Ditech Networks

| WHITE PAPER |

Cost Effective Transcoding for Carriers

September 2006

Contents

Executive Summary

Voice over IP (VoIP) technology has gained enormous popularity over the past few years. The reason for this recent and rapid increase in popularity is that VoIP is finally delivering on its promise of new services and cost savings through network convergence. All the pieces of the VoIP puzzle are mostly in place: quality of service (QoS), signaling protocols, media servers, feature servers, media gateways, softswitches, session border controllers, IP PBX systems, and IP Centrex services. An IDC report titled "U.S. Residential VoIP Services 2006-2010 Forecast and Analysis: Where There Is Smoke, Is There Fire?" predicts that the number of VoIP subscribers in the U.S. alone will grow from 10 million in 2006 to 44 million in 2010. If this forecast rings true, VoIP will be used in nearly two-thirds of broadband households by the end of the decade—and businesses are expected to adopt VoIP even more aggressively.

However, in one area there are too many pieces in the VoIP puzzle—specifically concerning codec standards. Codecs are an encoding/decoding technology used to convert analog voice to digital data streams, and vice versa. This allows for transport of voice in a circuit-switched time-division multiplex (TDM) network, or as packets in an IP network. Initially, there was a single, universal codec standard—ITU-T G.711 from the International Telecommunication Union—for the public switched telephone network (PSTN). Today, there are more than 20 codec standards used in various networks, both wireline and wireless. As innovation continues to improve, new codec technologies deliver better voice quality while using less bandwidth. As time progresses, the number of codecs will continue to increase.

Carriers are struggling to deal with this growing number of codec standards. To date, there have been only two options for providers, neither serving as optimal scenarios for commercial success.

The first option is to support all—or most—codec standards throughout the entire VoIP infrastructure, including customer endpoints as well as provider-based media gateways, softswitches, and media servers. This choice is prohibitively expensive because it requires equipping every media system with support for all codec types on all ports.

The second option is to standardize on a single codec type to help reduce costs and operational complexity by simplifying the media requirements. This choice limits the market potential of VoIP offerings, because calls using unsupported codec types would not be connected, and certain applications may lack sufficient bandwidth for a workable service.

Now there is a third, more cost-effective alternative for carriers to consider. Known as codec transcoding, this technique converts one type of encoding (codec standard) to another, and is the best option to achieve interoperability among all codec standards from different endpoints. Transcoding can be performed in a variety of different devices. However, the best approach is to deploy a purpose-built media processing system.

This paper provides an overview of codecs and the growing suite of codec standards and then assesses the effectiveness of the various transcoding options. Finally, it outlines the characteristics necessary for deploying a carrier-class transcoding solution. The material is intentionally non-technical, and additional information on codec transcoding and related topics is available at www.ditechnetworks.com.

The Growing Suite of Codec Standards

The need for voice codecs dates from the time when the fully digital PSTN began replacing the original analog plain old telephone system (POTS). The PSTN's standard DS0 channel established the data rate for a digitized voice channel. At the time, the most innovative analog-to-digital encoding available was a technique called Pulse Code Modulation (PCM). PCM operated at a rate of 8000 samples/second with 8 bits per sample, resulting in a bandwidth utilization of 64 kbps—a single DS0. Two different encoding algorithms were used. The first was µ-law, used in North America and Japan; the second was A-law, which was used in Europe and the rest of the world. To ensure global interoperability, in 1972 the International Telecommunication Union standardized these algorithms as ITU-T Recommendation G.711. The standard was most recently updated in 1989.

Much has changed since then. There are now over 20 codec standards in use today. All of these standards are designed to take advantage of advanced technologies that aim to balance three separate objectives:

1. Reduce the bandwidth utilization
2. Improve voice quality
3. Minimize the processing power required by the codec algorithm

Reducing Bandwidth Utilization

The advantages of reducing bandwidth consumption are compelling. With less bandwidth consumed by each voice session, the infrastructure is able to accommodate more calls. This allows carriers to maximize the statistical multiplexing of packet-switched networks—the factor primarily responsible for VoIP's low cost and growing popularity. Some codec standards pre-date VoIP, having been created to increase the number of conversations that could be carried in TDM networks by making better use of limited wire availability, particularly in the transatlantic cable. However, most of the newer standards are designed to optimize VoIP calls, particularly in wireless networks where bandwidth is not as generous.

Improving Voice Quality

Improving voice quality also remains a critically important characteristic. Although people have learned to accept marginal voice quality in order to gain mobility with cellular services, quality continues to be a key competitive differentiator among carriers. "Can you hear me now?" remains a memorable advertising slogan of a leading wireless carrier. With voice, quality is subjective. This means that what users perceive as good quality is far more important than what measurable network metrics, such as packet throughput and loss, define as good quality. Long ago the telephone industry established the Mean Opinion Score (MOS) to determine what constitutes acceptable quality to a user, and would pay test subjects to subjectively assess voice quality in a lab setting.

In VoIP networks, the quality of the received analog signal is a product of:

- The encoding at, or near, the sending end of the conversation
- Decoding at the receiving end
- The artifacts of the transport in between, including delay, jitter (variations in delay), packet loss and, of course, any transcoding.

For example, the voice quality of G.711 using Pulse Code Modulation (PCM) at 64 kbps is typically around MOS 4.4 (very good), whereas G.723.1 using Code Excited Linear Prediction (CELP) at 5.3 kbps is at MOS 3.0, which is unsatisfactory for most users. Another scale for measuring voice quality today is the R-factor, described in ITU-T Recommendation G.107. The wider R-factor scale, from 0 to 100, makes it more effective for assessing codec voice quality.

Minimizing Processing Power

The third and final objective is to minimize the processing power required for coding and decoding analog and digital signals. In general, the lower the data rate, the more complex the codec. For example, a G.711 (PCM) codec operating at 64 kbps requires as little as two-tenths of a MIPS (millions of instruction per second) of processing power, whereas a G.723.1 (Algebraic CELP) codec operating at 6.3 kbps might require up to 20 MIPS. Naturally, the requirement for more processing power increases costs, and can also introduce more latency. In this example, G.711 has a latency of less than one millisecond compared to G.723.1's of 30 milliseconds.

Balancing Objectives

The G.729 specification, popular in VoIP applications, serves as a good example of balancing these three objectives. G.729 consumes only 8 kbps of bandwidth—one-eighth that of G.711's 64 kbps. Extensions, 11.8 kbps or 6.4 kbps respectively, support marginally better or worse quality in exchange for bandwidth. The G.729a specification is a compatible extension that requires less computational power, but results in reduced voice quality. G.729b and G.729ab add "silence suppression," to improve throughput and enhance the subjective quality of calls; generating background "comfort noise" also keeps calls from being dropped prematurely during lulls in a conversation.

Many of the newer codec standards are targeted at specific applications. For example, some are optimized for wireless-to-wireless communications, such as the Enhanced Variable Rate Codec (EVRC) standard used in cellular Code Division Multiple Access (CDMA) networks. EVRC offers a "full rate" of 9.6 kbps, along with a "half rate" of 4.8 kbps and even an "eighth rate" as low as 1.2 kbps—which is really only intended for background "comfort noise". Overall average bandwidth utilization is approximately 6 kbps—still impressive at just 10% of G.711 bandwidth for acceptable voice quality (by cellular standards, that is). EVRC was later superseded by Selectable Mode Vocoder (SMV). SMV is now destined to be replaced by 4th-Generation Vocoder (4GV), which is a next-generation 3GPP2 standard based on the EVRC-B codec—an enhanced "B" version of EVRC. Such is the dynamic world of codecs.

Another example of application-specific codecs is the Internet Low Bit Rate Codec (iLBC), intended for packet-based communications. The most important characteristic of this codec is its lack of inter-packet dependency, which makes sessions quite tolerant of lost packets. For this reason, iLBC is suitable for audio streaming,

High Definition Voice

Among the reasons for VoIP's increasing popularity is its ability to support multiple codec types across different networks. With this codec flexibility, carriers can select different codec types for different customer applications.

One exciting example is with the use of wideband codecs. Wideband codecs offer users a "better than PSTN quality" experience to improve the total fidelity of a call. This is done by increasing the sampling rate to allow a larger frequency range to be transmitted between callers. High definition (HD) voice allows carriers to offer a new differentiated service to its customers.

What this means for codecs is a reversal of the recent trends to decrease bandwidth utilization and processing power demands. Instead, with high-definition voice, improving quality is the sole objective. The G.722 series of codec standards, for example, are notable for their ability to produce voice quality superior to that available in the PSTN. Other wideband codecs include Speex WB, iSAC, BV 32, and others.

archiving and messaging applications, in addition to VoIP. The encoding algorithm is a variant of block-independent linear predictive coding with a controlled response to packet loss, delay and jitter. It is standardized by the Internet Engineering Task Force (IETF) in RFC 3951. Each session utilizes between 13.33 and 15.2 kbps with a processing demand similar to G.729a. Both Skype and Google Talk employ iLBC codecs.

These few examples demonstrate the myriad of trade-offs between bandwidth, quality, and processing power. Combined with the need for certain application-specific designs, it is clear why there are more than 20 codec types now in use (see Current Codec Standards table). Most of these are official standards based on ITU-T specifications. Others are standards established by vendors or by multi-vendor organizations and initiatives. The list will continue to grow as DSP technologies advance and new VoIP applications emerge.

Current Codec Standards

Standard(s)	Bit Rate	Encoding Technique
G.711	64 kbps	μ-law Pulse Code Modulation (PCM)
G.711	64 kbps	A-law PCM
G.721	16/24/32/40 kbps	Superseded by G.726
G.722	48/56/64 kbps	SB-ADPCM (Adaptive Differential PCM)
G.722.1	24/32 kbps	MLT (Modified Lapped Transform)
G.722.2 (GSM-AMR WB)	6.60/8.85/12.65/ 14.25/15.85/18.25/ 19.85/23.05/23.85 kbps	ACELP (Algebraic Code Excited Linear Prediction) 3GPP Adaptive Multi Rate WideBand
G.723.1	5.3/6.3 kbps	MP-MLQ, and CELP (Code Excited Linear Prediction)
G.726	16/24/32/40 kbps	ADPCM
G.727	16/24/32/40 kbps	G.726 optimized for PCME (Packet-Circuit Multiplex Equipment)
G.728	16 kbps	LD-CELP (Low Delay CELP)
G.729a	8 kbps	CS-ACELP (Conjugate-Structure-Algebraic CELP)
G.729b/ab	8 kbps	With Silence Suppression
G.729e/g	8/11.8 kbps	μ-law or A-law PCM
Speex	2-44 kbps	CELP
IBLC (Internet Low Bit Rate Codec)	13.33/15.2 kbps	Block-independent Linear Predictive Coding
ISAC (Internet Speech Audio Codec)	10-32 kbps	Proprietary
GSM-HR (Half Rate) GSM-FR (Full Rate) GSM-EFR (Enhanced)	5.6 kbps 13 kbps 12.2 kbps	VSELP (Vector Sum Excited Linear Prediction) RPE-LTP (Regular Pulse Excitation Long Term Prediction) ACELP (Algebraic Code Excited Linear Prediction)
EVRC (Enhanced Variable Rate Codec)	1.2/4.8/9.6 kbps	RCELP (Relaxed Code Excited Linear Prediction)
SMV (Selectable Mode Vocoder)	2/4/8.5 kbps	eX-CELP (eXtended CELP)
4GV (4TH Generation Vocoder)	1.2/2.4/4.8/9.6 kbps	RCELP
AMBE	2.0-9.6 kbps	Advanced Multi-Band Excitation
BV16/32 (BroadVoice)	16/32 kbps	CELP optimized for Packet Cable

Transcoding Options—A Qualitative Cost Comparison

The growing list of codec standards is both good and bad news for carriers. The good news is that newer standards consume far less bandwidth, allowing the existing network infrastructure to accommodate more VoIP traffic. Other good news is that VoIP is being widely adopted across:

- Different devices—from PCs to multi-mode smartphones
- Different networks—LANs, WLANs, and cellular
- Different applications—such as voice messaging and high-definition voice (see High Definition Voice sidebar) for both residential and business subscribers

The bad news is the sheer number of codec standards involved. Fortunately, codec transcoding systems offer a solution to this growing problem. Codec transcoding involves the digital-to-digital conversion from one encoding technique (codec standard) to another. When properly implemented and deployed, it is transparent to users and interconnects with the rest of the VoIP network infrastructure.

To date, most carriers have chosen to "standardize" on G.711 for it's ease of deployment and low cost. Although this approach seems rational, the reality is that differing applications require different codec types for voice service and a "one-size fits all" approach will not satisfy the needs of the customer. In addition, this approach becomes limiting when carriers look to converge networks across circuit and packet networks in the wireline and wireless space. Codec transcoding options become even more important to minimize operational complexity and cost while maximizing revenue potential.

There are currently four possible options for performing codec transcoding in the IP network. These devices are located at points in the network where codec transcoding becomes a logical possibility for deployment.

- Media Gateways—The media gateway is designed to serve as the gateway function from TDM to IP networks. It naturally resides at the TDM to IP border and serves to transcode and packetize at this border. Typically, these devices have been designed around support for G.711 and are limited in their capabilities to support more MIPS-intensive codec types. Also, their proximity at the TDM to IP border may not be optimal given that an IP-to-IP call would need to traverse the network to the media gateway for transcoding treatment before being backhauled to the opposite VoIP caller. For these reasons, the media gateway is not the best option for universal network support of codec transcoding.
- Media Servers—Media servers reside in the backbone of VoIP networks and serve to provide announcements and conferencing services to the network. It is a centralized resource for support of these features to make them available network-wide. With their centralization in the IP network, media servers are plagued with backhaul issues when serving as a codec transcoding platform. Their architectures are also not optimized around performing this function so density and economics are not attractive to most carriers.
- Session Border Controllers—SBCs sit at the carrier's access and peering border and are located in an optimal spot to perform transcoding (at the ingress/egress point). Although their location in the network is optimal, their design and economics for providing this function are not. Since they are designed more for providing security and network

connectivity, they are limited in the available MIPS resources to provide scalable and economical codec transcoding functions. Also, adding transcoding functions to this device takes away from the density of SBC sessions that are needed. This can harm the economic proposition for the SBC functions as a result.

- Media Processing Platforms—Media processing platforms provide an optimal value proposition for both network location as well as scalability and economics. The media processing platform is deployed in the access or peering border and performs codec transcoding functions. The platform is purpose-built to provide a tremendous level of MIPS for transcoding functions thus is able to provide a very scalable platform with attractive economics. The media processing platform becomes the best of the transcoding options available to carriers.

Aggregation and peering boundaries are strategic deployment locations because carriers use these borders to control their own domains. They have little or no control over what occurs outside of their network. Also, the access and peering borders are where all VoIP sessions enter or exit the network, making this a low latency place to perform transcoding. This is why transcoding at the access and peering borders becomes beneficial for carriers.

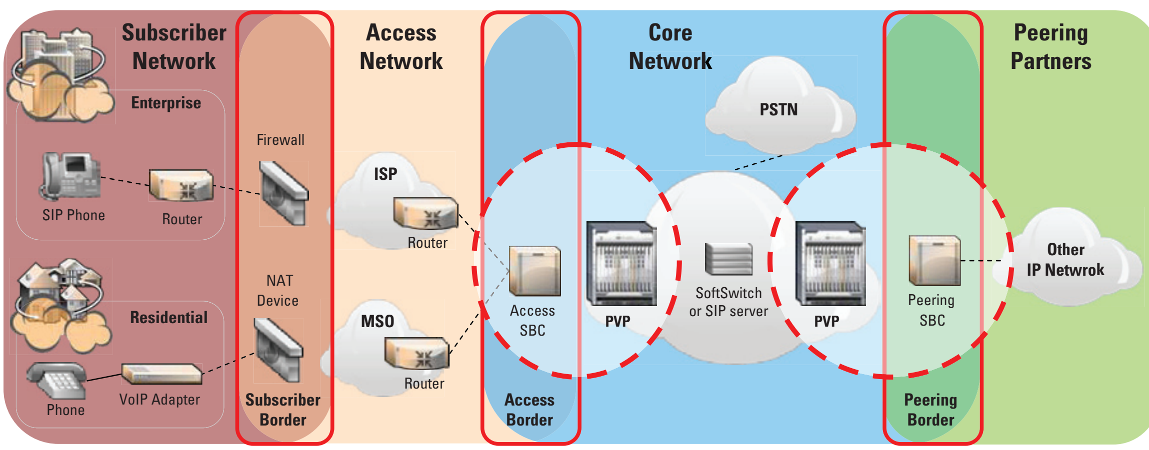

Figure : : Deployment Scenarios

The ideal location for transcoding is at the network's ingress and egress points. In a VoIP network, these points occur at local access or aggregation borders, and at the peering border interconnected at the IP backbone.

Requirements for Codec Transcoding Platforms

There are six critical requirements for a cost-effective transcoding system. Only a solution that supports all six requirements should be considered a true carrier-class transcoding solution. It must include:

1. Comprehensive list of vocoders—As the number of codecs continue to grow, carriers require a codec transcoding system that can support a number of codec types. This list should include the standard G.7xx series codecs as well as newer VoIP codecs such as iLBC, iSAC, Speex, and others. The system should also be able to accommodate new codec additions over time.
2. Complete "any-to-any" codec transcoding—This is an area where most SBCs and media server systems fall short of providing carrier-class capabilities. Anything less than any-to-any transcoding is certain to limit service offerings and, in turn, limit the return on investment in infrastructure.
3. An extensible design with DSP pooling for high performance—With a long and growing list of codecs, it is essential that the transcoding platform have an extensible design capable of supporting a framework for additional MIPS processing capabilities DSP pooling is the best way to achieve extensibility with full investment protection—and without sacrificing performance—as new codec types are added and additional densities are required.
4. High density and scalability—Continued strong growth in VoIP services is a certainty. The service providers prepared to accommodate this growth with competitive offerings will be the ones reaping the rewards in both the number of subscribers and average revenue per user (ARPU). To keep pace cost-effectively, the transcoding system should provide a high-density scalability of up to 50,000 sessions per rack. Ideally, the system is intelligent enough to process only those calls that require transcoding, allowing it to be over-subscribed and support even more sessions.

Making the Most of Transcoding with Voice Quality

While transcoding may be necessary, it may no longer be sufficient in a truly competitive VoIP service offering. Indeed, many VoIP services fail to achieve the level of quality people have learned to expect from a long history with the PSTN. Some of the challenge here involves processing real-time voice communications in packet-switched networks. Most of these challenges, however, can be addressed with a sufficiently robust converged infrastructure (ample bandwidth, class of service prioritization, etc.).

The primary source of VoIP quality problems today is often introduced from the user endpoints (e.g. softphones, speakerphones, and IP phones) and include a myriad of quality problems including echo, noise, static, and improper voice levels.

For this reason, the process of transcoding lends itself perfectly to additional conditioning of the VoIP traffic. Here are some examples of synergistic voice assurance capabilities offered on media processing systems:

- Canceling echo (both hybrid and acoustic)
- Matching and normalizing signal levels
- Removing noise and distortion
- Repairing damaged packets
- Reconstructing missing packets that have been dropped
- Improving voice intelligibility
- Monitoring and measuring actual and perceived voice quality in real-time

5. 99.999% availability for dependable operation—The PSTN has earned an enviable reputation for total dependability. At any time under any circumstances, users expect to pick up a phone and hear a dial tone. To compete, VoIP services need to be just as dependable. As a potential single point of failure that would affect tens of thousands of users, the transcoding system needs to offer PSTN-like resiliency and reliability.

6. Ability to provide additional functionality beyond transcoding—With pooled DSP technology, it makes both technical and economic sense to perform other VoIP processing functions on a transcoding platform. Most of these synergistic functions involve conditioning the encoded data streams in a variety of ways that together, are capable of enhancing voice quality by reducing quality impairments (see Making the Most of Transcoding with Voice Quality sidebar).

Another consideration may ultimately affect the transcoding options available to carriers. Whether defined by the IP Multimedia Subsystem (IMS), the Telecoms and Internet converged Services and Protocols for Advanced Networks (TISPAN) or one of their many predecessors, VoIP architectures are all fully distributed with partitioned functionality. Partitioning has served the industry well as a means of gradually and cost-effectively phasing in VoIP services. As certain functions mature, consolidating multiple functions onto a common platform often makes technical and economic sense. The role of the session border controller, for example, is now converging on a set of functions generally accepted throughout the industry. It is reasonable to expect that such consolidation may someday involve all alternatives presented here into a Border Services Platform.

Conclusion

In a VoIP network, transcoding has become necessary. In carrier VoIP networks, transcoding needs to be sufficiently comprehensive to serve the widest possible target market. Now there are more ways to maximize transcoding's usefulness.

A purpose-built media processing platform affords six important advantages—a large list of supported codecs; any-to-any transcoding; an extensible design; high performance; high density and scalability; 99.999% reliability; and an opportunity to enhance overall voice quality.

The Ditech Networks' Packet Voice Processor™ is purpose-built for transcoding and voice quality assurance for VoIP traffic. Its extensible platform supports the industry's most comprehensive list of codec types, enabling the system to deliver an industry-leading low total cost of ownership. To learn more about how your business can benefit from the Packet Voice Processor, visit Ditech Networks at www.ditechnetworks.com, or call 1-650-623-1300 (1-800-234-0884 in the USA and Canada).

About Ditech Networks

Ditech Networks supplies voice processing equipment for telecommunication networks around the world. Ditech Networks' technology solutions include voice, media processing, SIP, and security delivered on carrier-grade scalable platforms to enhance the delivery of communications services over mobile, Voice over IP, and wireline networks. Ditech Networks' customers are premier network operators including Verizon, AT&T, Orascom Telecom, and others that collectively serve more than 150 million subscribers

Corporate Headquarters

US Sales – West
825 East Middlefield Rd.
Mountain View, CA 94043
1-800-234-0884 (Toll Free)
1-650-623-1300 (Direct)
1-650-564-9599 (Fax)
NA-Sales@ditechnetworks.com

US Sales – East and Central
4975 Preston Road, Suite 675
Plano, TX 75093
1-972-985-6505 (Phone)
1-972-985-6900 (Fax)
NA-Sales@ditechnetworks.com

Canada Sales Office
2275 Lakeshore Blvd West,
Suite 500
Toronto, ON Canada
M8V 3Y3
1-416-255-7776 (Phone)
NA-Sales@ditechnetworks.com

Mexico Sales Office
Blvd. Avilla Camacho, 36 – 10th floor
Col Lomas de Chapultepec – CP 11000
Mexico DF
+52 55 9171-1630 (Phone)
+52 55 9171-1699 (Fax)
Mexcio-Sales@ditechnetworks.com

South America Sales Office
Praia de Botafogo, 501 – 1st Floor
Torre Pao de Acucar, Botafogo
Rio de Janeiro, 22250-040
Rio de Janeiro, Brazil
+55 21 2586-6353 (Phone)
+55 21 2586-6001 (Fax)
South-America-Sales@ditechnetworks.com

Europe Sales Office
Feringastrasse 6
85774 München, Germany
+49 (89) 99216-493 (Phone)
+49 (89) 99216-200 (Fax)
Europe-Sales@ditechnetworks.com

Eastern Europe Sales Office
Wisniowy Business Park – Building E II floor
ul. Hzecka 26, 02-153
Warsaw, Poland
+48 22 575-7034 (Phone)
+48 22 575-7001 (Fax)
E-Europe-Sales@ditechnetworks.com

Middle East/Africa Sales Office
21 El-Fawakeh St., 3rd Floor
Dokki, Giza 12311, Egypt
+20 2 336-5100 (Phone)
+20 2 761-8964 (Fax)
MEA-Sales @ditechnetworks.com

South Asia Sales Office
No. 8/11, Sarvapriya Vihar
New Delhi – 110016
India
+91 9810372555
South-Asia-Sales@ditechnetworks.com

South-East Asia Sales Office
Lippo Plaza 3rd Floor
Jl Jend. Sudirman Kav. 25
Jakarta 12920
+62 (21) 5291 3780 (Phone)
+62 (21) 522 1977 (Fax)
SE-Asia-Sales@ditechnetworks.com

Thailand Sales Office
23/F M Thai Tower – All Seasons Place
87 Wireless Road, Lumpini
Phatumwan, Bangkok 10220 Thailand
+66 (0) 2 627 9317 (Phone)
+66 (0) 2 627 9001 (Fax)
Thailand-Sales@ditechnetworks.com

China Sales Office
Unit 3010, No. 500
Xiangyang South Road
Xuhui District,
Shanghai
200031 PRC
+86 (21) 5456 0305
China-Sales@ditechnetworks.com