AVT Working Group
Internet-Draft
Intended status: Informational
Expires: February 25, 2010

M. Xu
CS. Zheng
T. Jiang
WX. Zhang
JP. Zhang
Digital Wave Co., Ltd
August 24, 2009

RTP Payload Format for DRA Audio
draft-xu-avt-dra-00

Status of this Memo

Copyright Notice

Abstract

The present document describes a RTP packaging scheme for DRA compressed audio data transmission, as well as the corresponding RTP payload format. According to the properties of DRA compressed audio frame and the Maximum Transmission Unit (MTU) of network, the scheme provides 3 packaging modes for different coding and transmission requirements.

Table of Contents

1. Introduction

DRA multi-channel digital audio coding technology is capable of preserving high fidelity sounds over limited storage and limited transmission bandwidth, suitable for a wide variety of application areas including digital audio broadcasting, digital TV accompany sound, home theater, Internet streaming media, and personal media player. In the fields of IP network transmission and real-time broadcasting, Real-Time Transport Protocol, as described in RFC 3550 [RFC3550], can be employed for transmission and synchronization of DRA audio streams.
This document describes an RTP payload format for transmitting audio data compressed by DRA technology .

2. Requirements Language

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 RFC 2119 [RFC2119].

3. Overview of DRA

DRA digital audio coding technology keeps 24-bit precision (except deliberating quantization loss)for all channels. Its supported channel configurations include commonly known stereo, 5.1 , 6.1, and 7.1 surround sound, besides a large space reserved for future audio development (maximum 64.3 surround sound). DRA can handle standard sampling frequencies spanning from 8kHz to 192kHz, including 44.1kHz and 48kHz. There is no explicit restriction on DRA encoding bitrate, which is determined by the channel bandwidth and audio quality requirement.

DRA has a fixed input frame length of 1024 samples for encoding, thus its time duration can be calculated according to sampling rate. Bit count for encoded frame is controlled by encoding configuration parameters, which are recorded in the frame header. Figure 1 illustrates the structure of a DRA encoded frame described in "Specification for multichannel digital audio coding technology" (China national standard) [GB/T 22726-2008].

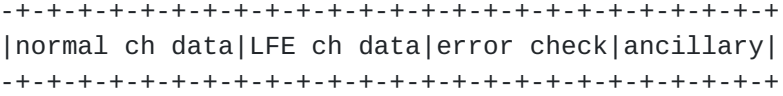

Figure 1: DRA Frame Structure

Every audio frame starts with the sync word 0x7fff, then frame header. Figure 2 illustrates parts of the header structure.

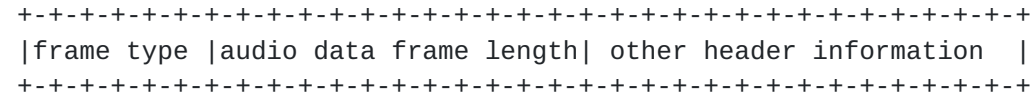

Figure 2: DRA Frame Header's Structure

Frame type: 1 bit, '0' for normal frame header, '1' for extension header.

Audio data frame length: 10 bits (normal frame header) or 13 bits (extension header), representing the length from the sync word to the error check word, counting in unit of 32-bit.

4. RTP Payload Format

DRA encoded frame length varies with sampling rate and bitrate, and possibly overruns the Maximum Transmission Unit (MTU) in high bitrate modes. In this case, a single audio frame is parted to several blocks. On the other hand, DRA supports Variable Bit-Rate (VBR) encoding, allocating bit resources according to perceptible information in each frame. Therefore, within the limit of MTU, several short audio frames are packaged into one RTP packet for higher transmission efficiency.

3 RTP DRA payload formats are employed:

- Single-Frame packet: RTP payload comprising a single complete DRA compressed frame;
- Multi-Frame Packet: RTP payload comprising multiple complete DRA compressed frame;
- Frag-Frame Packet: RTP payload comprising a fragment of one DRA compressed frame.

Subsection 4.2 has detailed description of the above 3 modes.

4.1. RTP Header Usage

The usage of RTP header information specified in this document observes the RTP header information defined inRFC 3550 [RFC3550]. Figure 3 gives the RTP header format for the reader's convenience.

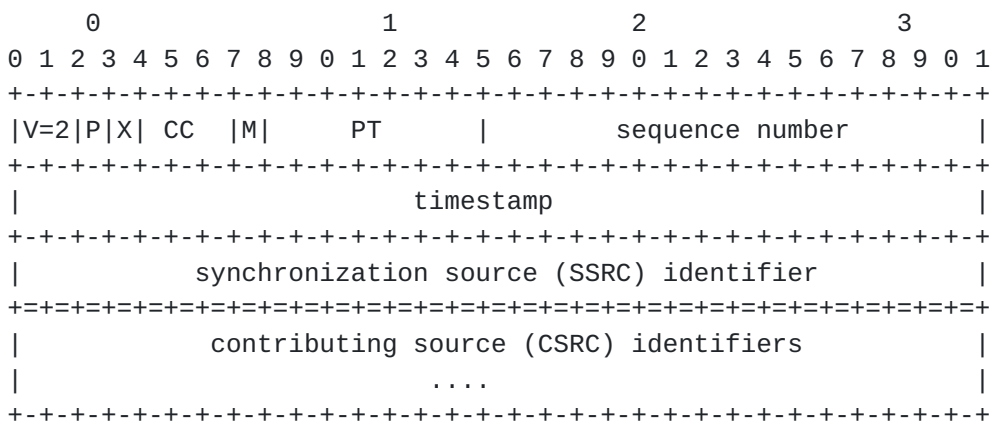

Figure 3: RTP Header according to RFC 3550

Marker bit (M): 1 bit

Set to '1' only if at least one complete audio frame or the last block of an audio frame contained in the current RTP packet.

Payload type (PT): 7 bits

DRA RTP packet payload type designation is not discussed in the present document. Its is set by the RTP profile of DRA RTP payload format, or dynamically set.

TimeStamp (timestamp): 32 bits

The timestamp of RTP is used to record the sampling time of the first audio frame in the current RTP packet. For DRA RTP payload, the clock frequency corresponding to the time stamp is equal to the audio sampling frequency. If the current RTP packet comprises more than one DRA audio frame, the sampling time for each audio frame can be derived from the time stamp plus the accumulated audio frame duration time. In Frag-Frame Packet mode, all RTP packets belong to the same audio frame should have the same time stamp.

4.2. Structure of the RTP DRA Payload Format

RTP DRA payload comprises a 16-bit payload header and audio frame data. Only the latter vary among the 3 RTP packaging modes. Figure 4 illustrates the payload structure of the Multi-Frame Packet mode.

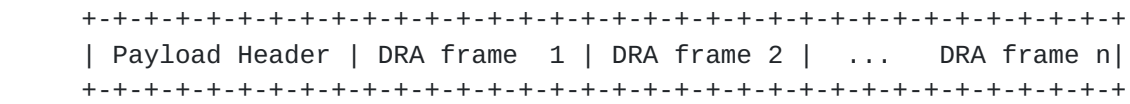

Figure 4: Multi-Frame Packet RTP payload structure

The payload header structure is given in figure 5.

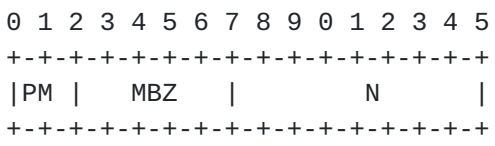

Figure 5: RTP payload header structure

Packaging Mode (PM): 2 bits

Indicating the current RTP packaging modes, as is shown in table 1:

PM value	Mode
0	Single-Frame Packet
1	Multi-Frame Packet
2	Frag-Frame Packet
3	Forbidden

Table 1: Packaging Mode

Must Be Zero (MBZ, reserved): 6 bits

Reserved for future use, all 6 bits must be set to '0'

Number (N, block serial number or frame number): 8 bits

If PM = 1, or Multi-Frame Packet mode, 'Number' is set to the number of DRA frames in the current RTP packet; if PM = 2, 'Number' is set to the serial number of the current block in a DRA audio frame; if PM = 0, 'Number' is set to '1'.

5. SDP related information for DRA Audio

5.1. Media Type Registration

This section uses the template provided in RFC 4288 [RFC4288] and follows RFC 4855 [RFC4855].

- MIME media type name: audio
- MIME subtype name: vnd.dra
- Required parameters:

  rate: The RTP timestamp clock rate is equal to one of the audio sampling rates. The sampling rates consist of 13 possible values as specified in Table 10 of GB/T 22726-2008 [GB/T 22726-2008]: 8000 Hz, 11025 Hz, 12000 Hz, 16000 Hz, 22050 Hz, 24000 Hz, 32000 Hz, 44100 Hz, 48000 Hz, 88200 Hz, 96000 Hz, 176400 Hz, 192000 Hz.

- Optional parameters: none
- Encoding considerations: Transport_provided

  This media type is framed and contains binary data.

- Security considerations:

  See section 7 of this document.

- Interoperability considerations: None
- Published specification:

  This specification and see GB/T 22726-2008 [GB/T 22726-2008].

- Applications which use this media type:

  Audio and audio for video can be multi-channel audio compressed with this media type.

- Additional information:

  Magic number(s): The first 16 bits of an DRA frame are fixed as synchronization word, which is 0x7FFF in hex expression.

- Person & email address to contact for further information:

  Jiang Tian <jiangt@digitalwave.cn>

- Intended usage: COMMON

- Author/Change controller: XU Mao, et al.

5.2. SDP Format for DRA audio

Only the channel number and the sampling rate information are needed besides raw DRA bitstream defined in GB/T 22726-2008 [GB/T 22726-2008] for decoding. Both will be present in SDP defined in RFC 4566 [RFC4566]to assist decoding. But there is no explicit configuration information embedded.

5.2.1. Attributes (a=)

a. First, a short item.a=maxptime:<maximum packet time>

   This is an optional parameter, found by:

   <maximum packet time>

   = <maximum packet size in byte>*8*1000 /<minimum bitrate in kbps>.

   The maximum packet size is confined by Maximum Transmission Unit (MTU).

b. a=rtpmap:<payload type> <encoding name> / <clock rate> [/<encoding parameters>]

   This is a required attribute, where

   <payload type> = 97, defined in RFC 3551 [RFC3551], dynamically assigned

   <encoding name> = "dra"

   <clock rate> = DRA audio sampling rate

   <encoding parameters> = channel number

c. a=fmtp:<format> <format specific parameters>

   This is a optional attribute used if and only if the < format specific parameters> is not empty, where

   <format> = <payload type> = 97,

   <format specific parameters> currently includes only bitrate=#

where "#" is an integer bitrate in bps.

5.2.2. Media Description(m=)

m=<media> <port>/<number of ports> <proto> <fmt>

This is a required field, where

<media> = "audio"

<port> = 58636 or any available even numbered port

<number of ports>

= number of RTP sessions, only used for multiple sessions

<proto> = "RTP/AVP"

<fmt> = <payload type> = 97.

5.2.3. Examples

This section provides two examples of the SDP data for DRA as follows:

a. 48000Hz, stereo DRA bitstream:

m=audio 58636 RTP/AVP 97

a=rtpmap:97 dra/48000/2

b. 320kbps, 44100Hz, 5.1 channel DRA bitstream, two sessions:

m=audio 58800/2 RTP/AVP 97

a=rtpmap:97 dra/44100/6

a=fmtp:97 bitrate=320000

6. IANA Considerations

A new media subtype has been assigned for DRA by IANA; see Section 5.1.

7. Security Considerations

The payload format described in this document is subject to the security considerations defined in RFC 3550 [RFC3550]. Confidentiality protection would have to be applied, so as to protect the users' provacy and contents with copyright.

8. References

8.1. Normative References

[GB/T 22726-2008] You et al., "Specification for multichannel digital audio coding technology", Dec, 2008.

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

[RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.

[RFC3551] Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video Conferences with Minimal Control", STD 65, RFC 3551, July 2003.

[RFC4288] Freed, N. and J. Klensin, "Media Type Specifications and Registration Procedures", BCP 13, RFC 4288, December 2005.

[RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.

[RFC4855] Casner, S., "Media Type Registration of RTP Payload Formats", RFC 4855, February 2007.

8.2. Informative References

[RFC2736] Handley, M. and C. Perkins, "Guidelines for Writers of RTP Payload Format Specifications", BCP 36, RFC 2736, December 1999.

[RFC4184] Link, B., Hager, T., and J. Flaks, "RTP Payload Format for AC-3 Audio", RFC 4184, October 2005.

Authors' Addresses

Mao Xu
Digital Wave Co., Ltd
Room 2002A, Building B, Cyber Tower, No.2, ZhongGuanCun South Avenue
Beijing, 100031
China

Phone: +86-10-5278-6263
Email: xumao@digitalwave.cn

ChuSheng Zheng
Digital Wave Co., Ltd
Room 2002A, Building B, Cyber Tower, No.2, ZhongGuanCun South Avenue
Beijing, 100031
China

Phone: +86-10-5278-6263
Email: zhengcs@digitalwave.cn

Tian Jiang
Digital Wave Co., Ltd
Room 2002A, Building B, Cyber Tower, No.2, ZhongGuanCun South Avenue
Beijing, 100031
China

Phone: +86-10-5278-6263
Email: jiangt@digitalwave.cn

WeiXiong Zhang
Digital Wave Co., Ltd
Room 2002A, Building B, Cyber Tower, No.2, ZhongGuanCun South Avenue
Beijing, 100031
China

Phone: +86-10-5278-6263
Email: zhangwx@digitalwave.cn

JingPing Zhang
Digital Wave Co., Ltd
Room 2002A, Building B, Cyber Tower, No.2, ZhongGuanCun South Avenue
Beijing, 100031
China

Phone: +86-10-5278-6263
Email: zhangjp@digitalwave.cn