



IMPLEMENTATION OF DTS AUDIO IN DYNAMIC ADAPTIVE STREAMING OVER HTTP (DASH)

Document No.: 9302K62400

Revision: A

Version: 1.0

Effective Date: July 2012

DTS, Inc.
5220 Las Virgenes Road
Calabasas, CA 91302
USA

www.dts.com



PATENTS

THE TECHNOLOGY AND/OR WORKS ASSOCIATED WITH THIS DOCUMENT ARE PROTECTED BY (A) APPLICABLE COPYRIGHT LAW AND/OR (B) EUROPEAN PATENT NUMBERS: _0864146, 182240, 183092, 183498, 1743326, U.S. PATENT NUMBERS: _5,956,674, 5,974,380; 5,978,762, 6,487,535; 6,226,616; 7,212,872, 7,003,467, 7,272,567, 7,668,723, 7,392,195, 7,930,184, 7,333,929, 7,548,853 AND OTHER U.S. AND INTERNATIONAL PATENTS BOTH PENDING AND ISSUED.

COPYRIGHT

Do Not Duplicate. Copyright © 2012 DTS, Inc. All Rights Reserved. Unauthorized duplication is a violation of State, Federal, and International laws.

This publication is copyrighted and all rights are reserved by DTS, Inc. Without the express prior written permission of DTS, no part of this publication may be reproduced, photocopied, stored on a retrieval system, translated, or transmitted in any form or by any means, electronic or otherwise.

TRADEMARKS

DTS, the Symbol, and DTS and the Symbol together are registered trademarks of DTS, Inc.

All other trademarks are the property of their respective owners



Contents

1	Introduction	5
1.1	Conformance.....	5
1.2	Conventions	5
2	Resources	5
3	Terms, Definitions and Abbreviations.....	6
4	MPEG DASH.....	6
4.1	Overview	6
4.2	Media Presentation Description	7
4.2.1	Representation Base Type.....	7
4.2.2	Audio Channel Configuration Descriptor	8
4.2.3	Representation.....	8
Annex A	Media Presentation Description Examples (Informative).....	10
A.1	Example MPD for ISO Base media file format On Demand profile.....	10
A.2	Example MPD for ISO Base media file format Live profile.....	12
A.3	Example MPD for MPEG-2 TS main profile	14



Tables

Table 1 - Common attributes	7
Table 2 - AudioChannelConfiguration attributes	8

DTS Audio and MPEG DASH

1 Introduction

This document describes the Media Presentation Description (MPD) requirements for delivering DTS-HD™ audio streams using MPEG DASH (Dynamic Adaptive Streaming over HTTP). In particular, this document addresses the use of DTS-HD in audio adaptation sets, providing examples to support for the following DASH profiles:

- ISO Base media file format On Demand profile
- ISO Base media file format live profile
- Mpeg-2 TS main profile

1.1 Conformance

The following words have reserved meanings that imply a level of requirement.

may - A key word that indicates flexibility of choice with no implied preference.

shall - A key word indicating a mandatory requirement. Designers are required to implement all such mandatory requirements.

1.2 Conventions

Within this document are references to a number of specifications. References to these specifications, or variable names used in these specifications, are followed by the resource tag. For example, **variableName** [TAG] is an indication that the details on **variableName** can be found in [TAG] which will be listed in Chapter 2.

2 Resources

The documents listed here are necessary in implementing this specification.

- [DTSISO] “Implementations of DTS Audio in Media Files Based on ISO/IEC 14496”, DTS Inc., Document #9302J81100, www.dts.com.
- [DTSMPG] “Implementations of DTS Audio in MPEG-2 Structures as defined in ISO/IEC 13818-1”, DTS Inc., Document #9302J85300, www.dts.com.
- [ISODSH] ISO/IEC 23009-1, First Edition (2012), “Information technology – Dynamic adaptive streaming over HTTP (DASH) - Part 1: Media presentation description and segment formats”, www.iso.org.

These additional documents may also be helpful in the usage of this specification.

- [DTSHD] ETSI TS 102 114 (2011-08), “DTS Coherent Acoustics Core and Extensions, with Additional Profiles”, www.etsi.org.



[ISOFF] ISO/IEC 14496-12, Third Edition (2008) and including corrigendum and amendments, “Information technology – Coding of Audio-Visual Object, part 12: ISO Based Media File Format”, www.iso.org.

[MPEG2] ISO/IEC 13818-1, Third Edition (2007), Information Technology – Generic Coding of moving picture and associated audio information: Systems”, www.iso.org.

3 Terms, Definitions and Abbreviations

The following terms, definitions and abbreviations are used throughout this document.

audio stream – A sequence of synchronized audio frames

audio frame – A component of an audio stream that corresponds to a certain number of PCM audio samples. Usually also an AU.

core or core substream: A DTS audio stream, or a component of a DTS audio stream, that conforms to [DTSHD] and always begins with the 32-bit Sync word of 0x7FFE8001.

duration – The time represented by one decoded audio frame; may be represented in audio samples per channel at a specific audio sampling frequency or in seconds.

extension – For DTS bitstreams, a component of an audio frame; could exist in sequence with other extension components or a core component.

extension substream – A DTS audio stream, or a component of a DTS audio stream, that conforms to [DTSHD] and always begins with the 32-bit sync word of 0x64582025.

LFE – Low Frequency Effects or subwoofer channel

4 MPEG DASH

4.1 Overview

Dynamic Adaptive Streaming over HTTP (DASH) provides a standard-based adaptive media streaming model where chunks of media streams and file segments are requested with HTTP and spliced together by a client that controls the media delivery. DASH reuses widely deployed HTTP servers and caches for efficient delivery over existing content distribution infrastructure components such as Content Distribution Networks (CDNs), NATs and firewalls. It provides a rich set of features to support on-demand, live streaming and time-shift applications and services to network-connected devices.

DASH is based on a hierarchical data model described by Media Presentation Description (MPD), which defines formats to announce resource identifiers for a collection of encoded and deliverable versions of media content. Media content is composed of single or multiple contiguous segments. The MPD provides sufficient information for a DASH client to provide a streaming service to the user by requesting segments from an HTTP web server and de-multiplexing, decoding and rendering the included media streams.

The segment formats specify the formats of the entity body of the HTTP response to an HTTP GET request or a partial HTTP GET with the indicated byte range using HTTP/1.1 to a resource identified in the MPD. DASH reuses the segment formats defined in ISO/IEC 14496-12 [ISOFF] and ISO/IEC 13818-1 [MPEG2]. Where values in this specification differ depending on the segment format used, the values to use are listed under the segment format name to which they correspond.

4.2 Media Presentation Description

4.2.1 Representation Base Type

The following common attributes apply to Adaptation Set, Representation and Sub Representation elements.

Table 1 - Common attributes

Attribute	Description
@codecs	This attribute specifies the codecs used to encode all representations within the adaptation set and the value shall be one of "dtsc", "dtsh", "dtsl" or "dtse" corresponding to the composition of the elementary stream. <u>ISO/IEC 14496 [ISOFF]</u> This value may be derived from the coding name used in DTSSampleEntry box [DTSISO]. <u>ISO/IEC 13818 [MPEG2]</u> For elementary streams signaled by the DTS audio descriptor, this value will always be 'dtsc'. For elementary streams signaled by the DTS-HD audio descriptor, asset_construction [DTSMPG] directly corresponds to StreamConstruction [DTSISO] where the value can be looked up directly.
@mimeType	<u>ISO/IEC 14496 [ISOFF]</u> For adaptation sets that conform to ISO/IEC 14496, this value shall be set to: "audio/mp4" (ISO Base Media File that contains a DTS audio track but no accompanying video track) [ISODSH] "video/mp4" (ISO Base Media File that contains a DTS audio track and one or more accompanying video tracks) [ISODSH] <u>ISO/IEC 13818 [MPEG2]</u> For adaptation sets that conform to ISO/IEC 13818, this value shall be set to: "video/mp2t"

@audioSamplingRate	Sampling rate shall be equal to the maximum sampling frequency of the audio encoded in the DTS-HD bitstream. This value shall be a whole decimal number representing the sampling frequency in Hz. <u>ISO/IEC 14496 [ISOFF]</u> This value may be derived from DTSSamplingFrequency [DTSISO] in the DTSSpecificBox. <u>ISO/IEC 13818 [MPEG2]</u> For elementary streams signaled by the DTS audio descriptor, the value may be derived from sample_rate_code [DTSMPG] in the descriptor. For elementary streams signaled by the DTS-HD audio descriptor, if an extension substream is present, this value may be derived from substream_0.sampling_frequency [DTSMPG], otherwise the value may be derived from substream_core.sampling_frequency [DTSMPG].
--------------------	---

4.2.2 Audio Channel Configuration Descriptor

Audio channel configuration is used to identify the audio channel configuration scheme employed.

Multiple **AudioChannelConfiguration** elements may be present, indicating that the

Representation supports multiple audio channel configurations.

Table 2 - AudioChannelConfiguration attributes

Attribute	Description
@schemeIdUri	Scheme as described in "urn:dts:dash:audio_channel_configuration:2012".
@value	AudioChannelConfiguration shall be set to the total number of discrete output channels represented in the stream, including LFE channels. The value shall be a whole decimal number in the range of 1 to 32. <u>ISO/IEC 14496 [ISOFF]</u> The channelcount [DTSISO] parameter in DTSSampleEntry [DTSISO] may be used to set AudioChannelConfiguration. <u>ISO/IEC 13818 [MPEG2]</u> For elementary streams signaled by the DTS audio descriptor, this value may be derived from surround_mode [DTSMPG], lfe_flag [DTSMPG] and extended_surround_flag [DTSMPG], defined in the descriptor. For elementary streams signaled by the DTS-HD audio descriptor, if an extension substream is present, this value may be derived from substream_0.channel_count [DTSMPG], otherwise the value may be derived from substream_core.channel_count [DTSMPG].

4.2.3 Representation

A Representation describes a deliverable encoded version of one or more media content components. A DASH client may switch from Representation to Representation within an Adaptation Set to adapt to varying network bandwidth conditions. For DTS, bit rates may differ across the same DTS stream type (e.g. DTS Express™ 512 kbps to DTS Express 192 kbps) in one Adaptation Set.

4.2.3.1 Coding Constraints for Seamless Stream Switching

Seamless stream switching shall enable a DASH client to switch from one DTS stream to another without interruption or muting between samples of the same encoded DTS audio content. To allow seamless stream switching between multiple DTS streams encoded with the same content within one Adaptation Set, the following parameters shall be the same in all representations:

- DTS audio coding name (@codecs) with the exception that for this purpose “dtsc” and “dtsh” may be considered as the same
- Audio sampling frequency
- Duration of synchronized frame
- Audio channel arrangement

4.2.3.2 Coding Constraints for Smooth Stream Switching

Smooth stream switching enables a DASH client to switch from one DTS stream to another by briefly fading out and then fading in (“V-fade”) without a decoder reset. To allow smooth stream switching between multiple DTS streams encoded with the same content within an Adaptation Set, the following parameters shall be the same in all representations:

- DTS audio coding name (@codecs) with the exception that for this purpose “dtsc” and “dtsh” may be considered as the same
- Audio sampling frequency
- Duration of synchronized frame

4.2.3.3 Consideration for Switching of Audio Channel Arrangement (Informative)

Switching of audio channel arrangement can be achieved via smooth stream switching with V-fade. As frequent switching of audio channel arrangement may be disruptive, a DASH client application may employ intelligent algorithms to decide on switching of audio channel arrangement depending on user experience requirements and other factors. For example, based on hysteresis and improving bandwidth conditions, a DASH client application may return to a higher audio channel arrangement configuration after monitoring and adjusting media rates over a certain period of time.

Annex A Media Presentation Description Examples (Informative)

A.1 Example MPD for ISO Base media file format On Demand profile

The following is an example of a static presentation with self-initializing Media Segments and multiple base URLs. It describes the content available from two sources (cdn1 and cdn2) with four representations of the DTS Express 5.1 audio provided at bitrates between 192 kbps and 510 kbps. The media presentation complies with the ISO base media file format On Demand profile, as defined in ISO/IEC 23009-1 [ISODSH].

```
<MPD
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns="urn:mpeg:DASH:schema:MPD:2011"
  xsi:schemaLocation="urn:mpeg:DASH:schema:MPD:2011 DASH-MPD.xsd"
  type="static"
  mediaPresentationDuration="PT0H0M21.35S"
  minBufferTime="PT1.5S"
  profiles="urn:mpeg:dash:profile:isoff-on-demand:2011">
  <BaseURL>http://cdn1.example.com/</BaseURL>
  <BaseURL>http://cdn2.example.com/</BaseURL>
  <Period start="PT0S" duration="PT0H0M21.35S">
    <!-- DTS Express 5.1 channel English -->
    <AdaptationSet
      mimeType="audio/mp4"
      codecs="dtse"
      lang="en"
      audioSamplingRate="48000"
      startWithSAP="1"
      subsegmentStartsWithSAP="1" >
      <AudioChannelConfiguration
        schemeIdUri="urn:dts:dash:audio_channel_configuration:2012"
        value="6"/>
      <ContentComponent id="100" contentType="audio"/>
      <Representation id="a1" bandwidth="192000">
        <BaseURL>dtse_192_dash.mp4</BaseURL>
        <SegmentBase indexRangeExact="true" indexRange="852-920"/>
      </Representation>
      <Representation id="a2" bandwidth="256000">
        <BaseURL>dtse_256_dash.mp4</BaseURL>
        <SegmentBase indexRangeExact="true" indexRange="853-921"/>
      </Representation>
      <Representation id="a3" bandwidth="384000">
        <BaseURL>dtse_384_dash.mp4</BaseURL>
        <SegmentBase indexRangeExact="true" indexRange="852-920"/>
      </Representation>
      <Representation id="a4" bandwidth="510000">
        <BaseURL>dtse_510_dash.mp4</BaseURL>
        <SegmentBase indexRangeExact="true" indexRange="853-921"/>
      </Representation>
    </AdaptationSet>
    <!-- Video -->
    <AdaptationSet
      mimeType="video/mp4"
      codecs="avc1.640028"
      lang="en"
      startWithSAP="1"
      subsegmentStartsWithSAP="1" >
```



```
<ContentComponent id="101" contentType="video"/>
<Representation id="v1" width="640" height="360" bandwidth="1020000">
  <BaseURL>avc_360p_dash.mp4</BaseURL>
  <SegmentBase indexRangeExact="true" indexRange="1176-1244"/>
</Representation>
<Representation id="v2" width="1280" height="720" bandwidth="5250000">
  <BaseURL>avc_720p_dash.mp4</BaseURL>
  <SegmentBase indexRangeExact="true" indexRange="1177-1245"/>
</Representation>
<Representation id="v3" width="1920" height="1080" bandwidth="8450000">
  <BaseURL>avc_1080p_dash.mp4</BaseURL>
  <SegmentBase indexRangeExact="true" indexRange="1178-1246"/>
</Representation>
</AdaptationSet>
</Period>
</MPD>
```

A.2 Example MPD for ISO Base media file format Live profile

The following is an example of a dynamic presentation with multiple languages, multiple base URLs and multiple video and audio bitrates. It describes the content available from two sources (cdn1 and cdn2) with DTS Express stereo available in English or French at rates between 64 kbps and 128 kbps. All the audio segments are aligned so language switching can be done seamlessly. The media presentation complies with the ISO base media file format live profile, as defined in ISO/IEC 23009-1 [ISODSH].

```
<?xml version="1.0" encoding="UTF-8"?>
<MPD
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns="urn:mpeg:DASH:schema:MPD:2011"
  xsi:schemaLocation="urn:mpeg:DASH:schema:MPD:2011 DASH-MPD.xsd"
  type="dynamic"
  minimumUpdatePeriod="PT2S"
  timeShiftBufferDepth="PT30M"
  availabilityStartTime="2012-06-25T12:30:00"
  minBufferTime="PT4S"
  profiles="urn:mpeg:dash:profile:isoff-live:2011">
  <BaseURL>http://cdn1.example.com/</BaseURL>
  <BaseURL>http://cdn2.example.com/</BaseURL>
  <Period>
    <!-- DTS Express stereo: English Audio -->
    <AdaptationSet
      mimeType="audio/mp4"
      codecs="dtse"
      lang="en"
      audioSamplingRate="48000"
      segmentAlignment="true"
      startWithSAP="1" >
      <AudioChannelConfiguration
        schemeIdUri="urn:dts:dash:audio_channel_configuration:2012"
        value="2"/>
      <SegmentTemplate
        timescale="48000"
        initialization="audio/en/$Bandwidth$/dtse_init.mp4a"
        media="audio/en/$Bandwidth$/dtse_$Time$.mp4a">
        <SegmentTimeline>
          <S t="0" d="96000" r="10"/>
        </SegmentTimeline>
      </SegmentTemplate>
      <Representation id="a0" bandwidth="64000" />
      <Representation id="a1" bandwidth="96000" />
      <Representation id="a2" bandwidth="128000" />
    </AdaptationSet>
    <!-- DTS Express stereo: French Audio -->
    <AdaptationSet
      mimeType="audio/mp4"
      codecs="dtse"
      lang="fr"
      audioSamplingRate="48000"
      segmentAlignment="true"
      startWithSAP="1" >
      <AudioChannelConfiguration
        schemeIdUri="urn:dts:dash:audio_channel_configuration:2012"
        value="2"/>
      <SegmentTemplate
        timescale="48000"
        initialization="audio/fr/$Bandwidth$/dtse_init.mp4a"
```



```
media="audio/fr/$Bandwidth$/dtse_$Time$.mp4a">
  <SegmentTimeline>
    <S t="0" d="96000" r="10"/>
  </SegmentTimeline>
</SegmentTemplate>
<Representation id="a3" bandwidth="64000" />
<Representation id="a4" bandwidth="96000" />
<Representation id="a5" bandwidth="128000" />
</AdaptationSet>
<!-- Video -->
<AdaptationSet
  mimeType="video/mp4"
  codecs="avc1.4D401F"
  frameRate="30000/1001"
  segmentAlignment="true"
  startWithSAP="1">
  <BaseURL>video/</BaseURL>
  <SegmentTemplate
    timescale="90000"
    initialization="$Bandwidth$/avc_init.mp4v"
    media="$Bandwidth$/avc_$Time$.mp4v">
    <SegmentTimeline>
      <S t="0" d="180180" r="10"/>
    </SegmentTimeline>
  </SegmentTemplate>
  <Representation id="v0" width="320" height="240" bandwidth="250000"/>
  <Representation id="v1" width="640" height="480" bandwidth="500000"/>
  <Representation id="v2" width="960" height="720" bandwidth="1000000"/>
</AdaptationSet>
</Period>
</MPD>
```

A.3 Example MPD for MPEG-2 TS main profile

The following is an example of a static presentation with multiple base URLs, multiple video and audio bitrates (DTS-HD Master Audio 5.1 at bit rates between 768 kbps and 1536 kbps). The media presentation complies with the MPEG-2 TS main profile, as defined in ISO/IEC 13818-1 [MPEG2].

```
<MPD
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns="urn:mpeg:DASH:schema:MPD:2011"
  xsi:schemaLocation="urn:mpeg:DASH:schema:MPD:2011 DASH-MPD.xsd"
  type="static"
  mediaPresentationDuration="PT0H0M21.35S"
  minBufferTime="PT2.50S"
  profiles="urn:mpeg:dash:profile:mp2t-main:2011">
  <BaseURL>http://cdn1.example.com/</BaseURL>
  <BaseURL>http://cdn2.example.com/</BaseURL>
  <Period start="PT0S" duration="PT21.34S">
    <!-- DTS-HD Master Audio 5.1 channel -->
    <AdaptationSet
      mimeType="video/mp2t"
      codecs="dtsh"
      lang="en"
      audioSamplingRate="48000"
      segmentAlignment="true"
      bitstreamSwitching="true"
      subsegmentAlignment="true">
      <AudioChannelConfiguration
        schemeIdUri="urn:dts:dash:audio_channel_configuration:2012"
        value="6"/>
      <ContentComponent id="36" contentType="audio"/>
      <Representation id="a1" bandwidth="768000">
        <SegmentTemplate
          timescale="1000"
          duration="10000"
          startNumber="0"
          media="dtsh_768_segment$Number$.ts" >
          <RepresentationIndex sourceURL="dtsh_768_dash_index.didx"/>
        </SegmentTemplate>
      </Representation>
      <Representation id="a2" bandwidth="960000">
        <SegmentTemplate
          timescale="1000"
          duration="10000"
          startNumber="0"
          media="dtsh_960_segment$Number$.ts" >
          <RepresentationIndex sourceURL="dtsh_960_dash_index.didx"/>
        </SegmentTemplate>
      </Representation>
      <Representation id="a3" bandwidth="1344000">
        <SegmentTemplate
          timescale="1000"
          duration="10000"
          startNumber="0"
          media="dtsh_1344_segment$Number$.ts" >
          <RepresentationIndex sourceURL="dtsh_1344_dash_index.didx"/>
        </SegmentTemplate>
      </Representation>
      <Representation id="a4" bandwidth="1536000">
        <SegmentTemplate
          timescale="1000"
          duration="10000"
          startNumber="0"
```

```
        media="dtsh_1536_segment$Number$.ts" >
        <RepresentationIndex sourceURL="dtsh_1536_dash_index.didx"/>
    </SegmentTemplate>
</Representation>
</AdaptationSet>
<!-- video -->
<AdaptationSet
    mimeType="video/mp2t"
    codecs="avc1.640015"
    lang="en"
    segmentAlignment="true"
    bitstreamSwitching="true"
    subsegmentAlignment="true">
    <ContentComponent id="38" contentType="video"/>
    <Representation id="v1" width="640" height="360" bandwidth="3250000">
        <SegmentTemplate
            timescale="1000"
            duration="10000"
            startNumber="0"
            media="avc_360p_segment$Number$.ts" >
            <RepresentationIndex sourceURL="avc_360p_index.didx"/>
        </SegmentTemplate>
    </Representation>
    <Representation id="v2" width="1280" height="720" bandwidth="7920000">
        <SegmentTemplate
            timescale="1000"
            duration="10000"
            startNumber="0"
            media="avc_720p_segment$Number$.ts" >
            <RepresentationIndex sourceURL="avc_720p_index.didx"/>
        </SegmentTemplate>
    </Representation>
    <Representation id="v3" width="1920" height="1080" bandwidth="14900000">
        <SegmentTemplate
            timescale="1000"
            duration="10000"
            startNumber="0"
            media="avc_1080p_segment$Number$.ts" >
            <RepresentationIndex sourceURL="avc_1080p_index.didx"/>
        </SegmentTemplate>
    </Representation>
    </AdaptationSet>
</Period>
</MPD>
```