Dialogic

White Paper

The Growing Importance of HD Voice in Applications

Executive Summary

A new excitement has entered the voice communications industry with the advent of wideband audio, commonly known as High Definition Voice (HD Voice). Although enterprises have gradually been moving to HD VoIP within their own networks, these networks have traditionally remained "islands" of HD because interoperability with other networks that also supported HD Voice has been difficult.

With the introduction of HD Voice on mobile networks, which has been launched on numerous commercial mobile networks and many wireline VoIP networks worldwide, consumers can finally experience this new technology firsthand. Verizon, AT&T, T-Mobile, Deutsche Telekom, Orange and other mobile operators now offer HD Voice as a standard feature. Because mobile users tend to adopt new technology rapidly, replacing their mobile devices seemingly as fast as the newest models are released, and because landline VoIP speech is typically done via a headset, the growth of HD Voice continues to be high and in turn, the need for HD-capable applications will further accelerate.

This white paper provides an introduction to HD Voice and discusses its current adoption rate and future potential, including use case examples which paint a picture that HD Voice upgrades to network infrastructure and applications will be seen as important, and perhaps as a necessity to many.

Table of Contents

What Is HD Voice?

Voice quality is poised to be a significant differentiator as competition becomes more intense among service providers, enterprises and mobile operators. This is where HD Voice comes in. HD Voice uses wideband audio coding technology (codecs) to more accurately reproduce the human voice, resulting in significantly more natural sounding speech. Some have likened hearing HD Voice to a sense of actually being in the same room with the person on the other end of the phone line. Many also feel that HD Voice gives them an increased ability to recognize people by their voices and to understand highly accented speech more easily.

One reason new users sense such a marked improvement in quality with HD Voice is that traditional telephony is constrained by dated standards. In fact, even today's digital telephony standards (e.g. ITU-T G.711) are based on 1960s digital circuit technology and 1930s microphone technology! Until the advent of HD Voice, G.711 was the standard of voice quality, with mobile telephony typically providing even less than G.711 quality.

What do these limitations "sound" like? A major issue is that older standards limit the range of audio frequencies to ~300Hz to 3400Hz. The drawbacks in adhering to this narrow range are:

- Difficulty recognizing "fricative" sounds like "s" and "f" (higher frequencies needed)
- Problems distinguishing "m" from "n" and "p" from "t" (again, higher frequencies needed)
- Inability to hear the fundamental resonances in spoken vowels (lower frequencies needed)

The resulting "telephone speech" sounds harsh and unnatural, but with the wider range of frequency with HD Voice, the limitations of the narrow range of frequencies are removed - even on mobile phones.

Where is HD Voice Being Deployed?

In the enterprise, HD Voice usage has been spreading, akin to business telephone system's adoption of VoIP technology. Although telephone handsets from Avaya, Cisco, Grandstream, Gigaset, Polycom, Snom, and others support wideband audio and incorporate a variety of higher-quality audio components, and because headsets connected to Unified Communications applications such as Microsoft Skype for Business support HD Voice, many enterprise IP telephony systems operate as islands of VoIP. They connect to calls on the traditional PSTN, which, as noted above, adheres to dated voice standards and operates within a narrow frequency range. As a result, HD Voice is often restricted to internal connections on the network.

On the wireless networks, Voice over LTE (VoLTE) has seen over 74 successful commercial operators launched in 41 countries[1]. As a standard feature of VoLTE, HD Voice is widely supported on Android and iOS mobile devices.

VoLTE and Wi-Fi calling (using consumer Wi-Fi to provide last-mile connections) extends HD Voice into previously hard-to-access areas and off-load traffic from congested cell towers.

Many wireline network operators are also offering HD Voice as a differentiator.

Web-based WebRTC applications are also expected to fuel the adoption of HD Voice. With wide-band OPUS, VP8 and VP9 codecs being native to WebRTC, virtually all applications leveraging WebRTC will provide a rich HD Voice experience.

Use Case Examples

Person to Person Calling

The first use case example is person-to-person calling. In the enterprise, if one person is calling another via the same PBX or Unified Communication application that supports HD Voice using either a headset or a desk phone that also supports HD Voice, then it will be an HD Voice conversation. If a call is within the same mobile network that supports HD Voice with mobile phones that also support HD Voice, then the conversation will be HD. If a call is from one network operator that supports HD Voice to another that supports a HD Voice peering arrangement, then it will likely result in an HD Voice call (For information on HD Voice peering, see the section on *Implementation Issues*).

[1] *New GSMA VoLTE Specification Removes Market Fragmentation – GSMA Blog, July 28, 2016*

Communications Applications

A second use case example is utilizing a communications application like conferencing, voice mail, IVR, call logging or contact center. As shown in Figure 1: In order for communications applications to provide a HD Voice experience for both wireline and wireless callers, the applications must support wide-band codecs, including G.722.2. AMR-WB, OPUS, EVS or others. In situations where the end-point and application do not support a common wide-band codec, the application must provide a means to convert (transcode) between the end-point's codec and the native codec supported by the application. Transcoding is a function facilitated by media servers, SBCs and some media gateways.

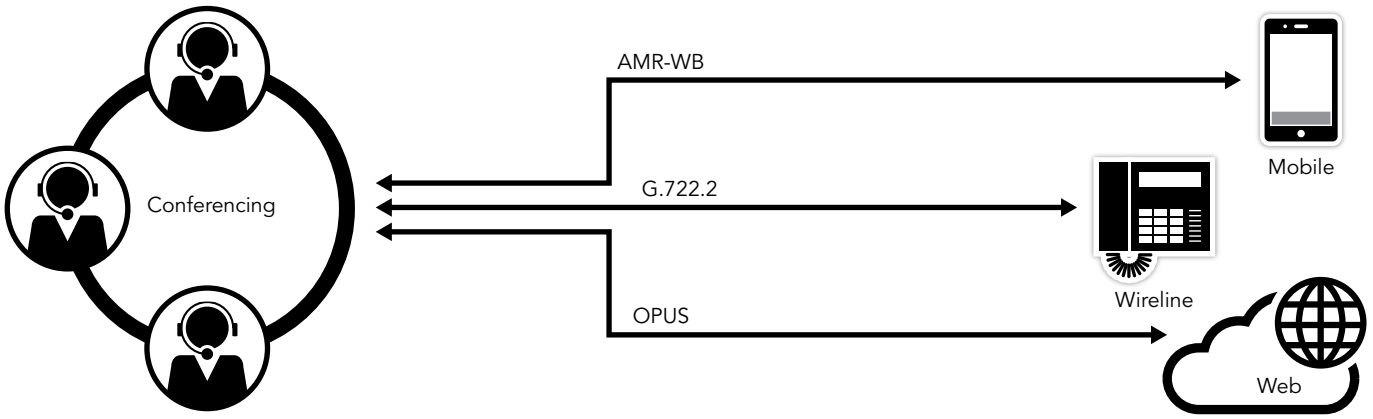

Figure 1: conferencing application with HD clients

Cloud Communications

A third use case example is supporting the previous use case example through a cloud implementation of the communications adjunct. Figure 2 shows this type of example. Given that many communications applications are moving to the cloud, it stands to reason that HD Voice applications should follow.

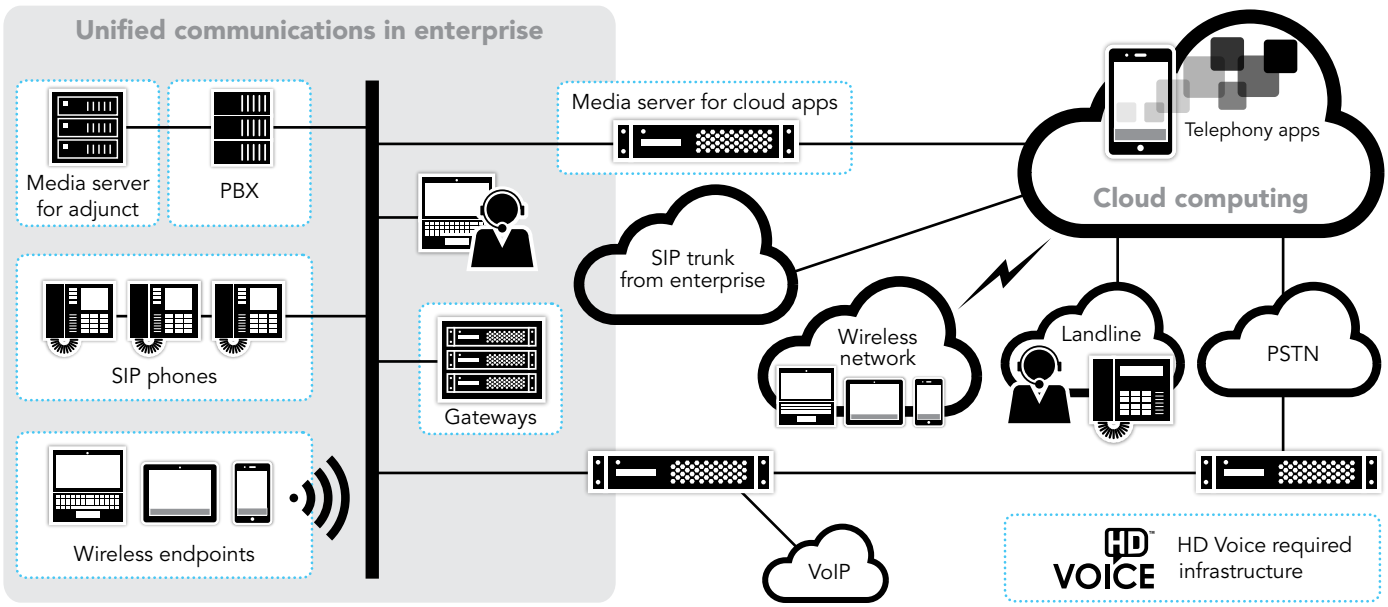

Figure 2: diagram showing an example of cloud implementation points of HD Voice

A fourth category of use case examples involve innovation. Historically, when new capabilities are rolled out, new innovative services which did not exist before tend to follow. In the case of HD Voice, it depends to some extent on the development community ecosystem's imagination, but types of innovation could be in areas such as voice recognition for the masses or improved outbound text to speech to mobile phones.

Implementation Issues

Figure 2 also gives a good overview of an example of implementation points that address HD Voice.

In the enterprise, HD Voice would need to be supported by the PBX, phones (if not using a headset), and media server (if using a voice mail, IVR, or contact center solution and the full HD Voice experience is desired for the caller). Wideband codecs used in the enterprise (G.722.1, G.722.1c, G.719, RTA, SILK) can vary.

In service provider networks, mobile wideband audio (HD Voice) service has been fully standardized by the 3GPP since 1998. The service uses the AMR wideband (AMR-WB or G.722.2) codec, which must be supported in handsets, the core network's gateways, and the GSM TRAU units (part of the Base Station Controller) to work effectively. Although AMR-WB requires more DSP processing than AMR-NB, the extra processing power is a small fraction of what a handset needs (because of radio requirements) and can typically be supported as a software upgrade on recently deployed core network equipment. A mobile operator's core network must also be able to support TFO and TrFO signaling. This signaling is already available in core network equipment, or can be implemented as a software upgrade for most systems deployed in recent years. What's more, added radio resources are not required, which means radio re-planning is not needed. Although AMR-WB doubles the audio bandwidth, its bit rates are similar to those of AMR-NB, especially when operating in the mandatory configuration (set 0) of the standard. The takeaway is that AMR-NB and AMR-WB can easily co-exist and operate side-by-side in today's radio infrastructure.

Given that it is likely that two different wideband codecs will be used when trying to connect the service provider and enterprise networks together, a gateway or border element would need to be used in order to transcode one codec to another.

And like the enterprise, a network-based media server needs to support HD Voice codecs for the subscriber to have an HD Voice experience, so that the conference servers, voice mail, IVR and other value-added services can also be supported by HD Voice.

Dialogic's Role

Dialogic has been an early HD Voice supporter, and actively supports efforts by the GSMA, GSA, wireline operators and enterprise developers to deliver rich high-fidelity HD Voice applications.

Dialogic currently supports HD Voice in:

- Dialogic® PowerMedia® XMS Media Server Software
- Dialogic® BorderNet™ Session Border Controller
- Dialogic ® IMG 2020 Integrated Media Gateway

For More Information

GSMA HD Voice Materials

The voice renaissance: VoLTE, HD Voice and Wi-Fi calling bring innovation to voice.

www.dialogic.com

For a list of Dialogic locations and offices, please visit: https://www.dialogic.com/contact.aspx

09/17 13313-02