SCTE · ISBE

STANDARDS

Digital Video Subcommittee

AMERICAN NATIONAL STANDARD

ANSI/SCTE 242-4 2018

Next Generation Audio Coding Constraints for Cable Systems: Part 4 – DTS-UHD Audio Coding Constraints

NOTICE

The Society of Cable Telecommunications Engineers (SCTE) / International Society of Broadband Experts (ISBE) Standards and Operational Practices (hereafter called "documents") are intended to serve the public interest by providing specifications, test methods and procedures that promote uniformity of product, interchangeability, best practices and ultimately the long-term reliability of broadband communications facilities. These documents shall not in any way preclude any member or non-member of SCTE•ISBE from manufacturing or selling products not conforming to such documents, nor shall the existence of such standards preclude their voluntary use by those other than SCTE•ISBE members.

SCTE•ISBE assumes no obligations or liability whatsoever to any party who may adopt the documents. Such adopting party assumes all risks associated with adoption of these documents, and accepts full responsibility for any damage and/or claims arising from the adoption of such documents.

Attention is called to the possibility that implementation of this document may require the use of subject matter covered by patent rights. By publication of this document, no position is taken with respect to the existence or validity of any patent rights in connection therewith. SCTE•ISBE shall not be responsible for identifying patents for which a license may be required or for conducting inquiries into the legal validity or scope of those patents that are brought to its attention.

Patent holders who believe that they hold patents which are essential to the implementation of this document have been requested to provide information about those patents and any related licensing terms and conditions. Any such declarations made before or after publication of this document are available on the SCTE•ISBE web site at http://www.scte.org.

140 Philips Road
Exton, PA 19341

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. Scope

This document is part four of a multi-part standard that specifies the coding constraints of Next Generation Audio system for cable television. In conjunction with SCTE 242-1 [1], this document defines the coding constraints on DTS-UHD for cable television. The carriage of the streams described in this specification is defined in SCTE 243-4 [8] in conjunction with SCTE 243-1[7].

2. Normative References

The following documents contain provisions, which, through reference in this text, constitute provisions of this document. At the time of Subcommittee approval, the editions indicated were valid. All documents are subject to revision; and while parties to any agreement based on this document are encouraged to investigate the possibility of applying the most recent editions of the documents listed below, they are reminded that newer editions of those documents might not be compatible with the referenced version.

2.1. SCTE References

[1] ANSI/SCTE 242-1 2017, Next Generation Audio Coding Constraints for Cable Systems: Part 1 – Introduction and Common Constraints

2.2. Standards from Other Organizations

[2] ATSC Standard A/342-1:2017, A/342 Part 1, Audio Common Elements
[3] ETSI TS 103 491 V1.1.1 (2017-04), DTS-UHD Audio Format; Delivery of Channels, Objects and Ambisonic Sound Fields
[4] Recommendation ITU-R BS.1770-4 (2015-10), Algorithms to measure audio programme loudness and true-peak audio level
[5] DVB BlueBook A038 (2017-12), Digital Video Broadcasting (DVB); Specification for Service Information (SI) in DVB systems (Final draft of ETSI EN 300 468 v 1.16.1)
[6] ISO/IEC 639-2:1998, "Codes for the representation of names of languages - Part 2: Alpha-3 code"

2.3. Published Materials

- No normative references are applicable.

3. Informative References

The following documents might provide valuable information to the reader but are not required when complying with this document.

3.1. SCTE References

[7] ANSI/SCTE 243-1 2017, Next Generation Audio Carriage Constraints for Cable Systems: Part 1 – Common Transport Signaling
[8] SCTE 243-4 2018, Next Generation Audio Carriage Constraints for Cable Systems: Part 4 – DTS-UHD Audio Carriage Constraints

3.2. Standards from Other Organizations

[9] ETSI TS 103 584 V1.1.1 (2018-01), DTS-UHD Point Source Renderer
[10] ATSC Doc A/85:2013, Techniques for Establishing and Maintaining Audio Loudness for Digital Television

3.3. Published Materials

- No informative references are applicable.

4. Compliance Notation

shall	This word or the adjective "***required***" means that the item is an absolute requirement of this document.
shall not	This phrase means that the item is an absolute prohibition of this document.
forbidden	This word means the value specified shall never be used.
should	This word or the adjective "*recommended*" means that there may exist valid reasons in particular circumstances to ignore this item, but the full implications should be understood and the case carefully weighted before choosing a different course.
should not	This phrase means that there may exist valid reasons in particular circumstances when the listed behavior is acceptable or even useful, but the full implications should be understood and the case carefully weighed before implementing any behavior described with this label.
may	This word or the adjective "*optional*" means that this item is truly optional. One vendor may choose to include the item because a particular marketplace requires it or because it enhances the product, for example; another vendor may omit the same item.
deprecated	Use is permissible for legacy purposes only. Deprecated features may be removed from future versions of this document. Implementations should avoid use of deprecated features.

5. Abbreviations and Definitions

5.1. Abbreviations

LFE	Low Frequency Effects.
ISBE	International Society of Broadband Experts
SCTE	Society of Cable Telecommunications Engineers

5.2. Definitions

This specification uses the definitions defined in ANSI/SCTE 242-1[1] and, by incorporation, ATSC A/342-1[2]. The following terms have definitions specific to DTS-UHD and shall apply to all clauses in this document.

Audio Chunk	block of data within an audio frame containing compressed audio samples
Audio Frame	unit of coded audio that, when decoded, will generate defined number

	of uncompressed Linear PCM audio samples for each wave form
BroadcastChunk	block of data within an audio stream containing data that maps audio components to preselections
Frame Duration	time represented by one decoded Audio Frame
Metadata Chunk	block of data within an audio frame containing metadata describing an audio presentation
Object	Audio Element as defined in ATSC A/342-1[2] and referenced in ANSI/SCTE 242-1[1]
Object Group	selected collection of audio objects to be played together
Presentation	selected collection of Channels, or Objects and Object Groups used together to generate the rendered output

6. DTS-UHD System Description

The DTS-UHD coding system is the third generation of DTS audio delivery formats. It is designed to both improve efficiency and deliver a richer set of features than the second generation DTS system. The first two generations of DTS codecs were designed primarily for Channel Based Audio (CBA). DTS-UHD is primarily designed to support audio objects, where a given object can represent a channel based presentation, sound field channels, or audio objects used in Object Based Audio (OBA).

6.1. Terminology

Table 1 lists terms defined in ATSC A/342-1[2] and maps them to corresponding terms defined in ETSI TS 103 491[3]

Table 1 - Common Terms Cross Reference

Common Term	DTS-UHD (TS 103 491) terms
Audio Element	Object
Audio Element Metadata	Metadata Chunk
Audio Presentation	Presentation
Audio Program	Audio Program
Audio Program Component	Object, Object Group or Presentation*
Elementary Stream	Elementary Stream
* In a DTS-UHD stream, an Audio Presentation points to a list of Components and contains additional metadata for rendering. The Components may be Objects, Object Groups or other Presentations. When multiple elementary streams are used to create an Audio Program, each DTS-UHD Presentation in those streams is an Audio Program Component.	

6.2. Overview

The DTS-UHD bitstream supports 32 pre-defined channel locations and definition of up to 224 objects plus 32 object groups. A DTS-UHD object is a set of coded waveforms plus an associated metadata

	of uncompressed Linear PCM audio samples for each wave form
BroadcastChunk	block of data within an audio stream containing data that maps audio components to preselections
Frame Duration	time represented by one decoded Audio Frame
Metadata Chunk	block of data within an audio frame containing metadata describing an audio presentation
Object	Audio Element as defined in ATSC A/342-1[2] and referenced in ANSI/SCTE 242-1[1]
Object Group	selected collection of audio objects to be played together
Presentation	selected collection of Channels, or Objects and Object Groups used together to generate the rendered output

6. DTS-UHD System Description

The DTS-UHD coding system is the third generation of DTS audio delivery formats. It is designed to both improve efficiency and deliver a richer set of features than the second generation DTS system. The first two generations of DTS codecs were designed primarily for Channel Based Audio (CBA). DTS-UHD is primarily designed to support audio objects, where a given object can represent a channel based presentation, sound field channels, or audio objects used in Object Based Audio (OBA).

6.1. Terminology

Table 1 lists terms defined in ATSC A/342-1[2] and maps them to corresponding terms defined in ETSI TS 103 491[3]

Table 1 - Common Terms Cross Reference

Common Term	DTS-UHD (TS 103 491) terms
Audio Element	Object
Audio Element Metadata	Metadata Chunk
Audio Presentation	Presentation
Audio Program	Audio Program
Audio Program Component	Object, Object Group or Presentation*
Elementary Stream	Elementary Stream
* In a DTS-UHD stream, an Audio Presentation points to a list of Components and contains additional metadata for rendering. The Components may be Objects, Object Groups or other Presentations. When multiple elementary streams are used to create an Audio Program, each DTS-UHD Presentation in those streams is an Audio Program Component.	

6.2. Overview

The DTS-UHD bitstream supports 32 pre-defined channel locations and definition of up to 224 objects plus 32 object groups. A DTS-UHD object is a set of coded waveforms plus an associated metadata

	of uncompressed Linear PCM audio samples for each wave form
BroadcastChunk	block of data within an audio stream containing data that maps audio components to preselections
Frame Duration	time represented by one decoded Audio Frame
Metadata Chunk	block of data within an audio frame containing metadata describing an audio presentation
Object	Audio Element as defined in ATSC A/342-1[2] and referenced in ANSI/SCTE 242-1[1]
Object Group	selected collection of audio objects to be played together
Presentation	selected collection of Channels, or Objects and Object Groups used together to generate the rendered output

6. DTS-UHD System Description

The DTS-UHD coding system is the third generation of DTS audio delivery formats. It is designed to both improve efficiency and deliver a richer set of features than the second generation DTS system. The first two generations of DTS codecs were designed primarily for Channel Based Audio (CBA). DTS-UHD is primarily designed to support audio objects, where a given object can represent a channel based presentation, sound field channels, or audio objects used in Object Based Audio (OBA).

6.1. Terminology

Table 1 lists terms defined in ATSC A/342-1[2] and maps them to corresponding terms defined in ETSI TS 103 491[3]

Table 1 - Common Terms Cross Reference

Common Term	DTS-UHD (TS 103 491) terms
Audio Element	Object
Audio Element Metadata	Metadata Chunk
Audio Presentation	Presentation
Audio Program	Audio Program
Audio Program Component	Object, Object Group or Presentation*
Elementary Stream	Elementary Stream
* In a DTS-UHD stream, an Audio Presentation points to a list of Components and contains additional metadata for rendering. The Components may be Objects, Object Groups or other Presentations. When multiple elementary streams are used to create an Audio Program, each DTS-UHD Presentation in those streams is an Audio Program Component.	

6.2. Overview

The DTS-UHD bitstream supports 32 pre-defined channel locations and definition of up to 224 objects plus 32 object groups. A DTS-UHD object is a set of coded waveforms plus an associated metadata

	of uncompressed Linear PCM audio samples for each wave form
BroadcastChunk	block of data within an audio stream containing data that maps audio components to preselections
Frame Duration	time represented by one decoded Audio Frame
Metadata Chunk	block of data within an audio frame containing metadata describing an audio presentation
Object	Audio Element as defined in ATSC A/342-1[2] and referenced in ANSI/SCTE 242-1[1]
Object Group	selected collection of audio objects to be played together
Presentation	selected collection of Channels, or Objects and Object Groups used together to generate the rendered output

6. DTS-UHD System Description

The DTS-UHD coding system is the third generation of DTS audio delivery formats. It is designed to both improve efficiency and deliver a richer set of features than the second generation DTS system. The first two generations of DTS codecs were designed primarily for Channel Based Audio (CBA). DTS-UHD is primarily designed to support audio objects, where a given object can represent a channel based presentation, sound field channels, or audio objects used in Object Based Audio (OBA).

6.1. Terminology

Table 1 lists terms defined in ATSC A/342-1[2] and maps them to corresponding terms defined in ETSI TS 103 491[3]

Table 1 - Common Terms Cross Reference

Common Term	DTS-UHD (TS 103 491) terms
Audio Element	Object
Audio Element Metadata	Metadata Chunk
Audio Presentation	Presentation
Audio Program	Audio Program
Audio Program Component	Object, Object Group or Presentation*
Elementary Stream	Elementary Stream
* In a DTS-UHD stream, an Audio Presentation points to a list of Components and contains additional metadata for rendering. The Components may be Objects, Object Groups or other Presentations. When multiple elementary streams are used to create an Audio Program, each DTS-UHD Presentation in those streams is an Audio Program Component.	

6.2. Overview

The DTS-UHD bitstream supports 32 pre-defined channel locations and definition of up to 224 objects plus 32 object groups. A DTS-UHD object is a set of coded waveforms plus an associated metadata

	of uncompressed Linear PCM audio samples for each wave form
BroadcastChunk	block of data within an audio stream containing data that maps audio components to preselections
Frame Duration	time represented by one decoded Audio Frame
Metadata Chunk	block of data within an audio frame containing metadata describing an audio presentation
Object	Audio Element as defined in ATSC A/342-1[2] and referenced in ANSI/SCTE 242-1[1]
Object Group	selected collection of audio objects to be played together
Presentation	selected collection of Channels, or Objects and Object Groups used together to generate the rendered output

6. DTS-UHD System Description

The DTS-UHD coding system is the third generation of DTS audio delivery formats. It is designed to both improve efficiency and deliver a richer set of features than the second generation DTS system. The first two generations of DTS codecs were designed primarily for Channel Based Audio (CBA). DTS-UHD is primarily designed to support audio objects, where a given object can represent a channel based presentation, sound field channels, or audio objects used in Object Based Audio (OBA).

6.1. Terminology

Table 1 lists terms defined in ATSC A/342-1[2] and maps them to corresponding terms defined in ETSI TS 103 491[3]

Table 1 - Common Terms Cross Reference

Common Term	DTS-UHD (TS 103 491) terms
Audio Element	Object
Audio Element Metadata	Metadata Chunk
Audio Presentation	Presentation
Audio Program	Audio Program
Audio Program Component	Object, Object Group or Presentation*
Elementary Stream	Elementary Stream
* In a DTS-UHD stream, an Audio Presentation points to a list of Components and contains additional metadata for rendering. The Components may be Objects, Object Groups or other Presentations. When multiple elementary streams are used to create an Audio Program, each DTS-UHD Presentation in those streams is an Audio Program Component.	

6.2. Overview

The DTS-UHD bitstream supports 32 pre-defined channel locations and definition of up to 224 objects plus 32 object groups. A DTS-UHD object is a set of coded waveforms plus an associated metadata

	of uncompressed Linear PCM audio samples for each wave form
BroadcastChunk	block of data within an audio stream containing data that maps audio components to preselections
Frame Duration	time represented by one decoded Audio Frame
Metadata Chunk	block of data within an audio frame containing metadata describing an audio presentation
Object	Audio Element as defined in ATSC A/342-1[2] and referenced in ANSI/SCTE 242-1[1]
Object Group	selected collection of audio objects to be played together
Presentation	selected collection of Channels, or Objects and Object Groups used together to generate the rendered output

6. DTS-UHD System Description

The DTS-UHD coding system is the third generation of DTS audio delivery formats. It is designed to both improve efficiency and deliver a richer set of features than the second generation DTS system. The first two generations of DTS codecs were designed primarily for Channel Based Audio (CBA). DTS-UHD is primarily designed to support audio objects, where a given object can represent a channel based presentation, sound field channels, or audio objects used in Object Based Audio (OBA).

6.1. Terminology

Table 1 lists terms defined in ATSC A/342-1[2] and maps them to corresponding terms defined in ETSI TS 103 491[3]

Table 1 - Common Terms Cross Reference

Common Term	DTS-UHD (TS 103 491) terms
Audio Element	Object
Audio Element Metadata	Metadata Chunk
Audio Presentation	Presentation
Audio Program	Audio Program
Audio Program Component	Object, Object Group or Presentation*
Elementary Stream	Elementary Stream
* In a DTS-UHD stream, an Audio Presentation points to a list of Components and contains additional metadata for rendering. The Components may be Objects, Object Groups or other Presentations. When multiple elementary streams are used to create an Audio Program, each DTS-UHD Presentation in those streams is an Audio Program Component.	

6.2. Overview

The DTS-UHD bitstream supports 32 pre-defined channel locations and definition of up to 224 objects plus 32 object groups. A DTS-UHD object is a set of coded waveforms plus an associated metadata

	of uncompressed Linear PCM audio samples for each wave form
BroadcastChunk	block of data within an audio stream containing data that maps audio components to preselections
Frame Duration	time represented by one decoded Audio Frame
Metadata Chunk	block of data within an audio frame containing metadata describing an audio presentation
Object	Audio Element as defined in ATSC A/342-1[2] and referenced in ANSI/SCTE 242-1[1]
Object Group	selected collection of audio objects to be played together
Presentation	selected collection of Channels, or Objects and Object Groups used together to generate the rendered output

6. DTS-UHD System Description

The DTS-UHD coding system is the third generation of DTS audio delivery formats. It is designed to both improve efficiency and deliver a richer set of features than the second generation DTS system. The first two generations of DTS codecs were designed primarily for Channel Based Audio (CBA). DTS-UHD is primarily designed to support audio objects, where a given object can represent a channel based presentation, sound field channels, or audio objects used in Object Based Audio (OBA).

6.1. Terminology

Table 1 lists terms defined in ATSC A/342-1[2] and maps them to corresponding terms defined in ETSI TS 103 491[3]

Table 1 - Common Terms Cross Reference

Common Term	DTS-UHD (TS 103 491) terms
Audio Element	Object
Audio Element Metadata	Metadata Chunk
Audio Presentation	Presentation
Audio Program	Audio Program
Audio Program Component	Object, Object Group or Presentation*
Elementary Stream	Elementary Stream
* In a DTS-UHD stream, an Audio Presentation points to a list of Components and contains additional metadata for rendering. The Components may be Objects, Object Groups or other Presentations. When multiple elementary streams are used to create an Audio Program, each DTS-UHD Presentation in those streams is an Audio Program Component.	

6.2. Overview

The DTS-UHD bitstream supports 32 pre-defined channel locations and definition of up to 224 objects plus 32 object groups. A DTS-UHD object is a set of coded waveforms plus an associated metadata

	of uncompressed Linear PCM audio samples for each wave form
BroadcastChunk	block of data within an audio stream containing data that maps audio components to preselections
Frame Duration	time represented by one decoded Audio Frame
Metadata Chunk	block of data within an audio frame containing metadata describing an audio presentation
Object	Audio Element as defined in ATSC A/342-1[2] and referenced in ANSI/SCTE 242-1[1]
Object Group	selected collection of audio objects to be played together
Presentation	selected collection of Channels, or Objects and Object Groups used together to generate the rendered output

6. DTS-UHD System Description

The DTS-UHD coding system is the third generation of DTS audio delivery formats. It is designed to both improve efficiency and deliver a richer set of features than the second generation DTS system. The first two generations of DTS codecs were designed primarily for Channel Based Audio (CBA). DTS-UHD is primarily designed to support audio objects, where a given object can represent a channel based presentation, sound field channels, or audio objects used in Object Based Audio (OBA).

6.1. Terminology

Table 1 lists terms defined in ATSC A/342-1[2] and maps them to corresponding terms defined in ETSI TS 103 491[3]

Table 1 - Common Terms Cross Reference

Common Term	DTS-UHD (TS 103 491) terms
Audio Element	Object
Audio Element Metadata	Metadata Chunk
Audio Presentation	Presentation
Audio Program	Audio Program
Audio Program Component	Object, Object Group or Presentation*
Elementary Stream	Elementary Stream
* In a DTS-UHD stream, an Audio Presentation points to a list of Components and contains additional metadata for rendering. The Components may be Objects, Object Groups or other Presentations. When multiple elementary streams are used to create an Audio Program, each DTS-UHD Presentation in those streams is an Audio Program Component.	

6.2. Overview

The DTS-UHD bitstream supports 32 pre-defined channel locations and definition of up to 224 objects plus 32 object groups. A DTS-UHD object is a set of coded waveforms plus an associated metadata

	of uncompressed Linear PCM audio samples for each wave form
BroadcastChunk	block of data within an audio stream containing data that maps audio components to preselections
Frame Duration	time represented by one decoded Audio Frame
Metadata Chunk	block of data within an audio frame containing metadata describing an audio presentation
Object	Audio Element as defined in ATSC A/342-1[2] and referenced in ANSI/SCTE 242-1[1]
Object Group	selected collection of audio objects to be played together
Presentation	selected collection of Channels, or Objects and Object Groups used together to generate the rendered output

6. DTS-UHD System Description

The DTS-UHD coding system is the third generation of DTS audio delivery formats. It is designed to both improve efficiency and deliver a richer set of features than the second generation DTS system. The first two generations of DTS codecs were designed primarily for Channel Based Audio (CBA). DTS-UHD is primarily designed to support audio objects, where a given object can represent a channel based presentation, sound field channels, or audio objects used in Object Based Audio (OBA).

6.1. Terminology

Table 1 lists terms defined in ATSC A/342-1[2] and maps them to corresponding terms defined in ETSI TS 103 491[3]

Table 1 - Common Terms Cross Reference

Common Term	DTS-UHD (TS 103 491) terms
Audio Element	Object
Audio Element Metadata	Metadata Chunk
Audio Presentation	Presentation
Audio Program	Audio Program
Audio Program Component	Object, Object Group or Presentation*
Elementary Stream	Elementary Stream
* In a DTS-UHD stream, an Audio Presentation points to a list of Components and contains additional metadata for rendering. The Components may be Objects, Object Groups or other Presentations. When multiple elementary streams are used to create an Audio Program, each DTS-UHD Presentation in those streams is an Audio Program Component.	

6.2. Overview

The DTS-UHD bitstream supports 32 pre-defined channel locations and definition of up to 224 objects plus 32 object groups. A DTS-UHD object is a set of coded waveforms plus an associated metadata