White Paper

Testing the EVS Codec

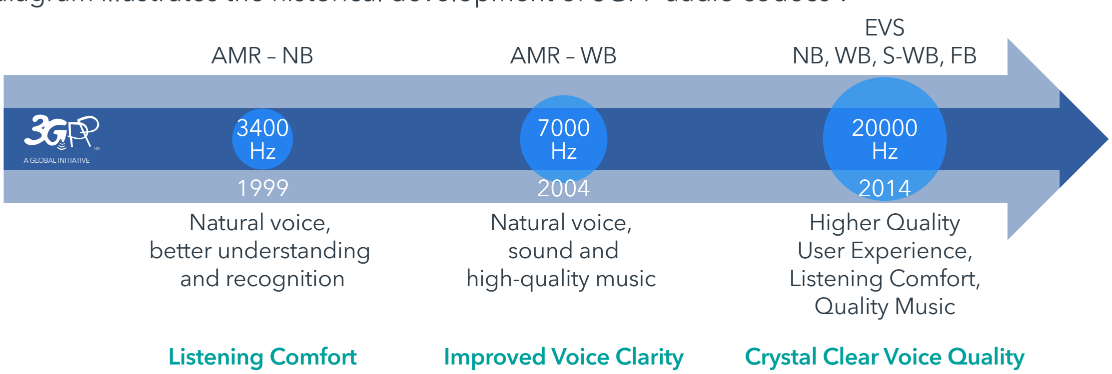
Addressing Complex Challenges to Ensure Superior Audio Quality Performance

A Brief History of Enhanced Voice Services (EVS)

Quality of voice is one of the most important aspects of customer satisfaction regarding mobile networks, and can exceed price as a factor in selecting a mobile operator[1]. The industry has therefore developed audio coding standards to continually improve voice quality and optimize network performance for the transmission of voice.

Audio Codec Evolution

The following diagram illustrates the historical development of 3GPP audio codecs[2]:



1. "Connecting with the Consumer", The Nielsen Company, 2014.

2. AMR - Adaptive Multi-Rate Wideband, the 3GPP codec which preceded EVS

The 3GPP Enhanced Voice Services (EVS) codec is beginning to appear in the wireless network arena: chipset implementations are ready; devices are under evaluation; and worldwide operators are starting to work on deployment plans. The EVS codec delivers unprecedented voice quality to users and can optimize network efficiency for carriers, but it also brings new testing challenges due to the extent and robustness of its feature set. This white paper discusses the benefits of the EVS codec and the required testing to ensure EVS-enabled devices perform according to the codec specifications and deliver a superior user experience.

The advancement of network standards corresponds to increases in bandwidth:

- Early mobile networks used Narrowband (NB) frequencies between 200Hz and 3,400Hz, which is typically associated with the quality level of telephone speech.
- The industry evolved to Wideband (WB; 50Hz–7,000Hz) and finally Full Band (FB; 20Hz–20,000Hz), which covers the general range of audio signals.
- With the advent of internet audio applications, other bandwidths have come into use, such as Super Wideband (S-WB; 50Hz–14,000Hz)

With every increase in bandwidth, there is a need to adapt signal processing and meet increasing demands for audio quality.

A general overview of the EVS codec is defined by 3GPP TS 26.441[3], with many additional specifications through TS 26.451 defining individual features such as Jitter Buffer Management, Backward Compatible Functions, EVS Codec Error Concealment of Lost Packets, and others listed below. The EVS specifications were approved for standardization in September 2014 and completed in December later that year.

EVS promises to change the way we think about mobile audio. It provides full HD voice quality at a bandwidth of up to 20,000Hz, covering the full frequency range of human hearing. EVS delivers audio at levels of quality similar to those available through streaming and downloading, and enables clarity for conversations even when there is significant background noise, as in the subway or walking down a busy street.

EVS is built around an all-IP LTE network to improve upon the high-definition audio promised with Voice over LTE (VoLTE), but can also enhance VoIP and circuit-switched systems. 3GPP recommends the EVS codec for VoLTE and WCDMA implementations and as an alternative to the AMR-WB codec. EVS is also the standard codec for S-WB and FB speech.

3. 3GPP specifications are available at http://www.3gpp.org

EVS Codec Features and Improvements

EVS encompasses a wide array of features that contribute to its superior performance, including:

- Wide range of audio bandwidths (NB, WB, S-WB, FB)
- Wide range of bit rates that vary to optimize audio performance
- Robustness improvements (VAD, CNG, DTX)
- Error concealment mechanism (with channel aware mode)
- Jitter buffer management (JBM)
- Enhanced interoperation with the AMR-WB codec
- Improved speech and music content

The following sections touch upon each of these features in greater detail.

Wide Range of Bandwidths and Bit Rates

EVS uses an innovative, flexible mechanism to switch between specialized coding setups for speech and music. Input and output sampling occurs at 8, 16, 32 and 48kHz, and an automatic bandwidth detector adapts to the actual input signal bandwidth. This switching can occur at every 20 millisecond frame, which permits rapid changes to optimize channel capacity.

Similarly, EVS supports a wide range of bit rates, which provides the optimal rate for each network and which delivers an increase in subjective quality. EVS allows a noticeable increase in voice quality when applied to FB and S-WB operations, beginning at 9.6 and 16.4 kbit/s respectively. The maximum bit rate of EVS is 24.4 kbit/s for NB and 128 kbit/s for all other audio bandwidths.

Robustness Improvements (VAD, CNG, DTX)

EVS contains a complete set of communication system functions including voice/sound activity detection (VAD), comfort noise generation (CNG), and discontinuous transmission (DTX). Transmission of noise is replaced by comfort noise generation in the decoder, which helps to improve voice activity detection; DTX allows optimization of battery life. To determine the application of DTX, EVS monitors (through VAD) the occurrence of active speech, active music, or the lack of both, when background noise peaks. Based on these conditions and the strength of background noise, EVS implements CNG in the decoder (either frequency domain or LP-based). The adjusted DTX mode is available at all relevant operating points up to 24.4 kbit/s.

Error Concealment Mechanism (with Channel Aware Mode)

EVS provides an error concealment mechanism to neutralize the effects of transmission errors. A constant issue for audio services is deterioration of quality due to radio channel characteristics and network conditions, leading to the loss, corruption, or unacceptable latency of audio packets. Codecs therefore perform frame error concealment (FEC) on a limited number of frames in order to produce a neutral signal and replace what otherwise sounds strange or missing. Previous audio codecs can manage approximately 3% of lost frames without critical artifacts (there is noticeable distortion). In comparison, EVS codec can perform FEC on more than 15% of lost frames without critical artifacts and still maintain relatively clear voice signals.

Channel Aware Mode

Channel aware mode is designed to improve frame/packet error resilience. EVS channel aware mode is specifically created for reliable speech transmit under especially bad network conditions such as high packet loss, jitter, etc. Channel aware is available in two configuration modes: EVS WB 13.2 and EVS S-WB 13.2.

Jitter Buffer Management (JBM)

Jitter buffer management reduces the impact of variable transmission time. Delay jitter is another form of quality loss. It is often caused by the dynamic adaptation of buffers built into VoLTE applications. The longer the buffers, the more compensation for packet jitter, which results in increased latency. Conversely, for shorter jitter buffers, there is less latency but a greater chance of packet loss. Network impairment such as packet loss, jitter, delay simulation, and packet ordering must be utilized in order to test full jitter buffer management capability.

The defined adaptive JBM solution reflects the above mentioned trade-offs. While attempting to minimize packet losses, the JBM algorithm in the receiver also keeps track of the delay in packet delivery as a result of the buffering. The JBM solution suitably adjusts the depth of the de-jitter buffer in order to achieve the trade-off between delay and late losses. The EVS codec therefore not only reaches the very limits of human audio perception but is also designed to cope with network impairments in IP-based communications.

Enhanced Interoperation with the AMR-WB Codec

The EVS codec is backwards compatible with all nine AMR-WB codec formats used throughout different telephony services, including circuit-switched networks. The EVS codec's AMR-WB interoperable operation modes are either identical to those in the 3GPP AMR-WB codec or may be different, but contain interoperable bitstreams.

Improved Speech and Music Content

EVS provides a unique method for handling voice and music content. It deploys bandwidth extensions using on-the-fly compression switching according to content type at a low algorithmic delay of 32 milliseconds. EVS employs either a time-domain bandwidth extension or an integrated frequency domain solution, depending on the content.

Other music augmentations include post-processing tools such as enhancer, inactive signal post processing, bass-boost filter, and formant post filter. Post processing generates significant improvements, particularly in noisy environments and for mixed content.

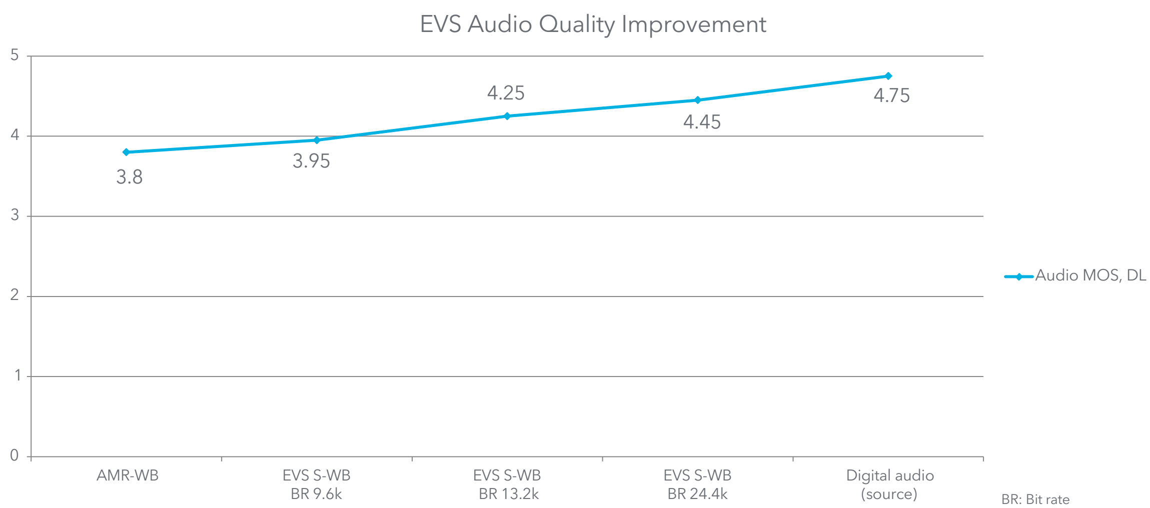
EVS vs AMR Codecs

The EVS codec enhances audio quality and improves coding efficiency for NB and WB audio, and provides a significant enhancement in voice quality with S-WB and FB operation.

EVS Audio Quality Measurement

POLQA[4], the successor to P.862/PESQ, is based on ITU-T Recommendation P.863[5]. POLQA provides voice quality analysis of HD Voice and Full HD Voice. It offers an advanced level of benchmarking accuracy, and adds significant new capabilities for WB, S-WB and EVS voice signals along with support for the most recent voice coding and VoIP transmission technologies. For these reasons, POLQA is the best choice for evaluating, optimizing and measuring the voice quality of next-generation networks and delivers accurate, high-resolution analysis results. POLQA is measured as a Mean Opinion Score (MOS). MOS is expressed as a single number in the range 1 to 5, where 1 is lowest perceived audio quality and 5 is the highest perceived audio quality measurement. Standard deviation of the mean distribution needs to be very small. It is important to test the uplink and downlink audio MOS measurement of devices. When tested in parallel, the MOS needs to remain stable.

The following graph illustrates a codec quality comparison between AMR and EVS at various bit rates, based on POLQA MOS:



Comparison of AMR to EVS with varying bit rates.

As the graph indicates, an audio quality comparison between EVS and AMR-WB reveals that EVS is superior at any bit rate. It delivers better quality at similar bit rates and provides incomparable quality improvements as bit rate increases. Starting from 13.2kbit/s, EVS operating with S-WB allows “direct source” quality levels, which is close to the quality of the source audio for both voice and music.

4. POLQA (Perceptual Objective Listening Quality Assessment) is an ITU-T standard for speech quality modeling using digital speech signal analysis. It produces objective measures to simulate quality scores derived from subjective, real speech listening tests.

5. ITU-T released the new ‘Implementer’s Guide on assessment of EVS coded speech” with Recommendation ITU-T P.863. EVS, also referred to as “Full-HD Voice”, is designed to deliver superior voice quality and more efficient usage of the available network bandwidth.

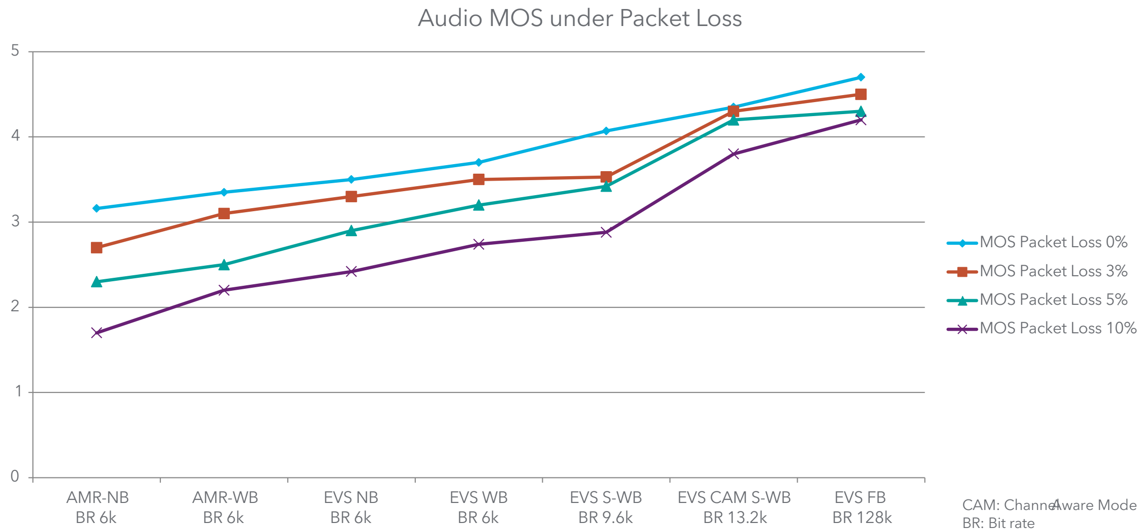
EVS Performance

Packet loss is seen by users to be the most noticeable disturbance in voice communications, and is more likely to occur in unmanaged networks, where Quality of Service (QoS) is not present. Managed networks have more capability than unmanaged networks, but also require a skilled administrator or engineer to make the most of them. A managed switch allows you have better control of your network and all the traffic moving through it. An unmanaged switch allows Ethernet devices to communicate with one another automatically using auto-negotiation to determine parameters such as the data rate and whether to use half-duplex or full-duplex mode.

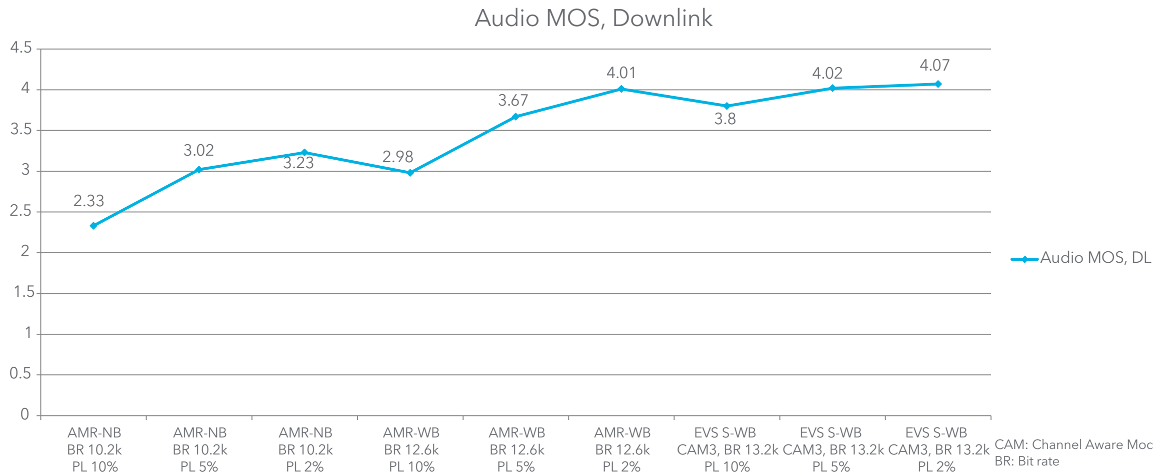
Features available on managed switches may vary between manufacturers and models, ability to implement quality of service (QoS), support for virtual LANs (VLANs), bandwidth rate limiting and port mirroring.

Comparison of EVS and AMR Codec Performance

The following graphs illustrate a comparison of EVS and AMR codec performance with different packet loss simulations. Audio MOS was evaluated via POLQA MOS. As the graphs indicate, an audio quality comparison between EVS channel aware, AMR-WB, and AMR-NB reveals that EVS is superior. EVS delivers better quality at similar bit rates and provides incomparable quality improvements as bit rate increases.



Extensive lab testing was performed to compare the EVS and AMR codec with different transmission packet losses of 3%, 5% and 10%. As shown in the prior graph, the EVS codec MOS is much higher than the AMR. The robustness of EVS using channel aware mode (as shown with 13.2k S-WB) delivers superior performance through all packet loss scenarios, with results nearing FB quality.



Again, with packet losses of 2%, 5%, and 10%, the audio downlink MOS of the EVS 13.2k S-WB codec with channel aware mode is much higher than the AMR-WB and AMR-NB.

The previous graphs clearly illustrate the benefits of the EVS codec and its ability to deliver voice performance superior to AMR codecs over varying bit rates and under packet loss conditions. The table below provides a comparison of the EVS feature set with its predecessors; the following section includes a detailed explanation of some necessary test considerations for evaluating the EVS codec.

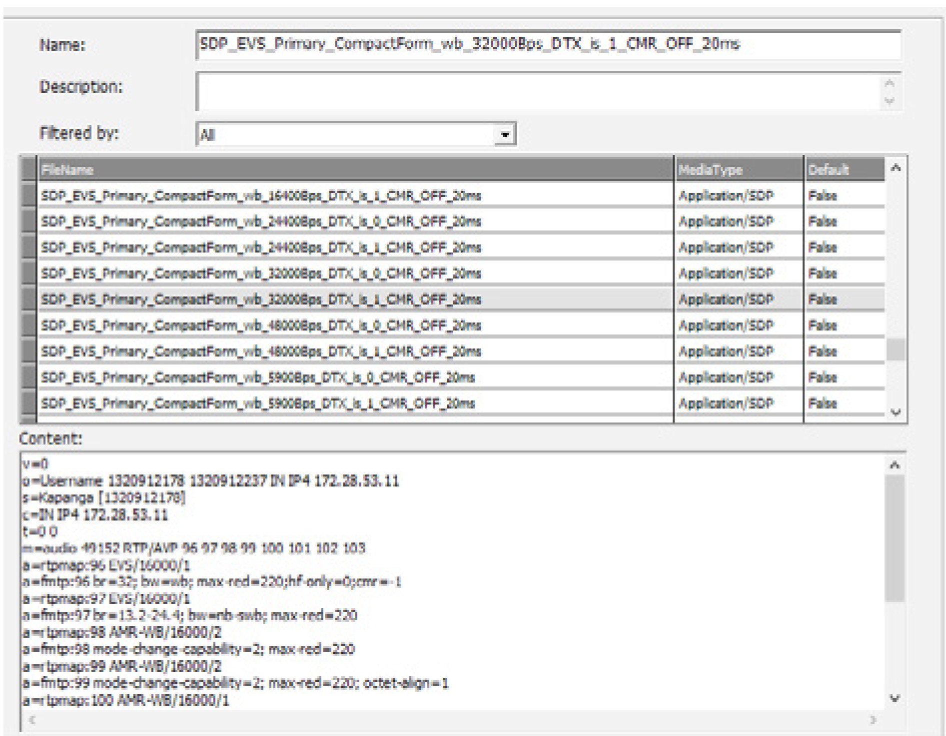
Feature	AMR	EVS	Benefit
Compact form	Not available	Yes	Reduces bandwidth
Codec change request	Always	Configurable	Rate switch and bandwidth reduction
Channel aware mode	Not available	Yes	Packet loss resilience and better quality with poor signals
Band	NB , WB	NB, WB, S-WB, FB	Option to switch between bands without involving signaling
Bit rate	4.75-23.85	5.9-128	Clearer speech
EVS switch mode	No	Yes	Backwards compatibility and mobility

EVS Testing Challenges

EVS testing and measurement is complex because there are so many codec configurations, features, parameters, and audio quality measurements to take into consideration, which is exacerbated by the sheer number of variables that affect performance.

EVS Codec Feature Testing via Session Description Protocol (SDP)

EVS uses Session Description Protocol (SDP) as part of the Session Initiation Protocol (SIP). Contained within the SDP are all the parameters for the EVS features that are supported. SIP arranges the capabilities, media streams between devices, and interactive features, such as the setup or exclusion of media elements. Management of coded media is directed through the transport protocol and through control of related data originating from the network. Voice-related media components use real-time transport (RTP) protocols. There are literally hundreds of SDP permutations, and each SDP variation must be tested, resulting in the need to build more than 200 SDP scripts to test the different types of EVS capabilities. Any voice application that implements SDP protocol must ensure that all SDP and media type parameters are interpreted correctly in order to ensure that the implementation can always determine if it is capable of communicating or not.



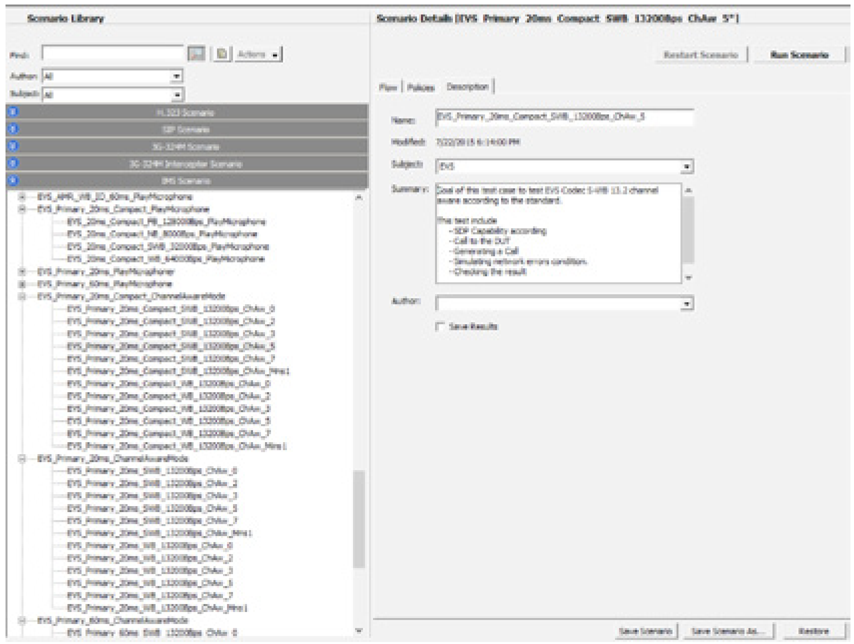
Sample Spirent ProLab SDP script.

EVS Feature Testing

3GPP Release 12 has provided an EVS feature list that must be evaluated comprehensively in order to meet performance requirements. Chipset and device manufacturers are implementing EVS features in accordance with the 3GPP specifications. They, along with operators, are defining the EVS functional and audio quality performance testing required. A subset is provided here:

Feature	Testing Considerations	What to Test
Encoder/ decoder bit rate	Checks compatibility of encoder and decoder bit rates. Source codec bit rates for EVS primary mode (SC-VBR): 5.9, 7.2, 8, 9.6, 13.2, 16.4, 24.4, 32, 48, 64, 96, and 128.	Choose bit rates that represent the full spectrum of bit rates that will be used in the final system. While the EVS is encoding, CPU usage and quality must be tracked.
Multi-rate audio codec	Checks and adjusts EVS primary mode audio bandwidth for sending and receiving directions during the session. Bandwidths include NB, WB, S-WB, and FB, as well as signals between NB and each of the wider bandwidths.	Different types of EVS modes with all the associated rates.
Robustness, HF-only	EVS looks for 0 and 1 values. If it detects 0 or no value, both compact and full header formats can be used for sending and receiving directions. If 1, only full header format (without zero padding for size collision avoidance) is used.	Different configurations to make sure the device is also sending compact and full headers.
Robustness, DTX	Checks for use of discontinuous transmission application (i.e., encoding of silence). EVS looks for 0 and 1 values. If 0, DTX is disabled in the session. If 1 or not present, DTX is enabled.	Must consider different types of audio content such as silence only, silence and noise, etc.
Channel aware mode	Checks for configuration and use of the receiving channel aware mode. This mode deals with packet loss resilience.	Simulate different types of packet loss to measure concealment by using different channel aware mode configurations.
Jitter buffer management	The codec contains a system for JBM to manage the delay of the received packets, duplicate packets, packet ordering, etc.	Simulate different types of network impairments in order to test the JBM including static and dynamic
Compact header handling (CPM)	Checks the ability to function with /without full headers (i.e., full vs. compact headers). Full headers are sometimes not sent due to bandwidth limitations (for EVS AMR-WB IO modes only).	Example: Test scenario with compact header with low rate.
EVS mode switch	Checks bandwidth differences. EVS primary mode is used at the start or update of the session for sending and receiving directions.	If bandwidth is different between encoder and decoder, EVS bit rate and bandwidth modes must be adjusted. Multiple combinations can be tested.
Codec mode request	Checks for the codec mode request (CMR) in the RTP protocol payload header. In EVS primary mode, CMR on the RTP payload header is disabled (EVS AMR-WB IO mode).	Different scenarios when the IMS gateway is supporting EVS and is not supporting EVS. Test different types of mode change and need for the device to support each mode.

Hundreds of test cases are required for comprehensive EVS codec capability testing that include SIP, SDP configuration, audio content, media encoder, decoder and network media impairment (see figure).



Sample list of Spirent ProLab IMS test scripts.

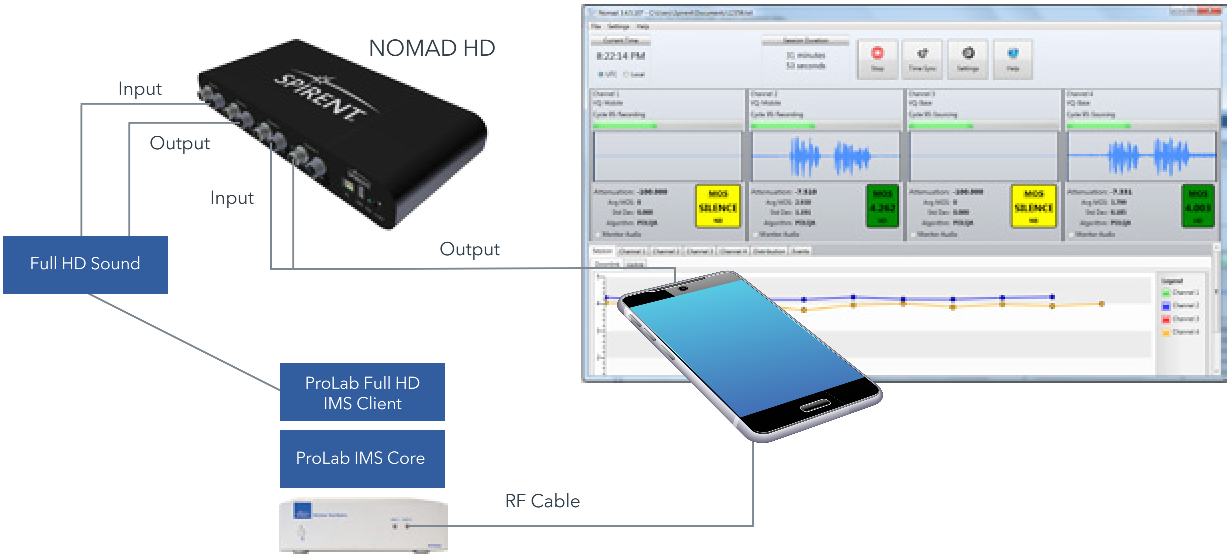
Chipset and Device Manufacturer Challenges

There is no question about the level of improvements in audio mobile communications that EVS is creating. However, many of the various features of EVS require more resources on the device side compared to previous codecs like AMR-WB. For instance, EVS requires 86 Weighted Millions of Operations per Second (WMOPS), compared to 38 WMOPS for AMR-WB. Similarly, AMR-WB requires 24K-300K memory, compared to 175K-285K for EVS. These characteristics translate into a significant drain on battery life and handset performance.

The exact influence of EVS on device performance is an essential aspect of EVS implementation. To understand the magnitude of EVS implementation issues, and to make effective technological adaptations, extensive and realistic testing is necessary.

The Spirent Solution

Spirent provides a complete solution for testing and measuring EVS audio codec and decoder performance using Elevate™, ProLab™ and Nomad, which cover functional, quality, and user experience requirements including audio quality (based on POLQA full-reference and no-reference). Spirent's testing suite enables device manufacturers and operators to deploy worldwide, commercial EVS at excellent voice and music audio quality levels while optimizing bandwidth.



Spirent addresses the complex needs of EVS system evaluation by testing the adherence of EVS systems to 3GPP standards in an IMS environment. Spirent employs a built-in SDP codec configuration to test different EVS configurations and Wave content to evaluate EVS capabilities. Due to Spirent's extensive testing record, evaluations are conducted against a background of hundreds of previous test cases that serve as functional, performance, and quality benchmarks.

Testing for EVS evaluates both uplink and downlink quality for full HD sound. Spirent uses a full HD IMS client, media server, and IMS core as part of its test topology. Various EVS quality measurements are available, including online/offline POLQA including capture (PCAP) analysis based on user experience, network impairment test cases to test jitter buffer management implementation, channel aware mechanism (for jitter buffer management and quality of experience), and offline PCAP quality of experience analysis.

About Spirent Communications

Spirent Communications (LSE: SPT) is a global leader with deep expertise and decades of experience in testing, assurance, analytics and security, serving developers, service providers, and enterprise networks.

We help bring clarity to increasingly complex technological and business challenges.

Spirent's customers have made a promise to their customers to deliver superior performance. Spirent assures that those promises are fulfilled.

For more information, visit: www.spirent.com

Operator Benefits

Service providers that implement EVS have the potential for increased customer satisfaction and retention due to enhanced HD voice quality, which can also lead to longer and more frequent calls. Other revenue opportunities include enterprise applications and fees for HD voice services. EVS can provide seamless interoperability with global legacy networks and does not require an increase in bandwidth. EVS also provides an opportunity to compete with OTT (over the top) VoIP services (such as Skype) because OTT is a "best-effort" service, whereas EVS voice traffic over VoLTE networks always gets priority over data traffic due to its QoS mechanisms. Key EVS benefits include:

1. Significantly improved quality over AMR-NB and AMR-WB by using similar bit rates, as shown earlier
2. Improved network capacity while maintaining the same quality
3. Unique packet concealment to improve quality during network error conditions without severe artifacts
4. Improved clear calls in full HD voice, even in noisy environments
5. Exceptional quality of music

Conclusion

The EVS codec has an extensive feature set that can help improve the overall performance of voice services over multiple networks and across multiple bands (NB,WB,S-WB, FB). Due to the scope of the EVS feature set and the variations within each feature, EVS-enabled devices must be tested thoroughly to ensure compliance with 3GPP standards and deliver the full potential of HD voice and superior QoE.

spirent™
Promise. Assured.

Contact Us

For more information, call your Spirent sales representative or visit us on the web at www.spirent.com/ContactSpirent.

www.spirent.com

Americas 1-800-SPIRENT
+1-800-774-7368 | sales@spirent.com

US Government & Defense
info@spirentfederal.com | spirentfederal.com

Europe and the Middle East
+44 (0) 1293 767979 | emeainfo@spirent.com

Asia and the Pacific
+86-10-8518-2539 | salesasia@spirent.com