

Network Working Group
Internet Draft
Document: [draft-ietf-avt-variable-rate-audio-00.txt](#)
Expires: April 2005

S. Wenger
C. Perkins

October 2004

RTP Timestamp Frequency for Variable Rate Audio Codecs

Status of this Memo

By submitting this Internet-Draft, I certify that any applicable patent or other IPR claims of which I am aware have been disclosed, or will be disclosed, and any of which I become aware will be disclosed, in accordance with [RFC 3668](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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 cite them other than as "work in progress".

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>

This document is a submission of the IETF AVT WG. Comments should be directed to the AVT WG mailing list, avt@ietf.org.

Abstract

This memo discusses the problems of audio codecs with variable external sampling rates. Historically, for audio codecs, the RTP timestamp frequency was chosen to match the sampling rate of the audio codec. However, this choice is nowadays more difficult to justify, because of the advent of audio codecs (and, even more important, practical use cases) that support multiple sample rates and the switch between the sample rates during the lifetime of an RTP session. This Internet draft addresses the problem by suggesting that RTP Payload RFCs for such codecs to utilize a

single, high, unified RTP timestamp frequency.

1. Introduction

Internet Draft

October 2004

One key property of audio codecs is the external input sample rate. For many of codecs, this sample rate is fixed. ITU-T G.711 [2], also known as a-law and mu-law, uses, for example, a sample rate of 8 kHz. Other audio codecs give the user a choice between different sample rates. However, until recently, applications never changed the sample rate during the lifetime of an RTP session, even if this is technically feasible and probably advantageous from both the user perception, and the network point-of-view. One example for such a codec is MPEG-1 audio, layers 1, 2, or 3 [3]. At the time RTP [1] and the AV-profile [4] was developed, it was a reasonable design choice to use an RTP timestamp frequency that is identical to the codec's input sample rate, as this facilitates sample exact synchronization and processing of media data in endpoints, mixers and translators, among other advantages. Although neither RTP [1] nor the audio-visual profile [4] require the codec sample rate being the same as the RTP timestamp frequency, this paradigm was observed in practice.

Recently, codecs have been developed which do not only support variable sample rates, but use unannounced (in-band only signaled) changes of the sample rate as one of their key mechanisms. Similarly, applications have emerged, that not only support variable sample rates, but, to some extent, rely on this feature. For most (if not all) of these codecs, it is true that the required bit rate and the user experience scales with the sample rate selected. This allows, in the future, a network-dictated scaling of the transmission bit rate of an audio codec -- a feature that was not available before -- which could turn out to be very useful in Internet environments, for example to support congestion control.

With the modern codecs mentioned, the current paradigm of RTP time stamp frequency equal to codec sample rate does not make much sense any more. The purpose of this draft is to provide guidance for the developers of RTP payload specs for codecs with variable sample rate to use a single, relatively high, RTP timestamp frequency, which is specified in this draft.

2. Audio codecs with variable sample rates: Examples

Examples for audio codecs with variable sample rates, that (at least in theory) could switch the sample rate on the fly without out-of-band signaling support, include:

- * AMR-WB+ [5] with a choice of 56 different sample rates
- * VBR-WB [6] with the choice of 8 kHz and 16 kHz sample rates

- * MPEG-4 AAC+ [7] with the choice of (need details here)
- * Any others?

All these codecs use in-band signaling of the sample rate.

3. Rounding

Wenger, Perkins
Internet Draft

Expires April 2005

Page 2
October 2004

It is possible (even likely) that no unified RTP timestamp frequency can be found that, on one hand, fulfills one key requirement spelled out later (namely: is low enough to make timestamp wrap-around during erasure periods unlikely for all practical application scenarios) and, on the other hand, is an integer multiple of all sampling frequencies the codecs support. It is well possible that, in the future, codecs be developed that can make sample rate choices in a granularity of 1 Hz or even finer. Considering this, it is required to specify a rounding algorithm for such cases where no sample-exact position of an audio frame can be found in the RTP timestamp numbering space. Specifying this rounding algorithm ensures that all equipment conforming to this draft use the same rounding algorithm. If that selected rounding algorithm guarantees that inaccuracies do not add up (as spelled out in the requirements later), then even frequent transcoding steps will not lead to an increase to inaccuracy of the timing beyond the unavoidable minimum.

4. Requirements discussion

4.1. Requirements for this draft (general)

- 1) This draft MUST specify a unified RTP timestamp frequency that fulfills the requirements of [section 4.2](#).
- 2) This draft MUST specify a rounding algorithm that can be used for non-sample exact alignment of samples stemming from more than one audio codec, at least one of which having a variable sample rate). The rounding algorithm MUST fulfill the requirements of [section 4.3](#).
- 3) This draft SHOULD state that its provisions MUST be used for the design of future RTP payload formats for audio codecs with variable sample rates
- 4) This draft SHOULD state that its provisions SHOULD be considered in the design of future RTP payload formats for non-audio codecs that have similar problems as variable sample rate audio codecs.
- 5) This draft SHOULD provide an application example for a well-understood variable sample rate codec.

4.2. Requirements for the unified RTP timestamp rate

- 6) The unified RTP timestamp rate (uRTR) MUST be sufficiently high to fulfill the requirements for timestamps in [RFC3550](#)[1]
- 7) The uRTR MUST be low enough to make wrap-arounds of the RTP

timestamp during erasure periods (packet loss bursts) unlikely in all reasonable application scenarios.

Informative note: Such scenarios include, for example, cell handovers in wireless cellular networks, where erasure periods of a few seconds can occur.

- 8) The uRTR SHOULD share the prime factors of the sample rates of the most commonly used fixed sample rate audio codecs, so to allow for sample exact mixing of streams coded by those fixed sample rate audio codecs.
- 9) The uRTR SHOULD be chosen to include a sufficiently high number of prime factors so to support as many future variable sample rate codec code points as possible for sample-exact mixing

4.3. Requirements for a rounding algorithm

- 10) The rounding algorithm MUST be applicable for all sample rates lower than the $0.5 * \text{uRTR}$ specified in this draft.
- 11) The rounding algorithm MAY specify a minimum and maximum sample rate, in units of $x * \text{uRTR}$. Only within this band it is a reasonable expectation that the application of the rounding algorithm does not lead to audible distortions for the common user.
- 12) The rounding algorithm MUST be simple enough to be implemented, without a serious cycle burden, in networking equipment.
- 13) The rounding algorithm SHOULD be implementable in fixed-point arithmetic
- 14) The rounding algorithm MAY, advantageously, be specified such that it does not require division operations
- 15) The rounding algorithm SHOULD be designed such that that multiple applications of the algorithm does not lead to the introduction of errors larger than one tick of the uRTR.

Informative Note: this is a much more difficult requirement as it seems at the first glance. Think of a transcoding scenario where variable goes to 44.1 kHz goes to variable, and the unified timestamp frequency does not share all prime factors of 44.1 kHz. One way out of this would be to rewrite all fixed rate payload specs that use timestamp frequencies that do not fit into the prime factors of the uRTR to be rewritten so to use the uRTR. Is it possible to do this for 44.1 -- or is this nailed down in [RFC3551](#)?

5. Open issues

- * Very general: is this a good idea?
- * What would be a good choice for the uRTR? 192 kHz?
- * Is it a good idea to require ALL future I-Ds on audio (not only the variable clock frequency ones) to use the uRTR?

- * Or only those that do not fit the uRTR (fit == subset of prime factors)?
- * Revisit CD 44.1. No variable sample rate needed? Are there proposals for an 88.2 CD audio codec?

6. Security Considerations

None

7. Congestion Control

None

8. IANA Consideration

None

9. Acknowledgements

None

10. Full Copyright Statement

Wenger, Perkins
Internet Draft

Expires April 2005

Page 4
October 2004

Copyright (C) The Internet Society (2004). This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

11. Intellectual Property Notice

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

12. References

12.1. Normative References

[1] RTP, [RFC 3550](#), STD 64

12.2. Informative References

- [2] G.711
- [3] ISO/IEC 11172 part 3
- [4] RTP AV profile, [RFC 3551](#), STD 65
- [5] AMR-WB+
- [6] VMR-WB
- [7] ISO/IEC 14496 part xxx, AAC+

13. Author's Addresses

Stephan Wenger
Nokia Research Center

Phone: +358-50-486-0637
Email: stewe@stewe.org

Wenger, Perkins Expires April 2005 Page 5
Internet Draft October 2004

P.O. Box 100
FIN-33721 Tampere
Finland

Colin Perkins <csp@csperkins.org>
University of Glasgow
Department of Computing Science
17 Lilybank Gardens
Glasgow G12 8QQ
United Kingdom

14. RFC Editor Considerations

none

Wenger, Perkins Expires April 2005 Page 6