

SCTE | **STANDARDS**

Digital Video Subcommittee

AMERICAN NATIONAL STANDARD

ANSI/SCTE 242-1 2022

**Next Generation Audio Coding Constraints for
Cable Systems: Part 1 – Introduction and
Common Constraints**

NOTICE

The Society of Cable Telecommunications Engineers (SCTE) 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, interoperability, interchangeability, best practices, and the long term reliability of broadband communications facilities. These documents shall not in any way preclude any member or non-member of SCTE 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 members.

SCTE 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.

NOTE: The user’s attention is called to the possibility that compliance with this document may require the use of an invention covered by patent rights. By publication of this document, no position is taken with respect to the validity of any such claim(s) or of any patent rights in connection therewith. If a patent holder has filed a statement of willingness to grant a license under these rights on reasonable and nondiscriminatory terms and conditions to applicants desiring to obtain such a license, then details may be obtained from the standards developer. SCTE 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 web site at <https://scte.org>.

All Rights Reserved
©2022 Society of Cable Telecommunications Engineers, Inc.
140 Philips Road
Exton, PA 19341

Document Types and Tags

Document Type: Specification

Document Tags:

- | | | |
|--|------------------------------------|--|
| <input type="checkbox"/> Test or Measurement | <input type="checkbox"/> Checklist | <input checked="" type="checkbox"/> Facility |
| <input type="checkbox"/> Architecture or Framework | <input type="checkbox"/> Metric | <input type="checkbox"/> Access Network |
| <input checked="" type="checkbox"/> Procedure, Process or Method | <input type="checkbox"/> Cloud | <input type="checkbox"/> Customer Premises |

Document Release History

Release	Date
SCTE 242-1 2017	10/10/2017
SCTE 242-1 2022	October 2022

Note: Standards that are released multiple times in the same year use: a, b, c, etc. to indicate normative balloted updates and/or r1, r2, r3, etc. to indicate editorial changes to a released document after the year.

Table of Contents

Title	Page Number
NOTICE.....	2
Document Types and Tags.....	3
Document Release History.....	3
Table of Contents.....	4
1. Introduction.....	5
1.1. Executive Summary.....	5
1.2. Scope.....	5
1.3. Benefits.....	5
2. Normative References.....	5
2.1. SCTE References.....	5
2.2. Standards from Other Organizations.....	5
2.3. Other Published Materials.....	5
3. Informative References.....	5
3.1. SCTE References.....	5
3.2. Standards from Other Organizations.....	6
3.3. Other Published Materials.....	6
4. Compliance Notation.....	6
5. Abbreviations and Definitions.....	6
5.1. Abbreviations.....	6
5.2. Definitions.....	7
6. System overview.....	7
6.1. Features of NGA Systems.....	7
6.2. NGA Concepts.....	7
6.2.1. Audio Program Components and Preselections.....	7
6.2.2. Audio Element Formats.....	8
6.2.3. Audio Rendering.....	8
7. Encoding Constraints.....	9
7.1. Sampling Rate.....	9
7.2. Audio Program Structure.....	9
7.3. General Elementary Stream Structure.....	10

List of Tables

Title	Page Number
Table 1 - Encoding of Broadcast Operating Profiles.....	10

1. Introduction

1.1. Executive Summary

This document is part of a suite documenting coding constraints of Next Generation Audio (NGA) systems for cable television. It is intended to be used in conjunction with the specific audio technologies described in subsequent Parts of this standard [SCTE 242-2], [SCTE 242-3], and SCTE 242-4].

1.2. Scope

This document specifies the common framework for Next Generation Audio (NGA) systems for cable television. It is intended to be used in conjunction with the specific audio technologies described in subsequent parts of this Standard [SCTE 242-2], [SCTE 242-3], and [SCTE 242-4].

1.3. Benefits

The Next Generation Audio (NGA) system audio system provides immersive and personalizable sound for television. It is not compatible with the audio system used in [SCTE 54]-era service.

2. Normative References

The following documents contain provisions which, through reference in this text, constitute provisions of this document. The editions indicated were valid at the time of subcommittee approval. 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

No normative references are applicable.

2.2. Standards from Other Organizations

[A 342-1] ATSC A/342:2021 Part 1: Audio Common Elements.

2.3. Other 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

[SCTE 242-2] SCTE 242-2 202x, Next Generation Audio Coding Constraints for Cable Systems: Part 2 – AC-4 Audio Coding Constraints

[SCTE 242-3] SCTE 242-3 202x, Next Generation Audio Coding Constraints for Cable Systems: Part 3 – MPEG-H Audio Coding Constraints

- [SCTE 242-4] SCTE 242-4 202x, Next Generation Audio Coding Constraints for Cable Systems: Part 4 – DTS-UHD Audio Coding Constraints
- [SCTE 243-1] SCTE 243-1 202x, Next Generation Audio Carriage Constraints for Cable Systems: Part 1 – Common Transport Signaling
- [SCTE 54] SCTE 54 2020, Digital Video Service Multiplex and Transport System Standard for Cable Television

3.2. Standards from Other Organizations

No informative references are applicable.

3.3. Other Published Materials

No informative references are applicable.

4. Compliance Notation

<i>shall</i>	This word or the adjective “ <i>required</i> ” means that the item is an absolute requirement of this document.
<i>shall not</i>	This phrase means that the item is an absolute prohibition of this document.
<i>forbidden</i>	This word means the value specified <i>shall</i> never be used.
<i>should</i>	This word or the adjective “ <i>recommended</i> ” means that there <i>may</i> exist valid reasons in particular circumstances to ignore this item, but the full implications <i>should</i> be understood and the case carefully weighed before choosing a different course.
<i>should not</i>	This phrase means that there <i>may</i> exist valid reasons in particular circumstances when the listed behavior is acceptable or even useful, but the full implications <i>should</i> be understood and the case carefully weighed before implementing any behavior described with this label.
<i>may</i>	This word or the adjective “ <i>optional</i> ” indicate a course of action permissible within the limits of the document.
deprecated	Use is permissible for legacy purposes only. Deprecated features <i>may</i> be removed from future versions of this document. Implementations <i>should</i> avoid use of deprecated features.

5. Abbreviations and Definitions

5.1. Abbreviations

AD	Audio Description ¹
CM	Complete Main (see ATSC [A 342-1])
Dx	Dialog element ‘x’ (mono) (see ATSC [A 342-1])
HOA	Higher Order Ambisonics (see ATSC [A 342-1])
M&E	Music and Effects (see ATSC [A 342-1])
NGA	Next Generation Audio
O	Other object (mono), i.e. PA feed (see ATSC [A 342-1])

VDS	Video Description Service ¹
-----	--

5.2. Definitions

Definitions of terms used in this document are provided in this section. Defined terms that have specific meanings are capitalized. When the capitalized term is used in this document, the term has the specific meaning as defined in this section.

This document uses the terminology specific to the ATSC 3.0 audio system as defined in ATSC [A 342-1] Clause 4. Additionally, the following definitions are used:

Audio Description	Defined in a regulation of the United States, 47 CFR §79.3(a)(3), as “The insertion of audio narrated descriptions of a television program’s key visual elements into natural pauses between the program’s dialogue.” It has sometimes been termed Video Description or VDS. Such usage is considered obsolete.
Audio Preselection	set of Audio Program Components representing a version of the Audio Program that may be selected by a user for simultaneous decoding. An Audio Preselection is a sub-selection from all available Audio Program Components of one Audio Program. An Audio Preselection can be considered the NGA equivalent of audio services in predecessor systems, whereby each audio service comprises a complete audio mix.

6. System overview

6.1. Features of NGA Systems

A description of the main features of Next Generation Audio systems is provided in ATSC [A 342-1] subclause 5.1.

6.2. NGA Concepts

Several concepts are common to all NGA systems (e.g., systems supported by ATSC 3.0). This section describes these common concepts.

6.2.1. Audio Program Components and Preselections

Audio Program Components are separate pieces of audio data that are combined to compose an Audio Preselection. A simple Audio Preselection may consist of a single Audio Program Component, such as a Complete Main Mix for a television program. Audio Preselections that are more complex may consist of several Audio Program Components, such as ambient music and effects, combined with dialog and audio description.

Audio Preselections are combinations of Audio Program Components representing versions of the audio program that may be selected by a user. For example, a complete audio with English dialog, a complete

¹ As Defined in a regulation of the United States, 47 CFR §79.3(a)(3), as “The insertion of audio narrated descriptions of a television program’s key visual elements into natural pauses between the program’s dialogue.” It has sometimes been termed Video Description or VDS. Such usage is considered obsolete.

audio with Spanish dialog, a complete audio (English or Spanish) with audio description, or a complete audio with alternate dialog may all be selectable Preselections for a Program.

The Components of a Preselection can be delivered in a single audio Elementary Stream or in multiple audio Elementary Streams.

6.2.2. Audio Element Formats

The NGA systems support three fundamental Audio Element Formats:

- 1) Channel Sets are sets of Audio Elements consisting of one or more Audio Signals presenting sound to speaker(s) located at canonical positions. These include configurations such as mono, stereo, or 5.1, and extend to include non-planar configurations, such as 7.1+4.
- 2) Audio Objects are Audio Elements consisting of audio information and associated metadata representing a sound's location in space (as described by the metadata). The metadata may be dynamic, representing the movement of the sound.
- 3) Scene-based audio (e.g., HOA) consists of one or more Audio Elements that make up a generalized representation of a sound field.

6.2.3. Audio Rendering

Audio Rendering is the process of composing an Audio Preselection and converting all the Audio Program Components to a data structure appropriate for the audio outputs of a specific receiver. Rendering may include conversion of a Channel Set to a different channel configuration, conversion of Audio Objects to Channel Sets, conversion of scene-based sets to Channel Sets, and/or applying specialized audio processing such as room correction or spatial virtualization.

6.2.3.1. Audio Description (AD)

Audio Description is an audio service carrying narration describing a television program's key visual elements. These descriptions are inserted into natural pauses in the program's dialog. Audio Description makes TV programming more accessible to individuals who are blind or visually impaired. The Audio Description may be provided by sending a collection of "Music and Effects" components, a Dialog component, and an appropriately labeled Audio Description component, which are mixed at the receiver. Alternatively, an Audio Description may be provided as a single component that is a Complete Mix, with the appropriate label identification.

6.2.3.2. Multi-Language

Traditionally, multi-language support is achieved by sending Complete Mixes with different dialog languages. For NGA systems, multi-language support can be achieved through a collection of "Music and Effects" streams combined with multiple dialog language streams that are mixed at the receiver.

6.2.3.3. Personalized Audio

Personalized audio consists of one or more Audio Elements with metadata, which describes how to decode, render, and output "full" Mixes. Each personalized Audio Preselection may consist of an ambience "bed", one or more dialog elements, and optionally one or more effects elements. Multiple Audio Preselections can be defined to support a number of options such as alternate language, dialog or ambience, enabling height elements, etc.

There are two main concepts of personalized audio:

- 1) Personalization selection – The bit stream may contain more than one Audio Preselection where each Audio Preselection contains pre-defined audio experiences (e.g., “home team” audio experience, multiple languages, etc.). A listener can choose the audio experience by selecting one of the Audio Preselections.
- 2) Personalization control – Listeners can modify properties of the complete audio experience or parts of it (e.g., increasing the volume level of an Audio Element, changing the position of an Audio Element, etc.).

7. Encoding Constraints

The following constraints are defined on all NGA systems for cable television.

7.1. Sampling Rate

The sampling frequency of Audio Signals shall be 48 kHz.

7.2. Audio Program Structure

An Audio Program shall consist of one or more Audio Preselections. One Audio Preselection shall be signaled as the default (main).

The default Audio Preselection shall have all of its Audio Program Components present in the broadcast stream.

Note: This constraint is intended to facilitate future applications in which additional Program Components are delivered by other means (e.g., hybrid use case in ATSC 3.0).

The main Audio Preselection is intended to be the default in cases where no other selection guidance (user-originated or otherwise) exists.

Audio Preselections shall consist of at least one Audio Program Component of any Audio Element Format.

Audio Program Components may be delivered in more than one Elementary Stream. Audio Preselections other than the default Preselection may include Audio Program Components from multiple Elementary Streams. Audio Preselections shall not utilize Audio Program Components from more than three Elementary Streams.

Audio program structures SHALL BE drawn from those listed in ATSC A/342-1 Table A.1.1, made normative here (and reproduced below for the benefit of the reader as Table 1):

Table 1 - Encoding of Broadcast Operating Profiles

channel_mode	Profile Type	Input Elements	Presentations/Presets	Elements Referenced by Presentation/Presets
1	Complete Main	2.0 CM	CM	CM
2		5.1 CM	CM	CM
3		HOA(6) CM	CM	CM
4		5.1.2 CM	CM	CM
5		7.1.4 CM	CM	CM
6		HOA(12) CM	CM	CM
7		O(15).1 CM	CM	CM
8	M&E + Objects	2.0 M&E + D	English	M&E + D
9			5.1 M&E + D1 (en) + D2 (es) + AD (en)	M&E Only
		English		M&E + D1
		English + AD		M&E + D1 + AD
		Spanish		M&E + D2
10		HOA(6) + D1 (en) + D2 (es) + AD (en)	M&E Only	M&E
			English	M&E + D1
			English + AD	M&E + D1 + AD
11		5.1.2 M&E + D1 (en) + D2 (es) + AD (en)	Spanish	M&E + D2
			M&E Only	M&E
			English	M&E + O + D1
			English + AD	M&E + D1 + AD
12		7.1.4 M&E + D1 (en) + D2 (es) + AD (en) + O	Spanish	M&E + O + D2
			M&E	M&E + O
	English		M&E + D1	
13	O(15).1 M&E + D1 (en) + D2 (es) + AD (en)	English + AD	M&E + D1 + AD	
		Spanish	M&E + D2	
		M&E Only	M&E	
14	HOA(12) M&E + D1 (en) + D2 (es) + AD (en) + O	English	M&E + O + D1	

Languages are NOT constrained to English and Spanish only, those are shown above as examples.

Further constraints specific to individual codecs are defined in subsequent Parts of this standard.

7.3. General Elementary Stream Structure

The carriage of the streams described in this specification is defined in [SCTE 243-1], with further specifications defined in subsequent Parts of this standard.