

Implementing High-Quality Voice Solutions: The Technology Behind a Carrier Class VoIP Solution

Freescale Semiconductor, Inc.

The ITU, through its standardization activities, has indeed served the burgeoning IP Telephony industry well by providing a common baseline through which the industry could interoperate and flourish. In fact, the ITU has been so successful in its efforts that one might be tempted to believe that all IP telephony implementations are created equal.

However, once a Voice over IP Telephony solution undergoes acceptance testing prior to deployment in a carrier-class environment, it becomes clear that not all voice solutions are created equal despite the apparent adherence to ITU standards.

Surf has deployed VoIP solutions for carrier class implementations — in fact, Surf has become one of the leading suppliers for this market. These opportunities have given Surf a unique insight into many subtle quality and performance issues surrounding packet-telephony implementations. The following will discuss a few of the differentiating aspects that determine the perceived quality of the voice channel and the end-user satisfaction with the deployed implementation.

The heart of the voice channel is the voice coder/decoder (vocoder) that includes, for example, the ITU G.711, G.726, G.729, and G.723 algorithms. Even though the vocoders perform the necessary compression and analog to digital transcoding well, they do not compensate for other critical factors that affect perceived voice quality. In other words, a vocoder is a necessary component but not a complete solution for a carrier-class voice deployment.

The basic components of a voice channel are (ref. Fig 1) the vocoder, echo canceller, call progress tone (CPT) monitor and discriminator, voice activity detector and comfort noise generator (VAD/CNG), the jitter buffer and RTP/RTCP protocols.

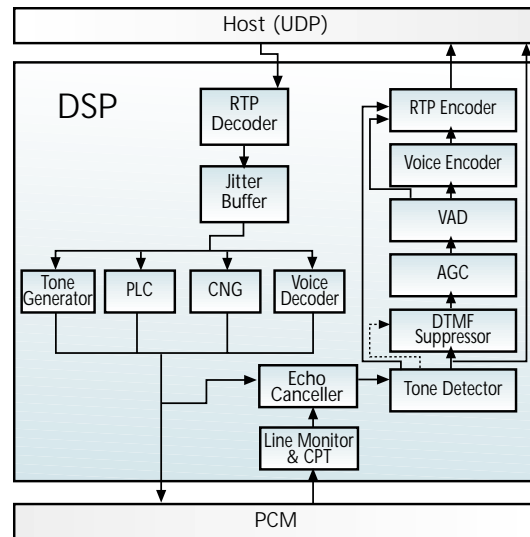


Figure 1: Basic Structure of Voice Channel

VAD/CNG

The purpose of VAD/CNG is to save network bandwidth by identifying silence in a phone conversation (voice activity detection) and temporarily suspending the vocoder's compression scheme by sending very low bandwidth comfort noise generation information instead. When considering a VoIP solution, the exact VAD/CNG implementation should be carefully scrutinized. The ITU's G.711 App II recommendation uses the Linear Prediction Coding algorithm to identify the transition from speech to silence, and vice versa, to maximize bandwidth savings and avoid undesired side-effects such as speech clipping. Implementing this algorithm correctly, along with fine-tuning the necessary parameters, can realize bandwidth savings in excess of 30% compared to other techniques such as spectral energy calculation. In environments of speech mixed with typical background noise, such as air-conditioning systems or computer ventilating fans,



Communication Solutions

Freescale Semiconductor, Inc.

non-ITU G.711 App II implementations will not identify correctly transitions from speech to silence.

The CNG is supposed to accurately reconstruct the background noises at the receiving end. Once again, simpler algorithms, as the spectral energy method, will distort the original background noise during the reconstruction procedure.

Therefore, simple spectral energy implementations of VAD/CNG will lead to excessive network bandwidth demands coupled with speech clipping and to distortions in the original noise that adversely affect perceived voice quality. The more complex LPC algorithm, recommended by the ITU, should be implemented to ensure true voice fidelity.

Surf's VAD/CNG implementation is compliant with the ITU G.711 App II standard.

are exacerbated in network topologies where the media gateway is deployed far away from the PSTN line-card leading to echo tails up to 128ms in length. Surf's low MIPS software implemented echo cancellation is based on proprietary signal processing techniques to ensure swift convergence, even in the presence of long 128ms echo tails and changing network topology.

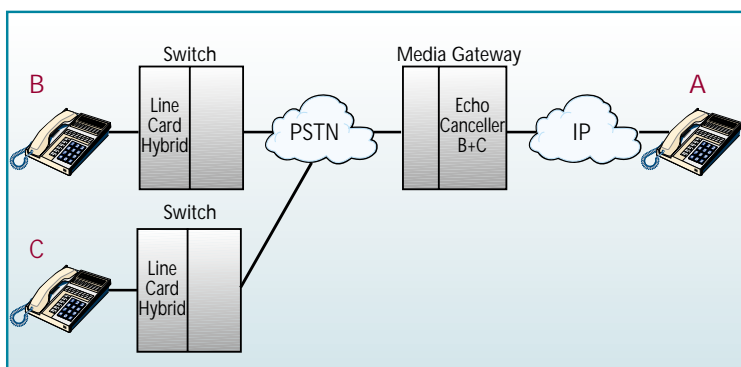


Figure 2: Changing Echo Topology as Subscribers A & B Have Subscriber C Join the Call

Echo Cancellation

While the ITU's G.168 echo cancellation requirements are well defined, the actual method of implementation can dramatically affect the perceived voice quality.

A brute force and straightforward implementation of the G.168 standard will result in very long echo cancellation filters that must converge in a reasonable amount of time. However, lengthy filters are slow to converge. This, of course, is unacceptable from a perceived voice quality standpoint. Furthermore, the echo canceller must be swift enough to react to instantaneous changes in the echo topology (arising, for example, from a 3rd caller entering into a conference call — ref. Fig. 2). These requirements

Adaptive Jitter Buffer and Packet Loss Concealment

A jitter buffer is required to overcome the inherent packet network conditions leading to variability ("jitter") in packet arrival time relative to transmission time. While a fixed jitter buffer may work adequately under deterministic conditions, it tends to perform poorly under the dynamic states commonly found in packet telephony networks.

The jitter buffer must be adaptive to provide the required depth for instantaneous network conditions without introducing needless delay emerging from over-buffering.

To maintain effective jitter buffer behavior, the incoming packets from the network must be synchronized with the outgoing packets. This is crucial when packets of varying sizes are handled. Recovery from a loss of voice packets, that minimizes the user's perceived voice quality degradation, mandates using a sophisticated Packet Loss Concealment (PLC) algorithm that should be implemented according to the ITU G.711 App I recommendation. Even though adhering to this standard increases the complexity of the solution and places higher computational demands on the DSP processor, other forms of PLC simply do not sound right in network environments, especially when the packet losses are beyond 1%. The recommended PLC algorithm uses aligned pitch phase interpolation that provides significantly better-perceived quality when compared to simpler algorithms. These simpler implementations include transmission of silence or repetition of previous packets.

Surf's voice implementation has a dynamically adaptive jitter buffer mechanism and implements PLC according to the ITU's G.711 App. I recommendation.

Call Progress Tone Monitoring, Discrimination and DTMF

The real challenge of call tone detection and generation is in meeting the rigorous and demanding conditions of the global telephone network. It is crucial to test the tone detection according to BELLCORE and ETSI standards, as a first step towards good tone and event detection. The next step requires actual field-tests to fine-tune the solution.

It is also critical to monitor both the network (IP) and the PSTN (TDM) sides of the voice channel to correctly detect and understand the tones.

Surf's solution has undergone BELLCORE and ETSI compliance testing, in addition to carrier-class field deployment acceptance testing.

An Integrated Solution

Although the various aspects of the complete voice channel have been discussed as separate components, one must emphasize the importance of correctly integrating all of the components into a unified and coherent voice channel solution. Without understanding the interworkings and cross-dependencies of the various voice channel components, it is impossible to achieve satisfactory perceived voice quality. For example, the echo cancellation and VAD/CNG algorithms must take each other into account to provide satisfactory performance.

Surf, the leader in converged access solutions, provides all of the voice channel components discussed above as a complete, integrated solution. Surf's solution adheres to the most rigorous ITU recommendations and has been field tested and deployed in carrier-class IP Telephony implementations.



Surf Communication Solutions Inc.
2 Clock Tower Place, Suite 620
Maynard, MA 01754, USA
Tel: (978) 461 1180 Fax: (978) 897 4101
Website: www.surf-com.com

Surf Communication Solutions Ltd.
Tavor Building
Yokneam Industrial, Park 20692, Israel
Tel: 972-4-909 5777 Fax: +972-4-959 4055
Website: www.surf-com.com

Surf Communication Solutions, Ltd. reserves the right to make changes to the product(s) or information contained herein without notice. No liability is assumed as a result of their use or application. No rights under any patent accompany the sale of any such product(s) or information.

Copyright © 2001 Surf Communication Solutions, Ltd.
All Rights Reserved