the output signal (User B's voice). In Fig. 3 H(t) is a mathematical representation of transformation of x(t) (User A's voice) in the tail circuit. This transformation produces an echo signal e(t) that exits the tail circuit together with signal y(t) (User B's voice). The echo canceler simulate the passing of x(t) through the tail circuit by using the H'(t) transformation. H'(t) is an estimation of the tail circuit's transformation formula.

$$e(t) = x(t) * H(t) \quad \text{(convolution)} \qquad (1)$$

$$e'(t) = x(t) * H'(t) \quad \text{(convolution)} \qquad (2)$$

$$y'(t) = y(t) + e(t) - e'(t) \qquad (3)$$

Echo canceller obtains the H'(t) formula through trial-and-error process, using the gradient descent algorithm to improve the coefficients of the finite impulse response filter (FIR). The echo canceler starts with all-zeros formula for H'(t) [3]. Through the period of adaptation the formula for H'(t) adapts in a controlled fashion based on the size of the error signal that leaves the echo canceler. Gradually, the error decreases and H'(t) becomes a better and better estimate of the H(t). This period is known as convergence period. This convergence period happens only when User B is silent. When User B talks, the formula continues to generate e'(t) estimates and subtract them from the incoming signal. If the estimation is accurate, User A hears User B talk with no echo from his own speech. Echo canceller has a defined threshold for determining the moment when User B starts to talk (double talk). But if the echo is too strong echo canceler will always interpret it as double talk and want go to the new estimation process [4].

2.2. Echo canceller coverage

If a signal enters the tail circuit (x(t) in Fig. 3), the echo signal (e(t) in Fig. 3) is indeed a series of delayed and attenuated echoes of the signal. Number of those echoes depends on the number of echo sources and delays between them. The time period that is needed for the entire echo signal to come out of the tail circuit is called *ringing time*. Echo canceller coverage is the time that the echo canceller keeps its estimation of the tail transformation formula, H'(t), in memory. It is the maximum echo delay that an echo canceller can eliminate. For successful echo cancellation it is necessary that the echo canceller coverage is as long as the ringing time of the tail circuit.

In addition, we can recognise two cases of uncancellable echo. First we have the echo that is too loud and requires more attenuation that a single echo canceller can provide. In the second case series of analog and digital links can delay an echo beyond the time of the echo canceller's coverage.

2.3. Measuring echo

There are three significant values that appear when measuring echo:

- echo return loss
- echo return loss enhancement
- acombined

Echo return loss (ERL) is the reduction in echo level generated by the tail circuit without the use of an echo canceller. Echo return loss enhancement (ERLE) is the reduction in echo level generated by the echo canceller. Acombined (ACOM) value is the sum of ERL and ERLE, or simply put, total echo return loss in the network (total reduction in echo signal level). These values are shown in Fig. 4.

3. THE IMPACT OF ECHO CANCELER

It is clear that the bigger the ERL is the smaller is the echo. Several types of problems can occur as the consequence of the insufficient ERL. If the ERL is small enough (less than 5 dB), ERL combined with ERLE (which is a constant value, usually between 20 and 30 dB) may not be enough to completely reduce the echo, so the echo is still audible. The other case is more rare but not less dangerous. We said in paragraph 2.1. that echo canceller stops improving the echo cancellation process (convergence) during the period of double talk (User B talks). We also mentioned a threshold by which echo canceller detects double talk. If ERL is small enough that echo reduced by ERL is still larger than the threshold value (usually 6 dB), even when User B is silent, the echo signal will be considered to be a part of the call and not an echo. The echo will be declared as double talk and the echo canceller will not have a chance to go through the convergence procedure and eliminate the echo.

Using the test network in Fig. 2, we tested the impact of using the echo canceler. User A is directly connected to the router GW1 through an analog port with matching impedances so it wouldn't cause any echo. User B is situated somewhere in the PSTN network to which the GW2 router

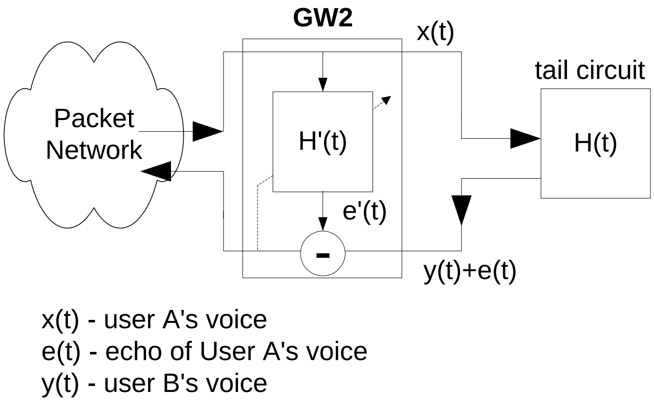

Fig. 3. Echo canceller operation

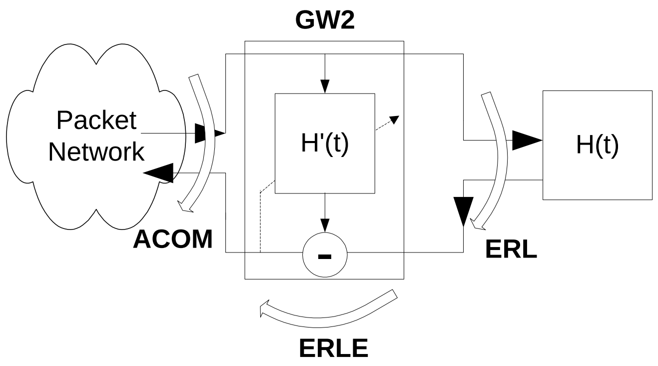

Fig. 4. Echo canceller operation

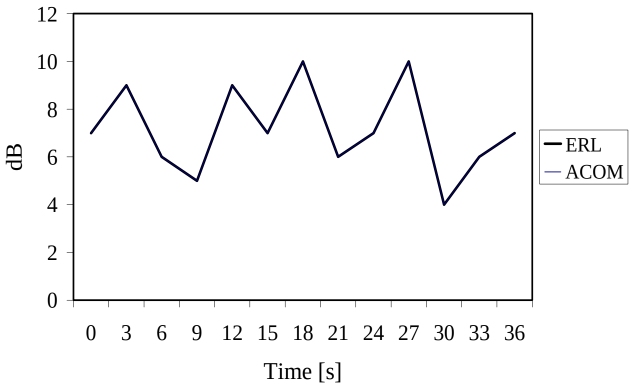

Fig.5. Level of ACOM=ERL without echo canceller

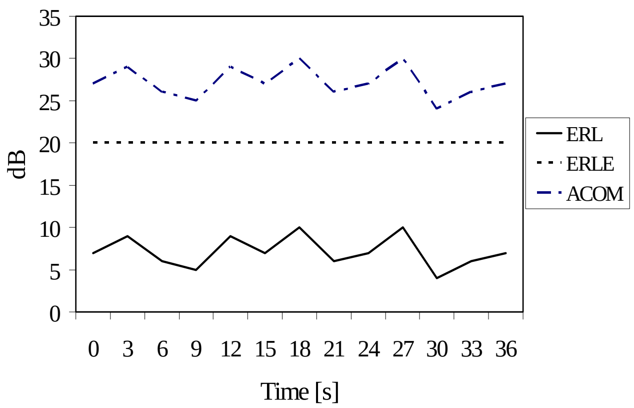

Fig.6. Level of ACOM with echo canceller

is connected. We put very slow links (32 kbit/s) between routers so we could get a delay, which will exacerbate the echo that is generated from the PSTN. In Fig. 5 we can see the amount of ACOM without the echo canceller, which is equal to the amount of ERL. In Fig. 6 we can see the amount of ACOM when using echo canceller. It clearly shows that only with the use of echo canceller it is possible to get the echo below the noticeable level (ACOM > 25 dB) [3-5].

We can see that with echo canceler, ELRE is constant, in the amount of 20 dB. When echo canceler on router GW2 is not used, ELR is from 4 to 10 dB. In the case with echo canceler, the amount is enough to make echo imperceptible to the user from which the speech was originated. This need for echo cancellers is very important in networks that VoIP links that connect several PSTN networks. In this case it is necessary for every gateway that faces the tail circuit to have echo cancellers enabled, for it is the only way to eliminate echo.

4. CONCLUSIONS

In order to fully eliminate the occurring echo when migrating from the PSTN to the VoIP network, every voice gateway that terminates the call towards the PSTN must use the echo cancelling techniques. If the echo is too long or too loud so the canceler understands it as doubletalk, the echo must be dealt with in the PSTN network elements (switches) so that the quality of a VoIP call can be comparable to the quality of a PSTN call. One more thing that needs a comment is the use of QoS (quality of service) techniques when dealing with echo [6]. QoS techniques can improve end-to-end network delay and the shorter the delay, the less annoying the echo becomes. However, it is never possible to reduce the delay so much that it will make echo inaudible, because even the minimum delay that is inherent in VoIP network is long enough for echoes to be noticeable. So, QoS techniques can help in other ways, but they can not eliminate echo.

REFERENCES

[1] M. Miller, "Voice over IP Technologies: Building the Converged Network", *Hungry Minds*, Indianapolis, IN, 2002.
[2] J. Davidson, "Voice over IP Fundamentals", *Cisco Press*, Indianapolis, IN, 2000.
[3] J. Davidson, "Deploying CISCO Voice over IP solutions", *Cisco Press*, Indianapolis, IN, 2002.
[4] S. Keagy, "Integrating Voice and Data Networks", *Cisco Press*, Indianapolis, IN, 2000.
[5] ITU-T Recommendation G.131 Appendix A: "Relation between echo disturbances under single talk and double talk conditions (evaluated for one-way transmission time of 100 ms)", 1999.
[6] B. Morgan, C. Dennis, "Cisco CCNP Remote Access Exam Certification Guide", *Cisco Press*, Indianapolis, IN, 2001.

AUTHORS: B.Sc. Vedran Srsen, Coting d.o.o., Unska 2b, 10000 Zagreb, Croatia, e-mail: vedran.srsen@coting.hr, Ph.D. Sandor Dembitz, Ph.D. Petar Knezevic, Faculty of Electronical Engineering and Computing, Unska 3, Zagreb, Croatia, e-mail: petar.knezevic@fer.hr.