

RTP Payload Format for nonlinear 12 bits Audio on DV

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of [Section 10 of RFC2026](#).

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 to cite them other than as "work in progress."

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

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

[1](#). Abstract

This document specifies the packetization scheme for encapsulating the 12 bits nonlinear audio data streams used in "DV" video into a payload of the Real-Time Transport Protocol (RTP).

[2](#). Introduction

This document provides the information of 12 bits nonlinear audio used in the DV format and specifies the encapsulation into the Real-time Transport Protocol (RTP), version 2 [[1](#),[2](#)]. Also, this document just specifies the differentiated part of 16 bit linear audio as L16 [[3](#),[4](#)]. Reader is recommended to consult the L16 document with this one.

[3.](#) The need for the RTP encapsulation for 12 bits nonlinear DV audio.

The HD Digital VCR Conference has published a digital video specification set entitled "Specification of Consumer-Use Digital VCRs using 6.3mm magnetic tape" [[1](#)]. The digital video format defined by that specification is commonly known as "DV" format. The original DV format treats whole of the data including audio and video as single bundled stream data. On the other hand, RTP recommends that different media data will transport different RTP streams, even if the both streams made by the same source. Therefore, RTP encapsulation format of DV stream also recommends audio and video streams transport different RTP streams with its corresponding RTP format. In the DV standard, audio data encodes PCM and three types of encoding format are defined, i.e. 16 bits linear 20 bits linear and 12 bits nonlinear.(20 bits linear has not been used yet.) The RTP encapsulation format for audio previously published supports 16 bits linear audio only [[3,4](#)].

The format of 12 bits nonlinear DV audio is congruent with 16 bits linear audio except the format of single sampled data element. An element of 12 bits nonlinear audio data can be obtained from the single sampled element of 16 bits linear one. It is not difficult to convert 12 bits nonlinear into 16 bits linear on the sender side and send it as L16 audio previously defined. However, the amount of the data size of 16 bits increases 33% compared with the 12 bits and it waste network bandwidth with meaningless data.

[4.](#) 12 bits nonlinear audio format in DV (DV12)

The data of 12 bits nonlinear DV audio is derived from the single sampled data of the 16 bit linear audio format. The conversion detail between 16 and 12 bits is shown in the Table. Three levels of sampling frequency are defined in the DV specification, i.e. 32kHz, 44.1kHz and 48kHz. All the values are included by the samplig rates listed in L16 documents. And other parameters, encapsulation format and also MIME description are discussed in L16 document. When 12 bits size sampled data is packed into payload, the most significant bit MUST be encodes first. The sample code for packing 12 bits DV audio into RTP payload is shown in Appendix. 12 bits length of a sampled data does not accord to the 8 bits byte boundary of RTP payload. When odd number of samples in the payload, four LSBs data of the last byte is unused.

16 bits linear (X)	12 bits nonlinear (Y)
-----	-----
32,767 (7FFFh) $Y = \text{INT}(X/64) + (600h)$	2,047 (7FFh)
16,384 (4000h)	1,792 (700h)

Internet Draft

June 25, 1999

16,383 (3FFFh) Y = INT(X/32) + (500h)	1,791 (6FFh)
8,192 (2000h)	1,536 (600h)
8,191 (1FFFh) Y = INT(X/16) + (400h)	1,535 (5FFh)
4,096 (1000h)	1,280 (500h)
4,095 (0FFFh) Y = INT(X/8) + (300h)	1,279 (4FFh)
2,048 (0800h)	1,024 (400h)
2,047 (07FFh) Y = INT(X/4) + (200h)	1,023 (3FFh)
1,024 (0400h)	768 (300h)
1,023 (03FFh) Y = INT(X/2) + (100h)	767 (2FFh)
512 (0200h)	512 (200h)
511 (01FFh) Y = X	511 (1FFh)
0 (0000h)	0 (000h)
-1 (FFFFh) Y = X	-1 (FFFh)
-512 (FE00h)	-512 (E00h)
-513 (FFFFh) Y = INT((X + 1)/2) - (101h)	-513 (DFFh)
-1,024 (FE00h)	-768 (D00h)
-1,025 (FBFFh) Y = INT((X + 1)/4) - (201h)	-769 (CFFh)
-2,048 (F800h)	-1,024 (C00h)
-2,049 (F7FFh) Y = INT((X + 1)/8) - (301h)	-1,025 (BFFh)
-4,096 (F000h)	-1,280 (B00h)
-4,097 (EFFFh) Y = INT((X + 1)/16) - (401h)	-1,281 (AFFh)
-8,192 (E000h)	-1,536 (A00h)
-8,193 (DFFFh) Y = INT((X + 1)/32) - (501h)	-1,537 (9FFh)
-16,384 (C000h)	-1,792 (900h)
-16,385 (BFFFh) Y = INT((X + 1)/64) - (601h)	-1,793 (8FFh)
-32,768 (8000h)	-2,048 (800h)

Table. Conversion between 16 bits to 12 bits [[1](#)]

6. Security Considerations

RTP packets using the payload format defined in this specification are subject to the security considerations discussed in the RTP specification [2], and any appropriate RTP profile. This implies that confidentiality of the media streams is achieved by encryption.

Internet Draft

June 25, 1999

Because the data compression used with this payload format is applied end-to-end, encryption may be performed after compression so there is no conflict between the two operations.

A potential denial-of-service threat exists for data encodings using compression techniques that have non-uniform receiver-end computational load. The attacker can inject pathological datagrams into the stream which are complex to decode and cause the receiver to be overloaded. However, this encoding does not exhibit any significant non-uniformity.

As with any IP-based protocol, in some circumstances a receiver may be overloaded simply by the receipt of too many packets, either desired or undesired. Network-layer authentication may be used to discard packets from undesired sources, but the processing cost of the authentication itself may be too high. In a multicast environment, pruning of specific sources may be implemented in future versions of IGMP [5] and in multicast routing protocols to allow a receiver to select which sources are allowed to reach it.

7. Full Copyright Statement

Copyright (C) The Internet Society (1999). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works.

However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be

revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS 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."

Kobayashi, et al

Expires December 1999

[Page 4]

Internet Draft

June 25, 1999

8. Authors' Addresses

Katsushi Kobayashi
Communication Research Laboratory
4-2-1 Nukii-kita machi, Koganei
Tokyo 184-8795
JAPAN
EMail: ikob@koganei.wide.ad.jp

Akimichi Ogawa
Keio University
5322 Endo, Fujisawa
Kanagawa 252
JAPAN
EMail: akimichi@sfc.wide.ad.jp

Stephen L. Casner
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
United States
EMail: casner@cisco.com

Carsten Bormann
Universitaet Bremen FB3 TZI
Postfach 330440
D-28334 Bremen, GERMANY
Phone: +49.421.218-7024
Fax: +49.421.218-7000
EMail: cabo@tzi.org

10. Bibliography

- [1] IEC 61834, Helical-scan digital video cassette recording system using 6,35 mm magnetic tape for consumer use (525-60, 625-50, 1125-60 and 1250-50 systems)

- [2] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson. RTP: A transport protocol for real-time applications. IETF Audio/Video Transport Working Group, January 1996. [RFC1889](#).
- [3] Schulzrinne, H., "RTP Profile for Audio and Video Conferences with Minimal Control", [RFC 1890](#), January 1996.
- [4] Salsman, J., "The Audio/L16 MIME content type", [RFC 2586](#), May 1999
- [5] Deering, S., "Host Extensions for IP Multicasting", STD 5,

Kobayashi, et al

Expires December 1999

[Page 5]

Internet Draft

June 25, 1999

[RFC 1112](#), August 1989.

[Appendix A](#). Sample code for packing and unpacking

```
int pack12(short[] s, unsigned char[] b1, int n) {
    unsigned char *b = b1;
    while (n >= 2) {
        n -= 2;
        int s1 = *s++;
        int s2 = *s++;
        *b++ = s1 >> 4;
        *b++ = s1 << 4 + ((s2 >> 4) & 0xF);
        *b++ = s2;
    }
    if (n == 1) {
        int s1 = *s++;
        *b++ = s1 >> 4;
        *b++ = s1 << 4;
    }
    return b - b1;
}

int unpack12(unsigned char[] b, short[] s1, int n) {
    short *s = s1;
    while (n >= 3) {
        n -= 3;
        *s++ = b[0] << 4 + b[1] >> 4;
        *s++ = b[1] << 8 + b[2];
        b += 3;
    }
    if (n == 2) {
        *s++ = b[0] << 4 + b[1] >> 4;
    } else if (n == 1) {
```

```
        error("alignment error");
    }
    return s - s1;
}
```