

Metadata 2.0 Specifications

Content Encoding Profiles 2.0 Specification

MD-SP-VOD-CEP2.0-I02-070105

ISSUED

Notice

This Metadata specification is the result of a cooperative effort undertaken at the direction of Cable Television Laboratories, Inc. for the benefit of the cable industry and its customers. This document may contain references to other documents not owned or controlled by CableLabs. Use and understanding of this document may require access to such other documents. Designing, manufacturing, distributing, using, selling, or servicing products, or providing services, based on this document may require intellectual property licenses from third parties for technology referenced in this document.

Neither CableLabs nor any member company is responsible to any party for any liability of any nature whatsoever resulting from or arising out of use or reliance upon this document, or any document referenced herein. This document is furnished on an "AS IS" basis and neither CableLabs nor its members provides any representation or warranty, express or implied, regarding the accuracy, completeness, noninfringement, or fitness for a particular purpose of this document, or any document referenced herein.

© Copyright 2004-2007 Cable Television Laboratories, Inc.
All rights reserved.

Document Status Sheet

Document Control Number:	MD-SP-VOD-CEP2.0-I02-070105			
Document Title:	Content Encoding Profiles 2.0 Specification			
Revision History:	I01 – Released January 7, 2004 as VOD-CEP			
	I02 – Released January 5, 2007 as VOD-CEP2.0			
Date:	January 5, 2007			
Status:	Work in Progress	Draft	Issued	Closed
Distribution Restrictions:	Author Only	CL/Member	CL/Member/Vendor	Public

Key to Document Status Codes:

- Work in Progress** An incomplete document, designed to guide discussion and generate feedback, that may include several alternative requirements for consideration.
- Draft** A document in specification format considered largely complete, but lacking review by Members and vendors. Drafts are susceptible to substantial change during the review process.
- Issued** A stable document, which has undergone rigorous member and vendor review and is suitable for product design and development, cross-vendor interoperability, and for certification testing.
- Closed** A static document, reviewed, tested, validated, and closed to further engineering change requests to the specification through CableLabs.

Trademarks:

DOCSIS®, eDOCSIS™, PacketCable™, CableHome®, CableOffice™, OpenCable™, CableCARD™, DCAS™, OCAP™, and CableLabs® are trademarks of Cable Television Laboratories, Inc.

Contents

1	INTRODUCTION	1
1.1	Purpose of Document	1
1.2	Organization of Document	1
2	REFERENCES	2
2.1	Normative References	2
2.2	Informative References.....	3
3	NOTATION, DEFINITIONS, AND TERMINOLOGY	4
3.1	Compliance Notation	4
4	TERMS AND DEFINITIONS	5
5	ABBREVIATIONS AND ACRONYMS	6
6	BASIC ENCODING	8
6.1	Metadata Specification	8
6.2	Video Encoding Specification	8
6.3	Usage of ANSI/SCTE 35.....	8
6.3.1	<i>ANSI/SCTE 35 Registration Descriptor Specification</i>	8
6.3.2	<i>Usage of the ANSI/SCTE 35 segmentation_descriptor()</i>	8
6.4	Closed Captioning / V-Chip Specification	10
6.5	Audio Encoding Specification.....	11
6.6	Subtitling Data Specification	11
6.7	MPEG-2 Systems Constraints	11
6.7.1	<i>Video PES Constraints</i>	11
6.7.2	<i>Transport Stream Constraints</i>	11
6.7.3	<i>Transport Bitrate Constraints</i>	12
6.7.4	<i>PSI Constraints</i>	12
6.7.5	<i>PID Value Constraints</i>	13
6.8	Recommended Video Compression Practices (Informative).....	13
6.9	Bitrate (Informative).....	13
6.9.1	<i>Transport Bitrate</i>	13
6.9.2	<i>Video Bitrate</i>	14
6.10	Handling Multiple Audio Streams (Informative)	14
7	STANDARD DEFINITION (SD) ENCODING	15
7.1	Metadata Specification	15
7.2	Video Encoding Specification	15
7.3	MPEG-2 Systems Constraints	15
7.3.1	<i>Transport Bitrate Constraints</i>	15
8	HIGH DEFINITION (HD) ENCODING	16
8.1	Metadata Specification	16
8.2	Video Encoding Specification	16
8.3	Audio Encoding Specification.....	16
8.4	MPEG-2 Systems Constraints	16
8.4.1	<i>Transport Bitrate Constraints</i>	16
8.5	Recommended Video Compression Practices (Informative).....	17
8.6	Transport Bitrate (Informative)	17

APPENDIX I ACKNOWLEDGEMENTS18

APPENDIX II REVISION HISTORY19

1 INTRODUCTION

This specification defines the video, audio, and related encoding parameters for both Standard and High-Definition content for contribution to cable television systems. The content may be VOD content or may be advertising content intended for insertion into VOD or other broadcast services. The term "contribution" is being used here to describe the process by which individual programs or ads of finite length are encoded and formatted as files or byte streams for storage on video servers or other headend equipment.

The encoding parameters described in this document are not constraints on the streams delivered by cable systems to a home or business. The streams delivered to a home or business are generally more complex: they may not be finite in length, they are often multiplexes of multiple feeds, and they may include dynamic transitions from program to program.

In general, this specification is used in conjunction with other related specifications from CableLabs to create a full "package" of content. Other standards (e.g., AVC, VC-1) are currently beyond the scope of this document.

1.1 Purpose of Document

This document defines the content specifications for use with encoding systems, asset management and distribution. It does not define a distribution method or define all aspects of the cable system infrastructure that content may encounter during distribution and playback.

1.2 Organization of Document

This document contains three major sections. Section 6 covers general encoding, Section 7 gives additional constraints specific to Standard Definition encoding, and Section 8 covers High Definition encoding.

2 REFERENCES

2.1 Normative References

In order to claim compliance with this specification, it is necessary to conform to the following standards and other works as indicated, in addition to the other requirements of this specification. Notwithstanding, intellectual property rights may be required to use or implement such normative references.

- [ATSC A/52B] ATSC A/52B: Digital Audio Compression (AC-3, E-AC-3), Revision B, 2005.
- [ATSC A/53E] ATSC A/53E: ATSC Digital Television Standard, Revision E with Amendment No. 1, 2005.
- [CEA 608-C] EIA/CEA-608-C: Line 21 Data Services, 2005.
- [CEA 708-C] EIA/CEA-708-C: Digital Television (DTV) Closed Captioning, 2006.
- [IEC 11172-3] ISO/IEC 11172-3:1998 (E), International Standard, Information Technology - Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbits/s – Part 3: Audio.
- [IEC 13818-1] ISO/IEC 13818-1:2000 (E), International Standard, Information Technology - Generic Coding of Moving Pictures and Associated Audio Information: Systems.
- [IEC 13818-2] ISO/IEC 13818-2:2000 (E), International Standard, Information Technology - Generic Coding of Moving Pictures and Associated Audio Information: Video.
- [IEC 13818-3] ISO/IEC 13818-3:1998 (E), International Standard, Information Technology - Generic Coding of Moving Pictures and Associated Audio Information: Audio.
- [IEC 13818-4] ISO/IEC 13818-4:1998/Cor-2:1998, International Standard, Information Technology - Generic Coding of Moving Pictures and Associated Audio Information: Conformance Testing, Technical Corrigendum 2.
- [IEC 61672-1] IEC 61672-1 Electroacoustics – Sound level meters – Part 1: Specifications.
- [ISO 639-2] ISO 639-2: 1998, Codes for the Representation of Names of Languages- Part 2: Alpha-3 Code.
- [ITU-R BS.1770] ITU-R BS.1770 Algorithms to measure audio programme loudness and true peak audio level.
- [SCTE 20] ANSI/SCTE 20: Methods for Carriage of Closed Captions and Non-Real Time Sampled Video, 2004.
- [SCTE 27] ANSI/SCTE 27: (Formerly DVS 026), Subtitling Methods for Broadcast Cable, 2003.
- [SCTE 43] ANSI/SCTE 43: Digital Video Systems Characteristics Standard for Cable Television, 2005.
- [SCTE 35] ANSI/SCTE 35: Digital Program Insertion Cueing Message for Cable, 2004.

- [VOD 1.1] MD-SP-VOD-CONTENT1.1-I05-060831: CableLabs Video-On-Demand Content 1.1 Specification, August 31, 2006, Cable Television Laboratories, Inc.
- [VOD 2.0] MD-SP-VOD-CONTENT2.0-I02-070105: Video-On-Demand Content 2.0 Specification, January 5, 2007, Cable Television Laboratories, Inc..

2.2 Informative References

This specification uses the following informative references.

- [CEA-CEB10-A] CEA-CEB10-A: EIA-708-B Implementation Guidance, 2002.
- [SCTE 30] ANSI/SCTE 30: Digital Program Insertion Splicing API, 2005.
- [SCTE 104] ANSI/SCTE 104: Automation System to Compression System Communications Applications Program Interface (API), 2004.

3 NOTATION, DEFINITIONS, AND TERMINOLOGY

3.1 Compliance Notation

Throughout this document, words are used that define the significance of particular requirements. These words are:

"MUST"	This word means that the item is an absolute requirement of this specification.
"MUST NOT"	This phrase means that the item is an absolute prohibition of this specification.
"SHOULD"	This word 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 weighed 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 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.

4 TERMS AND DEFINITIONS

The syntax and semantics of this specification conform to [IEC 13818-1].

The following terms warrant a definition as used in the context of this specification:

- Access unit** In the case of audio, an access unit is the coded representation of an audio frame. In the case of video, an access unit includes all the coded data for a picture and any stuffing that follows it. If the picture is preceded by a Sequence Header, the access unit begins with the first byte of the `sequence_header_code`. If the picture is preceded by a Group of Pictures Header and no Sequence Header, the access unit begins with the first byte of the `group_start_code`. If the picture is not preceded by a Sequence Header or Group of Pictures, the access unit begins with the first byte of the `picture_start_code`.
- Group of Pictures** Also referred to as a GOP, a Group of Pictures is a sequence of coded pictures, beginning with an I-Picture and ending just before the subsequent I-Picture.
- Video content** Video content is a single file composed of an MPEG-2 transport stream that is an encoding of a feature and optionally trailers and advertising. The video content can contain one or more audio tracks and one or more sets of subtitles. It may also contain private or additional data PID streams.

5 ABBREVIATIONS AND ACRONYMS

The following abbreviations and acronyms are used in this specification:

ANSI	American National Standards Institute
ATSC	Advanced Television System Committee
CBR	Constant Bitrate
CEA	Consumer Electronics Association
DTS	Decoding Time Stamp
DTV	Digital Television
DVB	Digital Video Broadcasting Group
DVS	Digital Video Specification
EIA	Electronic Industries Alliance
fps	Frames per second
GOP	Group of Pictures
HD	High Definition
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
JVT	Joint Video Team
Mbps	Megabits Per Second
MPEG	Moving Picture Experts Group
NTSC	National Television System Committee
PAT	Program Association Table
PCR	Program Clock Reference
PID	Program Identification
PES	Packetized Elementary Stream
PMT	Program Map Table
PTS	Presentation Time Stamp

QAM	Quadrature Amplitude Modulation
SCTE	Society of Cable Telecommunication Engineers
SD	Standard Definition
SPTS	Single Program Transport Stream
T-STD	Transport Stream System Target Decoder
UI	User Interface
VBR	Variable Bitrate
VOD	Video on Demand

6 BASIC ENCODING

This section describes general coding constraints that are applicable to both SD and HD content.

6.1 Metadata Specification

This section describes the normative specification of Metadata associated with encoded video content.

Metadata **MUST** be created in accordance with [VOD 2.0].

6.2 Video Encoding Specification

This section describes the normative specification of the encoded video content.

1. The first byte of the video elementary stream **MUST** be the first byte of a `sequence_start_code`.
2. The video elementary stream **MUST** contain an integral number of access units.
3. A Sequence Header and a Sequence Extension **MUST** precede each I-Picture.
4. All B-Pictures in the video bit stream **MUST** use prediction based on pictures present in the bit stream. Specifically, the first GOP in the video elementary stream **MUST NOT** begin with a B-Picture predicted from a reference picture that does not exist in the stream. That is, the first GOP **MUST** be closed.
5. The length of each GOP **SHOULD** be 15 frames long for 30 fps video-source material, and 12 frames long for 24 fps film-source material.
6. The number of consecutive B-Pictures between anchor pictures (I-Picture or P-Picture) **MUST** be two or less.
7. The 2-bit `picture_structure` field in the `picture_coding_extension()` of each picture **MUST** have the value of '11' in binary, indicating that the picture is encoded as a frame picture.
8. The video elementary stream **MUST NOT** use any MPEG-2 scalable extensions, which include the Sequence Scalable Extension, the Picture Temporal Scalable Extension, and the Picture Spatial Scalable Extension.
9. Black level of content **MUST** be at 0 IRE.

6.3 Usage of ANSI/SCTE 35

6.3.1 ANSI/SCTE 35 Registration Descriptor Specification

This section describes the normative specification of requirements for inclusion of the ANSI/SCTE 35 descriptor.

The ANSI/SCTE 35 Registration Descriptor, as specified in section 6.1 of [SCTE 35], **MUST** be present in the PMT.

6.3.2 Usage of the ANSI/SCTE 35 `segmentation_descriptor()`

This descriptor provides a standardized mechanism for placing declarations into the bitstream for content identification as well as providing a standardized mechanism for segmenting the content into chapters. While additional uses are possible, these two uses will be documented here.

The `segmentation_descriptor()` **MUST** be used only with the `time_signal()` and the `splice_null()` constructs of [SCTE 35]. The descriptor syntax and semantics are defined in section 8.3.3 of [SCTE 35].

6.3.2.1 Content Identification Declaration

This usage of the `segmentation_descriptor()` MUST place the descriptor within a `splice_null()` as defined in section 7.3.1 of [SCTE 35]. The use of this declaration covers content including all content types both live feeds as well as pre-encoded content of both long and short forms (includes programming and advertising).

The semantics of the fields within the `segmentation_descriptor()` for this purpose follow:

segmentation_event_id - A 32-bit segmentation event identifier, unique within the content duration.

segmentation_event_cancel_indicator - MUST be set to '0'.

program_segmentation_flag - MUST be set to '1'.

segmentation_duration_flag - MUST be set to '0'.

segmentation_upid_type - MUST be set to the relevant value found in table 8-7 of [SCTE 35] identifying the type of UPID.

segmentation_upid_length - MUST be set to the relevant value found in table 8-7 of [SCTE 35].

segmentation_upid() - MUST be set to the value assigned to the UPID and be consistent with table 8-7 in [SCTE 35] and the associated **segmentation_upid_type** and **segmentation_upid_length** fields.

segmentation_type_id - MUST be set to 0x10 (unless [SCTE 35] is changed to permit 0x00, when it MUST be set to 0x00, indicating "not used").

chapter and **chapter_count** - MUST be set to zero, indicating "not used."

Duplicate occurrences of a Content Identification Declaration message are permitted. To avoid unnecessary use of bandwidth, the minimum spacing of duplicates MUST be no less than one per minute. The minimum occurrence of a Content Identification Declaration message is once per program. Actual spacing of these messages is likely to be determined by the content provider's Legal and Content Protection staffs.

6.3.2.2 Segmenting Content

This usage of the `segmentation_descriptor()` MUST place the descriptor within a `time_signal()` as defined in section 7.3.4 of [SCTE 35]. Segments MUST have a logical hierarchy consisting of programs (highest level), chapters, and advertisements (refer to table 8-8 of [SCTE 35]). National and local advertisements share the lowest logical level and should not overlap.

Descriptors should normally be paired, the first for a given segment start and the second for segment end. Each segment end usage MAY be followed by another segment start. Overlapping segment definitions MAY be used. `segmentation_descriptor()` pairs are uniquely identified by `segmentation_upid()`, `segmentation_event_id`, and `chapter`. The `segmentation_upid()` MAY be omitted, but if present, MUST be the same between identifier pairs.

For the purposes of defining the `segmentation_descriptor()` semantics, the following definition applies:

Segment - MUST be either a **Program**, a **Chapter**, a **National Advertisement**, a **Local Advertisement**, or an **Unscheduled Event** as listed in table 8-8 of [SCTE 35]. The valid pairings are:

Program start/end – Program end can be overridden by program early termination

Chapter start/end

National advertisement start/end

Local advertisement start/end

Unscheduled_event_start/end

The semantics of the fields within the `segmentation_descriptor()` for segmenting content follow:

segmentation_event_id - A 32-bit segmentation event identifier, unique within the content duration. If a segment end is signaled, the `Segment end time_signal()` **segmentation_event_id** value MUST match the Segment start **segmentation_event_id** value.

segmentation_event_cancel_indicator - MUST be set to '0'.

program_segmentation_flag - MUST be set to '1'.

segmentation_duration_flag - May be set to '0' or '1'. If set to '1', a valid **segmentation_duration()** MUST be included in the descriptor.

segmentation_upid_type - MUST be set to the relevant value found in table 8-7 of [SCTE 35] identifying the type of UPID. A value of 0x00 may be used if desired.

segmentation_upid_length - MUST be set to the relevant value found in table 8-7 of [SCTE 35].

segmentation_upid() - MUST be set to the value assigned to the UPID and be consistent with table 8-7 in [SCTE 35] and the associated **segmentation_upid_type** and **segmentation_upid_length** fields. A null value may be provided if **segmentation_upid_type** is set to 0x00.

segmentation_type_id - MUST be set to a valid value from table 8-8 of [SCTE 35].

chapter MUST be set to non-zero values ranging from one to the value of **chapter_count**. For Program segments, this value MUST be set to one. This field may be used for Chapters and Advertisements as desired.

chapter_count MUST be set to a non-zero value, providing the number of chapters in the program. For Program segments, this value MUST be set to one.

6.4 Closed Captioning / V-Chip Specification

This section describes the normative specification of the encoding and transport of closed caption data in video picture user data.

1. Line 21 closed caption data, encoded in accordance with [CEA 608-C], MUST be transported in video picture user data in accordance with both [SCTE 20] and [ATSC A/53E], as modified by Amendment 1 to [ATSC A/53E].
2. DTV closed caption data, encoded in accordance with [CEA 708-C], MUST be transported in video picture user data in accordance with [ATSC A/53E], as modified by Amendment 1 to [ATSC A/53E].
3. When both [SCTE 20] and [ATSC A/53E] user data are present, the sequence of picture user data MUST be interleaved in accordance with [SCTE 43], section 5.2.2. **Informative Note:** [ATSC A/53E] user data may include both [CEA 608-C] and [CEA 708-C] formats.
4. V-Chip data, encoded in accordance with [CEA 608-C], MUST conform to the ratings and/or content advisory data values set in Metadata.

6.5 Audio Encoding Specification

This section describes the normative specification of the encoded audio content.

1. The audio compression format **MUST** conform to the bit stream syntax for the Digital Audio Compression (AC-3) Standard in accordance with [ATSC A/52B].
2. When available, the source audio **SHOULD** be encoded as 5.1 channel AC-3, i.e., the audio coding mode **SHOULD** be 3/2 and the low frequency effects channel **SHOULD** be on. If the source audio cannot be encoded as AC-3 5.1, then the audio coding mode **MUST** be 2/0, i.e., 2-channel stereo (Left & Right).
3. If the audio is encoded as 5.1 channel, then the encoded bitrate **MUST** be 384 kbps.
4. If the audio is encoded as 2-channel stereo (2/0), then the encoded bitrate **MUST** be 192 kbps.
5. The audio sample rate **MUST** be 48 kHz.
6. The first byte of the audio elementary stream **MUST** be the first byte of an audio access unit.
7. The audio elementary stream **MUST** contain an integral number of access units.
8. The dialogue normalization value (dialnorm) in each AC-3 elementary stream **MUST** be set to agree with (i.e., indicate) the level of average spoken dialogue within the encoded audio program. The dialogue level can be measured by means of an A-weighted integrated measurement (Leq(A) [IEC 61672-1] or [ITU-R BS.1770]).
Informative note: Receivers (i.e., Set-top Boxes, Home Theaters, etc.) use the dialnorm value to adjust the reproduced audio level upon decoding to normalize the dialogue level.¹

6.6 Subtitling Data Specification

This section describes the normative specification of the encoding and transport of subtitling data.

Subtitling data **MUST** be encoded and carried in the transport stream in accordance with [SCTE 27].

6.7 MPEG-2 Systems Constraints

This section describes the coding constraints that apply to the use of the MPEG-2 Systems specification, in creation of a single transport stream. Because this document is for stored content, there should not be any discontinuities or PSI version changes. Other applications, such as broadcast, may have PSI version changes and time base discontinuities. These uses are outside the scope of this document.

6.7.1 Video PES Constraints

This section describes the coding constraints that apply to the video Packetized Elementary Stream (PES).

1. Each video access unit **MUST** be completely contained within one PES packet, and the first byte of the PES packet payload **MUST** be the first byte of the video access unit.
2. Decoding and presentation time stamps (DTS and PTS) **MUST** be contained in the PES packet header of each PES packet that carries an I-Picture.

6.7.2 Transport Stream Constraints

1. The transport stream **MUST** comply with the definition of a transport stream as specified in [IEC 13818-1].
2. The transport stream **MUST** carry only a single program (SPTS).
3. The program in the transport stream **MUST** contain only a single video elementary stream.

¹ For further information see http://www.dolby.com/about/contact_us/contactus.cfm?goto=31

4. The program in the transport stream **MUST** contain at least one audio elementary stream.
5. The transport stream **MUST** consist of 188-byte transport packets.
6. The first byte of the transport stream **MUST** be the first byte of a transport packet.
7. The transport stream **MUST** contain an integral number of transport packets.
8. The transport stream **MUST NOT** contain continuity_counter discontinuities.
9. The transport stream **MUST** contain exactly one system time-base discontinuity (PCR) which **MUST** be signaled in the first PCR packet of the stream.
10. PCRs **MUST** have an accuracy of 5 ppm.
11. The first PCR packet of the stream **MUST** have the transport discontinuity_indicator flag set to '1'.
12. A PCR **MUST** be present in any transport packet containing the first byte of a video PES payload.
13. The audio T-STD **MUST** comply with section 3.6 of Annex A of [ATSC A/52B].
14. The random_access_indicator **MUST** be set to '1' in any transport packet containing the first byte of a video PES payload that carries an I-Picture.

6.7.3 Transport Bitrate Constraints

The transport stream **MUST** be constant bitrate.

6.7.4 PSI Constraints

1. A complete Program Association Table (PAT) **MUST** occur in the transport stream before the first byte of a Program Map Table (PMT).
2. A PMT that contains a complete program definition **MUST** occur in the transport stream before the first transport packet with an elementary stream PID.
3. The time interval in the transport stream between successive occurrences of the PAT **MUST** be less than or equal to 250 milliseconds. It is recommended that the time interval between successive occurrences of the PAT should be 125 milliseconds.
4. The time interval in the transport stream between successive occurrences of the PMT **MUST** be less than or equal to 250 milliseconds. It is recommended that the time interval between successive occurrences of the PMT **SHOULD** be 125 milliseconds.
5. The stream_type value assigned in the PMT to the video elementary stream **MUST** be 0x02 (MPEG-2) or 0x80 (User Private).
6. The stream_type value assigned in the PMT to all audio elementary streams **MUST** be 0x81 (AC-3).
7. Descriptors **MUST** be included in the PMT to comply with SCTE and ATSC standards. The following descriptors **SHOULD** be considered.

Table 1 - Descriptors

Descriptor	Defining Specification	Notes
Registration	ISO/IEC 13818-1	Optional per ANSI/SCTE 54 2004
ISO-639 language	ISO/IEC 13818-1	Not required after 1 March 2008 per ATSC A/53E
AC-3 Audio Stream	ATSC A/52B	After 1 March 2008 will include ISO 639 language descriptor per ATSC A/53E

8. Other private data PIDs are allowed in the PMT. These entries in the PMT SHOULD have registration_descriptors identifying their structure.
9. There MAY be PIDs in the transport stream that are not referenced in the PSI. The use and handling of these PIDs are beyond the scope of this document.
10. PMT sections SHOULD be no longer than 183 bytes in length and SHOULD be placed into a single transport stream packet. It is anticipated that future PMT sections may exceed this length due to PID demands for audio, ETV/OCAP and other applications.
11. All PATs in the file SHOULD be identical and should have a constant version_number.
12. All PMTs in the file SHOULD be identical and should have a constant version_number.

6.7.5 PID Value Constraints

This specification uses a fixed PID allocation for PSI, video and audio streams. In the future, these PIDs may not be fixed. In anticipation of this change, users and implementers should not assume these fixed values will always be used, and instead should determine the PIDs based on the contents of the PSI.

1. The program_map_PID for the program MUST have the value 0x1E0 (decimal 480).
2. The elementary_PID assigned to the video elementary stream MUST have the value 0x1E1 (decimal 481).
3. The PCR_PID of the program MUST have the value 0x1E1 (decimal 481).
4. The elementary_PID assigned to the first, or primary, audio elementary stream listed in the PMT MUST have the value 0x1E2 (decimal 482).
5. If one or more audio elementary streams are present in addition to the primary audio elementary stream, the elementary_PID assigned to the Nth audio elementary stream listed in the PMT MUST have the value $0x1E0 + N + 1$.

6.8 Recommended Video Compression Practices (Informative)

This section is for informative purposes only.

1. Film-source material SHOULD be encoded using "reverse" or "inverse" telecine, resulting in a coded frame rate of 23.97 Hz.
2. Each stream within the program SHOULD start without any significant leader (such as black video frames) and end without any significant trailer to facilitate the seamless back-to-back splicing of separate programs.
3. Quality Control cleanup MAY be performed, if necessary, using noise reduction and bandwidth limiting.

6.9 Bitrate (Informative)

This section describes bitrate calculations and concerns for video content. The examples here use SD values but the calculations, with appropriate values, are applicable to HD as well.

6.9.1 Transport Bitrate

There are many concerns, constraints and issues that determine the optimal bitrate for a given situation. For example, success has been widely achieved using the 3.75 Mbps transport bitrate for SD; other rates are possible. However, users and implementers should be aware that installed and legacy systems have constraints on bandwidth and system resource management that do not currently support widely-varied bitrates – especially within a single QAM multiplex as may be encountered in actual use. Thus, 3.75 Mbps is a "safe harbor" for SD until planned system improvements occur.

Depending on MPEG-2 encoder quality, desire to maintain high consumer picture quality, etc., it has been suggested that low-motion, low-complexity content can be successfully encoded at 3.37 Mbps, while sports content may require a rate exceeding that permitted by this standard (e.g., 4.125 Mbps). However, optimal picture quality can be obtained by changes to raw bitrate and/or changes in encoder resolution and/or filtering. Considering the "safe harbor" noted above, it is desired to change filtering first and bitrate as required on a secondary basis.

6.9.2 Video Bitrate

The maximum bitrate that is available for video is dependent on the number of audio services and their bitrates. For example, in the SD case where the transport stream is limited to 3.75 Mbps, the video rate must be lower when 5.1 audio is included at 384 kbps than when stereo audio at 192 kbps is used. If multiple audio streams are included, the video must leave room for the highest bitrate audio. For example, if both stereo and 5.1 audio are included, then the video and the 384 kbps audio must fit into 3.75 Mbps.

It is important to take overhead into account when determining the video rate. The video rate is just that – the rate of the video elementary stream. This does not include the PES and transport overhead. The transport overhead will add about 2.2% and the PES and PCR will add 4.4 kbps. These numbers are estimates.

For example, assuming a 5.1 audio at 384 kbps, what is the max video rate for SD? The audio rate is increased by the PES and transport overhead to 396 kbps. The PSI (PAT and PMT) at ten times per second add another 30 kbps. This leaves 3.324 Mbps for the video. Reducing by the transport overhead (2%), and subtracting the PES and PCR overhead leaves 3.25 Mbps. This calculation is meant to show how the maximum video rate is affected by the audio and how the transport and other overhead comes into play. The actual maximum video rate will depend on the encoding and multiplexing system in use.

Complex SD sequences, such as sports or action scenes, will require rates over 3 Mbps. It may be possible to use a much lower rate in some circumstances, such as for slide show type sequences of pictures. Because of the wide variety of source material, no limits are placed on the video elementary stream rate by this specification.

6.10 Handling Multiple Audio Streams (Informative)

This section is for informative purposes only.

Content encoded with additional or alternative audio elementary streams will have PID assignments pursuant to Section 6.7.5, and will not exceed the nominal transport bitrate specified in Section 6.7.3. It is expected that the provisioning of the content conceptually occurs as in the following narrative (in this example, Spanish audio is the "first alternate" language and, thus, is encoded on PID 0x1E3):

1. Consumer is presented with a selection of content that has alternative (Spanish) audio available (it is beyond the scope of this narrative to discuss the language of Metadata and/or UI).
2. Consumer selects the alternative language via the UI.
3. The server source / consumer sink session streams with Video PID=0x1E1 and Audio PID=0x1E3 (alternate audio). Primary audio PID=0x1E2 is dropped by the server and does not consume any QAM bandwidth. In the example that primary audio was Dolby 5.1 (384 kb/s) and the secondary audio was Dolby 2/0 (192 kb/s), the transport stream is 192 kb/s lower than the nominal bitrate in Section 6.7.3.
4. For clarity: It is important to note that the server did **not** re-map the alternative audio from 0x1E3 to 0x1E2.

7 STANDARD DEFINITION (SD) ENCODING

This section is applicable to "Standard Definition" (SD) MPEG-2 encoding of video content. SD video compression is constrained to the so-called "NTSC 480i" format. Although based on the same core SCTE parameters, this SD encoding specification contains parameters that were (and may continue to be) applicable to certain "legacy" implementations, and are not relevant to HD-compatible set-top devices and/or 2-way compatible host devices.

Further, there are some "fixed" or "set" values within this section that are applicable to the "state of the industry" as of this writing. For example, although MPEG video encoding continues to improve, some of the bitrate parameters specified herein are necessary due to limitations in QAM and system bandwidth resource management systems.

7.1 Metadata Specification

This section describes the normative specification of Metadata associated with SD encoded video content.

At a minimum, the Metadata **MUST** include a Video Content Asset in which the encodingProfile attribute has the value "MPEG2SD" as defined by [VOD 2.0].

7.2 Video Encoding Specification

This section describes the normative specification of the SD encoded video content.

1. The video compression format **MUST** conform to the requirements of [IEC 13818-2]. The allowable parameters **MUST** be bounded by the upper limits specified for the Main Profile at Main Level.
2. The video bit stream **MUST** meet the constraints and extensions described in [SCTE 43] for a coded vertical size of 480, aspect ratio of 4:3, and interlaced scan, i.e., it **MUST** be constrained to the entries in table 3 of [SCTE 43] with the vertical_size_value equal to 480, the aspect_ratio_information equal to 2 (4:3 display aspect ratio), the frame_rate_code equal to 4 (29.97 Hz), and the progressive_sequence equal to 0 (interlaced scan).
3. The video elementary stream **MUST** be encoded at a constant bitrate (CBR). In the future, variable bitrate (VBR) encoding may be allowed.

7.3 MPEG-2 Systems Constraints

This section describes the coding constraints that apply to the use of the MPEG-2 Systems specification in creation of a single transport stream containing SD encoded video content.

1. No transport packet **MUST** have an adaptation_field_length equal to zero, i.e., the adaptation field, when present, in any transport packet **MUST** have a length greater than one byte. **Informative Note:** This constraint exists to support legacy set-top boxes.
2. The number of bytes between the last byte of the start code preceding each Picture Start Code to the first byte of the Picture Start Code must be a multiple of four (commonly referred to as "quad-byte alignment"). **Informative Note:** This constraint exists to support legacy set-top boxes.

7.3.1 Transport Bitrate Constraints

The aggregate transport bitrate for PID 0, the PMT PID, the video PID, and any one audio PID **MUST NOT** exceed 3.75 Mbps.

8 HIGH DEFINITION (HD) ENCODING

This section is applicable to "High Definition" (HD) MPEG-2 encoding.

8.1 Metadata Specification

This section describes the normative specification of Metadata associated with HD encoded video content.

1. The Metadata **MUST** include a Video Content Asset in which the encodingProfile attribute has the value "MPEG2HD" as defined by [VOD 2.0].
2. Alternatively, the Metadata **MUST** include an "HDContent" tag that has a value of "Y" to indicate HD content encoded to this specification as defined by [VOD 1.1].

8.2 Video Encoding Specification

This section describes the normative specification of the HD encoded video content.

1. The video compression format **MUST** conform to the syntax of [IEC 13818-2], and **MUST** be subject to the constraints specified in Annex A of [ATSC A/53E]. The allowable parameters **MUST** be bounded by the upper limits specified for the Main Profile at High Level.
2. For video produced using the 1080i production format, the video bit stream **MUST** meet the constraints and extensions described in table 3 of [SCTE 43] for a coded vertical size of 1080, coded horizontal size of 1920, aspect ratio of 16:9, the frame_rate_code equal to 4 (29.97 Hz), and the progressive_sequence equal to 0 (interlaced scan).
3. For video produced using the 720P production format, the video bit stream **MUST** meet the constraints and extensions described in table 3 of [SCTE 43] for a coded vertical size of 720, coded horizontal size of 1280, aspect ratio of 16:9, the frame_rate_code equal to 7 (59.94 Hz), and progressive scan (progressive_sequence equal to 1).

8.3 Audio Encoding Specification

This section describes the normative specification of the audio associated with the HD encoded video content.

Alternatively, for applications of this standard outside of North America, the audio compression format **MAY** conform to either [IEC 11172-3] or [IEC 13818-3], subject to constraints and restrictions that are to be determined.

Informative Note: This feature exists to support DVB systems.

8.4 MPEG-2 Systems Constraints

This section describes the coding constraints that apply to the use of the MPEG-2 Systems specification, in creation of a single transport stream containing HD-encoded video content.

8.4.1 Transport Bitrate Constraints

The aggregate transport bitrate for PID 0, the PMT PID, the video PID, and any one audio PID **MUST NOT** exceed 19 Mbps.

8.5 Recommended Video Compression Practices (Informative)

This section applies to HD encoding of video content and is for informative purposes only.

The video elementary stream MAY be encoded at a variable bitrate (VBR).

8.6 Transport Bitrate (Informative)

This section applies to HD encoding of video content and is for informative purposes only.

There are many concerns, constraints and issues that determine the optimal bitrate for a given situation. Success has been widely achieved using the 15 Mbps transport bitrate, which is a multiple integer of the Standard Definition transport bitrate, using a variety of encoding systems and a variety of content types; other rates are possible.

However, users and implementers should be aware that installed and legacy systems have constraints on bandwidth and system resource management that do not currently support widely varied bitrates – especially within a single QAM multiplex as may be encountered in actual use. Thus, 15 Mbps is a "safe harbor" until planned system improvements occur. Improvements in both system resource management as well as MPEG encoding will result in successful accommodation of varied (and lower) bitrates – however, it does not appear that deployed systems can benefit from those improvements at this time.

Appendix I Acknowledgements

We wish to heartily thank the vendor participants contributing directly to this document:

Home Box Office – Author, Version 1

Manzanita Systems, Inc. – Author, Version 2

Contributors: Bigband Networks, Broadcom, Comcast, Concurrent, InDemand, C-COR, Harmonic, Tandberg, SA/Cisco, StarzEncore, Seachange, Turner, TVN, Time Warner Cable, and others.

Appendix II Revision History

The following ECNs was incorporated into version I02 of this specification:

ECN Identifier	Author	ECN Accepted Date	Description
VOD-CEP-N-06.0051-3	Greg Vines	1/3/07	VOD CEP omnibus