

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 02, 2015

A.B. Roach
Mozilla
July 01, 2014

WebRTC Video Processing and Codec Requirements
draft-ietf-rtcweb-video-00

Abstract

This specification provides the requirements and consideration for WebRTC applications to send and receive video across a network. It specifies the video processing that is required, codecs and their parameters, and types of RTP packetization that need to be supported.

Status of This Memo

This Internet-Draft is submitted 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 January 02, 2015.

Copyright Notice

Copyright (c) 2014 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	2
2. Terminology	2
3. Pre and Post Processing	2
3.1. Camera Source Video	3
3.2. Screen Source Video	3
4. Codec Considerations	3
4.1. VP8	3
4.2. H.264	3
4.3. VP9	4
4.4. H.265	4
5. Dealing with Packet Loss	4
6. Mandatory to Implement Video Codec	4
6.1. Temperature of Working Group	4
7. Security Considerations	5
8. IANA Considerations	5
9. Acknowledgements	6
10. References	6
10.1. Normative References	6
10.2. Informative References	7
Author's Address	7

1. Introduction

One of the major functions of WebRTC endpoints is the ability to send and receive interactive video. The video might come from a camera, a screen recording, a stored file, or some other source. This specification defines how the video is used and discusses special considerations for processing the video. It also covers the video-related algorithms WebRTC devices need to support.

2. Terminology

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

3. Pre and Post Processing

This section provides guidance on pre- or post-processing of video streams.

Unless specified otherwise by the SDP or Codec, the color space SHOULD be TBD.

TODO: What color space is our default?

3.1. Camera Source Video

To support a quality experience with no application level adjustment from the Javascript running in the browsers, WebRTC endpoints are REQUIRED to support:

- o Automatic focus, if applicable for the camera in use
- o Automatic white balance
- o Automatic light level control

TODO: What other processing should be specified here?

3.2. Screen Source Video

If the video source is some portion of a computer screen (e.g., desktop or application sharing), then the considerations in this section also apply.

TODO: What do we need to specify here?

4. Codec Considerations

WebRTC endpoints are not required to support all the codecs in this section.

However, to foster interoperability between endpoints that have codecs in common, if they do support one of the listed codecs, then they need to meet the requirements specified in the subsection for that codec.

All codecs MUST support at least 10 frames per second (fps) and SHOULD support 30 fps. All codecs MUST support a minimum resolution of 320X240.

TODO: These are strawman values. Are they adequate?

4.1. VP8

If VP8, defined in [RFC6386], is supported, then the endpoint MUST support the payload formats defined in [I-D.ietf-payload-vp8]. In addition it MUST support the 'bilinear' and 'none' reconstruction filters.

4.2. H.264

If [H264] is supported, then the device MUST support the payload formats defined in [RFC6184]. In addition, they MUST support Constrained Baseline Profile Level 1.2, and they SHOULD support H.264 Constrained High Profile Level 1.3.

TODO: What packetization modes MUST be supported?

4.3. VP9

If VP9, as defined in [I-D.grange-vp9-bitstream], is supported, then the device MUST support the payload formats defined in TODO.

TODO: The grange-vp9-bitstream draft does not really specify VP9 at all, is there a better reference?

4.4. H.265

If [H265] is supported, then the device MUST support the payload formats defined in [I-D.ietf-payload-rtp-h265].

5. Dealing with Packet Loss

This section provides recommendations on how to encode video to be robust to packet loss.

TODO: What do we want to require in terms of FEC, RTX, interleaving, etc?

6. Mandatory to Implement Video Codec

Note: This section is here purely as a placeholder and there is not yet WG Consensus on Mandatory to Implement video codecs. The WG has agreed not to discuss this topic until September 29, 2014 so that the WG can focus on getting other work done. Please, save your comments on this topic until that time.

The currently recorded working group consensus is that all implementations MUST support a single, specified mandatory-to-implement codec. The remaining decision point is a selection of this single codec.

6.1. Temperature of Working Group

To capture the conversation so far, this section summarizes the result of a straw poll that the working group undertook in December 2013 and January 2014. Respondants were asked to answer "Yes," "Acceptable," or "No" for each option. The options were collected from the working group at large prior to the initiation of the straw poll.

	Yes	Acc	No
	---	---	---
1. All entities MUST support H.264	48%	11%	41%
2. All entities MUST support VP8	41%	17%	42%
3. All entities MUST support both H.264 and VP8	9%	38%	53%
4. Browsers MUST support both H.264 and VP8, other entities MUST support at least one of H.264 and VP8	11%	34%	55%
5. All entities MUST support at least one of H.264 and VP8	10%	16%	74%
6. All entities MUST support H.261	5%	23%	72%
7. There is no MTI video codec	12%	30%	58%
8. All entities MUST support H.261 and all entities MUST support at least one of H.264 and VP8	4%	28%	68%
9. All entities MUST support Theora	7%	26%	67%
10. All entities MUST implement at least two of {VP8, H.264, H.261}	5%	30%	65%
11. All entities MUST implement at least two of {VP8, H.264, H.263}	5%	25%	70%
12. All entities MUST support decoding using both H.264 and VP8, and MUST support encoding using at least one of H.264 or VP8	7%	20%	73%
13. All entities MUST support H.263	6%	19%	75%
14. All entities MUST implement at least two of {VP8, H.264, Theora}	6%	27%	67%
15. All entities MUST support decoding using Theora	1%	15%	84%
16. All entities MUST support Motion JPEG	1%	25%	74%

7. Security Considerations

This specification does not introduce any new mechanisms or security concerns beyond what the other documents it references. In WebRTC, video is protected using DTLS/SRTP. A complete discussion of the security can be found in [I-D.ietf-rtcweb-security] and [I-D.ietf-rtcweb-security-arch]. Implementers should consider whether the use of variable bit rate video codecs are appropriate for their application based on [RFC6562].

8. IANA Considerations

This document requires no actions from IANA.

9. Acknowledgements

The authors would like to thank <GET YOUR NAME HERE - PLEASE SEND COMMENTS>. Thanks to Cullen Jennings for providing text and review. This draft includes text from draft-cbran-rtcweb-codec.

10. References

10.1. Normative References

- [H264] ITU-T Recommendation H.264, "Advanced video coding for generic audiovisual services", April 2013.
- [H265] ITU-T Recommendation H.265, "High efficiency video coding", April 2013.
- [I-D.grange-vp9-bitstream]
Grange, A. and H. Alvestrand, "A VP9 Bitstream Overview", draft-grange-vp9-bitstream-00 (work in progress), February 2013.
- [I-D.ietf-payload-rtp-h265]
Wang, Y., Sanchez, Y., Schierl, T., Wenger, S., and M. Hannuksela, "RTP Payload Format for High Efficiency Video Coding", draft-ietf-payload-rtp-h265-04 (work in progress), May 2014.
- [I-D.ietf-payload-vp8]
Westin, P., Lundin, H., Glover, M., Uberti, J., and F. Galligan, "RTP Payload Format for VP8 Video", draft-ietf-payload-vp8-11 (work in progress), February 2014.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4175] Gharai, L. and C. Perkins, "RTP Payload Format for Uncompressed Video", RFC 4175, September 2005.
- [RFC4421] Perkins, C., "RTP Payload Format for Uncompressed Video: Additional Colour Sampling Modes", RFC 4421, February 2006.
- [RFC6184] Wang, Y.-K., Even, R., Kristensen, T., and R. Jesup, "RTP Payload Format for H.264 Video", RFC 6184, May 2011.

- [RFC6386] Bankoski, J., Koleszar, J., Quillio, L., Salonen, J., Wilkins, P., and Y. Xu, "VP8 Data Format and Decoding Guide", RFC 6386, November 2011.
- [RFC6562] Perkins, C. and JM. Valin, "Guidelines for the Use of Variable Bit Rate Audio with Secure RTP", RFC 6562, March 2012.

10.2. Informative References

- [I-D.ietf-rtcweb-security-arch]
Rescorla, E., "WebRTC Security Architecture", draft-ietf-rtcweb-security-arch-09 (work in progress), February 2014.
- [I-D.ietf-rtcweb-security]
Rescorla, E., "Security Considerations for WebRTC", draft-ietf-rtcweb-security-06 (work in progress), January 2014.

Author's Address

Adam Roach
Mozilla
\
Dallas
US

Phone: +1 650 903 0800 x863
Email: adam@nostrum.com

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 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 18, 2014.

Copyright Notice

Copyright (c) 2014 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.

Table of Contents

1. Introduction	3
2. Terminology	3
3. Definitions	3
4. Rationale for additional WebRTC codecs	3
5. Additional suitable codecs for WebRTC	4
5.1. AMR-WB	5
5.1.1. AMR-WB General description	5
5.1.2. WebRTC relevant use case for AMR-WB	5
5.1.3. Guidelines for AMR-WB usage and implementation with WebRTC	5
5.2. AMR	5
5.2.1. AMR General description	5
5.2.2. WebRTC relevant use case for AMR	6
5.2.3. Guidelines for AMR usage and implementation with WebRTC	6
5.3. G.722	6
5.3.1. G.722 General description	6
5.3.2. WebRTC relevant use case for G.722	6
5.3.3. Guidelines for G.722 usage and implementation	7
5.4. [Codec x] (tbd)	7
5.4.1. [Codec X] General description	7
5.4.2. WebRTC relevant use case for [Codec X]	7
5.4.3. Guidelines for [Codec X] usage and implementation with WebRTC	7
6. Security Considerations	7
7. IANA Considerations	7
8. Acknowledgements	7
9. References	7
9.1. Normative references	7
9.2. Informative references	8
Authors' Addresses	8

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.

- o 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.
- o 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 narrowband 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 cordless 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 over 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