

Network Working Group
Internet-Draft
Intended status: Informational
Expires: August 29, 2013

X. Marjou
S. Proust
France Telecom Orange
K. Bogineni
Verizon Wireless
R. Jesske
B. Feiten
Deutsche Telekom AG
L. Miao
Huawei
E. Enrico
Telecom Italia
E. Berger
Cisco
February 25, 2013

**WebRTC audio codecs for interoperability with legacy networks.
draft-marjou-rtcweb-audio-codecs-for-interop-01**

Abstract

This document presents use-cases underlining why WebRTC needs AMR-WB, AMR and G.722 as additional relevant voice codecs to satisfactorily ensure interoperability with existing systems. It also presents a way forward that takes into consideration the concerns expressed against the addition of codecs besides Opus and G.711.

It is especially recognized that unjustified additional costs on browsers must be avoided. Therefore, the proposed solution intends to fully rely on the codecs already supported on the devices implementing the browsers and for which license and implementation costs have been already paid.

It is expected that this way forward will significantly limit the costs and technical impacts on browsers while greatly improving interoperability with legacy systems and overall quality. It intends to be considered as a good compromise beneficial to all parties and to the whole industry: the user quality experience will be optimized as a whole at limited additional costs without incurring high costs for both networks to support transcoding and browsers to support additional codecs.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on August 29, 2013.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	4
2.	Terminology	5
3.	Definitions	5
4.	Use cases	5
4.1.	AMR-WB	5
4.1.1.	Use case	5
4.1.2.	Problem	5
4.1.3.	Concerns from the browser manufacturers	7
4.2.	AMR	7
4.2.1.	Use case	7
4.2.2.	Problem	7
4.2.3.	Concerns from the browser manufacturers	8
4.3.	G.722	8
4.3.1.	Use case	8
4.3.2.	Problem	8
4.3.3.	Concerns from the browser manufacturers	9
5.	The proposed way-forward	9
6.	Security Considerations	11
7.	IANA Considerations	11
8.	Acknowledgements	11
9.	References	11
9.1.	Normative references	11
9.2.	Informative references	12
	Authors' Addresses	12

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.

In order to reach the agreement to select OPUS and G.711 as mandatory to implement codecs it was also agreed to consider possible additional codecs in order to take into account the concerns expressed on interoperability issues. The discussion is consequently currently taking place regarding the additional voice/audio codecs that need to be supported. It is mainly questioned whether Opus and G.711 are sufficient to properly address the interoperability issues with legacy systems or if additional codecs need to be supported.

This document presents some use cases highlighting that Opus and G.711 are not sufficient to properly cover all interoperability requirements. In [Section 4](#), important interoperability use-cases are presented describing the interoperability issues that would be encountered if only Opus and G.711 were supported. It therefore advocates for the addition of other codecs while addressing concerns raised against such support.

In [section 5](#), a way forward is proposed that intends to be a real compromise taking into consideration the concerns expressed against additional codecs. It is especially recognized that unjustified additional costs on browsers must be avoided. Therefore the proposed solution intends to strongly limit the cost and technical impact on browsers for the support of additional codecs (including license costs) while improving interoperability with legacy systems.

Regarding audio codecs, it is a common misconception that other existing voice networks only support G.711. Actually existing networks use circuit switched networks as well as voice-over-IP networks like H.323 and SIP-based networks, which means that audio codecs are not limited to G.711. For instance, from use cases described in [[I-D.ietf-rtcweb-se-ucases-and-requirements](#)], it can be foreseen that interoperability with mobile telephony systems will often happen. In such mobile systems, the UE must support the Adaptive Multi-Rate (AMR) speech codec, and if wideband speech communication is offered, the UE must support AMR wideband (AMR-WB) codec. An increasing number of customers are now experiencing high quality voice with HD Voice services over mobile, fixed networks or over the internet. For those customers, any fall back to the G.711 narrow band quality for interoperability purpose could be perceived as a strong and unacceptable quality degradation. Support of G.711 as the only codec for legacy interoperability purposes is currently

not sufficient.

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. Use cases

4.1. AMR-WB

4.1.1. Use case

The market of voice personal communication is driven by mobile terminals and WebRTC technology is expected to be increasingly used on smartphones. Furthermore "HD voice" is gaining momentum and more and more personal communication devices will support wideband communications. Customers are now getting used to the high quality offered by HD Voice mobile devices, CAT-iq fixed HD devices and eventually, HD Voice via WebRTC and OPUS over the internet.

Hence, many communications are expected to be held between a user of a WebRTC endpoint on a device integrating an AMR-WB module who wants to communicate with another user that can only be reached on a mobile device that only supports AMR-WB (and AMR); both endpoints support HD voice quality.

4.1.2. Problem

For this use case, the best situation will be to have the AMR-WB supported by both sides of the connection. Indeed, as mentioned in the introduction, AMR-WB is specified by 3GPP as the mandatory codec to be supported by wideband mobile terminals for a wide range of communication services as described in [[AMR-WB](#)]. This includes the massively deployed circuit switched mobile telephony services and new multimedia telephony services over IP/IMS and 4G/VoLTE as specified by GSMA as voice IMS profile for VoLTE in IR92. Hence, AMR-WB is

strongly increasing with deployment in more than 60 networks from 45 countries and more than 130 types of terminals (c.f. [\[Information-Papers\]](#)).

In that use case, if OPUS and G.711 remain the only codecs supported by the WebRTC endpoints, a gateway must then transcode these codecs into AMR-WB, and vice-versa, in order to implement the use-case. As a consequence, a high number of calls are likely to be affected by transcoding operations producing a degradation of the user quality experience for many customers. This will have a very significant business impact for all service providers on both sides, not only with respect to the transcoding costs but mainly with respect to user experience degradation.

The drawbacks of transcoding are recalled below:

- o Cost issues: transcoding places important additional costs on network gateways for example codec implementation and license costs, deployments costs, testing/validation costs etc... However these costs can be seen as just transferred from the terminal side to the network side. The real issue is rather the degradation of the quality of service affecting the end user perceived quality which will be harmful to all concerned service providers.
- o Intrinsic quality degradation: Subjective test results show that intrinsic voice quality is significantly degraded by transcoding. The degradation is around 0.2 to 0.3 MOS for most of transcoding use cases with AMR-WB at 12.65 kbit/s. It should be stressed that if transcoding is performed between AMR-WB and G.711, wideband voice quality will be lost. Such bandwidth reduction effect clearly degrades the user perceived quality of service leading to shorter and less frequent calls (see `ref_gsma`). Such a switch to G.711 will not be accepted anymore by customers. If transcoding is performed between AMR-WB and OPUS, wideband communication could be maintained. However, as the WB codecs complexity is higher than NB codecs complexity, such WB transcoding is also more costly and degrades the 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.
- o Degraded interactivity due to increased latency: Transcoding means 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.

In addition to these drawbacks related to transcoding, the following issue must be considered:

- o Efficiency over mobile radio access: AMR-WB has been designed and extensively tested for optimized and robust usage over mobile radio access which results in enhanced capacity and efficiency. The mobile radio bearer is optimized for such a codec with channel coding protecting its most sensitive bits. Furthermore, AMR-WB is more efficient than OPUS at regular bit rates used for mobile communication of 12.65 kbit/s with fall back modes down to 6.6 kbit/s. Finally, hardware optimized implementation may allow for less battery consumption

As a consequence, re-using AMR-WB would be beneficial for the specific usage of WebRTC technology over mobile networks. With the strong increase of the smartphone market the capability to use such a mobile codec could strongly enforce and extend the market penetration of the Web RTC technology.

4.1.3. Concerns from the browser manufacturers

The browser manufacturers are concerned by the additional costs that the implementation of AMR-WB would put on browsers which include integration and test costs and codec license costs. The proposed way forward in [Section 5](#) takes carefully into account this concern.

4.2. AMR

4.2.1. Use case

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 a high number of legacy terminals supporting only AMR (terminals with no wideband / HD Voice capabilities) are still used.

4.2.2. Problem

For this use case, the best solution will be to have the AMR supported by both sides of the connection. Indeed, AMR is specified by 3GPP as the mandatory codec to be supported by any mobile terminal

for a wide range of communication services. This includes the massively deployed circuit switched mobile telephony services and new multimedia telephony services over IP/IMS and 4G/VoLTE as specified by GSMA as voice IMS profile for VoLTE in IR92. Hundreds of millions of terminals are consequently currently supporting AMR and are not supporting OPUS nor G.711.

In that use case, if OPUS and G.711 remain the only codecs supported by the WebRTC endpoints the same problem as described in 4.1.1 will be experienced because of transcoding impacts (costs, quality degradation and increased latency) and lower efficiency over mobile radio access.

As a consequence, re-using AMR would be beneficial for the specific usage of WebRTC technology over mobile networks. With the strong increase of the smartphone market the capability to use such a mobile codec could strongly enforce and extend the market penetration of the Web RTC technology.

4.2.3. Concerns from the browser manufacturers

Same as in [Section 4.1.3](#).

4.3. G.722

4.3.1. Use case

As mentioned in [Section 4.1.1](#), HD Voice is gaining momentum and more and more personal communication devices support wideband communications. Customers get used to high quality voice and WebRTC aims at providing high voice quality over internet. In this use case, a user of a WebRTC endpoint on a device integrating G.722 module wants to communicate with another user that can only be reached on a device that only supports G.722 as a wideband codec, G.722 being specified by ETSI as the mandatory wideband codec for New Generation DECT (e.g. CAT-iq compliant).

4.3.2. Problem

For this use case, the best solution will be to have the G.722 supported by both sides of the connection.

Indeed, G.722 has been chosen by ETSI DECT to greatly increase the voice quality by extending the bandwidth from narrow band to wideband. Besides providing high wideband quality, it has low complexity and very low delay. It is widely used in HD fixed services in both hard and soft endpoints.

In this use case, if OPUS and G.711 remain the only codecs supported by the WebRTC endpoints, a gateway must then transcode them from and into G.722 in order to implement the use-case. As in [Section 4.1.2](#), it should be stressed that if transcoding is performed between G.722 and G.711, wideband quality will be lost with fall back to narrow band. This will be perceived as a strong and unacceptable quality degradation by customers experiencing more and more wideband voice calls. It is also important to recall that wideband audio can help persons with hearing impairments to use voice communication over distance and drafts regulations dealing with this requires wide band audio wherever there is voice communication pointing at G.722 as the common codec at least to assure interoperability with wide-band audio between providers.

On the other hand, transcoding with OPUS will greatly increase the complexity, now, as mentioned above, G.722 low complexity was a key factor in many applications mandating G.722.

4.3.3. Concerns from the browser manufacturers

Unlike AMR and AMR-WB, G.722, as G.711, is royalty free. Some concerns about the availability of G.722 PLC were raised. Indeed G.722 and G.711 were initially designed without Packet Loss Concealment (PLC); nevertheless, ITU-T did standardize such functionality as appendices to these recommendations to extend the capabilities of current systems to support new applications and to follow the market demand (for instance when these standards have been widely used in VoIP applications). It has been argued that, unlike the main recommendations, there are non-RF IPR declarations for these PLC appendices (G.711 [Appendix I](#), G.722 Appendices III and IV). It should be recalled that these appendices are only examples and implementers are free to use any PLC solution. For instance in the G.722 case, there are publicly available PLC in the ITU-T Software Tool Library.

5. The proposed way-forward

It is proposed that the browser manufacturers re-use AMR, AMR-WB and G.722 codecs whenever they are already supported on the device on which the browsers are implemented.

AMR and AMR-WB are already supported by millions of devices with license costs and technical costs (implementation, tests...) already paid for these codecs optimized for mobile usage.

Android now provides the APIs needed to give access to all the voice and audio features implemented on the hardware that are required to

develop a voice service. This especially includes AMR and AMR-WB encoding and decoding, adaptive gain control, echo cancellation and noise reduction.

It is therefore technically feasible that browsers offer AMR and AMR-WB as additional RTC Web codecs without re-implementing these codecs and so with very limited additional costs:

- o For implementation, no codec software code has to be developed, tested or integrated directly in the browser (neither for the encoding/decoding nor for related audio functions).
- o For licensing, fees are already paid for the hardware implementation. Since no additional codec is implemented on the device with only one active voice call using the codec at a time it cannot be argued that additional license fees have to be paid in that case. However, in order to fully guarantee this, it is made explicit in the proposed text that the support of AMR and AMR-WB is conditional on the fact that no additional license fee is required.

This proposed way forward is expected to be a good compromise beneficial to all parties. It will optimize the user experience with limited additional costs: excluding high costs for networks to support transcoding and high additional costs on browsers to support additional codecs.

It is consequently proposed the following wording to be added to [Section 3](#) (codec requirements) of [I-D.ietf-rtcweb-audio-codecs]:

3. Audio Codecs

3.1. Required Codecs

To ensure a baseline level of interoperability between WebRTC clients, a minimum set of required codecs are specified below.

WebRTC clients are REQUIRED to implement the following audio codecs.

- * Opus [[RFC6716](#)], with any ptime value up to 120 ms
- * G.711 PCMA and PCMU with one channel, a rate of 8000 Hz and a ptime of 20 - see [section 4.5.14 of \[RFC3551\]](#)
- * Telephone Event - [[RFC4733](#)]

3.2. Additional Codecs

To ensure an enhanced level of interoperability between WebRTC clients, AMR-WB, AMR and G.722 codecs SHOULD be implemented by WebRTC end-points to avoid transcoding costs and quality degradations towards legacy fixed and mobile devices and allow interworking with enhanced voice quality (rather than fall back to G.711 narrow band voice).

WebRTC browsers on devices for which the implementation of AMR is mandatory for voice services MUST allow AMR to be negotiated and used at WebRTC level provided it is ensured that no additional license fees are required.

WebRTC browsers on wide-band devices for which the implementation of AMR-WB is mandatory for voice services MUST allow AMR-WB to be negotiated and used at WebRTC level provided it is ensured that no additional license fees are required.

WebRTC browsers devices for which the implementation of G.722 is mandatory for voice services MUST allow G.722 to be negotiated and used at WebRTC level.

Note: the wording of section "3.2. Additional Codecs" is a first proposal and example to try to capture the general principle explained in this document to improve interoperability while limiting the cost impact on browsers. It is subject to further modifications to reach the best possible compromise.

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.

[RFC2616] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", [RFC 2616](#), June 1999.

9.2. Informative references

[AMR-WB] GSMA, "AMR-WB", 2011.

[I-D.ietf-rtcweb-audio]
Valin, J. and C. Bran, "WebRTC Audio Codec and Processing Requirements", [draft-ietf-rtcweb-audio-01](#) (work in progress), November 2012.

[I-D.ietf-rtcweb-overview]
Alvestrand, H., "Overview: Real Time Protocols for Browser-based Applications", [draft-ietf-rtcweb-overview-06](#) (work in progress), February 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-10](#) (work in progress), December 2012.

[Information-Papers]
GSMA, "Information Papers", 2013.

Authors' Addresses

Xavier Marjou
France Telecom Orange
2, avenue Pierre Marzin
Lannion 22307
France

Email: xavier.marjou@orange.com

Stephane Proust
France Telecom Orange
2, avenue Pierre Marzin
Lannion 22307
France

Email: stephane.proust@orange.com

Kalyani Bogineni
Verizon Wireless

Email: Kalyani.Bogineni@VerizonWireless.com

Roland Jesske
Deutsche Telekom AG

Email: R.Jesske@telekom.de

Bernhard Feiten
Deutsche Telekom AG

Email: R.Jesske@telekom.de

Lei Miao
Huawei

Email: lei.miao@huawei.com

Marocco
Telecom Italia

Email: enrico.marocco@telecomitalia.it

Espen Berger
Cisco

Email: espeberg@cisco.com

