

SMPTE RECOMMENDED PRACTICE

VC-1 Bitstream Transport Encodings



Intellectual property notice

Copyright 2010 THE SOCIETY OF MOTION PICTURE AND TELEVISION ENGINEERS

3 Barker Ave.
White Plains, NY 10601
+1 914 761 1100
Fax +1 914 761-3115
E-mail eng@smpte.org
Web <http://www.smpte.org>

The user's attention is called to the possibility that compliance with this document may require use of inventions covered by patent rights. By publication of this document, no position is taken with respect to the validity of these claims or of any patent rights in connection therewith. The patent holders have, however, filed statements of willingness to grant a license under these rights on fair, reasonable and nondiscriminatory terms and conditions to applicants desiring to obtain such a license. Contact information may be obtained from the SMPTE. No representation or warranty is made or implied that these are the only licenses that may be required to avoid infringement in the use of this document.

	Page
Foreword.....	4
Introduction	5
1 Scope.....	6
2 Normative References	6
3 Conformance Notation.....	6
3.1 Conforming Implementations.....	7
4 Background and Overview	7
4.1 Background.....	7
4.2 Overview	7
4.3 Abbreviations	7
4.4 Definitions	8
4.4.1 Data Access Unit Definition	8
5 MPEG-2 Transport Stream Encoding.....	9
5.1 SIGNALING OF VC-1 ELEMENTARY STREAMS.....	9
5.1.1 Stream Type	9
5.1.2 Registration Descriptor	9
5.1.3 Profile and Level Subdescriptor.....	10
5.1.4 Alignment Subdescriptor	12
5.1.5 Buffer Size Subdescriptor.....	13
5.1.6 Other ISO/IEC 13818-1 Elementary Stream Descriptors.....	13
5.1.7 PMT Descriptors for Programs or Services	14
5.1.8 Still Picture Subdescriptor	14
5.2 ENCAPSULATION OF VC-1 ELEMENTARY STREAMS.....	15
5.2.1 PES packet length	15
5.2.2 Stream ID.....	15
5.2.3 Data Alignment Indicator	15
5.2.4 Time Stamps (PTS and DTS).....	16
5.2.5 Extension Flags	16
5.2.6 Stream ID Extension.....	16
5.2.7 Discontinuity Indicator	16
5.2.8 Random Access Indicator.....	18
5.2.9 Elementary Stream Priority Indicator.....	18
5.2.10 Splice Countdown.....	18
5.2.11 Seamless Splice Flag	18
5.2.12 Splice Type	19
5.3 PES PACKET PAYLOAD FORMAT HEADER.....	19
5.3.1 VC-1 Simple and Main Profile	19
5.3.2 VC-1 Advanced Profile	20
5.4 T-STD BUFFER MODEL FOR VC-1 STREAMS	20
5.4.1 Transport System Target Decoder	20
5.4.2 Mode of Operation	21
5.4.3 Operation of Buffer TBn.....	21
5.4.4 Operation of Buffer MBn:.....	22
5.4.5 Operations of Buffer EBn.....	23
5.4.6 Decoding and Presentation	23
5.4.7 Decoding of first Access Unit following an Entry Point start code.....	24
5.4.8 Rates and Buffer Sizes.....	24
6 MPEG-2 Program Stream Encoding	24
6.1 SIGNALING OF VC-1 ELEMENTARY STREAMS.....	24
6.2 ENCAPSULATION OF VC-1 ELEMENTARY STREAMS.....	24
6.3 PES PACKET PAYLOAD FORMAT HEADER.....	24
6.4 P-STD BUFFER MODEL FOR VC-1 STREAMS	24
6.4.1 Program Stream Target Decoder	24
6.4.2 Mode of Operation	26

6.4.3	Operation of Buffer Bn	26
6.4.4	Decoding and Presentation.....	26
6.4.5	Decoding of first Access Unit following an Entry Point start code	26
6.4.6	Buffer sizes	26
Annex A	Graphics of Syntax for this SMPTE Recommended Practice (Informative).....	27
A.1	Introduction	27
A.2	PES packet	27
A.3	TS Program Map Section.....	28
A.4	Program Stream Map.....	29
Annex B	Bibliography (Informative)	30
Revision Notes		31

Foreword

SMPTE (the Society of Motion Picture and Television Engineers) is an internationally-recognized standards developing organization. Headquartered and incorporated in the United States of America, SMPTE has members in over 80 countries on six continents. SMPTE's Engineering Documents, including Standards, Recommended Practices and Engineering Guidelines, are prepared by SMPTE's Technology Committees. Participation in these Committees is open to all with a bona fide interest in their work. SMPTE cooperates closely with other standards-developing organizations, including ISO, IEC and ITU.

SMPTE Engineering Documents are drafted in accordance with the rules given in Part XIII of its Administrative practices.

SMPTE Recommended Practice RP 227 was prepared by Technology Committee C24.

Introduction

The purpose of this recommended practice is to define a minimum set of rules for the carriage of a VC-1 elementary stream in MPEG-2 Systems protocols. In particular, this design is intended to provide a generic method for carrying a VC-1 video elementary stream over specific MPEG-2 transport streams, such as the ones defined by ATSC, DVB and SCTE. It is also intended to provide a generic means of carrying a VC-1 video elementary stream in an MPEG-2 Program Stream as used by the DVD Forum or the Blu-Ray consortium. One of the particular aspects of this Recommended Practice is to make sure that any private fields and private values needed in the mapping is properly scoped by an identifier to avoid any collisions or ambiguity in any of these delivery networks or media.

VC-1 Bitstream Transport Encodings

1 Scope

This document describes the protocols for carrying video elementary streams conforming to the SMPTE VC-1 Simple, Main and Advanced Profile over MPEG-2 Transport Streams and MPEG-2 Program Streams. This document does not describe the video bit stream syntax. It describes the encapsulation and signaling of the video bit stream using the respective transport protocols.

The document builds on the synchronization layer specified by MPEG-2 Systems, which is the fundamental mechanism for ensuring Audio/Video/Data synchronization. However, this document does not cover the policies that must be adopted at a transmission system head-end to implement a given quality level of Audio/Video/Data synchronization.

Any field of the MPEG-2 Transport Stream packet header or MPEG-2 PES packet header that is not mentioned in this document shall keep its default meaning as defined in ISO/IEC 13818-1:2000.

2 Normative References

The following standards contain provisions which, through reference in this text, constitute provisions of this standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this standard are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

SMPTE Standard 421M¹: VC-1 Compressed Video Bitstream Format and Decoding Process

ISO/IEC 13818-1:2000 Information Technology – Generic Coding of Moving Pictures and Associated Audio Information: Systems (2nd Edition).

ISO/IEC 13818-1:2000/Cor.2:2001(E) Information Technology – Generic Coding of Moving Pictures and Associated Audio Information: Systems Corrigendum 2, "Buffer Model Extension".

ISO/IEC 13818-1:2000/AMD2, "Support of IPMP on MPEG-2 Systems", 2003.

3 Conformance Notation

Normative text is text that describes elements of the design that are indispensable or contains the conformance language keywords: "shall", "should", or "may". Informative text is text that is potentially helpful to the user, but not indispensable, and can be removed, changed, or added editorially without affecting interoperability. Informative text does not contain any conformance keywords.

All text in this document is, by default, normative, except: the Introduction, any section explicitly labeled as "Informative" or individual paragraphs that start with "Note:"

The keywords "shall" and "shall not" indicate requirements strictly to be followed in order to conform to the document and from which no deviation is permitted.

The keywords, "should" and "should not" indicate that, among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required; or that (in the negative form) a certain possibility or course of action is deprecated but not prohibited.

¹ 421M and VC-1 are used interchangeably in this document to refer to SMPTE 421M.

The keywords "may" and "need not" indicate courses of action permissible within the limits of the document.

The keyword "reserved" indicates a provision that is not defined at this time, shall not be used, and may be defined in the future. The keyword "forbidden" indicates "reserved" and in addition indicates that the provision will never be defined in the future.

3.1 Conforming Implementations

A conformant implementation according to this document is one that includes all mandatory provisions ("shall") and, if implemented, all recommended provisions ("should") as described. A conformant implementation need not implement optional provisions ("may") and need not implement them as described.

4 Background and Overview

4.1 Background

This document defines the encapsulation and signaling of VC-1 elementary streams in MPEG-2 Transport and Program Streams. In particular, this document is not concerned with the carriage of audio elementary streams or data elementary streams that may be multiplexed with the VC-1 elementary stream in the same MPEG-2 Program Stream or the same MPEG-2 Transport Stream. The synchronization of audio access units in an audio elementary stream with VC-1 video access units in a video elementary stream should be achieved by means of the synchronization layer specified by ISO 13818-1. Likewise, the synchronization of data access units in a synchronized or synchronous data elementary stream should be achieved by means of ISO 13818-1. The synchronization layer in MPEG-2 Systems is based on the reconstruction of a System Time Clock in the receiver and referencing instants on this timeline by means of Decoding Time Stamp and Presentation Time Stamps.

This document relies on ISO 13818-1 as a baseline layer and defines a series of allowed extensions that enable the encapsulation and signaling of VC-1 elementary streams. This results in implementations compatible with ISO 13818-1.

4.2 Overview

This document is organized in two main sections, the first one covering the carriage of VC-1 elementary streams in MPEG-2 Transport Streams and the second covering the carriage of VC-1 elementary streams in MPEG-2 Program Streams. Each section starts with a sub-section dealing with the signaling aspects. For example, in a Transport Stream, Programs are signaled using a collection of tables transmitted cyclically and known as PSI (Program Specific Information). Specifically, a Program Map Table (PMT) provides the Program details and specifics necessary for finding and decoding the component Elementary Streams. Also, as MPEG-2 Systems requires PES encapsulation of video data in order to generate packetized video elementary streams, the next sub-section deals with the packetization specifics for Simple, Main and Advanced Profiles. The last sub-section specifies the Hypothetical Buffer Model definition and defines the operations of this Buffer Model.

This document adopts a particular method for aggregating VC-1 elementary stream data (SMPTE 421M) and VC-1 metadata (see Annex I and J in SMPTE 421M) to interface with the underlying MPEG-2 Systems packetization and synchronization layer. The mapping of VC-1 elementary streams to MPEG-2 Systems uses this interface but it is not necessarily representative of the interfacing methodology that should be adopted for other transport protocols.

4.3 Abbreviations

This section provides a list of acronyms used in this document. This list is a supplement to the set of acronyms defined in section 2.1 of ISO 13818-1 and in section 4 of SMPTE 421M.

AV	Audio and Video
CRC	Cyclic Redundancy Check
DTS	Decoding Time Stamp
DTV	Digital Television
ES	Elementary Stream
ESCR	Elementary Stream Clock Reference
PAT	Program Association Table
PMT	Program Map Table
PCR	Program Clock Reference
PES	Packetized Elementary Stream
PID	Packet Identifier
PSI	Program Specific Information
PTS	Presentation Time Stamp
PU	Presentation Unit
SCR	System Clock Reference
SDO	Standards Development Organization
STDS	System Target Decoder
TS	Transport Stream

4.4 Definitions

4.4.1 Data Access Unit Definition

The following definition for an access unit shall be applicable for both MPEG-2 Transport Streams and MPEG-2 Program Streams.

An access unit is defined to be a coded representation of a single picture in a VC-1 elementary stream.

In the case of the Advanced Profile, an access unit shall include all the coded data for a picture and any flushing bits that follows it, up to but not including the start code of the next access unit (or any stuffing bytes before it). Coded picture data represents a video frame regardless whether the frame has been encoded as progressive or interlaced progressive mode, interlaced frame interlaced mode or interlaced field interlaced mode. The start of the next access unit shall be the byte of a sequence start code, an entry point start code or a frame start code. Also, if the frame is not preceded by a sequence start code and a sequence header or an entry point start code and an entry point header, the access unit shall begin with a frame start code. Otherwise, the access unit shall start with the first byte of the first of these structures (excluding any stuffing bytes) before the frame start code. An access unit shall also include any user data start code and user data bytes at the sequence, entry point, frame or field level. The rules defining the scope of the user data constructs are described in Annex F of SMPTE 421M.

In the case of Simple Profile or Main Profile, an access unit shall include all the coded data for a video frame, including the VC-1_SPMP_PESpacket_PayloadFormatHeader() bytes that precedes it as well as

any flushing bits present to ensure byte alignment that follows it, up to but not including the start code of the next access unit (or any stuffing bytes before it). The start of the next access unit shall be the first byte in the next VC-1_SPMP_PESpacket_PayloadFormatHeader() which is either a sequence start code or a frame start code.

5 MPEG-2 Transport Stream Encoding

For the description of coded bit streams this document follows the conventional approach used by ISO 13818-1. Data fields are represented using a different font format (e.g. **data_element**). The bit stream syntax is clearly defined in separate tables that use a C-like pseudo-code notation to expose the sequential order of data fields in the stream. These tables are normally followed by precise semantic field definitions.

5.1 SIGNALING OF VC-1 ELEMENTARY STREAMS

MPEG-2 Elementary Streams carrying VC-1 video elementary stream data shall be signaled in a Program Map Table as defined in ISO 13818-1 with Corrigendum 2 and Amendment 2.

The value of the fields in sections 5.1.1 and 5.1.2 are specified for VC-1 elementary streams. The sub-descriptor structures introduced in section 5.1.3 through 5.1.5 may only appear in an MPEG-2 Program Element conveying VC-1 elementary stream data.

5.1.1 Stream Type

The **stream_type** value in the PMT describing a VC-1 Elementary Stream should be set to 0xEA. This value indicates that this Elementary Stream is considered to be a privately-managed stream. The scope of this private value shall be captured by means of the Registration Descriptor defined below.

5.1.2 Registration Descriptor

At least one MPEG-2 **registration_descriptor()** shall be present in the inner descriptor loop of the MPEG-2 Program Element listed in the **TS_program_map_section** corresponding to the VC-1 Elementary Stream. The syntax and semantics for this descriptor appears in Table 1 and in the subsequent text.

Table 1 – Syntax for the registration_descriptor()

Syntax	Bits	Format
registration_descriptor() {		
descriptor_tag	8	0x05
descriptor_length	8	uimsbf
format_identifier	32	0x56432D31
for(i =0; i < K; i++){		
subdescriptor()	N*8	uimsbf
}		
}		

descriptor_tag – This 8-bit field shall be set to 0x05 to identify this data structure as an MPEG-2 registration descriptor.

descriptor_length – This 8-bit field specifies the number of bytes of the descriptor immediately following the **descriptor_length** field.

format_identifier – The value for this 32-bit field shall be set to 0x56432D31

subdescriptor() – A data structure of up to 251 bytes of additional information whose syntax is determined in a case-by-case basis. The first field of a **subdescriptor()** shall be an 8-bit value known as the **subdescriptor_tag** which identifies syntax and semantics for a particular sub-descriptor. Table 2 defines values for the **subdescriptor_tag**.

Table 2 – List of values applicable to the subdescriptor_tag

Assigned Value	Description
0x00	Indicates a null sub-descriptor. A null sub-descriptor consists only of an 8-bit subdescriptor_tag field.
0x01	Profile and Level subdescriptor defined as part of this document.
0x02	Alignment subdescriptor defined as part of this document.
0x03	Buffer size subdescriptor defines as part of this document
0x04	Still Picture subdescriptor defined as part of this document
0x04 - 0xFE	SMPTE Reserved for future applications.
0x05 - 0xFE	SMPTE Reserved for future applications.
0xFF	Indicates a null sub-descriptor. A null sub-descriptor consists only of an 8-bit subdescriptor_tag field.

In the MPEG-2 Registration Descriptor specified in Table 1 above, the sub-descriptors listed in Table 2 above shall always appear in increasing order of their **subdescriptor_tag** value.

Note: Whenever there is a need to add a new sub-descriptor to the list in Table 2 above, it is highly recommended that the sub-descriptor tag value assigned to the new sub-descriptor be assigned the lowest available tag value. This constraint added with the aforementioned constraint that the sub-descriptor tag values appear in increasing order ensures that the most recent sub-descriptors always appear at the bottom of the MPEG-2 Registration Descriptor defined in Table 1. In the event that a new sub-descriptor is defined, such discipline ensures that the oldest sub-descriptors always appear at the beginning of the sub-descriptor list and the most recent ones at the bottom of the list. Consequently, a decoder is always capable of parsing and processing all the sub-descriptors it knows and if an unknown sub-descriptor is found, that sub-descriptor and subsequent sub-descriptors in the list, if any, can be ignored.

5.1.3 Profile and Level Subdescriptor

A Profile and Level Subdescriptor shall be used to indicate the Profile and Level applicable to the VC-1 elementary stream. Syntax and semantics for this sub-descriptor are defined in Table 3 and in the subsequent text. The **sd_profile_level()** subdescriptor signals the profile and level for the associated VC-1 Elementary Stream.