February 18, 2008

High performance echo canceller for Asterisk VoIP systems

This article explains the general causes of echo and the methods of echo cancellation available for Asterisk. It introduces the High Performance Echo Canceller (HPEC) and compares it with the legacy Asterisk echo canceller.

By Scott Kurtz, Adaptive
Digital Technologies

For more on this topic, see our Asterisk VoIP PBX guide.

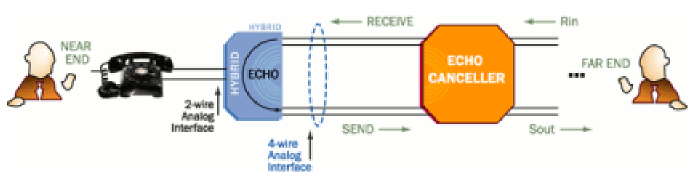
One hundred years ago, long before it became mainstream to talk about conservation and long before the so-called "carbon profile," a presumably wise engineer decided that voice signals could be carried in both directions of a telephone call over a single pair of wires. The ecological consequence of this is that we have saved uncountable tons of copper in the subsequent one hundred years. But ever since the time when the two-wire interface was introduced into the telephone local loop connection, we have had to deal with reflections in the telephone network.

Echo in the telephone network clearly preceded the introduction of the Asterisk VoIP system. Yet there are some aspects of Asterisk and VoIP in general that differentiate them from traditional telephone systems when dealing with echo.

This article begins with a brief overview of echo in the telephone network. Then we'll turn to echo cancellation as it pertains to Asterisk in particular. Then we will introduce the High Performance Echo Canceller (HPEC) that is now being deployed in Asterisk systems and compare it with the legacy Asterisk echo canceller. While comparing the legacy Asterisk echo canceller with the HPEC, we introduce many terms that need further explanation. The final portion of the article is therefore devoted to describing these features and the echo issues that necessitate the use of these features.

Echo in the telephone network

The reflection of a person's speech back into his or her earpiece is caused by hybrid circuits that convert between two-wire analog interfaces and four-wire analog interfaces, as seen in Figure 1.



(Click to enlarge)

Figure 1. The hybrid circuit in the telephone network that causes the voice signal to be reflected. Also shown is an echo canceller that is used to cancel the reflection.

At the four-wire side of the hybrid, one pair of wires carries voice signals toward the hybrid (sometimes called the receive path), and a second pair of wires carries voice signals away from the hybrid (sometimes called the send path). On the two-wire side of the hybrid, a single pair of wires carries voice signals in both directions. The echo comes about because hybrid circuits are not perfectly matched. As a result, some of the four-wire receive signal is leaked back into the four-wire send signal.

In the early days of telephony, the reflection was not yet an echo. For a person to perceive echo, the reflection must be delayed by more than a trivial amount before reaching the person's ear. As the delay approaches ten milliseconds, it begins to sound like a reverberation sound effect. When the delay exceeds twenty milliseconds, the reflection begins to sound like echo.

Echo during a phone call becomes intolerable very quickly. As the delay increases, it becomes less tolerable. But delay is not the only factor. The level of the echoed signal relative to the level of the person's speech (referred to as ERL or Echo Return Loss) is a second factor. A person may be able to tolerate an ERL of 6 dB if the delay is 15 milliseconds, but at 30 milliseconds, the ERL will have to be much better (higher) to be considered tolerable. (More specifically, ERL is defined to be the ratio of the person's speech to the level of the echo, expressed in dB. An ERL of 6 dB, for example, means that the amplitude of the echo is 6 dB below the level of the original speech.)

Referring back to Figure 1, the solution to the echo problem is the use of an echo canceller. The echo canceller in Figure 1 removes the reflection from the send signal without removing the near end speaker's voice.

Echo in Asterisk systems
For the purpose of this discussion we will discuss six types of phone connections supported by Asterisk:

- FXS – FXS
- FXS – FXO
- FXO – FXO
- FXS – Digital
- Digital – FXO
- Digital – Digital

They are shown collectively in Figure 2.

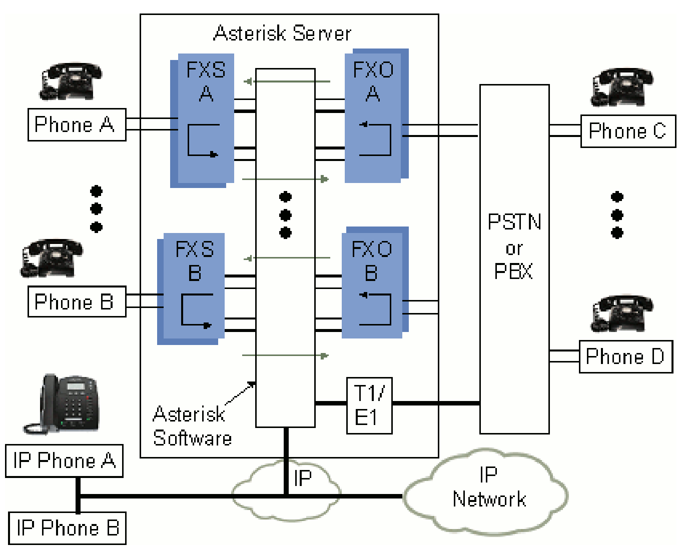


Figure 2. Asterisk fits into a telephone system with a variety of different interfaces.

Figure 2 shows three types of interfaces: FXS, FXO, and Digital. An FXS (Foreign Exchange – Subscriber) interface is a two-wire analog interface that connects to a telephone. An FXO (Foreign Exchange – Office) interface is a two-wire interface that connects to a telephone company central office. The primary differences between FXS and FXO are:

- FXS sources a battery feed. FXO sinks a battery feed.
- FXS provides ring signal. FXO detects ring signal.
- FXS detects off hook, pulse dialing, hook flash. FXO sources off-hook, pulse dialing, hook flash.

Both the FXS and FXO interfaces are sources of echo because both include a hybrid circuit that performs two-wire to four-wire conversion.

Digital interfaces include both IP and T1/E1 interfaces. Digital interfaces do not cause echo, but it is possible that an echo-producing hybrid may be located on the opposite side of a digital interface. For example, the echo could be on the opposite side of the PSTN (Public Switched Telephone Network) or on the opposite side of an IP network.

FXS to FXS
Let's start with a simple FXS to FXS connection. Analog phone A is connected to FXS port A via a 2-wire interface. The FXS port converts to a four-wire interface and connects with the Asterisk software via a digital link. The Asterisk software connects back to FXS port B, which converts the digital signal back to analog. FXS port B connects to analog phone B via a 2-wire interface.

In this case, there are two sources of echo, both at the FXS ports. Since both FXS ports are located in the same Asterisk server, it is possible that the round-trip echo delay is short enough not to be perceived. In such a scenario, one could argue that an echo canceller is not required.

FXS to FXO
Next, let's take the FXS to FXO connection. The path is similar to that in an FXS-FXS connection, but in the case of the FXO, the analog interface is connected to the PSTN. The delay through the PSTN can be short or long. The person at the opposite side of the PSTN may therefore hear echo caused by the FXS port's hybrid circuit. This echo must be cancelled by the Asterisk system, either in software on the server or using hardware on the FXS board. It is ironic that the local hardware (the Asterisk server) is responsible for taking care of the echo perceived at the other side of the network: the person paying for the Asterisk system does not benefit from this portion of the echo cancellation.

There is another possible source of echo in this scenario. The PSTN is supposed to cancel the echo that may occur on the right side of the PSTN (Figure 2). But what if it does not? Perhaps the PSTN is replaced by a PBX. Suppose further that the delay caused by the PBX did not warrant the use of an echo canceller. Then, by connecting the Asterisk system, the cumulative delay in the PBX plus the Asterisk system could become large enough to warrant echo cancellation on the right side of the diagram.

Although one could argue that the PBX ought to cancel that echo, a user would perceive that the Asterisk system, being the new hardware in the equation, introduced the echo and therefore should cancel it. The truth is that the Asterisk only pushed the delay into the perceptible range. Nevertheless, the customer is always right and the Asterisk echo canceller may have to do double-duty.

FXO to FXO
The FXO to FXO case is similar to the FXS to FXO case, except that in this case, the call is routed between two PSTN phones, introducing two legs of delay in the PSTN.

Digital Interfaces
The digital interfaces, including IP and T1/E1 interfaces, do not by themselves introduce echo. There may be echo at the opposite side of a digital interface if there is a hybrid at that side. We can argue that the equipment near the source of the echo, not the Asterisk server, should cancel that echo. But on rare occasions, using a similar argument to that used in the FXS to FXO case, the Asterisk system may need to take care of the external echo.

Hardware vs. software echo cancellation
The echo canceller can be run either in software on the Asterisk server, or in hardware on a DSP (Digital Signal Processor) chip on the FXS, FXO, or Digital interface boards. There are advantages and disadvantages to both approaches. The software echo canceller is ideal for systems that have a relatively low number of analog ports, and when the Asterisk server has sufficient processing horsepower left over for echo cancellation. By doing the echo cancellation on the Asterisk server, the cost of the

interface boards (FXS and FXO boards) is reduced.

The hardware based echo canceller is better when the port count is high, or when the Asterisk server CPU is needed for other compute-intensive operations.

Legacy Asterisk echo canceller vs. high performance echo canceller
The legacy Asterisk echo canceller is a basic textbook echo canceller that works reasonably well under relatively benign, controlled echo conditions. Echo cancellation is more of an art than a textbook exercise, and echo conditions are rarely controllable. The legacy echo canceller therefore sometimes requires additional tuning or training, practices that are not considered acceptable in the telecommunications community or by the International Telecommunications Union's (ITU) G.168 recommendation—the standard that governs the performance of echo cancellers. In fact, the legacy Asterisk echo canceller is not G.168 compliant.

The legacy echo canceller served well in the earlier days of Asterisk. But Asterisk has grown in ways that nobody could have predicted. It is being used all over the world, in numerous applications, and in conjunction with many different vendors' analog interface cards. Furthermore, it is interoperating with a wide range of telephone equipment, both infrastructure and handsets. The legacy echo canceller has become a victim of the overall success of the Asterisk product.

On the other hand, the HPEC is an industry-standard, carrier-class G.168 echo canceller. The HPEC qualified as toll-quality at AT&T's Voice Quality Assessment Labs. AT&T's Voice Quality Assessment Lab evaluated the HPEC using its stringent series of performance tests including AT&T's Mean Opinion Score (MOS) subjective tests as well as the standardized set of G.168 objective tests. The subjective and objective performance of the HPEC surpassed even the performance of AT&T's benchmark lab echo cancellers.

The importance of subjective testing should not be overlooked. Although the ITU's G.168 recommendation is well thought out, this type of objective test is no guarantee that an echo canceller will sound good to a human listener. Doing both subjective (MOS) and objective (G.168) testing ensures that the HPEC will deliver excellent voice quality.

The following table compares features included in the legacy Asterisk echo canceller and the HPEC.

Feature	Legacy EC	HPEC
G.168 Compliant		X
Maximum Tail Length	32 msec	128 msec
Tandem Free Operation		X
Nonlinear Processor	X	X
Dynamic Nonlinear Processor		X
Comfort Noise Generator		X
Rapid Convergence		X
Automatic Tail Search		X
Cancels multiple reflectors		X
Tone Disabler	X	X
SS7 Tone Detection		X
Double-talk Detection	X	X
Stationary Signal Detector		X

Table 1. Comparison of feature sets of the legacy echo canceller and the HPEC.

The overall importance of echo cancellation should also not be overlooked. From the user's perspective, echo is arguably the worst type of impairment that can be encountered during a telephone conversation. Typically, if a circuit has echo, the two parties agree to hang up and dial again because it is so difficult to speak in the presence of one's own echo. All Asterisk users would be well advised to use the Linux-based HPEC in conjunction with cards that do not provide hardware (DSP) based echo cancellation.

Echo characteristics and the echo canceller features that address them
Echo tail

As discussed earlier in this article, echo in the telephone network is caused by hybrid circuits that convert between two-wire analog interfaces and four-wire analog interfaces, as seen in Figure 1.

If we characterize a hybrid in terms of its impulse response (shown in Figure 3), we see that the impulse response tends to be non-zero initially, followed by a non-zero period that can be as long as 16 milliseconds. The impulse response is often referred to as the echo tail, and the duration of the echo tail is often referred to as the tail length. The tail length of the echo tail shown in Figure 3 is 8 millseconds.

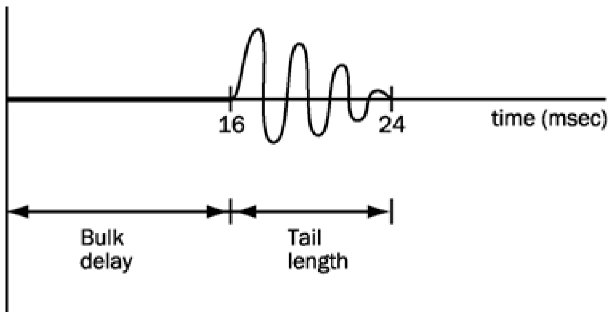


Figure 3. A time domain representation of an echo tail.

There are situations where the echo canceller is not located at the same place as the hybrid. It may be separated by one or more T-1 links or other types of links that cause the tail length to appear to be even longer from the point of view of the echo canceller. Furthermore, there may be more than one hybrid in the path—in an overall impulse response that is the concatenation of multiple impulse responses. In this case, each tail is sometimes referred to as a reflector, and the situation where there are multiple hybrids in a circuit is referred to as one where there are multiple reflectors.

Bulk delay
In rare situations, an echo canceller located on the opposite side of a VoIP link from the hybrid, may be expected to cancel the far-end hybrid's echo. In this case, there is considerable delay between the echo canceller and the hybrid in both directions. In this case, the echo tail appears to begin with a segment of relatively long segment of zero followed by the hybrid impulse response as shown in Figure 3. The duration of the segment of zeros is referred to as the bulk delay.

There are two ways to handle bulk delay. One is to place an artificial delay into the far end input to the echo canceller to effectively remove the bulk delay from the point of view of the echo canceller. This technique falls short for two reasons. First of all, it requires a-priori knowledge about the amount of bulk delay that will be encountered. Second, it does not allow for the situation in which there may be both a local reflector and a remote reflector. Because of both of these reasons, it may be preferable to use a second technique in which the entire possible delay window is analyzed, and any reflectors within that window are cancelled.

Echo cancellation
The primary job of an echo canceller is to remove the echo of the receive path that has bled through the hybrid into the send path. This is done by modeling the echo path with an adaptive filter, using that adaptive filter to predict the echo, and subtracting the predicted echo from the send signal. This is shown in Figure 4. But, as we will explain, there is far more to an echo canceller than meets the eye.

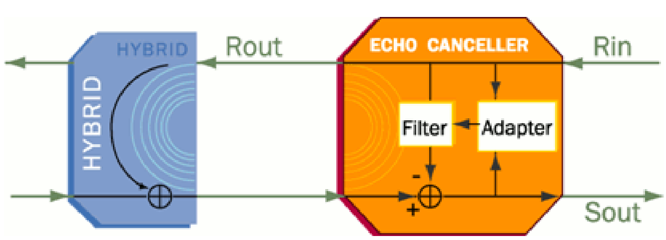


Figure 4. Role of an adaptive filter in echo cancellation.

The ideal situation for an echo canceller to model the echo tail is when the far end speaker is speaking and the near end speaker is silent. This condition is referred to as single-talk. The reason this is ideal is that the receive signal is used as a reference signal for the echo canceller for comparison with the send signal. If both the far end speaker and the near end speaker are speaking at the same time (a condition known as "double-talk"), the near end speech will be added to the echo thereby making it more difficult for the echo canceller to compare it to the reference. In fact, a double-talk condition can cause an echo canceller's adaptive filter to diverge. In order to prevent this from happening, echo cancellers employ double-talk detectors. When double-talk is detected, the echo canceller temporarily stops adapting its filter to prevent divergence.

Echo cancellers use adaptive filters that do not perfectly model the echo tail. As a result, the echo cancellers have to perform post-processing (Nonlinear Processing or NLP) to remove the residual echo caused by the imperfections. Similarly, the echo path may have some nonlinearities that the adaptive filter cannot perfectly match. The NLP is intended to take care of this also.

The NLP has to be designed carefully in order to minimize unwanted effects. The nonlinear processor is free to suppress residual echo at will in a single-talk situation. But if it is too aggressive during a double-talk situation it may suppress the near end speech—an undesirable effect.

Another more subtle effect of an NLP occurs when there is background noise at the near end. This background noise will be heard by the far end speaker when the NLP is not engaged, but the noise will be suppressed when the NLP is engaged. The appearance and disappearance of the background noise can be annoying. In order to take care of this, an echo canceller will often replace the residual signal with comfort noise while the NLP is engaged rather than blindly suppressing the signal. When an echo canceller has this feature, it is referred to as a comfort noise generator (CNG).

Beyond nonlinearities in the echo path, there can be other impairments that an echo canceller may need to deal with.

- PCM Slips: A PCM slip (as can occur on a T-1 or E-1 link) shifts the echo tail by one sample period (125 microseconds). This causes the echo canceller to reconverge its adaptive filter. During the period of reconvergence, echo will temporarily reappear.

- Residual Acoustic Echo: There may be some residual echo on the line if the near end speaker is using a hands-free phone that does not have a good acoustic echo canceller. In this case, the network echo canceller has an opportunity to attack this source of echo. There can also be some residual echo when using a non-hands-free phone. This echo is caused by acoustic coupling between the handset earpiece and microphone. The level of this echo is usually small.

- Packet Loss: If an echo canceller is placed on the opposite side of a packet network with respect to the hybrid, a lost packet causes an interruption in the echo path. This can cause an echo canceller to diverge.

The features that handle PCM slips and packet loss are available, but not currently enabled, in the HPEC.

Tandem free operation

Yet another scenario that an echo canceller should deal with is one in which there is another echo canceller in a circuit that is closer to the hybrid than the HPEC. In this case, the other echo canceller should remove the echo leaving our echo canceller the appearance that there is no hybrid in the circuit. Under these circumstances, an echo canceller could actually create echo rather than leave the echo-free signal alone. Being able to handle this situation properly is referred to as tandem free operation. A similar situation is one in which an echo canceller is placed on a circuit that does not have a hybrid—a digital circuit, for example.

Additional echo canceller requirements

The telephone system carries more than just voice signals. It can carry fax and modem signals, and signaling tones (such as DTMF tones and inter-office signaling tones). Passing these tones properly imposes additional constraints on an echo canceller. For example, an echo canceller must detect the presence of certain modems by identifying their answer tones (as specified by ITU G.165, V.8). When these modems are present, the echo canceller must disable itself for the duration of the modem connection. This feature is known as a tone disabler. Similarly, when certain interoffice signaling tones known as SS7 tones are present, the echo canceller must disable itself temporarily. Finally, an echo canceller is at risk of diverging in the presence of tonal or otherwise periodic, stationary signals. An echo canceller should be able to detect such conditions and temporarily stop adapting its filter in order to avoid diverging.

Summary

Echo cancellation is a necessary and important part of almost all voice communication systems. The Asterisk system brings with it some unique network topologies. With those topologies come some non-trivial echo cancellation requirements. While the legacy Asterisk echo canceller served well in the early days of Asterisk, that same echo canceller has fallen victim to the overall success of the Asterisk product. The Asterisk product has become part of mainstream telecommunication and, as a result, Asterisk systems now need higher performance, carrier class echo cancellation.

About the author

Scott Kurtz, Vice President, Adaptive Digital Technologies, Inc., is an expert in the field of Digital Signal Processing and Digital Communications. Beginning his 24 year career at RCA Corporation's Government Communications Systems Division, Scott moved on to InterDigital Communications Corporation Inc. where he was instrumental in developing the first digital wireless local loop telephone system which was the precursor to today's digital cellular telephone. In 1994 he joined Adaptive Digital, where he has since focused on building an engineering team and developing DSP based software products and technology to accelerate and improve product development. Scott holds fourteen patents in the areas of echo cancellation, and digital communication. He has published several articles and written many papers pertaining to the same. Scott earned his BS in Electrical Engineering from Lehigh University and his MS in Electrical Engineering from Drexel University.