Network Working Group
Internet-Draft
Intended status: Informational
Expires: August 18, 2014

S. Proust
Orange
E. Berger
Cisco
B. Feiten
Deutsche Telekom
B. Burman
Ericsson
K. Bogineni
Verizon Wireless
M. Lei
Huawei
E. Marocco
Telecom Italia

February 14, 2014

Additional WebRTC audio codecs for interoperability with legacy networks.
draft-proust-rtcweb-audio-codecs-for-interop-00

Abstract

To ensure a baseline level of interoperability between WebRTC clients, [I-D.ietf-rtcweb-audio] requires a minimum set of codecs. However, to maximize the possibility to establish the session without the need for audio transcoding, it is also recommended to include in the offer other suitable audio codecs that are available to the browser.

This document provides some guidelines on the suitable codecs to be considered for WebRTC clients to address the most relevant interoperability use cases.

Status of This Memo

This Internet-Draft will expire on August 18, 2014.

Copyright Notice

Table of Contents

1. Introduction

As indicated in [I-D.ietf-rtcweb-overview], it has been anticipated that WebRTC will not remain an isolated island and that some WebRTC endpoints will need to communicate with devices used in other existing networks with the help of a gateway. Therefore, in order to maximize the possibility to establish the session without the need for audio transcoding, it is recommended in [I-D.ietf-rtcweb-audio] to include in the offer other suitable audio codecs that are available to the browser. This document provides some guidelines on the suitable codecs to be considered for WebRTC clients to address the most relevant interoperability use cases.

2. Terminology

In this document, the key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" are to be interpreted as described in RFC 2119 [RFC2119].

3. Definitions

Legacy networks: In this draft, legacy networks encompass the conversational networks that are already deployed like the PSTN, the PLMN, the IMS, H.323 networks.

4. Rationale for additional WebRTC codecs

The mandatory implementation of OPUS [RFC6716] in WebRTC clients can guarantee the codec interoperability (without transcoding) at the state of the art voice quality (better than narrow band "PSTN" quality) only between WebRTC clients. The WebRTC technology is however expected to have more extended usage to communicate with other types of clients. It can be used for instance as an access technology to 3GPP IMS services or to interoperate with fixed or mobile VoIP legacy HD voice service. Consequently, a significant number of calls are likely to occur between terminals supporting WebRTC clients and other terminals like mobile handsets, fixed VoIP terminals, DECT terminals that do not support WebRTC clients nor implement OPUS. As a consequence, these calls are likely to be either of low narrow band PSTN quality using G.711 at both ends or affected by transcoding operations. The drawbacks of such transcoding operations are recalled below:

o Degraded user experience with respect to voice quality: voice quality is significantly degraded by transcoding. For instance, the degradation is around 0.2 to 0.3 MOS for most of transcoding

use cases with AMR-WB at 12.65 kbit/s and in the same range for other wideband transcoding cases. It should be stressed that if G.711 is used as a fall back codec for interoperation, wideband voice quality will be lost. Such bandwidth reduction effect down to narrow band clearly degrades the user perceived quality of service leading to shorter and less frequent calls. Such a switch to G.711 is less than desirable or acceptable choice for customers. If transcoding is performed between OPUS and any other wideband codec, wideband communication could be maintained but with degraded quality (MOS scores of transcoding between AMR-WB 12.65kbit/s and OPUS at 16 kbit/s in both directions are significantly lower than those of AMR-WB at 12.65kbit/s or OPUS at 16 kbit/s). Furthermore, in degraded conditions, the addition of defects, like audio artifacts due to packet losses, and the audio effects resulting from the cascading of different packet loss recovery algorithms may result in a quality below the acceptable limit for the customers.

- Degraded user experience with respect to conversational interactivity: the degradation of conversational interactivity is due to the increase of end to end latency for both directions that is introduced by the transcoding operations. Transcoding requires full de-packetization for decoding of the media stream (including mechanisms of de-jitter buffering and packet loss recovery) then re-encoding, re-packetization and re-sending. The delays produced by all these operations are additive and may increase the end to end delay beyond acceptable limits like with more than 1s end to end latency.

- Additional costs in networks: transcoding places important additional costs on network gateways mainly related to codec implementation, codecs license, deployments, testing and validation costs. It must be noted that transcoding of wideband to wideband would require more CPU and be more costly than between narrowband codecs.

5. Additional suitable codecs for WebRTC

The following codecs are considered as relevant suitable codecs with respect to the general purpose described in section 4. This list reflects the current status of WebRTC foreseen use cases. It is not limitative and opened to further inclusion of other codecs for which relevant use cases can be identified.

5.1. AMR-WB

5.1.1. AMR-WB General description

The Adaptive Multi-Rate WideBand (AMR-WB) is a 3GPP defined speech codec that is mandatory to implement in any 3GPP terminal that supports wideband speech communication. It is being used in circuit switched mobile telephony services and new multimedia telephony services over IP/IMS and 4G/VoLTE, specified by GSMA as voice IMS profile for VoLTE in [IR.92]. More detailed information on AMR-WB can be found in [IR.36]. [IR.36] includes references for all 3GPP AMR-WB related specifications including detailed codec description and Source code.

5.1.2. WebRTC relevant use case for AMR-WB

The market of voice personal communication is driven by mobile terminals. AMR-WB is now implemented in more than 200 devices models and 85 HD mobile networks in 60 countries with a customer base of more than 100 million. A high number of calls are consequently likely to occur between WebRTC clients and mobile 3GPP terminals. The use of AMR-WB by WebRTC clients would consequently allow transcoding free interoperation with all mobile 3GPP wideband terminal. Besides, WebRTC clients running on mobile terminals (smartphones) may reuse the AMR-WB codec already implemented on these devices.

5.1.3. Guidelines for AMR-WB usage and implementation with WebRTC

Guidelines for implementing and using AMR-WB and ensuring interoperability with 3GPP mobile services can be found in [TS26.114]. In order to ensure interoperability with 4G/VoLTE as specified by GSMA, the more specific IMS profile for voice derived from [TS26.114] should be considered in [IR.92].

5.2. AMR

5.2.1. AMR General description

Adaptive Multi-Rate (AMR) is a 3GPP defined speech codec that is mandatory to implement in any 3GPP terminal that supports voice communication, i.e. several hundred millions of terminals. This include both mobile phone calls using GSM and 3G cellular systems as well as multimedia telephony services over IP/IMS and 4G/VoLTE, such as GSMA voice IMS profile for VoLTE in [IR.92]. In addition to impacts listed above, support of AMR can avoid degrading the high efficiency over mobile radio access.

5.2.2. WebRTC relevant use case for AMR

A user of a WebRTC endpoint on a device integrating an AMR module wants to communicate with another user that can only be reached on a mobile device that only supports AMR. Although more and more terminal devices are now "HD voice" and support AMR-WB; there is still a high number of legacy terminals supporting only AMR (terminals with no wideband / HD Voice capabilities) are still used. The use of AMR by WebRTC client would consequently allow transcoding free interoperation with all mobile 3GPP terminals. Besides, WebRTC client running on mobile terminals (smartphones) may reuse the AMR codec already implemented on these devices.

5.2.3. Guidelines for AMR usage and implementation with WebRTC

Guidelines for implementing and using AMR with purpose to ensure interoperability with 3GPP mobile services can be found in [TS26.114]. In order to ensure interoperability with 4G/VoLTE as specified by GSMA, the more specific IMS profile for voice derived from [TS26.114] should be considered in [IR.92].

5.3. G.722

5.3.1. G.722 General description

G.722 is an ITU-T defined wideband speech codec. [G.722] was approved by ITU-T in 1988. It is a royalty free codec that is common in a wide range of terminals and end-points supporting wideband speech and requiring low complexity. The complexity of G.722 is estimated to 10 MIPS [EN300175-8] which is 2.5 to 3 times lower than AMR-WB. Especially, G.722 has been chosen by ETSI DECT as the mandatory wideband codec for New Generation DECT with purpose to greatly increase the voice quality by extending the bandwidth from narrow band to wideband. G.722 is the wideband codec required for CAT-iq DECT certified terminal and the V2.0 of CAT-iq specifications have been approved by GSMA as minimum requirements for HD voice logo usage on "fixed" devices; i.e., broadband connections using the G.722 codec.

5.3.2. WebRTC relevant use case for G.722

G.722 is the wideband codec required for DECT CAT-iq terminals. The market for DECT cordeless phones including DECT chipset is more than 150 Millions per year and CAT-IQ is a registered trade make in 47 countries worldwide. G.722 has also been specified by ETSI in [TS181005] as mandatory wideband codec for IMS multimedia telephony communication service and supplementary services using fixed broadband access. The support of G.722 would consequently allow

transcoding free IP interoperation between WebRTC client and fixed VoIP terminals including DECT / CAT-IQ terminals supporting G.722. Besides, WebRTC client running on fixed terminals implementing G.722 may reuse the G.722 codec already implemented on these devices.

5.3.3. Guidelines for G.722 usage and implementation

Guidelines for implementing and using G.722 with purpose to ensure interoperability with Multimedia Telephony services overs IMS can be found in section 7 of [TS26.114]. Additional information of G.722 implementation in DECT can be found in [EN300175-8] and full codec description and C source code in [G.722].

5.4. [Codec x] (tbd)

5.4.1. [Codec X] General description

tbd

5.4.2. WebRTC relevant use case for [Codec X]

tbd

5.4.3. Guidelines for [Codec X] usage and implementation with WebRTC

tbd

6. Security Considerations

7. IANA Considerations

None.

8. Acknowledgements

Thanks to Milan Patel for his review.

9. References

9.1. Normative references

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

9.2. Informative references

[EN300175-8]
ETSI, "ETSI EN 300 175-8, v2.5.1: "Digital Enhanced Cordless Telecommunications (DECT); Common Interface (CI); Part 8: Speech and audio coding and transmission".", 2009.

[G.722] ITU, "Recommendation ITU-T G.722 (2012): "7 kHz audio-coding within 64 kbit/s".", 2012.

[I-D.ietf-rtcweb-audio]
Valin, J. and C. Bran, "WebRTC Audio Codec and Processing Requirements", draft-ietf-rtcweb-audio-04 (work in progress), January 2014.

[I-D.ietf-rtcweb-overview]
Alvestrand, H., "Overview: Real Time Protocols for Brower-based Applications", draft-ietf-rtcweb-overview-08 (work in progress), September 2013.

[I-D.ietf-rtcweb-use-cases-and-requirements]
Holmberg, C., Hakansson, S., and G. Eriksson, "Web Real-Time Communication Use-cases and Requirements", draft-ietf-rtcweb-use-cases-and-requirements-14 (work in progress), February 2014.

[IR.36] GSMA, "Adaptive Multirate Wide Band", 2013.

[IR.92] GSMA, "IMS Profile for Voice and SMS", 2013.

[RFC6716] Valin, JM., Vos, K., and T. Terriberry, "Definition of the Opus Audio Codec", RFC 6716, September 2012.

[TS181005]
ETSI, "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); Service and Capability Requirements V3.3.1 (2009-12)", 2009.

[TS26.114]
3GPP, "IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction", 2011.

Authors' Addresses

Stephane Proust
Orange
2, avenue Pierre Marzin
Lannion 22307
France

Email: stephane.proust@orange.com

Espen Berger
Cisco

Email: espeberg@cisco.com

Bernhard Feiten
Deutsche Telekom

Email: Bernhard.Feiten@telekom.de

Bo Burman
Ericsson

Email: bo.burman@ericsson.com

Kalyani Bogineni
Verizon Wireless

Email: Kalyani.Bogineni@VerizonWireless.com

Miao Lei
Huawei

Email: lei.miao@huawei.com

Enrico Marocco
Telecom Italia

Email: enrico.marocco@telecomitalia.it