EVS – Enhanced Voice Services

Next Generation in Speech Quality

ETSI STQ Workshop, Nov 2012

Dr. Imre Varga
Qualcomm Inc.
EVS – Next Gen 3GPP Speech Coding for Improved User Experience

AMR
4.75 kbps → 12.2 kbps
AMR-WB
6.6 kbps → 23.85 kbps
EVS
5.9 kbps → 128 kbps
What is EVS?

- EVS – Enhanced Voice Services
  - Next generation 3GPP speech coding
  - Following the successful FR, HR, EFR, AMR, AMR-WB codecs
  - Designed for packet-switched networks / mobile VoIP
  - VoLTE is a key target application
  - Application in other networks
  - AMR-WB interoperable mode
  - Rel-12 Work Item in 3GPP preceded by a Study Item TR 22.813

- Key features
  - Super-wideband speech (32 kHz sampling) – improved speech quality
  - Source-controlled variable bit-rate operation – improved capacity
  - Designed for VoIP – improved robustness
  - Improved music performance
  - Wide bit-rate range and all bandwidths for maximum flexibility
  - Backward interoperable mode to AMR-WB

- Standardization process
  - Qualification phase – currently on-going rigorous testing
  - Selection phase
  - Characterization phase
3GPP EVS is the Next Generation Speech Coder – Value for the Ecosystem

- Speech quality determines user experience
- Operators are very concerned about voice quality
  - Ensuring voice quality on new VoLTE deployments
  - EVS addresses all networks – mobile VoIP with QoS, best effort VoIP, CS
- 3GPP goals of Enhanced Voice Services (EVS) standardization
  - Feature-rich coder
    - Designed for VoIP applications such as MTSI in TS 26.114
    - It is further desirable that the codec fulfills needs in other networks such as CS
    - NB, WB, SWB bandwidths, FB and stereo optional, high robustness mode
    - Bit rates: 7.2, 8, 9.6, 13.2, 16.4, and 24.4 kb/s gross rates that comply with LTE TBSs; 32, 48, 64, 96, 128 kb/s net source rates.
  - Quality improvements – improving user experience
    - Better quality in VoLTE and UMTS (with no new RAB)
    - Evolution path: EVS provides SWB at around 13 kbps – lower rate and lower delay SWB than other industry coders without sacrificing quality
    - Better quality for music and mixed content in conversational applications
  - Capacity improvements – increasing system efficiency
    - VBR at 5.9 kbps provides high capacity mode
  - Robustness improvements – optimized behavior in VoIP applications
    - More robust NB/WB through significantly better error resilience
    - High robustness mode
# EVS Features

## Jitter Buffer Management

<table>
<thead>
<tr>
<th></th>
<th>NB</th>
<th>WB</th>
<th>SWB</th>
<th>FB</th>
<th>Stereo</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>New EVS Modes</strong>&lt;br&gt;(CBR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speech</td>
<td>7.2-13.2 kb/s</td>
<td>6.6-23.85 kb/s</td>
<td>7.2-128 kb/s</td>
<td>13.2-128 kb/s</td>
<td>7.2-128 kb/s</td>
</tr>
<tr>
<td>Music</td>
<td>5.9 kb/s average</td>
<td>12-128 kb/s</td>
<td>5.9 kb/s average</td>
<td>12-128 kb/s</td>
<td>7.2-128 kb/s</td>
</tr>
<tr>
<td><strong>New EVS Modes</strong>&lt;br&gt;(VBR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speech</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Music</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Why Deploy EVS?

- **Better Capacity**
  - Same quality as legacy NB/WB

- **Better Music**
  - Near AAC Quality at much lower delay

- **Better Quality**
  - Same capacity as legacy NB/WB
EVS From User’s Perspective

• More natural sounding speech for enhanced user experience during voice call
  • Super wideband speech coding
  • Better wideband and narrowband quality at same bit rates as legacy coders

• Consistent voice quality during call
  • Improved error resilience
  • Customizations for VoIP deployment

• Better in-call music quality
  • Improved coding of music for better sounding
    • Ring back and music on hold
    • Hear what I hear,
    • Remote music education and collaboration
## EVS Design Requirements

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Superwideband</strong></td>
<td>(0-16 kHz) Coding of Speech better than AMR-WB</td>
</tr>
<tr>
<td><strong>Wideband</strong></td>
<td>(0-8 kHz) Coding of Speech better than AMR-WB</td>
</tr>
<tr>
<td><strong>Narrowband</strong></td>
<td>(0-4 KHz) Coding of Speech better than AMR</td>
</tr>
<tr>
<td><strong>Improved Error Resilience</strong></td>
<td>for both Circuit Switched and Packet Switched Communication and VoIP Capability</td>
</tr>
<tr>
<td><strong>Source Controlled Variable Rate Coding</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Improved Coding of Music</strong></td>
<td>for In-call Music (Music on hold and Ringback)</td>
</tr>
<tr>
<td><strong>Constraints on Frame Length, Max. Algorithmic Delay, Complexity, JBM, Rate Switching, PLC, RTP Payload Format, VAD/DTX/CNG</strong></td>
<td></td>
</tr>
</tbody>
</table>
EVS Variable Bit-Rate Operation

- Improved system efficiency through lower average bit-rate
- Illustration of Source-Controlled Variable Bit-Rate Operation in the Figure
# EVS Requirements in SWB at Low Rates

<table>
<thead>
<tr>
<th>Category</th>
<th>Bitrate (kbit/s)</th>
<th>FER</th>
<th>DTX</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clean speech</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-26,-16,-36dBoV</td>
<td>13.2</td>
<td>0%</td>
<td>On†/Off</td>
<td>NWT G.722.1C @ 32</td>
</tr>
<tr>
<td></td>
<td>16.4</td>
<td></td>
<td></td>
<td>NWT G.722.1C @ 48</td>
</tr>
<tr>
<td></td>
<td>24.4</td>
<td></td>
<td></td>
<td>NWT G.718B @ 36</td>
</tr>
<tr>
<td><strong>Clean speech</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-26 dBoV</td>
<td>13.2</td>
<td>x=3%, 6%</td>
<td>Off</td>
<td>On† for 13.2</td>
</tr>
<tr>
<td></td>
<td>16.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>24.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Noisy Speech (Car, Office, Street)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-26 dBoV</td>
<td>13.2</td>
<td>0%</td>
<td>On‡/Off</td>
<td>NWT G.722.1C @ 24 when EVS DTX off</td>
</tr>
<tr>
<td></td>
<td>16.4</td>
<td></td>
<td></td>
<td>NWT AMR-WB @19.85 DTX on when EVS DTX on</td>
</tr>
<tr>
<td></td>
<td>24.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Noisy Speech (Car, Office, Street)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-26 dBoV</td>
<td>13.2</td>
<td>x=3%, 6%</td>
<td>Off</td>
<td>On‡ for 13.2</td>
</tr>
<tr>
<td></td>
<td>16.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>24.4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
EVS Standardization Process

- **Requirements phase** – design constraints and performance requirements

- **Candidate coders**
  - 13 companies submitted a candidate by 16 November 2012
    - Ericsson, Fraunhofer, Huawei, Motorola, Nokia, NTT, NTTDoCoMo, Orange, Panasonic, Qualcomm, Samsung, VoiceAge, ZTE
  - Standardization by competition

- **Qualification phase**
  - Aim is to keep the most promising candidates for selection (at most 5)
  - Extensive testing
    - 12 experiments, each candidate is tested in-house and in another listening lab
    - Global Analysis Lab performs collection and analysis of test results
  - Qualification meeting in March 2013

- **Selection phase**
  - Aim is to select the best candidate out of the max. 5 kept in qualification
  - Codec selection is based on extensive testing in neutral listening labs

- **Characterization phase**
  - Aim is to test the coder performance for all conditions and special signals / conditions

- **Approval of EVS Specifications and Technical Report**
Schedule

- 3GPP Standardization Schedule

Aug ’10 → March ’13 → 2013 → June ’14

13 Candidates 
EVS Qualification → 5 Candidates 
EVS Selection → EVS characterization

- Rel-12 Work Item closing in June 2014
EVS Deployment

- EVS targets – VoLTE and other networks
  - EVS Rel-12 standardization timeline matches VoLTE mass deployment plans
  - Provide EVS in WCDMA and best-effort VoIP at the same time

- VoLTE trials
  - Goal is to make EVS available for VoLTE trials and deployment
  - Pre-commercial phase during 2013

- Qualcomm is key partner in EVS deployment
  - Serve ecosystem by making EVS codec available in mobile chipsets
  - Pre-standard version available for VoLTE trials and branded voice services

- Deployment of standardized version can begin at availability of specifications

- Qualcomm supports best-in-class voice quality for IMS based voice service deployments on LTE with a complete suite of tools and features, including EVS, IMS client, and voice enhancement.
THANK YOU!