

IETF 84 – RTCWEB

Mandatory To Implement
Audio Codec Selection

Problem Statement

- We have consensus to specify a MTI (Mandatory To Implement) audio codec
 - Goal: prevent negotiation failure
- Need to decide which one(s)
 - Fewer the better
- *Not* trying to decide which codecs are recommended
 - Implementations MAY support as many codecs as they want, but goal of MTI is to address basic interop

Criteria for Consideration

- **Quality**
- **Versatility**
- **Licensing**
- Standardization
- Implementations
- Deployment
- Other(s)

AMR-NB

- Quality: Good narrowband speech at low bitrates
- Versatility: Limited (narrowband only, small number of pre-defined bitrates)
- Licensing: Well-known, not royalty-free
- Standardization: 3GPP
- Implementations: Optimized implementations available, only basico ps source available freely
- Deployment: Very well-deployed in mobile devices/networks (virtually all GSM, UMTS devices)

G.729

- Quality: Acceptable narrowband speech at 8 kb/s
- Versatility: Poor (narrowband-only, one bitrate)
- Licensing: Well-known, not royalty-free
- Standardization: ITU-T
- Implementations: Optimized implementations available, only basicops source available freely
- Deployment: Lots of gateways

AMR-WB (G722.2)

- Quality: Reasonable wideband speech at 12-24 kb/s
- Versatility: Limited (wideband only, small number of pre-defined bitrates)
- Licensing: Well-known, not royalty-free
- Standardization: 3GPP & ITU-T
- Implementations: Optimized implementations available, only basicops source available freely
- Deployment: Not widely deployed
 - GSM Association recently finished “HD Voice” description using AMR-WB

G.722.1C / G.719

- Quality: Good super-wideband/fullband speech starting at 48 kb/s, borderline music quality
- Versatility: Poor (super-wideband-only/fullband-only)
- Licensing: Currently royalty-free, but not open-source compatible
- Standardization: ITU-T
- Implementations: Only basicops version available freely
- Deployment: Video conferencing (Polycom, Ericsson)
- Other: Low-complexity, relatively high delay (40 ms)

AAC-LD

- Quality: Good quality stereo music at sufficiently high rates
- Versatility: Poor (fullband-only, no special speech support)
- Licensing: MPEG-LA, not royalty-free
- Standardization: MPEG
- Implementations: No freely-available implementation of any kind
- Deployment: Video conferencing

G.711

- Quality: Poor (narrowband-only at 64 kb/s)
- Versatility: Poor (narrowband-only at 64 kb/s)
- Licensing: None
- Standardization: ITU-T
- Implementations: Trivial
- Deployment: Everywhere
- Other: Trivial complexity

Speex

- Quality: Average (slightly worse than AMR-*)
- Versatility: Narrowband and wideband, speech-only
- Licensing: Royalty-free, open-source compatible
- Standardization: None (Xiph.Org)
- Implementations: Optimized, open-source C code
- Deployment: Adobe, Apple, Google, Microsoft, Asterisk, gstreamer, etc.

G.722

- Quality: Poor wideband at high rates
- Versatility: Poor (wideband-only, only 3 bitrates supported)
- Licensing: None (patents expired)
- Standardization: ITU-T
- Implementations: Optimized, open-source C code (as well as basicops)
- Deployment: ISDN video conferencing, desktop IP phones

iLBC

- Quality: Good narrowband speech at 13-15 kb/s
- Versatility: Poor (narrowband-only, only two bitrates supported)
- Licensing: Royalty-free, open-source compatible
- Standardization: IETF Experimental RFC
- Implementations: Optimized, open-source C code
- Deployment: Chrome, many gateways and switches

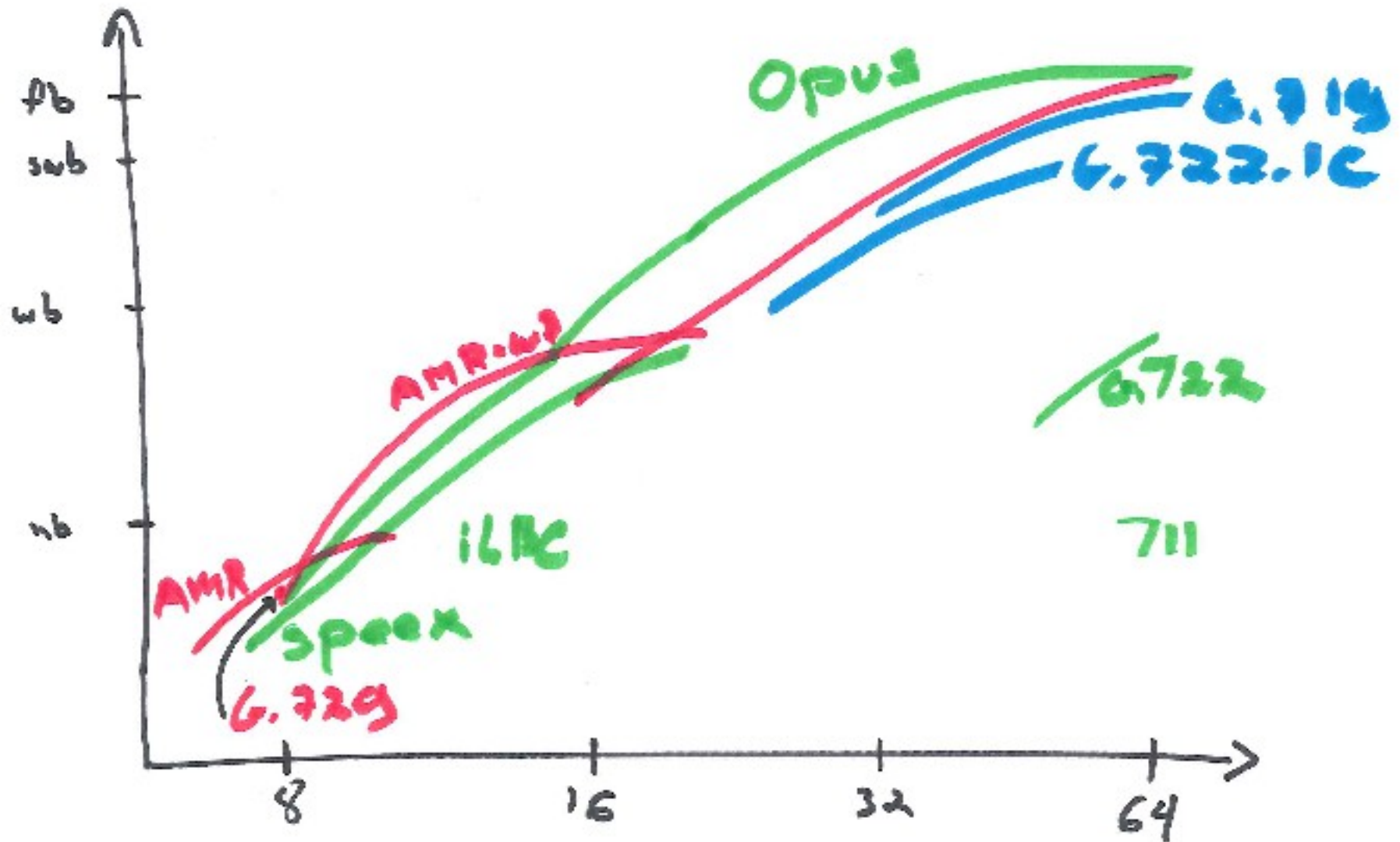
iSAC

- Quality: Okay wideband/super-wideband speech at 12-52 kb/s
- Versatility: Okay (wideband and super-wideband, adaptive bitrate, 30 and 60 ms frame sizes)
- Licensing: Royalty-free, open-source compatible
- Standardization: None (Google)
- Implementations: Optimized, open-source C code
- Deployment: Chrome, old Skype clients

Opus

- Quality: Equal or better than state of the art at vast majority of bitrates and audio bandwidths
- Versatility: Narrowband to fullband, 6-512 kb/s, mono, stereo, speech, music, arbitrary bitrates, variable frame sizes, seamless switching
- Licensing: Royalty-free, open-source compatible
- Standardization: IETF Standards-track
- Implementations: Optimized, open-source C code
- Deployment: Underway (Mozilla, Opera, Skype, Cisco, Asterisk, gstreamer, etc.)
- Other: Competitive with archival storage formats (Vorbis, AAC)

Mono Speech Quality Landscape



Proposal

- Opus
 - Handles all use cases
 - Does them as good or better than state-of-the-art
 - Freely implementable
- G.711
 - Addresses basic legacy interoperability
 - ~Zero added cost to implement
- And nothing else
 - Sufficient to avoid negotiation failure between WebRTC end-points
 - Mandating more codecs won't eliminate negotiation failure with non-WebRTC end-points