SMPTE timecode in Dolby audio bitstreams

11 January 2018
Confidential information

Copyright

For information, contact:

Dolby Laboratories, Inc.
1275 Market Street
San Francisco, CA 94103-1410 USA
Telephone 415-558-0200
Fax 415-863-1373
http://www.dolby.com

Trademarks

Dolby and the double-D symbol are registered trademarks of Dolby Laboratories. Following are trademarks of Dolby Laboratories:

Dolby®	Dolby Home Theater®
Dolby Atmos®	Dialogue Intelligence™
Dolby Audio™	Dolby Digital Plus Home Theater™
Dolby Cinema™	MLP Lossless™
Dolby Theatre®	Pro Logic®
Dolby Vision™	Surround EX™
Dolby Voice®	Dolby Digital Plus Advanced Audio™
Feel Every Dimension in Dolby™	Dolby Fidelio™
Feel Every Dimension™	Dolby AccessLink™
Feel Every Dimension in Dolby Atmos™	Dolby CaptiView™
Dolby Digital Plus™	Dolby CineAsset™
Dolby Advanced Audio™	Dolby CineAsset Player™

All other trademarks remain the property of their respective owners.

Confidential information

Contents

1 Introduction

This document describes the methods used to transmit SMPTE timecode within Dolby Digital, Dolby Digital Plus and Dolby TrueHD bitstreams.

- Overview
- About this documentation
- Resources
- Contacting Dolby

1.1 Overview

The ability to identify specific video frames regardless of the frame rate makes timecode essential for video editing. Likewise, timecode is also important for editing audio elementary streams, especially when simultaneously editing video and its associated audio. In addition to editing, timecode is also important in many multiplexing applications. Many file-based, streaming, and disc formats utilize timecode to maintain audio and video synchronization during the multiplexing process.

Timecode systems assign a number (or timestamp) to each video frame to allow each frame to be uniquely identified. Timecode systems support common video frame rates, including drop-frame and non-drop-frame variants of non-integer NTSC frame rates. Typically timecode systems encode time data as timestamps in the form of hours, minutes, seconds, and frames (HH:MM:SS:FF). This format provides a range of time data that is dependent on the frame rate of the video. For example, the timecode for a video stream with a frame rate of 30 frames per second (fps) has a range of 00:00:00:00 to 23:59:59:29. However, for encoded audio streams such as Dolby Digital Plus and Dolby TrueHD bitstreams, timecode does not identify individual audio samples, but rather represents how the first sample in a group of encoded samples relates to an associated video frame.

Timestamps can be inserted into Dolby Digital, Dolby Digital Plus and Dolby TrueHD bitstreams to provide an explicit relationship between the audio bitstream and accompanying video frames. Although all Dolby audio formats use the same coded timestamp format, the location and frequency of the timestamps differs between Dolby Digital, Dolby Digital Plus and Dolby TrueHD, due to each technology having different bitstream organizations.

1.2 About this documentation

This document:

- Describes the basic organization of Dolby Digital, Dolby Digital Plus and Dolby TrueHD bitstreams.
- Presents the syntax of the timestamps used in audio bitstreams with timecode.
- Describes how timestamps are carried as part of Dolby Digital, Dolby Digital Plus and Dolby TrueHD bitstreams.
- Describes the timing relationship between the audio bitstream and the video frames identified by the timestamps.
- Specifies how to use timestamps in multiplexing, packaging and audio interface transmission.

- Specifies when timestamps need to be removed from audio bitstreams to ensure compatibility with content distribution formats.

This document assumes a basic knowledge of the Dolby Digital, Dolby Digital Plus and Dolby TrueHD audio coding technologies. Only topics related to embedding timecode information into the audio bitstreams are covered here.

Related Information

Dolby audio bitstream organization on page 7

SMPTE timestamp specification on page 12

SMPTE timestamp location and frequency on page 15

Using SMPTE timestamps from Dolby audio bitstreams on page 19

1.3 Resources

Table 1 lists documents that supplement the information in this document. These documents include specifications of Dolby technologies.

Table 1: Supplemental resources

Document/Specification	Relevant Sections or Content Description
ETSI TS 102 366 v.1.4.1, Digital Audio Compression (AC-3, Enhanced AC-3) Standard	Dolby Digital (AC-3), and Dolby Digital Plus (E-AC-3) bitstream syntax
Meridian Lossless Packing Technical Reference for FBA and FBB streams – Version 1.0	Meridian Lossless Packing (MLP/Dolby TrueHD) bitstream syntax

1.4 Contacting Dolby

Support services are available to address any questions and to provide advice about implementing Dolby technology.

For any questions regarding the design, porting, or testing of an Implementation, contact Dolby at implementationsupport@dolby.com. By utilizing Dolby expertise, especially during the design process, many problems that might require design revisions before an Implementation is fully approved can be prevented.

If you have comments or feedback about this document, send us an email at documentation@dolby.com.

2 Dolby audio bitstream organization

Although utilized in several common applications, for example Blu-ray Disc, because Dolby Digital and Dolby Digital Plus offer high compression ratios, whereas Dolby TrueHD is a lossless compression format, the bitstream organizations for these technologies are very different.

- Dolby Digital bitstreams
- Dolby Digital Plus bitstreams
- Dolby TrueHD bitstreams

2.1 Dolby Digital bitstreams

Dolby Digital uses the AC-3 coding system to encode audio data. The encoded Dolby Digital bitstream consists of a series of **syncframes**. Each syncframe contains the **syncinfo** and **bsi** fields, six **audio blocks**, and the **auxdata**, and **errorcheck** bitfields. The audio blocks contain the encoded audio data for the channels which they convey. The basic organization of the bitstream is shown in Figure 1.

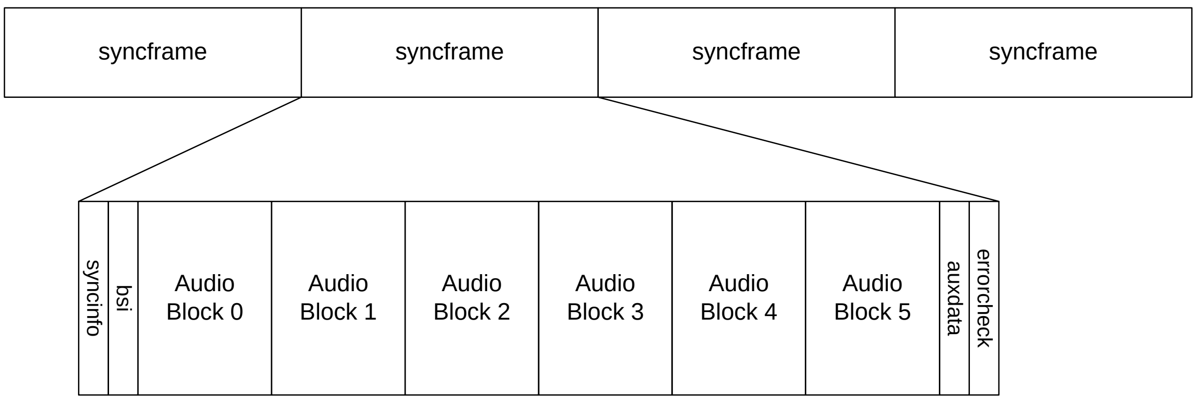

Figure 1: Section of a Dolby Digital bitstream

2.2 Dolby Digital Plus bitstreams

Dolby Digital Plus uses the Enhanced AC-3 coding system to encode audio data. The encoded Dolby Digital Plus bitstream consists of a series of **syncframes**. In its most basic form, syncframes contain the **syncinfo**, **bsi**, **audfrm**, **auxdata**, and **errorcheck** bitfields. The **audfrm** section of the syncframe contains one, two, three or six **audio blocks.** These audio blocks contain the encoded audio data for the channels which they convey. The basic organization of the bitstream is shown in Figure 2.

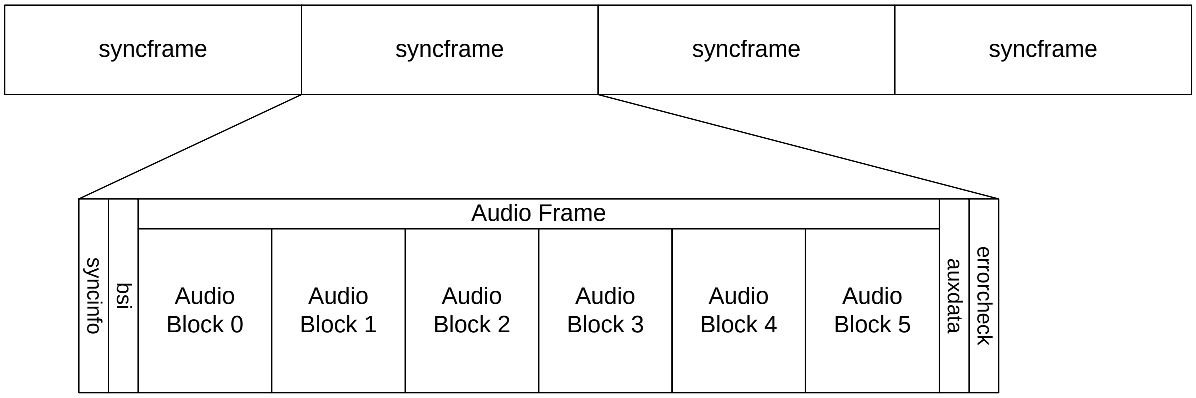

Figure 2: Section of a Dolby Digital Plus Bitstream (six audio blocks per syncframe shown)

Although the above figure shows six audio blocks, a Dolby Digital Plus syncframe can carry either one, two, three, or (most commonly) six audio blocks per syncframe. A decoder can decode individual syncframes regardless of the number of audio blocks contained within a frame. However, some applications require six audio blocks to be available before processing can begin. Depending on the number of audio blocks per syncframe, several syncframes may need to be buffered before these applications allow processing to begin.

2.2.1 Dolby Digital Plus syncframe grouping

As previously discussed, many applications require Dolby Digital Plus syncframes to be buffered and managed before processing can begin. In most cases, six audio blocks from each substream must be buffered. Table 2 lists the terms used in this document to describe various groupings of syncframes.

Table 2: Dolby Digital Plus bitstream and syncframe grouping terms

Term	Definition
Audio Block	One block of Dolby Digital Plus-encoded audio (256 PCM samples). Each audio block has a duration of 5.333 ms at a sampling rate of 48 kHz.
Syncframe	A syncframe as defined in the ETSI TS 102 366 specification (Other Resources). One Dolby Digital Plus syncframe may contain 1, 2, 3, or 6 audio blocks, corresponding to a duration of 256, 512, 768 or 1,536 samples (5.333, 10.667, 16 or 32 ms, at 48 kHz).
Time-slice	A time-slice consists of one syncframe from every substream in the Dolby Digital Plus bitstream. The duration of a time-slice is defined by the number of audio blocks in the syncframes. For example, a time-slice that consists of syncframes that contain three audio blocks has a duration of 768 samples (16 ms at 48 kHz).
Dolby Digital Plus Access Unit	A Dolby Digital Plus access unit consists of all of the syncframes necessary to deliver six blocks of audio data from every substream in the Dolby Digital Plus bitstream. If the number of audio blocks per syncframe is less than six, an access unit consists of more than one time slice. For example, if the syncframes in the bitstream each contain two audio blocks, an access unit from that bitstream consists of three time-slices. An access unit always represents 1,536 PCM-samples worth of audio data from every channel in every substream.

2.2.2 Multiple substreams

Dolby Digital Plus uses substreams to support transmission of additional channels as well as additional programs. There are two types of substreams: independent and dependent.

- Independent - An independent substream is the foundation of the audio program. It can be decoded independently by any Dolby Digital Plus decoder. Any standard 5.1-channel (or less) Dolby Digital Plus audio program is an independent substream. To carry multiple audio programs, multiple independent substreams are carried in the bitstream. A Dolby Digital Plus bitstream can contain up to eight independent substreams.
- Dependent - A dependent substream must be decoded in parallel with a corresponding independent substream to produce an extended-channel output program. Dependent substreams are not allowed to be decoded separately from the independent substream with which they are associated. Each independent substream can have up to eight dependent substreams.

  Note: In a multiple substream Dolby Digital Plus bitstream, all syncframes must have the same number of audio blocks per syncframe.

Like the basic Dolby Digital Plus bitstream, a multiple substream Dolby Digital Plus bitstream contains a series of syncframes. However, in a multiple substream bitstream there is a series of syncframes for each substream present. The substreams are interleaved on a syncframe-by-syncframe basis, with syncframes from dependent substreams immediately following the syncframes from independent substream with which they are associated. The use of multiple substreams allows for great flexibility, and many bitstream scenarios exist with Dolby Digital Plus. Here are a few examples:

Common Dolby Digital Plus bitstream and substream configurations

The following examples show configurations of Dolby Digital Plus bitstreams that are commonly used for content delivery applications.

- The simplest Dolby Digital Plus bitstream contains only one independent substream (I0) with a single program. This situation is identical to the basic bitstream shown in Figure 2.

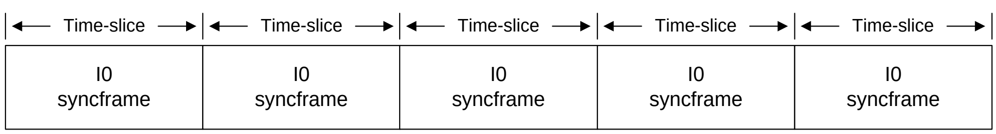

Figure 3: Dolby Digital Plus bitstream consisting of I0 syncframes only

- A bitstream carrying a single 7.1-channel program. The bitstream consists of one independent substream (I0) and one dependent substream (D0).

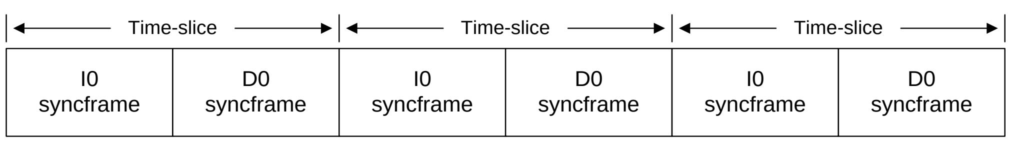

Figure 4: Dolby Digital Plus bitstream with I0 and D0 syncframes

- A bitstream carrying a single 7.1-channel primary audio program on a Blu-ray Disc. The bitstream consists of one independent substream (I0) and one dependent substream (D0). Due to the requirements of the Blu-ray Disc specification, the independent substream consists of Dolby Digital syncframes, and the dependent substream consists of Dolby Digital Plus syncframes. In this case, the number of audio blocks per syncframe is always six.

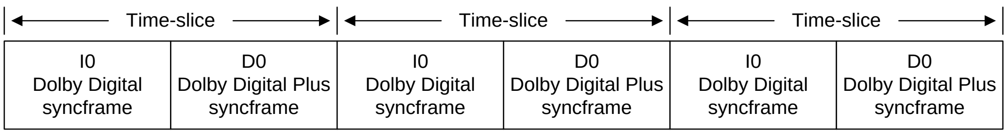

Figure 5: Dolby Digital Plus primary audio bitstream on a Blu-ray Disc with I0 Dolby Digital syncframes and D0 Dolby Digital Plus syncframes

Less common Dolby Digital Plus bitstream and substream configurations

The following examples show configurations of Dolby Digital Plus bitstreams that are supported by the Dolby Digital Plus format, but are not commonly used for content delivery applications.

- A bitstream carrying a single program with more than 7.1 channels. The bitstream consists of one independent substream (I0) and two dependent substreams (D0 and D1).

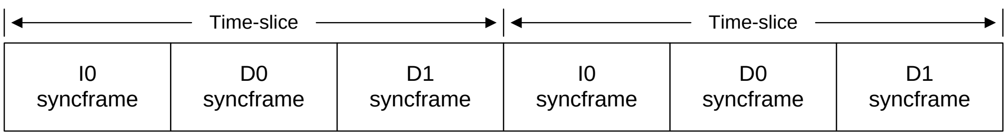

Figure 6: Dolby Digital Plus bitstream with I0, D0, and D1 syncframes

- A bitstream carrying two programs. Program 0 is carried in one independent substream (I0) and two dependent substreams (D0 and D1). Program 1 is carried in one independent substream (I1) and one dependent substream (D0).

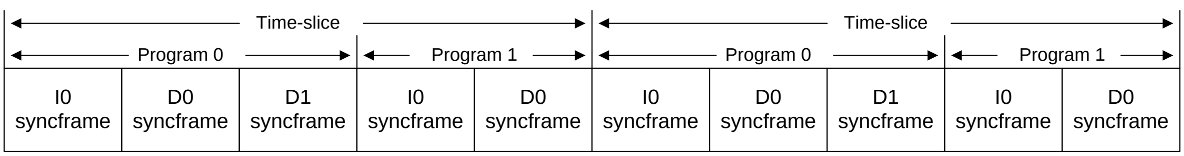

Figure 7: Dolby Digital Plus bitstream carrying two programs with I0, D0, D1, I1 and D0 syncframes

2.3 Dolby TrueHD bitstreams

Dolby TrueHD utilizes the lossless audio coding system MLP to encode audio data. The encoded Dolby TrueHD bitstream consists of a series of **access units**. Each access unit contains an **MLP Sync** followed by a segment from each substream. The substreams contain the encoded audio data for the channels which they convey. The Dolby TrueHD bitstream in inherently variable rate as the size of each access unit is not constant. The basic organization of the bitstream is shown in Figure 8.

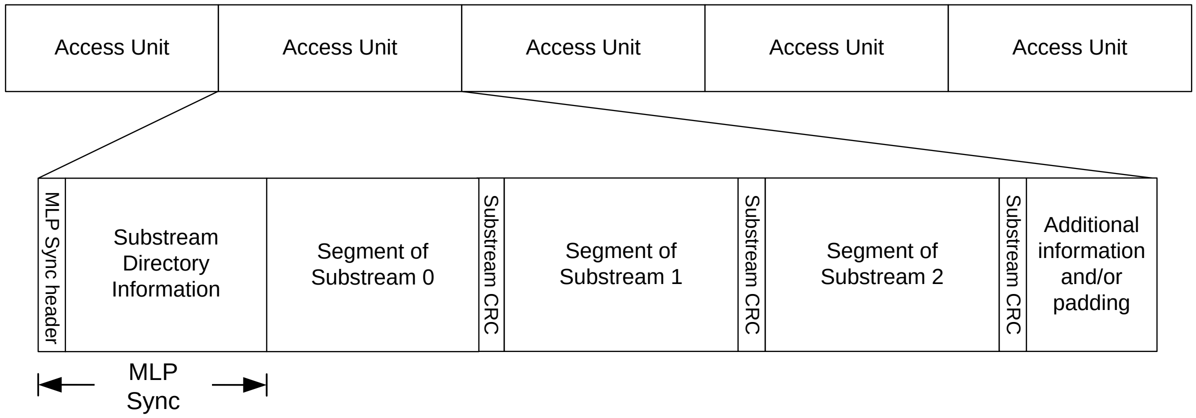

Figure 8: Section of a Dolby TrueHD (MLP) Bitstream

There are two types of MLP Sync: the **major sync** and the **minor sync**. Each sync contains a header and the substream directory information. The major sync has an expanded header containing all the information required to start full decoding of the stream, whereas a minor sync's header consists of no more than size and bitstream timing data. To minimize data rate overheads, most access units begin with a minor sync. Access units that begin with a major sync occur typically once per 128 access units. All minor syncs are associated with the first major sync immediately preceding it. The use of major and minor syncs means that not all access units are independently decodable.

In general the relationship between the sample rate of the encoded audio, video frame rate, and the frequency of access units is non-integer. Although editing systems define edit points in terms of timecode, an edit of a Dolby TrueHD bitstream must occur at an access unit boundary. To allow accurate conversion to/from timecode and access unit number, the Dolby TrueHD bitstream contains high resolution timing information that is embedded in the substream segments of access units that begin with a major sync. This information is serialized across multiple access units and allows for precise editing of the bitstream regardless of the frame rate of the associated video.

3 SMPTE timestamp specification

This section presents the syntax and semantics of the SMPTE timestamp.

- Timestamp syntax language description
- Description of timestamp() Elements

3.1 Timestamp syntax language description

The SMPTE timestamp syntax is expressed in a notation derived mainly form the C programming language.

Entries in bold, for example

```
    start_syncword                                   v(16)        3.3.1
```

signify the occurrence of the corresponding variable in the bitstream. The small print on the right-hand side of the page specifies the encoding (according to the notation in Bitfield encoding) and gives (where applicable) a reference to a section containing more information about the variable.

Substructures are defined using the **C** procedure syntax:

```
name(<optional arguments>) {
   .
   .
}
```

The substructure is expanded by the C call syntax name(<arguments>). Iteration is expressed using the usual **C** for, while and do...while statements. Conditional expansion uses the if (. . .) ... and if (. . .) ... else ... constructs.

The frequently required procedure of reading a flag from the stream and performing an <action> if the flag is TRUE (binary '**1**') is expressed as

```
    if(flag)                                         b(1)
       <action>
```

which can be regarded as shorthand for

```
    flag                                             b(1)
    if(flag)
       <action>
```

Flags in while statements obey similar rules.

> **Note:** In some cases flags may not represent an actual bitfield Boolean value, but may be presented as pseudo-code that has a value of TRUE or FALSE.

3.1.1 Bitfield encoding

The encoded data can be thought of as a serial bitstream containing items of data that occupy fields of varying width, e.g. three bits to specify a number of channels in the range [0. . .7]. The bits occupying each field are presented most significant bit first.

Values are encoded in signed or unsigned bitfields and are referred to in the syntax description using the following notation:

b(1) Boolean 1-bit: 1 → TRUE, 0 → FALSE

v(n) Bitfield taking n bits, with arbitrary representation

u(n) Unsigned integer taking n bits ($n \geq 0$, u(0)=0) bcd(n) Packed binary coded decimal format taking n bits (see Packed Binary Coded Decimal format for more information)

reserved bitfields are encoded as v(n) and should be ignored by timestamp readers.

3.1.2 Packed Binary Coded Decimal format

Many SMPTE timestamp parameters are delivered in packed binary coded decimal (BCD) format. This includes the hours, minutes, seconds, and frames parameters.

Note: Since none of the SMPTE parameters ever exceeds 99, only the two least significant nibbles are used.

The following formula encodes an input integer between 0 and 99 into packed BCD:

```
packed_bcd = ((input * 0.1 & 0xF) << 4) + ((input % 10) & 0xF);
```

Likewise, the following formula decodes a packed BCD to an integer between 0 and 99:

```
output = ((packed_bcd >> 4) & 0xF) * 10 + (packed_bcd & 0xF);
```

3.2 Timestamp syntax specification

The syntax of the SMPTE timestamp is shown below.

Syntax	Encoding	Section
timestamp() {		
start_syncword	v(16)	3.3.1
hours	bcd(16)	3.3.2
minutes	bcd(16)	3.3.3
seconds	bcd(16)	3.3.4
frames	bcd(16)	3.3.5
samples	u(16)	3.3.6
reserved1	v(10)	3.3.7
framerate	v(4)	3.3.8
reserved2	v(1)	3.3.9
dropframe	b(1)	3.3.10
reserved3	v(16)	3.3.11
}		

3.3 Description of timestamp() Elements

This section describes the timestamp elements. Each variable is followed by its encoding format. Many SMPTE timestamp parameters are delivered in packed binary coded decimal (BCD) format. This includes the **hours**, **minutes**, **seconds**, and **frames** parameters.

3.3.1 start_syncword v(16)

A 16-bit synchronization word, **start_syncword** signifies the start of the timestamp and is provided at the beginning so that the timestamp can be recognized without excessive navigation. The value is always 0x0110, or '0000 0001 0001 0000'.

3.3.2 hours bcd(16)

This field indicates the hours setting for the associated SMPTE video frame.

3.3.3 minutes bcd(16)

This field indicates the minutes setting for the associated SMPTE video frame.

3.3.4 seconds bcd(16)

This field indicates the seconds setting for the associated SMPTE video frame.

3.3.5 frames bcd(16)

This field indicates the frame number setting for the associated SMPTE video frame.

3.3.6 samples u(16)

This field indicates the sample offset between the first audio sample of the syncframe (in the case of a Dolby Digital Plus bitstreams) or bitstream (in the case of a Dolby TrueHD bitstream), and the audio sample that corresponds to the start of the corresponding SMPTE video frame. The units of the samples field are audio samples at the sampling rate of the Dolby audio bitstream.

Related information

SMPTE video frame and audio frame association on page 16

3.3.7 reserved1 v(10)

This field is reserved for future use.

3.3.8 framerate v(4)

This field indicates the frame rate for the associated SMPTE video frame. Table 3 lists the possible frame rates. If a reserved code is indicated, the timestamp should be considered invalid.

Table 3: framerate code value and corresponding video frame rate

framerate Code	Video Frame Rate
'0000'	reserved
'0001'	23.98 fps
'0010'	24.00 fps
'0011'	25.00 fps
'0100'	29.97 fps
'0101'	30.00 fps
'0110'	50.00 fps
'0111'	59.94 fps
‘1000’	60.00 fps
‘1001’ – ‘1111’	reserved

3.3.9 reserved2 v(1)

This bit is reserved for future use.

3.3.10 dropframe b(1)

If this bit is a 1, the timestamp uses SMPTE drop-frames. If it is a 0, SMPTE drop-frames are not used.

3.3.11 reserved3 v(16)

This reserved field is always 0x8000, or ‘1000 0000 0000 0000’.

4 SMPTE timestamp location and frequency

In order to create a bitstream with timecode information, Dolby Digital, Dolby Digital Plus and Dolby TrueHD bitstreams must have timestamps added to them. Due to differences in bitstream organization, timestamps are placed differently in Dolby Digital, Dolby Digital Plus and Dolby TrueHD bitstreams. This section describes how timestamps are added to each different bitstream type.

- Timestamps in Dolby Digital bitstreams
- Timestamps in Dolby Digital Plus bitstreams
- Timestamps in Dolby TrueHD bitstreams
- SMPTE video frame and audio frame association

4.1 Timestamps in Dolby Digital bitstreams

In a Dolby Digital bitstream, the timestamp is placed immediately before each syncframe in the bitstream. It is beneficial to place timestamps before every syncframe in a bitstream to make the timecode more robust and resistant to splicing and errors.

Timestamp	Dolby Digital syncframe	Timestamp	Dolby Digital syncframe	Timestamp	Dolby Digital syncframe

Figure 9: Timestamp location in a Dolby Digital Bitstream

4.1.1 Syntax of Dolby Digital bitstream with timestamps

The syntax for a Dolby Digital bitstream containing timestamps is as follows:

```
Syntax                                                Encoding     Section
bitstream() {
   while (Dolby Digital syncframes exist) {
      timestamp()                                                  3.2
      Dolby Digital syncframe()                                    2.1
   }
}
```

Related Information

Dolby Digital bitstreams on page 7

SMPTE timestamp specification on page 12

4.2 Timestamps in Dolby Digital Plus bitstreams

In a Dolby Digital Plus bitstream, the timestamp is transmitted immediately before the first syncframe of every time-slice in the bitstream. It is beneficial to place timestamps before every time-slice in a bitstream to make the timecode overall more robust and resistant to splicing and errors.

Timestamp	Dolby Digital Plus Time-slice	Timestamp	Dolby Digital Plus Time-slice	Timestamp	Dolby Digital Plus Time-slice

Figure 10: Timestamp location in a Dolby Digital Plus Bitstream

4.2.1 Syntax of Dolby Digital Plus bitstream with timestamps

The syntax for a Dolby Digital Plus bitstream containing timestamps is as follows:

Syntax	Encoding	Section
`bitstream() {`		
`  while (Dolby Digital Plus time-slices exist) {`		
`    timestamp()`		`3.2`
`    Dolby Digital Plus time-slice ()`		`2.2`
`  }`		
`}`		

Related Information

Dolby Digital Plus bitstreams on page 7

SMPTE timestamp specification on page 12

4.3 Timestamps in Dolby TrueHD bitstreams

In a Dolby TrueHD bitstream, the timestamp is transmitted immediately before the first access unit in the bitstream. Unlike Dolby Digital Plus streams where timestamps precede each time-slice, the variable data rate of the Dolby TrueHD bitstream, combined with high resolution timing information embedded in the bitstream, makes repeated timestamps unnecessary and only a single timestamp to provide an initial reference time is included.

Timestamp	Dolby TrueHD Bitstream

Figure 11: Timestamp location in a Dolby TrueHD Bitstream

4.3.1 Timestamp Syntax in Dolby TrueHD bitstreams

The syntax for a Dolby TrueHD bitstream containing a timestamp is as follows:

Syntax	Encoding	Section
`bitstream() {`		
`    timestamp()`		`3.2`
`    Dolby TrueHD bitstream ()`		`2.3`
`  }`		
`}`		

Related Information

Dolby TrueHD bitstreams on page 10

SMPTE timestamp specification on page 12

4.4 SMPTE video frame and audio frame association

Each timestamp specifies the start time of the first sample in the encoded audio frame (that is a syncframe for a Dolby Digital bitstream, a time-slice for a Dolby Digital Plus bitstream, and an access unit for a Dolby TrueHD bitstream) in relation to the SMPTE video frame indicated by the timestamp. Because encoded audio frame sizes and SMPTE video frame rates generally do not align with an integer relationship, a sample offset is required to specify a sample-accurate start time. The samples timestamp field specifies this sample offset.

4.4.1 Dolby Digital and Dolby Digital Plus

For Dolby Digital and Dolby Digital Plus bitstreams, the sample offset indicated by the value of the samples field in the timestamp specifies the number of audio samples, at the sampling rate of the Dolby Digital/Dolby Digital Plus bitstream, between the first audio sample of the syncframe (or time-slice in the case of a multiple substream Dolby Digital Plus bitstream), and the audio sample that corresponds to the start of the video frame indicated by the SMPTE timestamp. This relationship is illustrated in Figure 12.

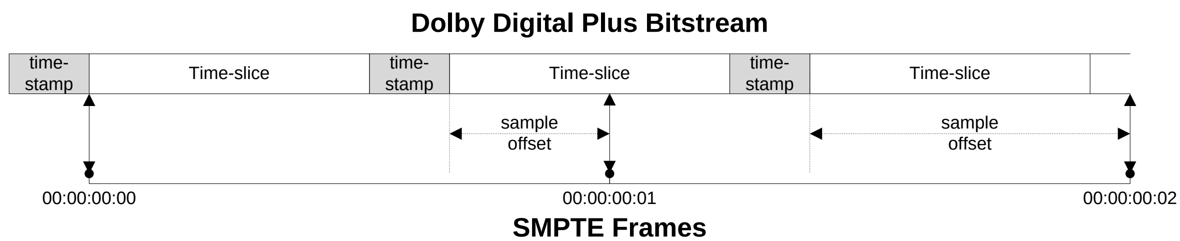

Figure 12: Timestamps and sample offset relationship in a Dolby Digital Plus bitstream

The maximum possible sample offset is dependent on the number of audio blocks in the syncframe/time-slice. Table 4 shows the maximum possible values of the sample offset indicated by the samples field per the number of audio blocks per syncframe/time-slice.

Table 4: Maximum sample offset value for different audio block numbers

Number of audio blocks per syncframe / time-slice	Maximum sample offset
One	255
Two	511
Three	767
Six	1535

Related Information

Dolby Digital Plus bitstreams on page 7

4.4.2 Dolby TrueHD

For Dolby TrueHD bitstreams, the sample offset indicated by the value of the samples field in the timestamp specifies the number of samples, at the sampling rate of the Dolby TrueHD bitstream, between the first audio sample of the Dolby TrueHD bitstream, and the audio sample that corresponds to the start of the video frame indicated by the SMPTE timestamp. This relationship is illustrated in Figure 13.

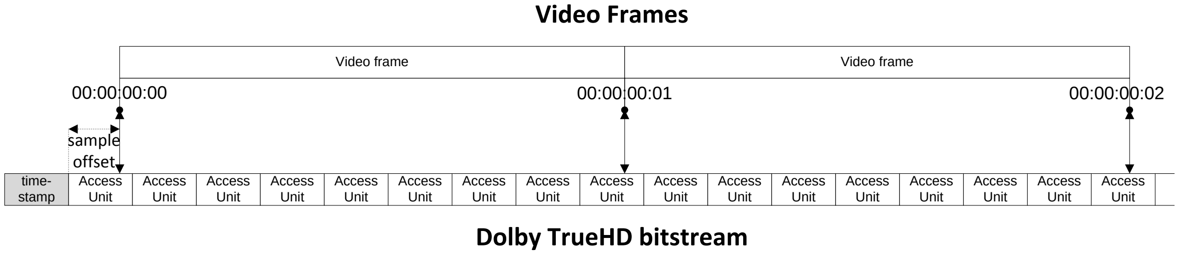

Figure 13: Timestamp and sample offset relationship in a Dolby TrueHD bitstream

The maximum possible sample offset is dependent on the sampling frequency of the Dolby TrueHD. Table 5 shows the maximum possible values of the sample offset indicated by the samples field per the sampling frequency of the Dolby TrueHD bitstream.

Table 5: Maximum sample offset value for different sampling rates

Sampling frequency of Dolby TrueHD bitstream	Maximum sample offset value
44.1 / 48 kHz	39
88.2 / 96 kHz	79
176.4 / 192 kHz	159

Related Information

Dolby TrueHD bitstreams on page 10

5 Using SMPTE timestamps from Dolby audio bitstreams

This section describes ways in which the timestamps within Dolby audio bitstreams can be used in the content production and delivery process.

- Use of timestamps when multiplexing/packaging Dolby audio bitstreams
- Transmitting Dolby audio bitstreams with SMPTE timestamps via AES3/SMPTE 337 interfaces

For all uses described in this section, the values of the timestamp(s) present within the Dolby audio bitstream are used by the application that is multiplexing or packaging the Dolby audio bitstream. However, the timestamps themselves must also be removed from the Dolby audio bitstreams during the multiplexing or packaging process, or before AES3/SMPTE 337 transmission, so that the Dolby audio bitstream within the final content delivery format no longer includes the timestamp(s).

5.1 Use of timestamps when multiplexing/packaging Dolby audio bitstreams

Timestamps can be used to ensure highly accurate alignment and synchronization between Dolby audio bitstreams and corresponding video bitstreams when multiplexing/packaging these bitstreams into formats such MPEG-2 Transport Streams or ISO Base Media (MP4) Files.

The values of the hours, minutes, seconds and frames fields in the timestamp can be used to ensure that audio is aligned with the correct video frames.

In addition, the samples field in the timestamp can be used by the multiplexer/packager to ensure that the audio sample referenced by the samples field value is perfectly aligned with the start of the corresponding video frame.

5.1.1 Usage in MPEG-2 Transport Stream multiplexing

When multiplexing Dolby audio with accompanying video into an MPEG-2 transport stream, for example at the final stages of Blu-ray Disc authoring the value of the samples field in the timestamp is used to specify offset between the start of the audio bitstream, and the start of the corresponding video bitstream. This ensures sample-accurate alignment between the audio and video.

To achieve this, the MPEG-2 TS multiplexer converts the samples field value into the units of the 90 kHz Presentation Time Stamp (PTS) clock, using the following equation:

$$\mathrm{PTS(samples)} = \mathrm{ceil}\left(\frac{\mathrm{samples} \times 90000}{\mathrm{sampling\_rate}}\right)$$

In this equation, `ceil(x)` represents the smallest integer not less than x.

For example, a samples value of 32 at a sampling rate of 48 kHz equals a `PTS(samples)` value of 60 PTS ticks. This means that the PTS value of the first audio frame of the bitstream should be 60 PTS ticks earlier than the PTS value of the first video frame of the bitstream. This is shown in Figure 14.

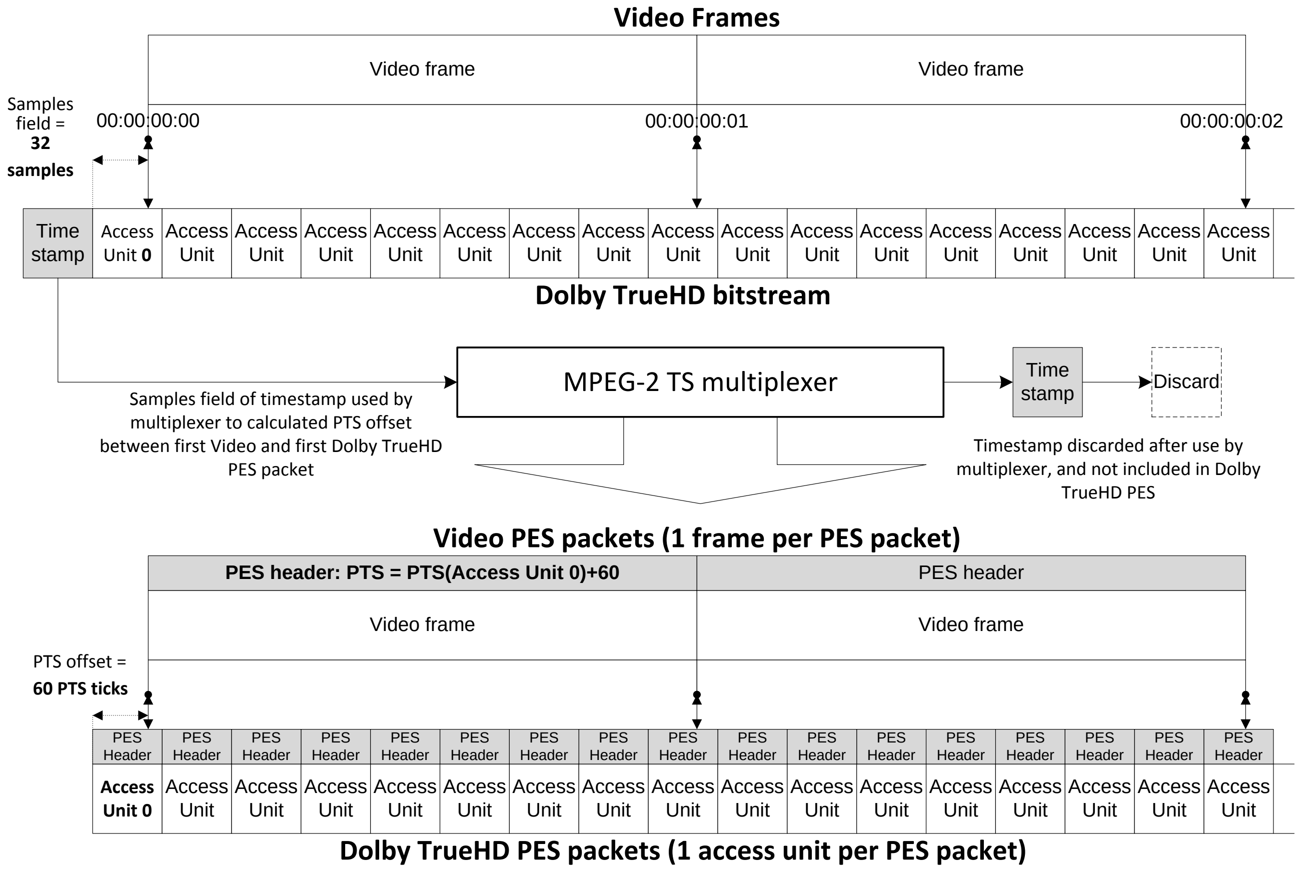

Figure 14: Use of SMPTE timestamp to specify initial A/V offset in MPEG-2 TS

In Blu-ray Disc authoring applications, Dolby Digital and Dolby TrueHD bitstreams containing the same audio program are multiplexed together with accompanying video. It is possible for the value of the samples field in the timestamp of the Dolby TrueHD bitstream to be different from the value of the samples field in the timestamp of the corresponding Dolby Digital frame. It is strongly recommended that the multiplexer uses each samples value to apply an individual PTS offset against the corresponding video frame for each audio format. This is shown in Figure 15.

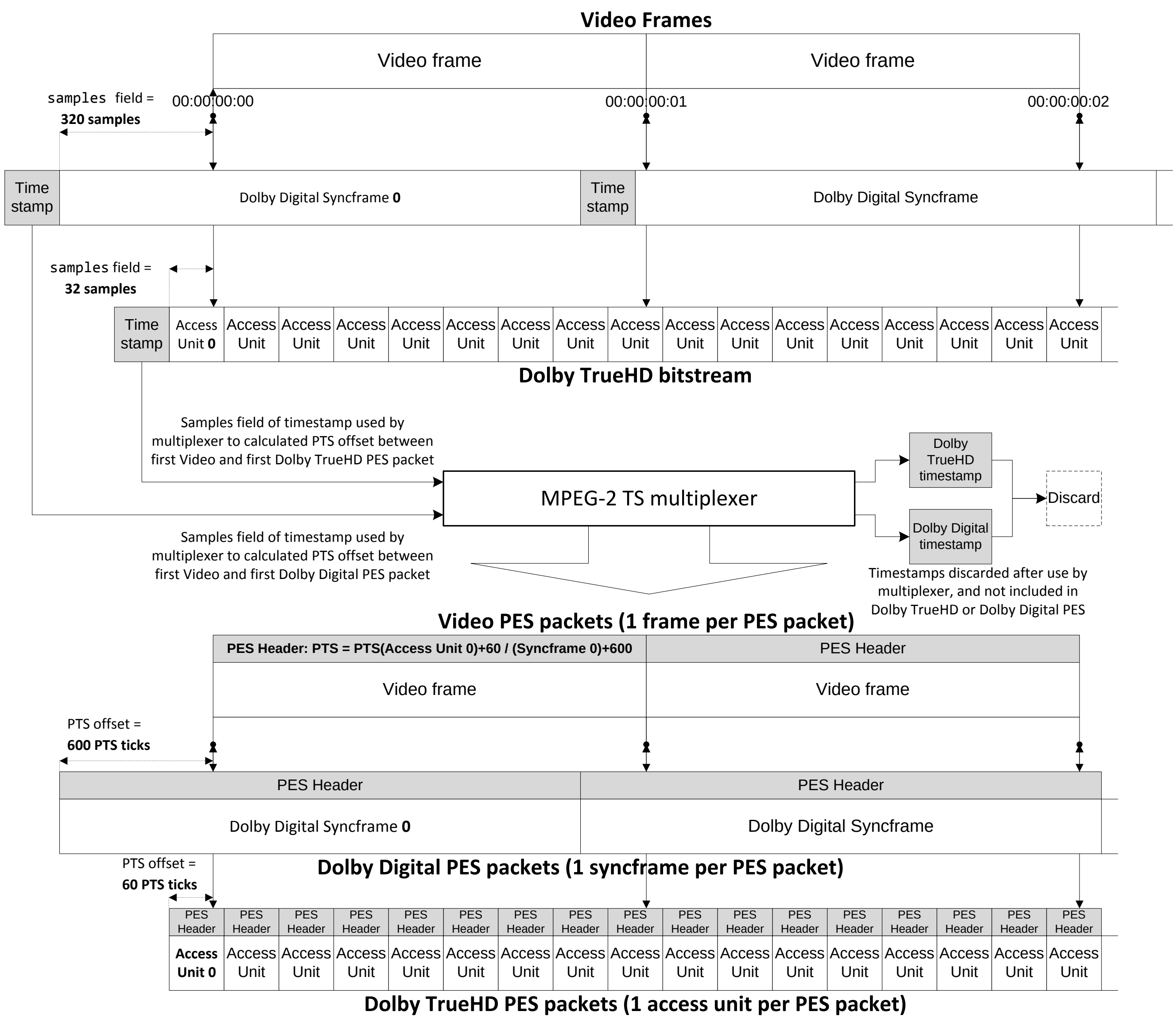

Figure 15: Use of SMPTE timestamps to specify initial A/V offset for BDMV MPEG-2 TS

If this is not possible due to the limitation of the multiplexer, the samples value from the Dolby TrueHD bitstream should be used.

Related Information

SMPTE video frame and audio frame association on page 16

5.1.2 Usage in ISO Base Media File Format packaging

To achieve maximum accuracy of A/V synchronization when packaging a Dolby audio bitstream with an accompanying video bitstream into an ISO base media file, the value of the samples field is used to specify the offset between the start of the audio track, and the start of the corresponding video track. To achieve this, samples value is used directly by the ISO base media file packager to specify a Media Time offset for the audio track by including an Edit Box containing an Edit List Box in the Track Box of the Dolby audio track.

For example, if the first video frame of a presentation is aligned with the 288th sample of the first audio frame of the Dolby audio bitstream, the Track Box of the Dolby audio track will contain an Edit Box that itself contains a single Edit List Box with the following values:

Table 6: Examples of Edit List Box parameters to enable correct alignment between audio and video tracks

Parameter name	Value
Segment-duration	0
Media-Time	288
Media-Rate	1

As the timebase of a Dolby audio track is equal to the sampling rate of the audio bitstream, the units of media time are in samples.

Similarly to MPEG-2 TS multiplexing, the timestamp used by the ISO base media file packager to calculate audio/video alignment, and any additional timestamps present in the Dolby audio bitstream, are discarded after being read, and are not included in the Dolby audio track within the ISO base media file.

Related Information

SMPTE video frame and audio frame association on page 16

5.2 Transmitting Dolby audio bitstreams with SMPTE timestamps via AES3/SMPTE 337 interfaces

The SMPTE 337 standard specifies a method for delivering coded audio bitstreams such as Dolby Digital and Dolby Digital Plus over an AES3 interface, using the data space normally occupied by PCM samples to carry the coded audio data. SMPTE 340 specifies the formatting of Dolby Digital and Dolby Digital Plus bitstreams for delivery according to SMPTE 337.

When transmitted according to SMPTE 337, each Dolby Digital syncframe and each Dolby Digital Plus access unit is carried within a data burst consisting of preamble data to enable synchronization, the coded audio data, and zero padding data.

Note: Transmission of Dolby TrueHD bitstreams is not supported by the SMPTE 337 standard.

5.2.1 Transmitting Dolby Digital bitstreams with SMPTE timestamps via AES3/SMPTE 337 interfaces

If a SMPTE 337 formatter processes a Dolby Digital bitstream that includes a timestamp before each syncframe, to conform to the SMPTE 340 standard, the following requirements must be met:

- The timestamp must be removed from the Dolby Digital syncframe and must not be included in the Dolby Digital data burst.
- The timestamp must be transmitted in its own data burst immediately before the Dolby Digital data burst.
- The data burst containing the timestamp must conform to SMPTE 339.

The formatting of a Dolby Digital bitstream containing timestamps for SMPTE 337 transmission is shown in Figure 16.

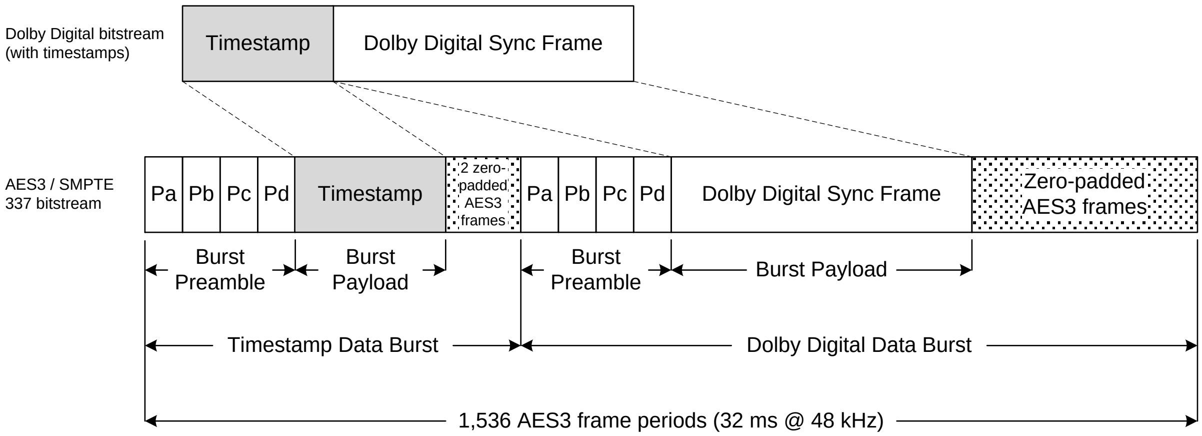

Figure 16: Formatting Dolby Digital bitstream with SMPTE timestamps for SMPTE 337 transmission

Related Information

Dolby Digital bitstreams on page 7

5.2.2 Transmitting Dolby Digital Plus bitstreams with SMPTE timestamps via AES3/SMPTE 337 interfaces

If a SMPTE 337 formatter processes a Dolby Digital Plus bitstream that includes a timestamp before each timeslice, to conform to the SMPTE 340 standard, the following requirements must be met:

- The timestamp must be removed from the Dolby Digital Plus timeslice and must not be included in the Dolby Digital Plus data burst.
- The timestamp must be transmitted in its own data burst immediately before the Dolby Digital Plus data burst.
- If the Dolby Digital Plus bitstream is encoded using 1, 2 or 3 blocks per frame, the time-slices are grouped into a Dolby Digital Plus Access Unit for SMPTE 337 transmission. In this case, only the timestamp that is associated with the first time-slice of the access unit is transmitted, and all timestamps associated with the additional time-slices in the access unit are discarded.
- The data burst containing the timestamp must conform to SMPTE 339.
- As both the timestamp and data-burst must be delivered within a time period of 1,536 AES3 frames, the maximum length of the Dolby Digital Plus burst-payload is limited to 1,528 AES frames.

The formatting of a Dolby Digital Plus bitstream containing timestamps for SMPTE 337 transmission is shown in Figure 17. In this example, the bitstream consists of sync frames containing three audio blocks per frame, requiring two time slices to be included in a single burst payload. The timestamp associated with the first timeslice is transmitted as a SMPTE 339 data burst that immediately precedes the Dolby Digital Plus data burst. The timestamp associated with the second timeslice is discarded and is not transmitted.

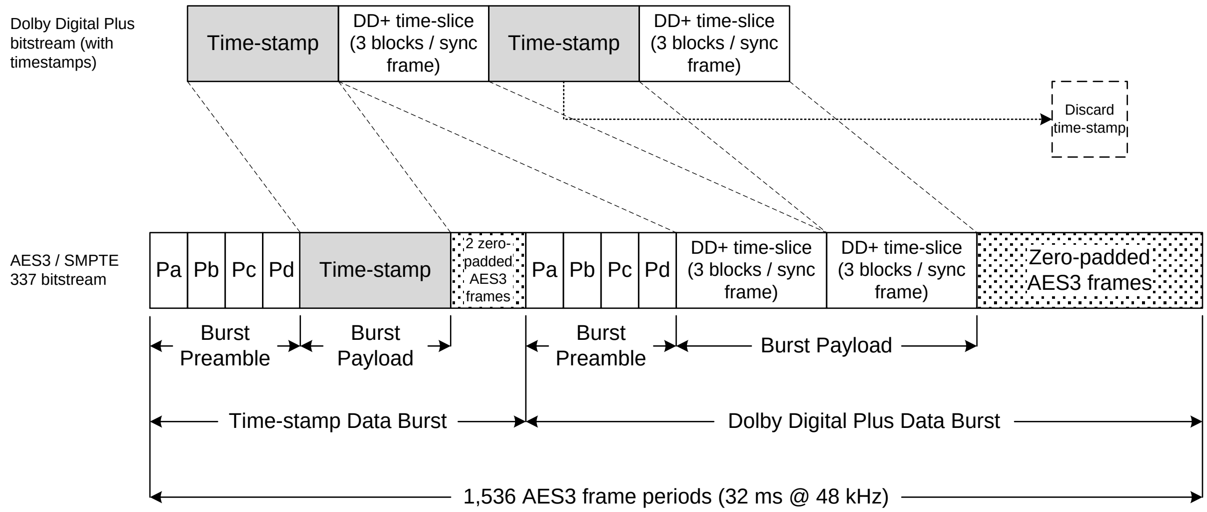

Figure 17: Formatting Dolby Digital Plus bitstream (3 blocks / syncframe) with SMPTE timestamps for SMPTE 337 transmission

Related Information

Dolby Digital Plus bitstreams on page 7

6 Glossary

BDMV

MPEG-2 Transport Stream format used for Blu-ray Disc that supports Dolby audio formats

MLP

Meridian Lossless Packing

PTS

Presentation Time Stamp

Timecode

Unique bitstream frame identifier, coded as HH:MM:SS:FF