COVERING NOTE

GENERAL SECRETARIAT OF THE INTERNATIONAL TELECOMMUNICATION UNION

Geneva, 20 May 2006

ITU – TELECOMMUNICATION STANDARDIZATION SECTOR

Subject: Erratum 1 (05/2006) to

ITU-T Recommendation J.181 (06/2004), *Digital program insertion cueing message for cable television systems*

Add Appendix I which was omitted when ITU-T Rec. J.181 (06/2004) was published:

Appendix I

Recommended practices and interpretation guide

I.1 Introduction

This appendix is to serve as an informational enhancement to this Recommendation, which is necessarily brief in many areas in order to maintain conciseness and accuracy.

I.2 Informative References

[1] Proposed Recommendation J.sapi, *Digital Program Insertion Splicing API.*

[2] ITU-T Recommendation H.222.0 (2000) | ISO/IEC 13818-1:2000, *Information technology – Generic coding of moving pictures and associated audio information: Systems.*

[3] ITU-T Recommendation H.262 (2000) | ISO/IEC 13818-2:2000, *Information technology – Generic coding of moving pictures and associated audio information: Video.*

[4] ISO/IEC 13818-4:1998, *Information technology – Generic coding of moving pictures and associated audio information – Part 4: Conformance testing.*

[5] SMPTE 312M (2001), *Television – Splice Points for MPEG-2 Transport Streams.*

[6] SCTE 40 2003 (Formerly SCTE DVS/313), *Digital Cable Network Interface Standard.*

[7] SCTE DVS/209 (Original Issue – Feb 1, 1999), *DPI System Physical Diagram.*

I.3 Glossary of terms and acronyms

I.3.1 Definitions

Throughout this appendix, the terms used have specific meanings. Because some of the terms that are defined in ITU-T Rec. H.222.0 | ISO/IEC 13818-1 [2] have very specific technical meanings,

the reader is referred to the original source for their definition. For terms used in this appendix, brief definitions are given below.

I.3.1.1 access unit: The coded representation of a video picture or an audio frame [2].

I.3.1.2 analog cue tone: In an analog system, a signal which is usually either a sequence of DTMF tones or a contact closure that denotes to ad insertion equipment that an advertisement avail is about to begin or end.

I.3.1.3 avail: Time space provided to cable operators by cable programming services during a program for use by the CATV operator; the time is usually sold to local advertisers or used for channel self promotion.

I.3.1.4 break: Avail or an actual insertion in progress.

I.3.1.5 cipher block chaining (CBC): This is a specific method of encryption. It is one of the methods used in DES.

I.3.1.6 component splice mode: A mode of the cueing message whereby the program_splice_flag is set to '0' and indicates that each PID/component that is intended to be spliced will be listed separately by the syntax that follows. Components not listed in the message are not be spliced.

I.3.1.7 cyclic redundancy check (CRC): A method to verify the integrity of a transmitted message.

I.3.1.8 cueing message: See message.

I.3.1.9 data encryption standard (DES): A method for encrypting data with symmetric keys.

I.3.1.10 digital video broadcasting (DVB): An international consortium for the development of digital television systems.

I.3.1.11 DPI cue message: See message.

I.3.1.12 electronic code book (ECB): This is a specific method of encryption. It is one of the methods used in DES.

I.3.1.13 entitlement control message (ECM): These are private conditional access information messages which specify control words and possibly other, typically stream-specific, scrambling and/or control parameters.

I.3.1.14 entitlement management message (EMM): These are private conditional access information messages which specify the authorization levels or the services of specific decoders. They may be addressed to single decoders or groups of decoders.

I.3.1.15 event: A splice event or a viewing event as defined below.

I.3.1.16 in point: A point in the stream, suitable for entry, that lies on an access unit boundary.

I.3.1.17 message: In the context of this document a message is the contents of any splice_info_section.

I.3.1.18 out point: A point in the stream, suitable for exit, that lies on an access unit boundary.

I.3.1.19 payload_unit_start_indicator: A bit in the transport packet header that signals, among other things, that a section begins in the payload that follows [2].

I.3.1.20 packet identifier (PID): A unique 13-bit value used to identify elementary streams of a program in a single or multi-program Transport Stream [2].

I.3.1.21 PID stream: A stream of packets with the same PID within a transport stream.

I.3.1.22 pointer_field: The first byte of a transport packet payload, required when a section begins in that packet [2].

I.3.1.23 presentation time: The time that a presentation unit is presented in the system target decoder [2].

I.3.1.24 program: A collection of video, audio, and data PID streams which share a common program number within an MPTS [2].

I.3.1.25 program in point: A group of PID stream In Points that correspond in presentation time.

I.3.1.26 program out point: A group of PID stream Out Points that correspond in presentation time.

I.3.1.27 program splice mode: A mode of the cueing message whereby the program_splice_flag is set to '1' and indicates that the message refers to a Program Splice Point and that all PIDs/components of the program are to be spliced.

I.3.1.28 program splice point: A Program In Point or a Program Out Point.

I.3.1.29 registration descriptor: Carried in the PMT of a program to indicate that, when signalling splice events, splice_info_sections shall be carried in a PID stream within this program. The presence of the Registration Descriptor signifies a program's compliance with this Recommendation.

I.3.1.30 reserved: The term "reserved", when used in the clauses defining the coded bit stream, indicates that the value may be used in the future for extensions to the standard. Unless otherwise specified in this Recommendation, all reserved bits shall be set to '1'.

I.3.1.31 splice event: An opportunity to splice one or more PID streams.

I.3.1.32 splice immediate mode: A mode of the cueing message whereby the splicing device shall choose the nearest opportunity in the stream, relative to the splice_info_table, to splice. When not in this mode, the message gives a "pts_time", which is a presentation time, for the intended splicing moment.

I.3.1.33 splice point: A point in a PID stream that is either an Out Point or an In Point.

I.3.1.34 viewing event: A television program or a span of compressed material within a service; as opposed to a splice event, which is a point in time.

I.3.2 Abbreviations

This Recommendation uses the following abbreviations:

ATSC	Advanced Television Systems Committee
CBC	Cipher Block Chaining
CBR	Constant Bit Rate
CRC	Cyclic Redundancy Check
DES	Data Encryption Standard
DVB	Digital Video Broadcasting
ECB	Electronic Code Book
ECM	Entitlement Control Message
EMM	Entitlement Management Message
MPTS	Multi Program Transport Stream
PID	Packet Identifier
PMT	Program Map Table
PTS	Presentation Time Stamp

SPTS Single Program Transport Stream

T-STD Transport Stream System Target Decoder

uimsbf Unsigned integer, most significant bit first

VBR Variable Bit Rate

I.4 Overview

This Recommendation supports the splicing of MPEG-2 transport streams for the purpose of Digital Program Insertion, which includes insertion of advertisement and other content types. An in-stream messaging mechanism is defined to signal splicing and insertion opportunities. A splicing device is free to ignore splicing events signalled by the DPI cue message because the message is not a command to splice, but is an indicator of the presence of an ad avail. The taking of an avail is optional.

As shown in Figure I.1, DPI cue messages are received and acted upon, in the cable system headends by splicer and server devices, to affect the insertion of local advertisements by splicing the ad bit stream (typically containing the commercial content) into the bit stream of programming content. This Recommendation does not differentiate between a splicing device and a server, as does Recommendation J.sapi [1]. When this Recommendation uses the terms "splicer" or "splicing device" the meaning of the sentence may apply to a splicer/server combination as well. In actual practice it is common for ad servers (and not splicers) to parse, interpret and initiate action upon DPI cue messages. Since splicer and server devices can be combined into one, this appendix often uses the term server/splicer to denote a device, or set of devices, that together perform both functions. Figure I.1, a modified version of the diagram originally shown in DVS/209 [7], describes the overall functionality and interoperability associated with the headend systems that accomplish this.

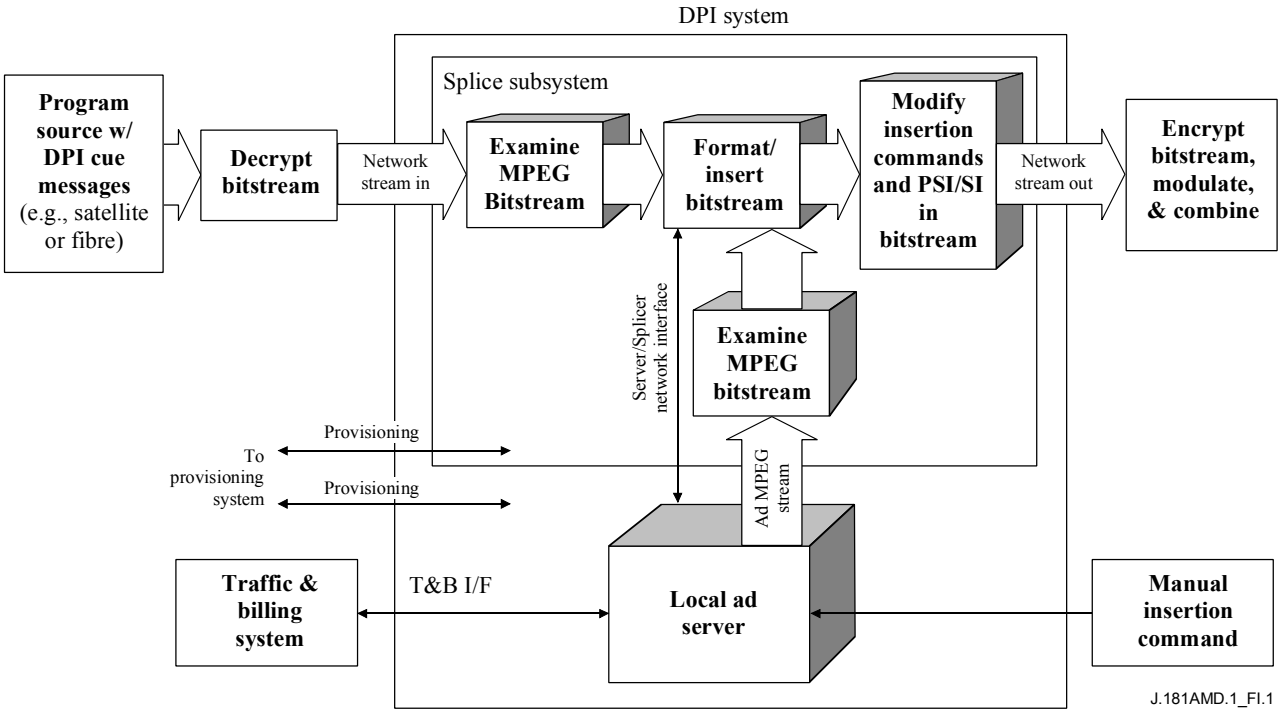

Figure I.1/J.181 – Headend system overview

In Figure I.1, a compliant MPEG-2 transport stream (either Multi Program Transport Stream or Single Program Transport Stream) is assumed for the network stream. No further constraints

beyond the inclusion of the defined cueing messages are placed upon the stream. It is expected that transport packet boundary splicing, as intended by ITU-T Rec. H.222.0 | ISO/IEC 13818-1 [2], and by SMPTE 312M [5], will not be suitable in cable plants due to the use of statistical multiplexing (VBR) and progressive refresh (no I-frames), both of which may require the removal of the transport layer.

This Recommendation specifies a technique for carrying notification of upcoming splice points in the transport stream. A splice information table is defined for notifying downstream devices of splice events, such as a network break, or return from a network break. The splice information table, which pertains to a given program, is carried in a separate PID referred to by that program's Program Map Table (PMT). In this way, splice event notification can pass through most transport stream remultiplexers without need for special processing. However, remultiplexers may need to obey certain constraints when they carry the DPI cue message. These constraints are addressed in this Recommendation and are elaborated on within this appendix.

This Recommendation does not address constraints on splicing devices and the splice_info_table syntax never suggests picture or splice quality. This Recommendation is not intended to guarantee seamless splicing.

I.4.1 Scope

The information within this appendix is intended as guideline information. Where this appendix contradicts the main portion of the Recommendation, the Recommendation shall take precedence.

I.4.2 Purpose

The purpose of this appendix is to aid splicing equipment designers, ad insertion equipment designers and purchasers and users of such equipment. Also expected to be interested are the networks that will originate DPI cue messages from their uplink sites and the manufacturers of the equipment to do this. This appendix is also expected to aid in the system integration of advertising-related equipment, both at the message origination end, and at the message reception end.

There may be crucial information within this appendix for manufacturers of equipment that pass the DPI cue message as part of the MPEG stream. An example of such equipment is a rate altering remultiplexer, which performs complex processing of the stream. When the stream is demultiplexed and processed and then remultiplexed, it is very important to place the DPI Cue Message in the proper position relative to the video service, and relative to nearby time base discontinuities. Such equipment may also be required to alter the message before retransmission.

I.5 Application guidelines

I.5.1 Practical boundaries for splice_time() in splice_insert()

How far ahead of the splice must a splice_insert message be sent, relative to the picture it refers to, in order to be safely responded to by an ad insertion system?

The "arm time" denotes the time a DPI cue message must precede that actual insertion. The arm time should be in the range of 5-8 seconds. This is in line with the pre-roll time for analog cue-tones. The arm time must not be so short that the avail passes by before the ad insertion system has time to respond. A minimum of 4 seconds is believed to be required for safe operation.

More thought about the possible consequences is required if it is desirable to extend the arm time beyond the recommended eight seconds. To specify a maximum arm time, therefore, seems premature.

The splice_info_section itself does not have any PTS or DTS associated with it. Therefore, it is not defined when the message should be decoded or when it should be presented, i.e., when it should

take effect. By choosing the minimum required arm time as 4 seconds, the problem of defining how this time should be measured can be avoided. Any reasonable measuring method will suffice.

I.5.2 Splice time accuracy

Although precise splice times can be specified using both the Program Splice Mode and Component Splice Mode, it is not required that the server/splicer performs splices at these precise times. This is true, especially when splicing at precisely the specified time would lead to a splice of lower quality than necessary. Instead, a server/splicer may use the splice times as guidelines.

For example, in the case of video, a splicer may insert a few black frames for GOP alignment or, instead of using a B-frame as an Out Point, it may choose a nearby I- or P-frame. In the case of audio, the video presentation time will not, in general, align with the audio presentation time(s) and, therefore, corresponding audio splices will be executed at times that are "close" to the nearest video presentation time. Thus, the intelligent choice of actual splice points will be one factor that differentiates splicers.

A cue message creator may be intelligent enough to never specify the presentation time of a B-frame as an Out Point from the network feed. Then there are very few good reasons for a splicer/server to change the splice_time specified for the start of a break. Similarly, a good practice by the message creator could be to choose an I-frame as the In Point at which the network feed should resume. However, each remote site may store local MPEG-2 ad files of imprecise length, structure and nature, e.g., when different encoder settings are used. Therefore, a message creator cannot ensure proper alignment between a given In Point and all the individual Out Points at the remote sites without additional constraints.

I.5.3 Splice_event_id usage and uniqueness

Cue messages can be created by several means: from information embedded in the original audio/video source, by uplink event trigger systems, or by headend event trigger systems. When all of these sources are combined within a service, there is potential for collisions in the splice_event_id value chosen for the cue message. The 32-bit splice_event_id should, therefore, be partitioned in the following way to allow each source a unique range of splice_event_id values:

Syntax	Bits	Type
Event_source	4	uimsbf
Event_number	28	uimsbf

Event_source – A user assigned number for the source of the cue message.

Event_number – A number chosen by the event source to identify an instance of the cue message.

The splice_event_id serves to identify a particular instance of an opportunity to change the multiplex. At least two distinct splice_events are required to perform one insertion, one to initiate the insertion and one to end it (unless the duration is included in the message signalling the splice-out).

Each event must have a unique splice_event_id. A splice_event_id cannot be reused before the splicing opportunity it describes (fully or in part) has completely passed. A splice_event_id is regarded to be in-use from the first cue message where it appears until about one second after the splice time that it is associated with. It may be reused for a new splice event immediately afterwards. It is not possible to use the same splice_event_id to signal an avail's start and stop if the stop is signalled before the start has been executed. However, it is possible to reuse the same splice_event_id if the first stop message is sent after the start has been executed.

The advertising industry uses additional information besides the splice_event_id to identify the actual material being played and the time of its insertion. This information includes: service name, time, ad-agency and spot number.

As long as the Event_source is unique for each point at which cue messages can be inserted in the following chain, see Figure I.2, the event identifiers will not collide:

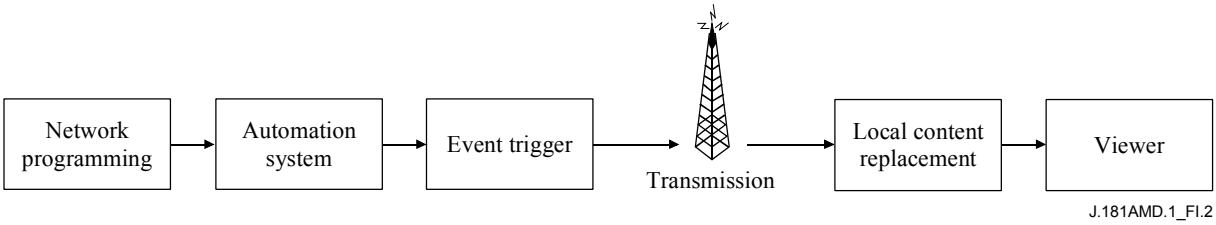

Figure I.2/J.181 – Cue message insertion points

The following values for Event_source are suggested:

Table I.1/J.181 – Event_source values

Source of cue message	Event_source value	Example splice_event_id ranges
Cue embedded in original source material	0	0x00000000, 0x00000001 .... 0x0fffffff
Cue created by automation system switching	4	0x40000000, 0x40000001 .... 0x4fffffff
Cue created by live event trigger system	6	0x60000000, 0x60000001 .... 0x6fffffff
Cue created by local content replacement system	12	0xC0000000, 0xC0000001 .... 0xCfffffff

I.5.4 Use of splice_schedule() command

The splice_schedule() command is intended for announcements of large schedules. It is not intended for the announcement of single events. Current implementations of DPI focus on the support of single events that do not rely on the use of the splice_schedule() command.

I.5.5 Component splice mode

The primary reason for the component splice mode is to allow the replacement of, or the passage of, individual elementary streams of any type in a manner consistent with the program content and the advertiser's intent.

For example, during a local commercial break, it might be useful to allow informational data (e.g., market returns) that is being streamed from the network provider to continue, although the viewer is being shown a locally inserted commercial. Conversely, with an advanced set-top box implementation, one might offer the viewer information to be downloaded for later review as part of the commercial. In this case, if the data content exceeds the duration of the commercial, the advertising data might be continued, although the program has resumed.

The methodology defined in this Recommendation allows some data types to be passed, while allowing others to be blocked, or replaced, during a break. The decision which to pass and which to block is made by the message inserter. A splicing device may choose to behave differently if it knows better, or is commanded to do so by an entity within the headend.

It is the intention of the standard, in both the Program Splice Mode and the Component Splice Mode, that the splicer should have a great deal of freedom in choosing the actual access units (both

video and audio) on which to splice. Some equipment will provide frame accurate splices and some may not.

The signalled splice_time may, or may not, be on an anchor frame or I-frame and the splicer may need to round off to a frame that makes sense for its capabilities.

There appears to be industry consensus that, in Program Splice Mode, it makes the most sense to give a presentation time of a video frame and then let the splicer choose whichever audio frame is closest or which makes sense for its methods. This was not, however, specified in the Recommendation. It is believed that the industry will come up with the best way to use the tools (especially if it differs from what is stated here).

In the Component Splice Mode, things work the same way. A single splice_time should be used (called the default splice_time). It is optional to give a splice_time for each component. There is some concern in the industry that the creator of a message will try to "help" the splicer by giving the exact time of all components and possibly complicate the job of the splicer (or the splicer may need to ignore the splice_time for some components).

It is understood that SMPTE 312M [5] requires a time-per-component to define the exact access unit for each component to be spliced. The cable industry is looking at a different usage for the time-per-component approach, however. This Recommendation allows individual components to start or end at very different times from the normal beginning or ending of the break. For instance, a "data" component such as a Java applet may have to be downloaded to the set-top box a number of seconds prior to the ad in order to support the ad. This is the primary usage of the unique splice_time for each component.

I.5.5.1 Erroneous component splice commands

When a splice message (in component mode) specifies invalid component tags, the splice should be executed for all correctly identified components. The intent is to let insertion succeed if possible, even if the complete chain of devices results in a violation of the standard. Several potential cases are given below.

In case a device removes one of several audio streams after the splice message has been inserted, but fails to update the splice messages (either because it is unaware of them or other reasons), the combination of stream and message at the server/splicer will be invalid. It is, however, still possible for the server/splicer to perform a valid insertion.

In case a device adds a component (e.g., a data channel), but fails to update the splice message (same possible reasons as above), it is still possible to perform a valid insertion.

Whenever a discrepancy between actual stream and splice message exists, a server/splicer should perform the insertion, in order to add error resilience to the chain.

I.5.6 Pre-roll functionality – Accomplishing a pre-roll function

A Pre-Roll function was necessary in the days of analog ad insertion to allow time for a tape machine to get up to speed by the beginning of the ad avail. So analog cue tones were often sent about 5 to 8 seconds prior to the avail (each network used its own chosen time and the consistency was varying). The early arrival of the cue tone has remained unchanged during the days of hybrid ad insertion where tape machines are replaced by MPEG servers and decoders, but the insertion is still analog. This early cue tone is useful for digital ad insertion as well. It gives the ad insertion equipment time to access its ad database, to determine which ad to play next, to start accessing its disk drives and to fill the MPEG decoder pipeline.

The splice_insert() command is constructed such that a pre-roll time can be used. It is also possible to repeat a splice_insert() command to increase the error resilience of the system.

The simplest variant is to send the splice_insert() command once about 8 seconds prior to an avail. This is similar to the analog insertion case.

An enhancement is to send the splice_insert() command 3 times: at about 8 sec, 6 sec and 4 sec prior to the avail. This increases the redundancy and prevents lost insertion opportunities. A lost avail nationwide is a very expensive error. In the case of multiple messages denoting a single avail, the splice_event_id field in all such messages must be identical. It is allowable that the splice time be different from message to message in the case where the first splice time is approximate and subsequent messages supply a more accurate splice time. Only the latest time is acted upon by the splicer.

I.5.7 Conditional access and cue encryption

I.5.7.1 What to encrypt

The format of the DPI section allows for the payload to be optionally encrypted. This includes all data from the "splice_command_type" through to the "E_CRC_32". This mechanism allows for any current, or future, commands to be protected. The reasons for protection can range from anti-piracy (e.g., commercial killers), malicious tampering of the data, and privacy when using cascaded ad insertion devices.

The specification requires that an encrypted CRC be present so that any receive device is able to verify that the encrypted data has not been changed since origination. This could be due to noise, but normally this type of corruption is detected by the standard CRC. The encrypted CRC has the primary purpose as a signature to detect that the receiver is authorized to receive the message. (i.e., that they have the correct control word). It is also used to detect willful modification of the encrypted data. This CRC is not present when the section is sent in-the-clear.

I.5.7.2 Operation in a cue insertion device

A cue insertion device is used to create splice_info_sections and insert them into a transport stream. When encryption is desired, the cue inserter uses a fixed key entered by an operator, plus the algorithm selected, to encrypt the section before being transmitted.

The cue insertion device should be able to maintain multiple simultaneous keys for the same program. The reason is that each ad inserter in a cascade could require one or more different keys to decrypt the splice_info_section.

A cue insertion device is only required to implement one of the standard encryption methods. In recommended practice, the cue insertion device should implement all standard algorithms.

I.5.7.3 Operation in an ad insertion device

An ad insertion device consumes splice_info_sections. When a splice_info_section is encrypted, the ad inserter will only act upon the section if the decryption is successful. This is determined by checking the encrypted CRC before interpreting the data. A failure will usually mean that the device is simply not authorized to interpret the information that the section contains.

The ad inserter is expected to allow the encrypted splice_info_section to pass through it to the next device. This is true whether the decryption was successful or not. The ad inserter never releases the contents of a decrypted section for security reasons.

Once the section has been decrypted, the device may act on the information it contains, or discard it. This is the same operation that would be performed if the section had been sent in-the-clear.

The ad inserter device should be able to hold 256 different decryption keys. The cw_index field in the section indicates which of these decryption keys should be used. The method of key interchange is outside the scope of this appendix.

For normal operation, it is expected that one or two keys are used for any one pair of send/receive devices. There is provision for a large number of keys to allow an ad inserter to connect to multiple programs in a transport stream, or to allow cascaded ad inserters to use different ranges of cw_indexes when connecting to the same program.

An Ad Insertion device should implement all defined decryption algorithms in this Recommendation. This allows it to receive encrypted sections from any cue insertion device. Any cue insertion device may choose any standard algorithm to protect the section, and the Ad Inserter must be able to decrypt any message it is authorized to receive.

I.5.7.4 Theory of operation

I.5.7.4.1 Encryption versus scrambling

The DPI WG is exploring the issue of scrambling the Digital Program Insertion Stream. In this context, scrambling means the use of the "standard" DES or DVB Common Scrambling algorithm that is being used for video and audio services.

To resolve this issue, one must consider that the Ad Insertion functionality is typically not located in a settop box. There has been provision made for this model, but it is not the primary model. Most systems will employ an Ad Insertion device in a digital headend (or some other distribution point). These standalone devices are typically computers. Using a descrambler would require a custom circuit designed to descramble the stream. Also, since the entire transport packet is scrambled, the ad insertion device has no access to the header data as it does with the current model. Some of the information, such as pts_adjustment, may need to be modified even when the section is not descrambled. There are similar considerations when creating the stream.

The model used for the splice_info_section is closer to the ECM model, than the Elementary Stream model. An ECM section is independently encrypted (and decrypted), and authorized by some external mechanism. In the case of the ECM, it is usually the EMM that controls authorization. In the case of the splice_info_section, it is a manual authorization.

The fixed key model was chosen for simplicity. The key distribution was left undefined and may use any mechanism that gets the key to the decryptor in a secure and timely manner. It is conceivable that a committee may standardize on some type of ECM and/or EMM to distribute the keys in the transport stream.

I.5.7.4.2 Standard encryption methods

The standard provides for three types of algorithms. All three algorithms use the DES decryptor as the basic building block. For Triple DES, the DES encryptor is also required. Software implementations of the DES algorithm are readily available.

I.5.7.4.2.1 Section stuffing

All DES type algorithms (block encryptors rather than stream encryptors) require the length of the data to be an exact multiple of 8 bytes. To achieve this, stuffing is often required. The alignment_stuffing field may be present whether the section is encrypted or not. The number of bytes of stuffing can be determined during the interpretation of the data. The stuffing bytes are never used, so they may take on any value.

To determine the length of the stuffing, one uses the fact that the total length is known from the section_length field. We also know the exact size of the header, to determine the start of a splice_command. Every command has a size completely determined by the syntax within the command itself. We therefore know that the length is:

<section_length> +3 – <end_of_command> – <length_trailing>

where <length_trailing> is four for non-encrypted sections and eight for encrypted sections.

Algorithmically, it is easier to simply ignore the stuffing. Work forward to decrypt and interpret the command, and work backward from the end to find the CRCs.

I.5.7.4.2.2 DES Electronic Codebook (ECB) algorithm

DES ECB requires a 56-bit key plus 8 parity bits to make up a complete 64-bit key, which is then used by the algorithm to encrypt or decrypt a stream. The DES algorithm is symmetric. The same key is used in both the encryptor and the decryptor. The actual encryption and decryption algorithms are slightly different to allow this symmetry to work.

It is suggested that the full 64-bit key be distributed between the two devices. The key will be entered as a 16-digit hexadecimal number. For example, 0x123456789ABCDEF0 would be distributed. In engineering notation, the leftmost digit is the most significant digit, and the MSB of this nibble is the most significant bit of the key (bit 63). Cipher algorithms generally use FIPS notation to represent keys. In this case, the leftmost bit of the key is numbered Bit 1, through to bit 64 on the right. Therefore, bit 63 of the distributed key value will be loaded into Bit 1 of the initial key register. Similarly, bit 0 of the key value is loaded into Bit 64 of the key register.

The Electronic Codebook method of encryption uses the *same* key for every 8-byte block of the original message. Figure I.3 gives an example of a simple 3-block message being encrypted. The arrows represent operations. Across the top, the key is being loaded into the key register. Side to side, the data is being shifted into the encryption register, and across the bottom, the DES algorithm is being applied.

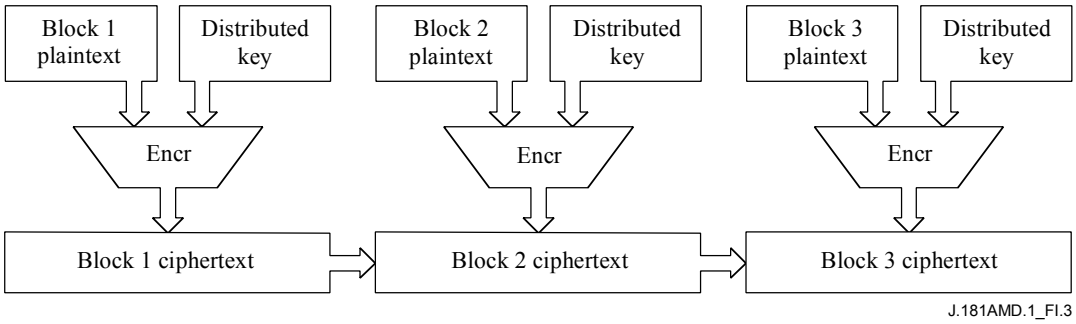

Figure I.3/J.181 – DES ECB example

I.5.7.4.2.3 DES Cipherblock Chaining (CBC) algorithm

DES CBC requires a 56-bit key plus 8 parity bits to make up a complete 64-bit key, which is used by to the algorithm to encrypt or decrypt a stream. The same bit ordering rules and key distribution methods can be used for CBC that were used for the ECB algorithm described in I.5.7.4.2.2.

The Cipherblock Chaining method of encryption uses a different key for every 8-byte block of the original message.

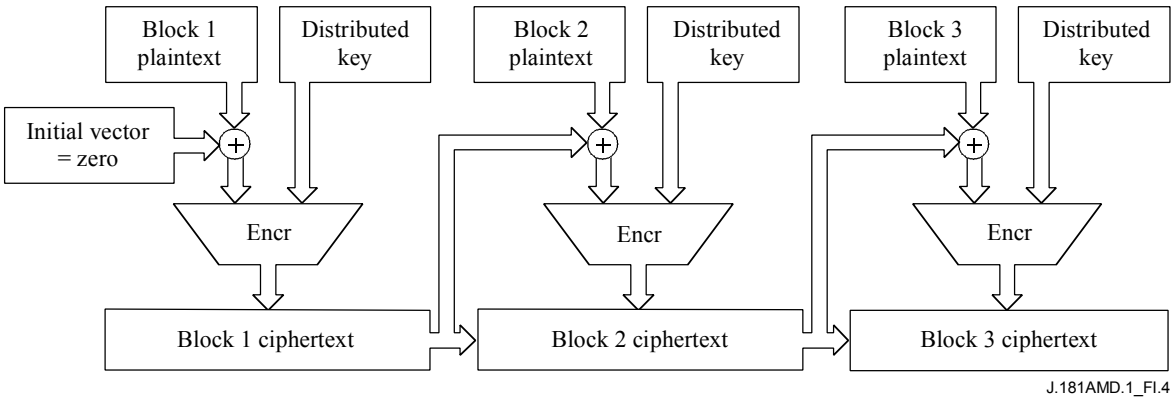

Figure I.4/J.181 – DES CBC encryption example

From Figure I.4, we can see that the plaintext data is modified by the result of the previous encryption block. For the first block, we need to have an "initial vector" to apply to the exclusive-or block. It turns out that the initial vector provides no extra security, whether it is known or not. So, for this system, the initial vector can be zero, which removes the need for it to be distributed with the key.

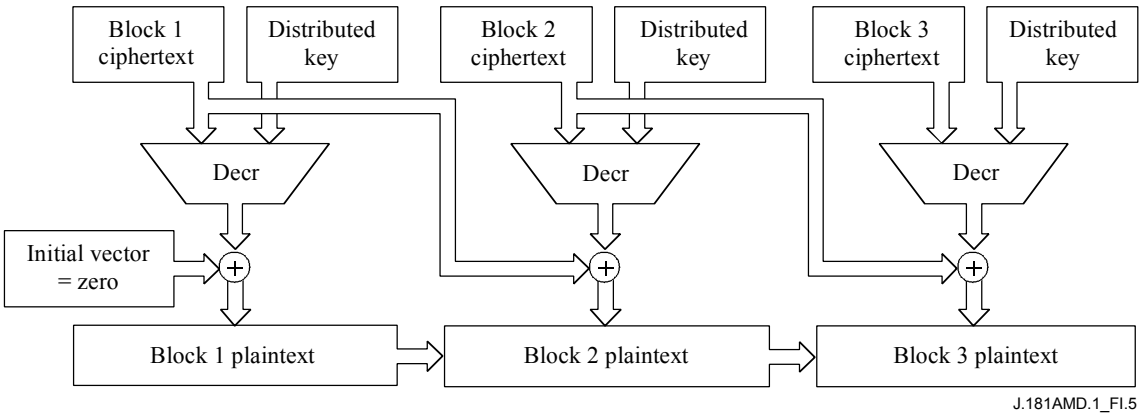

Figure I.5/J.181 – DES CBC decryption example

From Figure I.5 we can see the DES CBC decryption is slightly different from encryption. The exclusive-or block requires the previous block of encrypted data to arrive at the proper cipher key for the DES algorithm. In the decryptor, this encrypted block is derived from the transmitted data.

I.5.7.4.2.4 Triple-DES ECB Mode (EDE method) algorithm

Triple DES uses the standard DES algorithm, but applies the algorithm 3 times for each block of input data. This provides a significant security enhancement. The disadvantage is that one must distribute a 192-bit key (actually three 64-bit keys).

Using the combination of two algorithms and three passes (encryption and decryption are different), it is possible to generate eight different Triple DES variants[1]. The variants are designated by a three-letter code. Of these variants, the most common is selected for use in a splicing system. This variant is designated the EDE method, referring to the fact that pass one uses encryption, pass two

[1] Triple DES also has two key variants. This can be simulated by making KEY A == KEY C.

uses decryption, and pass three uses the encryption algorithm again. For simplicity, the DES ECB mode of operation is used for the Triple-DES algorithm[2].

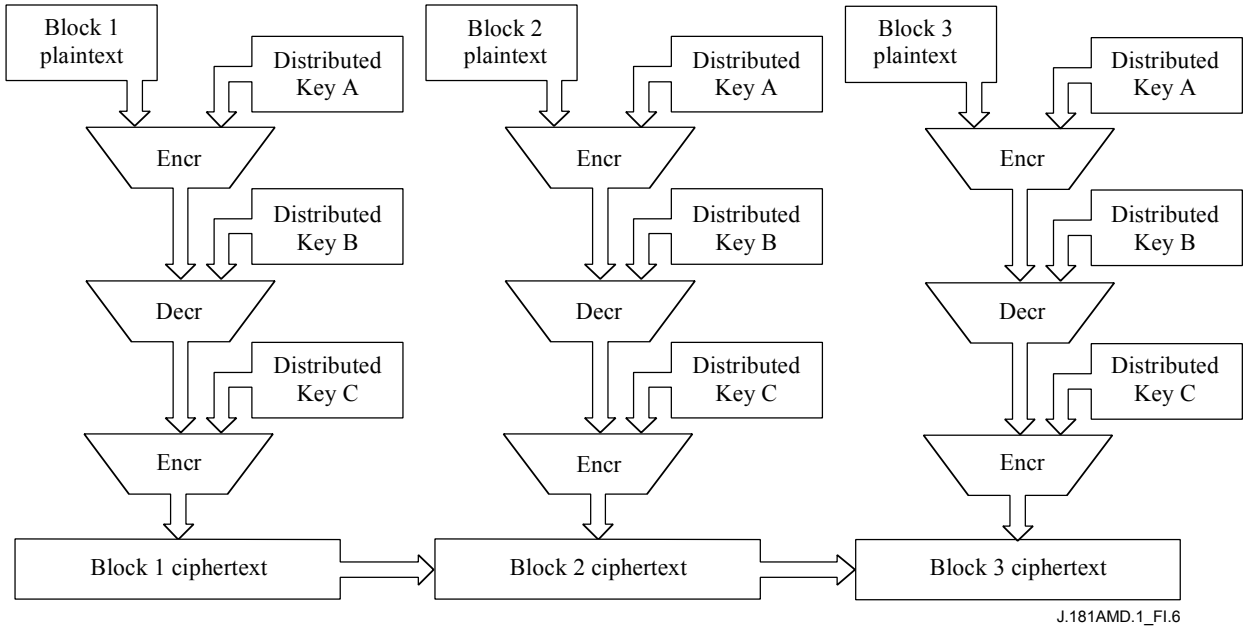

Figure I.6/J.181 – Triple-DES ECB encryption example

Triple-DES decryption requires that the algorithm be applied in the reverse of the encryption algorithm. As a result, the block diagram for decryption is slightly different again. It can be found in Figure I.7.

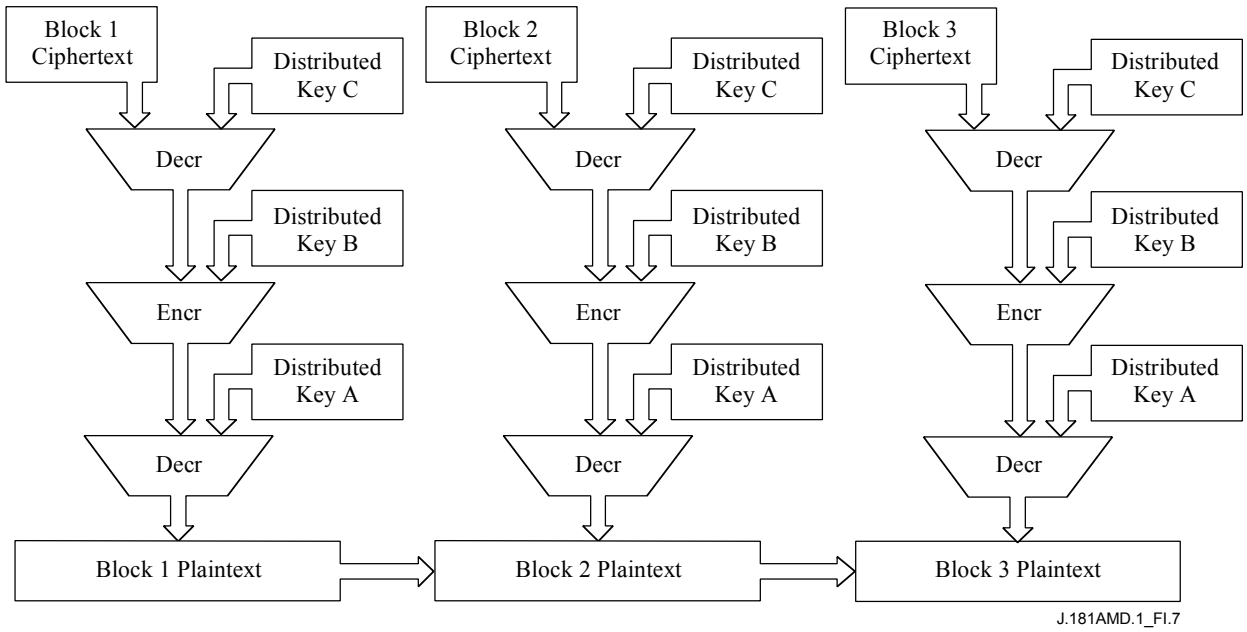

Figure I.7/J.181 – Triple-DES ECB decryption example

[2] Triple DES can also use the cipherblock chaining mode, but this mode is not one of the standard methods.

Key distribution for triple-DES is in the form of three 64-bit keys. Each of the keys is ordered the same way that the single key is ordered for single DES methods. Each of the keys is labeled A, B and C for reference. Using the block diagrams, we can see that key A is applied first, then B and C, during the encryption process. During decryption, the keys must be applied in the reverse order.

I.5.7.4.3 Private encryption methods

Approximately half of the encryption_algorithm values have been reserved for private use. These values will never be allocated by the Recommendation. Private data of this nature is prone to misinterpretation, due to its very nature. As a result, it is impossible to standardize a method of selecting user private algorithms without some type of registration authority. Since there is no registration authority, the method of coordinating algorithms has been left undefined.

The problem arises when two independent entities use the same encryption_algorithm value for different algorithms. When a cue insertion device from entity A encrypts the section, and that section is received by entity B, the decryption will not work. The equipment believes it is not authorized because of the CRC failure, even though the same key has been entered at both ends.

I.5.8 Usage of unique_program_id

I.5.8.1 What is a "Program"?

Network content is typically divided into a series of programs. A program may be of almost any duration. Most programs have a duration of thirty or sixty minutes. However, some programs may be only a few minutes in length, such as news reports or sports updates, while others may be several hours long, such as movies, sporting events, or special live awards programming.

I.5.8.2 What is a "program_id"?

A program_id is a "shorthand" way of identifying a specific program on a network. It is provided by the network.

I.5.8.3 Why should programs be identified and differentiated?

Some programs are of greater value to an advertiser than other programs. The relative value of the program is unrelated to its length or its scheduled time; an advertiser values a program on the basis of its ability to attract a particular audience.

While some advertisers may purchase commercial time on the basis of day parts (i.e., 10 A.M.-4 P.M., 8 P.M.-11 P.M., etc.) other advertisers specify that their commercials must run in a specific program or even a specific break within a program. The programming may be scheduled regularly (such as news or episodic shows), or it may be special, one-time programming, such as baseball games, awards programs or live interviews.

Frequently, special programs are purchased at rates in excess of surrounding, regular programs. Advertisers that have agreed to pay these higher rates expect their commercial messages to run within the special programming that they bought. They will not pay for commercials that run outside of these programs.

I.5.8.4 Why does the time at which a program is scheduled not identify it?

Networks may change the program schedule at the last minute. Sometimes these changes are communicated to the affiliates before the schedules of the commercials are released to the headends; occasionally, this information is not available until after program and commercial content have been scheduled.

The changes to program schedules most frequently are the result of last minute unplanned changes in live programming such as sporting events, awards shows and live news event coverage. These programs may be scheduled at the last minute, may run longer or shorter than originally anticipated, or may be cancelled as the result of weather or other conditions.

If scheduled on the basis of "time" alone, an advertiser who was scheduled to run in a special program may instead run in "regular" programming. In this event, the advertiser will not pay for the commercial message. In a worst-case scenario, a local affiliate may inadvertently bill the advertiser for the commercial as though it did run in the special programming, even though it did not. This may result in serious repercussions with regard to the relationship with that client.

I.5.8.5 How will a unique_program_id alleviate problems?

If a splice-event is identified by a unique_program_id, vendors of sales/traffic systems (those computer systems that schedule commercials specified by the ad sales department for their clients) will be able to specify with each commercial the program within which it was intended to be played. They will also be able to specify commercials that can be used to substitute (at no loss in revenue) those commercials if the program is not broadcast at the time anticipated.

By providing this information, the vendors of the ad insertion systems will be able to substitute an appropriate advertising message if the unique_program_id of the splice event does not match the unique_program_id of the commercial message scheduled for that time period. The addition of this field and the implementation of the functionality will require some effort on the part of ad insertion system vendors. This capability was never a part of older generation analog systems, but the functionality is extremely important.

I.5.9 Avail fields usage

I.5.9.1 What is an avail?

An avail is an opportunity provided by the network to a local affiliate to insert a commercial event into a program. This start of an avail is indicated as a splice event in the programming stream. The duration of the avail may vary from a few seconds to several minutes.

Usually, avails are sixty seconds long, although avails of ninety seconds or two minutes are not uncommon. Since most commercial messages are thirty seconds in length, two or more of these are normally inserted into each avail.

I.5.9.2 How many avails occur within a program?

A program's length is the most common predictor of the avails that will occur within it. Usually, there is one sixty-second avail within any thirty-minute program. The exact number of avails is known from the program schedule provided by the network. While the number and length of the avails may vary, based on the program, the program schedule allows the local advertising affiliate to know in advance how many avails to expect within a program.

I.5.9.3 Why is it important to identify the avails within a program?

In the simplest of all instances, an individual advertiser may have purchased a specific "position" within a program; that is, it may be important for that message to run in the first, or the third, or the sixth position within that program. An advertiser might specify (and pay a premium for) the "last avail" within a basketball game, on the assumption that in a close game, viewers will be less likely to tune away.

In other instances, it may be that a variety of advertising sources (local or national/regional interconnect) have arranged to divide the avails within a program on a predetermined fashion between them. In this case, it is important for each advertising entity to know which avail is associated with a splice event so that they can accurately keep their insertion schedules "in-sync" with the program.

I.5.9.4 How does the "avail" field provide for this?

If each avail within a program is uniquely identified, then sales and traffic systems, as well as ad insertion equipment vendors, can associate the avail identification with those commercials that are required to run in a specific avail.

If an avail splice event is "missed" for any reason (whether a failure by the network or the local advertising affiliate), the subsequent splice event, with its specific avail field id, will allow the commercial insertions to be resynchronized to the correct event.

I.5.9.5 What does the avail_count field do?

This field provides a count of the total number of avails that are anticipated within the complete program. Providing this information allows the ad insertion device to guarantee that its anticipated count of the number of avails matches with the count being executed by the network.

The value in the avail field could be larger than the value in the avail_count field. This would occur, for instance, if a sporting event ran longer than its allocated duration. If the network content provider sent cue messages during this "over-run" period, the avail field would have a number greater than that in the avail_count field.

I.5.10 Cueing usage

The use of splice commands to update, modify, cancel and terminate events is clarified in the clauses below.

I.5.10.1 Starting a break

A DPI cue message can be used to indicate to the server/splicer combination an opportunity to either splice out of the network into an ad, or splice into network, out of an ad. The out_of_network indicator in the "splice insert" command is set to "1" when exiting the network, and "0" when entering the network. The splice point is derived from the splice_time() structure in splice_info_section(). When splicing out of network, the DPI cue message must reach the server/splicer combination at least four seconds prior to the splice time. This can be referred to as having a four second "pre-roll". The server/splicer combination decides on the best splice point related to the splice time.

A cue message that indicates a splice out of network, and that is still outside the specified "pre-roll" window, can be cancelled. A "cancel" is issued by sending a DPI cue message with the cancel_event_indicator flag set to "1". If a "cancel" is received during the "pre-roll" or after the break has started, it is ignored.

A splice out of network event can be updated by an update message without the need for a cancellation of a previously transmitted message for the same splice event. An event can be updated several times. The latest message that adheres to the timing constraints given in terms of the "pre-roll" should be considered valid by the splicer, superseding all earlier messages for the same splice event.

If the need to stop a break arises, and the timing is uncertain, sending a "cancel" and a "stop" or "terminate" is acceptable. The splicer should react to whichever of the two messages is valid and ignore the other. The "stop" and "terminate" are described in clause I.5.10.2.

I.5.10.2 Ending a Break

There are two ways to end a break that is under way (in a playing state). They are described below. If the process of switching back to network is already in progress, or finished, or if the break is not in "pre-roll" yet, then the "terminate" and/or "stop" should be ignored by the server/splicer combination.

- A "terminate" command can be issued by sending a message with the splice_immediate flag set to "1" and the out_of_network_indicator flag set to "0" in splice_info_section(). The server/splicer combination switches back to network immediately (as quickly as possible).
- A "stop" can be issued by sending a message with the splice_immediate flag set to "0" and the out_of_network_indicator flag set to "0" in splice_info_section(). The splice_time() is used as the point in which to splice back to the network.

I.5.10.3 Spot sharing within a break

A DPI cue message can be used to indicate an opportunity to insert multiple ads to several servers. In this scenario, the splicer receives the DPI cue message in the network stream, and passes it to each of the servers. The servers are responsible for the scheduling and arbitration. Each server should handle the DPI cue message in such a manner that it takes the proper spot position within the break. The end result is a split break, where one server may take the first spot position in the break, and the second server takes the next spot position, or some other combination of the two servers taking spot positions within the break.

I.5.11 Creation and usage of private splice descriptors

I.5.11.1 What are descriptors

Descriptor is a term from the MPEG-2 standard. A descriptor is used to introduce new syntax into an existing standard. It does so in a way that allows existing equipment to skip the new syntax. A descriptor may optionally be included inside a section of information. Since a descriptor has a known format, it may be skipped inside receiving equipment without causing a loss of synchronization during the parsing process.

Another use of descriptors is to provide optional syntax. Like the new syntax, any existing equipment that does not understand the optional information is able to skip the data.

I.5.11.1.1 The problem

Standard descriptors have a problem when users create private descriptors for their own use. The problem is that different users can use the same number for a descriptor tag, but have a completely different syntax in the payload of the descriptor. This is not a problem if the receive equipment does not understand the descriptor, but it is a problem when it tries to understand the wrong descriptor.

MPEG and DVB have solved this problem introducing a descriptor specifically for solving this problem. MPEG never formally described how to use the descriptor, but DVB did. Basically, the private_data_descriptor has a 32-bit identifier that changes the "mode of understanding" in the receiver. When the device sees an identifier that it understands, it will start accepting user private descriptor tags. When it sees an identifier it does not understand, or before it sees any identifier, it stops accepting any user private tags.

This method works, but it suffers from a flaw in that the private_data_descriptor was made optional originally. So, many companies made devices that did not send (or receive) private data and the original problem that should have been solved is not being used consistently. Also, it is not used at all in MPEG since there was no rule enforced for the descriptor.

I.5.11.1.2 The solution

The 32-bit identifier found in the splice_info_descriptors serves the same purpose. But, by having it inside the descriptor header, it is made mandatory. The disadvantage is that it triples the size of the header. Splice_info_sections are expected to be quite small, and to fit easily inside a single transport packet, so the extra header bytes should not be a detriment.

I.5.11.2 Registration

The identifier in the descriptor header is a self-registered value. The assumption is that with 2^{32} combinations, no two companies should choose the same value at random. Any company that wishes to use a private descriptor will create a 32-bit value for its own use. A simple "random" mechanism is to use the 32-bit number as a 4 character string. Then, use an abbreviation of a company name to derive the 4 characters. For example, the standard registration_descriptor identifier is 0x43554549 (ASCII "CUEI").

Of course, if all companies creating private descriptors use the same identifier (like 0x00000001), then the original problem will still exist.

Once an identifier has been chosen, then the descriptor tag is available for use. This gives any company up to 256 private descriptors for their own use, before they need to create a new identifier.

I.5.11.3 Creating compatible private descriptors

Assuming that an identifier has been chosen, a private descriptor can be defined. All descriptors, whether they are private or new descriptors that become part of the standard, have the same six bytes in the header. Table I.2 represents the outline for all descriptors contained within a splice_info_section. This includes all current and future descriptors that are defined by the Recommendation.

Table I.2/J.181 – Splice_descriptor()

Syntax	Bits	Description
splice_descriptor() {		
splice_descriptor_tag	**8**	**uimsbf**
descriptor_length	**8**	**uimsbf**
identifier	**32**	**uimsbf**
for(i=0; i<N; i++) {		
private_byte	**8**	**uimsbf**
}		
}		

This splice_descriptor is not a real descriptor – it is a template for all real descriptors. Within the standard, there is only one real descriptor that has been defined. That is the avail_descriptor, and it has been reproduced in Table I.3 as an example of how to create a new descriptor.

Table I.3/J.181 – Avail_descriptor()

Syntax	Bits	Description
avail_descriptor() {		
splice_descriptor_tag	**8**	**uimsbf**
descriptor_length	**8**	**uimsbf**
identifier	**32**	**uimsbf**
provider_avail_id	**32**	**uimsbf**
}		

Step 1: Choose an identifier. In this case, the identifier is "CUEI".

Step 2: Choose a splice_descriptor_tag. Since this was the first descriptor defined, a value of zero was chosen.

Step 3: Define the private syntax. This is the body of the descriptor. In the template, the payload of the descriptor is represented as a generic loop containing bytes of data. To any device that does not understand the descriptor, this is exactly how it looks. By using the descriptor_length field, the device is able to skip the payload and continuing processing the next descriptor.

In this example, the payload consists of exactly one field, a field that contains 32 bits.

The only restriction within the payload of a descriptor is that the payload byte is an exact number of bytes long (a multiple of 8-bits), and that it comprises less than 250 bytes total[3]. Fields in the payload of a descriptor may contain any number of bits. If the syntax that is required does not have a multiple of eight bits, then reserved bits must be inserted to pad the total to a multiple of eight.

I.5.11.4 Using the avail_descriptor

The provider_avail_id field is a 32-bit uimsbf value that may be utilized in multiple ways. The current analog DTMF systems use 4 characters for cue messages. They have a three-digit code to identify the break. The fourth character is usually an '*' for a start tone indicating a prespecified preroll duration or a '#' for an immediate stop tone. An example would be "635*".

Current analog systems use these messages to indicate different types of cue information such as duration, time of hour, breaking news or private break.

It is recommended to utilize the other features of this Recommendation that identify the start vs. stop nature of the cue message and then to simply insert the identification code as a 32-bit integer number. For example, if a network used the character sequence "017*" as the ID for an analog cue tone, you would insert the value 17 into the provider_avail_id field and set the out_of_network_indicator bit in the cue message. The end message should use the same provider_avail_id as the start message.

I.5.12 Handling time base discontinuities

It is stated in I.6.1.1 that a cue message that is sent just before a Time Base Discontinuity (TBD), is not allowed to carry a splice_time expressed in the new time base that follows after the TBD. One reason for this requirement is that it would otherwise be very difficult, in a remultiplexing operation, to preserve the validity of the splice_time field. Without this requirement, a remux or restamping device could find itself in a position where it needs the value of a PTS which will not arrive until hundreds of milliseconds later.

It was, therefore, decided that the splice_time field should always be expressed in the time base currently in effect, i.e., where the cue-message packet containing this field is inserted in the stream. Consequently, the cue-message inserter must behave as if, even in the rare cases when it really knows otherwise, there will not be a TBD during the "arm time". The "arm time" is the time between the first cue message referring to an event, and the event itself.

The splice_time field might need to point to a picture that is located after the TBD. That picture will be associated with a PTS expressed in the new time base. The cue-message is not allowed to use that PTS value. Instead, it must use the PTS value that the picture would have been associated with if the TBD did not exist, or had been removed, e.g., by restamping of all time stamps following the TBD to the old time base.

This simplifies the work of a simple remux which passes the input timing through to its output (no restamping of PTS and PCR and hence no removal of time base discontinuities present on its input). Simple remultiplexers are responsible for preserving the validity of the splice time carried in the cue-message that they pass. It is, therefore, up to a simple remultiplexer to guarantee that a cue message is not allowed to cross over a TBD boundary, since this would destroy the validity of the splice time in the cue message.

---

[3] The convention in MPEG, DVB and ATSC is that the total length of a descriptor be 256 bytes, so the descriptor length is limited to make this true.

Remultiplexers that continuously restamp the PCR and PTS in their output, and thereby automatically remove time base discontinuities, must also preserve the cue message splice_time validity. They must be aware of which side of an arriving TBD the cue message is on in order to properly map the cue message into the device's output (and restamp the splice_time field accordingly).

For both simple and restamping remultiplexers, it is strongly recommended that the arm time is never decreased while processing the stream and embedded cue messages (i.e., never add delay to the cue message relative to the System Time Clock in the stream). Decreased arm time for a cue message that already is at the lower limit of arm time recommendations or requirements, might result in a violation of these recommendations or requirements.

The splice_time field is carried in the possibly encrypted part of the cue-message, making it potentially inaccessible to a restamping device. If encryption is employed, the restamping device must instead update the pts_adjustment field, which is provided in the non-encrypted part of the syntax for exactly this purpose. When encryption is not employed, the restamping device may update the splice_time field directly, or may choose to modify the pts_adjustment field.

When there is a TBD during the arm-time, a device receiving and acting upon the cue-message, e.g., a splicer/server, is responsible for the translation between the old and the new time base. That means translation of the splice_time carried in the cue-message (adjusted according to the value of the pts_adjustment field) to the new time base to obtain the PTS of the picture intended to splice at.

I.5.13 Cascaded splicing devices

This Recommendation allows for cascaded splicer/servers since a splicer is just a form of a remultiplexer. Figure I.8 illustrates the scenario where two splicer/server combinations are cascaded to perform insertion. The outer boundaries of the system are the same as in the single-splicer scenario. No direct communication between the two servers is required.

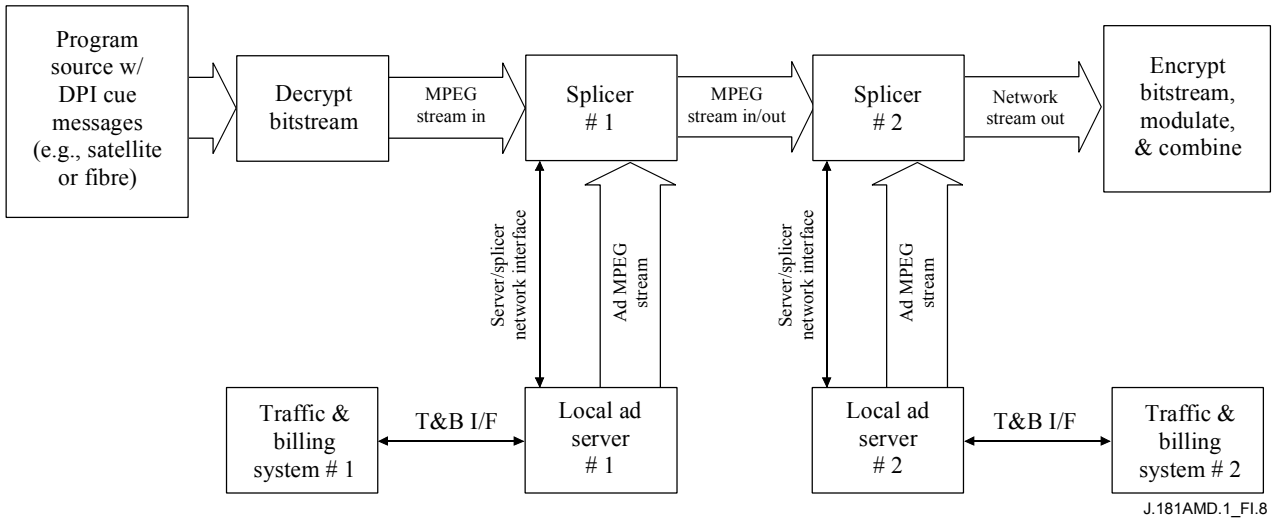

Figure I.8/J.181 – Cascading of splicer/server devices

There are a few issues to be considered when cascading devices. These are addressed below.

I.5.13.1 Restamping cue messages

Splicers may handle their output timing (PCR, PTS) in several ways. If a splicer always passes the time domain of the network feed thru to its output, even when an advertisement is replacing the network feed, it does not need to do any restamping of a passed-on cue message. If, however, a splicer always, or momentarily, restamps PCRs and PTSs (at the moment of a cue message), the

splicer is required to make the passed-on cue message once again truthful. This Recommendation has an unencrypted field, called pts_adjustment, that allows the cue packet to be effectively restamped by altering this adjustment field rather then having to access the pts_time() variable(s) and restamping each individually. This was originally included for encrypted operation where an intermediate multiplexer/splicer would not be able to decrypt the message, but it would be restamping cue packets. This field, though, makes it simpler for intermediate muxes in all cases. The splicer has to add the pts_adjustment field to all pts_time() fields to get the actual splice time(s).

I.5.13.2 Cue propagation

Splicers should have some form of configuration variable to allow for passing a time corrected cue message, and they should also have the option to block the cue message from further propagation. It is up to the network and the operator to decide how far a cue message may propagate. The cue message could go all the way to the set-top since the cable plant will most likely encrypt it. The set-top would then be able to decrypt it as part of a closed system and this should prevent some forms of commercial killers.

I.5.13.3 Delay

Splicers typically cause some amount of delay in the MPEG stream. This is more of a concern when splicers are cascaded since delays will accumulate. The only applications that would be very dependent on delay would be simulcast operations and time-critical variables (Real-time stock quotes). It is thought that the simulcast operations will not really be needed with high quality audio now in the system. Current data delivered over the broadcast TV mechanism are typically time-delayed already, and that fact is usually stated clearly on the screen with the data. In either condition, current splicers typically add one to two seconds of delay and this does not seem to concern the cable operators at this point.

I.5.13.4 Logical cascading

The potential need for cascaded splicers can be taken care of by the Ad-Server/Splicer API [1]. Using this API, one physical splicer can handle multiple ad servers doing Local/Regional/National advertising. This logical cascading eliminates the need for having three physically cascaded splicers with their additional delay.

I.6 Additional information

I.6.1 Considerations for evaluation of MPEG-2 splicing devices

I.6.1.1 Overview

The purpose of this clause is to provide a means for understanding the performance of MPEG-2 splicing products with respect to current MPEG-2 standards. This is an informational clause intended solely for clarification.

MPEG-2 splicers represent an emerging technology whose detailed capabilities and limitations tend to be quite dependent on their operating environment and are not generally well understood. In order to specify what capabilities a splicer should have, and what level of performance it should achieve, it is necessary to know what factors typically affect splicer performance and identify the context in which a splicer is intended to operate. For example, it is meaningless to discuss issues such as frame accuracy without clearly identifying the conditions under which the accuracy is to be measured. It is also useful to know what general technological approach is used by a splicer, since this can provide insight into its general capabilities and limitations.

I.6.1.2 Splicer technology

Currently, there are several approaches to MPEG-2 splicing, each with its own advantages and disadvantages. The following is an attempt to categorize these approaches.

I.6.1.2.1 Transport Stream (TS) splicing

TS splicers operate at the MPEG-2 transport stream level and simply switch from one transport packet stream to another. A TS splicer will typically perform PID remapping, but will not modify the VBV state of the stream associated with PTS/DTS restamping.

A TS splicer does not have knowledge of the state of the elementary streams it is splicing. In order to produce a result that maintains decoder integrity at all times, a TS splicer must operate on streams that have been conditioned to ensure that the splice points chosen meet certain requirements (e.g., by adhering to SMPTE 312M [5]).

Because they operate only on TS data, and assume that the stream has been properly conditioned (and that they are only instructed to splice at valid points), TS splicers are the simplest form of splicer to implement.

I.6.1.2.2 Elementary Stream (ES) splicing

PES splicers operate at the MPEG-2 elementary stream level and modify the elementary stream data as necessary to perform a splice. This enables them to modify the VBV state of video streams as necessary, and to properly handle situations such as splicing 3:2 pull down material. It also enables them to mute audio streams at splice points to avoid the popping typically associated with hard edits.

Because they have the ability to modify elementary stream data on the fly, PES splicers do not have the content restrictions that TS splicers have in order to achieve the same level of performance.

I.6.1.2.3 Picture level splicing

I.6.1.2.3.1 Partial recoding

Picture level splicers move a level deeper than PES splicers by operating on the picture data contained within the video elementary stream. Picture level data is not decompressed and recompressed, but certain other operations, such as requantization, can be performed to manage the bit rate and VBV state of the stream. This type of splicer is more complex than a PES splicer but can outperform it in certain situations because of the finer degree of bit stream control it possesses. Additionally, special consideration must be given to avoiding picture quality loss.

I.6.1.2.3.2 Complete recoding

Recoding splicers actually decode the MPEG data, perform the splice in baseband, and re-encode the result. These splicers embody an MPEG decoder and encoder per channel and are the most expensive to implement. Additionally, special consideration must be given to avoiding picture quality loss.

I.6.1.2.3.3 I-frame generation

I-frame generation is not a type of splicer itself, as much as it is a possible splicer feature. Splicers with this capability can produce I-frames at any point in the stream. TS splicers certainly do not have this capability, whereas recoders certainly do. Both PES and picture level splicers may have this capability.

I.6.1.3 Environment

In order to characterize splicer behaviour, it is necessary to identify the environment(s) in which a splicer is expected to operate. Different splicer architectures and implementations will have different levels of performance in different situations.

In the context of this discussion, the environment describes the type of MPEG-2 program content that the splicer is expected to process, and how this content is delivered to, and produced by, the splicer.

Although the mechanism used to control the splicer can have a direct effect on its performance in certain areas (frame accuracy), control issues have been left outside the scope of this discussion.

I.6.1.3.1 Video elementary stream

There are several issues relating to the composition of a video elementary stream that can have a significant impact on splicer behaviour.

I.6.1.3.1.1 Hierarchical subsets of MPEG-2 streams – Profile @ level

The profile and level of a video elementary stream determine how different types of splicers behave. Because they do not operate on elementary streams, TS splicers do not care what profile or level stream is being spliced. Other types of splicers may care to varying degrees. For example, a PES splicer may scale from MP@ML to MP@HL without significant hardware impact, whereas a picture-level or recoding splicer may be substantially more expensive.

The profile/level combinations that are likely to be of most interest are: MP@ML (Main Profile @ Main Level), 422P@ML (4:2:2 Profile at Main Level), and MP@HL (Main Profile @ High Level).

It is worth noting that the spatial resolution may change between clips being spliced together and, while this should not be a problem for good splicer implementations, decoders may not be able to properly handle the switch, resulting in a non-seamless splice.

I.6.1.3.1.2 Data stream structure – GOP structure

There are three issues relating to GOP structure that have an impact on splicer behaviour. The first is whether the GOP is open or closed. An open GOP can be problematic for splicers because it starts with B-frames that reference the previous GOP; if a splice occurs at a GOP boundary, the anchor frame referenced by the B-frames will not exist in the output stream. This may be important to certain types of splicers.

The second issue relating to GOP structure is the number of B-frames used between anchor frames. If a splicer is commanded to splice between anchor frames, this number determines how far the nearest anchor frame is from the splice point. Some splicers may need to ensure that either an in- or out-point (or both) occur on anchor frames, so this issue can affect splicer behaviour.

Another issue is whether "progressive refresh" is used. In progressive refresh systems, no I-frames appear in the stream (I-macroblocks are used instead). The absence of an I-frame can pose a serious problem for splicers that are not able to generate I-frames on demand.

I.6.1.3.1.3 Bit rate

Streams can be encoded at a constant bit rate (CBR) or a variable bit rate (VBR). Splicers may be expected to splice between CBR sources with the same bit rate; CBR sources with different bit rates; or VBR sources.

Splicing between VBR sources is significantly more complex than splicing between similar bit rate CBR sources. A splicer designed only to handle CBR sources may have problems handling VBR data.

I.6.1.3.1.4 Splice points

Different splicer architectures may place different constraints on the placement of splice points. A transport stream splicer may require that streams be SMPTE 312M [5] conditioned, while other splicers may require that in-points and out-point occur on I-frames or anchor frames.

When instructed to splice at a point not meeting their requirements, different splicers may behave very differently. Some may have no problem under any conditions; others may adjust the position of the splice point in the stream, insert a transition sequence between clips, or create a non-seamless splice.

I.6.1.3.2 Audio elementary streams

Audio elementary streams are considerably simpler than their video counterparts, but producing clean audio splices that are properly synchronized with video can be challenging. Issues such as the encoding type, bit rate, and sample rate all affect how well splicers (and set-tops) can do their job.

I.6.1.3.2.1 Stream type

The two most widely used audio types are MPEG-1 Layer II and Dolby AC-3. AC-3 uses a greater frame duration that may have an impact on A-V synchronization accuracy in some splicers. Splicing between dissimilar stream types can be problematic.

I.6.1.3.2.2 Bit rate/sample rate

The sample rate and bit rate can be different between clips. These changes affect the selection of the splice point and must be properly accounted for by the splicer.

I.6.1.3.3 Data streams

The handling of data streams associated with a program can be difficult. In a generic sense, splicers can simply choose to switch a data stream at a PES boundary closest to the splice point, or choose not to switch a data stream at all. However, certain types of data streams may be related to the A-V content in such a way that a "hard-cut", without performing type-specific processing on the stream, may be inadequate.

I.6.1.3.4 Multiplex type

Once the characteristics of the programs being processed by a splicer have been identified, it will be necessary to look at the characteristics of the multiplex that contains these programs, since splicers will typically be used to operate on one or more programs within a multiplexed stream.

I.6.1.3.4.1 Statistical multiplexing

Multiplexes can be created with fixed bit rate allocations for programs they contain, or they can be created using statistical multiplexing, where the bit rate of the programs they contain is allowed to vary. In the latter case, a statistical multiplexer (stat-mux) is used to ensure that the aggregate bit rate of all programs contained within a multiplex does not exceed a certain bound.

On the input side, if a splicer can handle VBR streams, it can handle a statistically multiplexed input. On the output side, statistical multiplexing can provide a significant challenge for a splicer.

Statistical remultiplexing can be a complex operation whose behaviour and performance can potentially be more difficult to characterize than splicing itself. Different stat-mux architectures take different approaches to limiting the aggregate bit rate of a multiplex, and a single stat-mux is likely to employ multiple techniques depending on the state of multiplex it is processing.

Some stat-muxes may be designed to limit transient overages in aggregate bit rate, while others may be designed to perform well when presented with long-term overages. The evaluation of image quality in a variety of situations is likely to be a key issue, with stat-muxes tending to introduce time-varying spatial or temporal artifacts (or both) when dealing with more severe bit rate overages.

A classification of stat-mux types is outside the scope of this appendix. However, it is at least useful to note whether a given splicer is capable of producing a statistically multiplexed output or not. (Splicers that do not directly produce a statistically multiplexed output can still be effectively used in a statistically multiplexed environment through the use of an external stat-mux.)

I.6.1.3.4.2 Multichannel

Multiplexes carrying HD signals may contain a mixture of HD and SD programs. In addition to splicing between HD programs, a splicer may be called upon to splice between a single HD program to multiple SD programs. In this situation, the splicer conditions the programs so that a downstream decoder can switch between the HD program and one of multiple SD programs. This type of behaviour is also useful in an SD-only world when dealing with certain kinds of targeted advertising schemes.

In addition to specifying the MPEG-2 profiles and levels that a splicer can process, it is also useful to note whether splicing between profiles and/or multichannel splicing is supported by a particular splicer.

I.6.1.4 Splicer performance

There are numerous metrics that can be used to evaluate the performance of MPEG-2 processing equipment. Many of these metrics are commonly applied to devices such as encoders and remultiplexers, and they can be applied to splicers as well (e.g., PCR jitter, etc.). This clause concentrates on those metrics that have particular relevance to splicers.

I.6.1.4.1 Seamlessness

Seamless splicing is taken to mean different things by different people. Because of this, it is important to define more clearly what seamless means, and to acknowledge that there are in fact different levels of "seamless" splicing behaviour. The following definitions are recommended by this appendix:

I.6.1.4.1.1 near-seamless splice n.: A synonym for **non-seamless splice**.

I.6.1.4.1.2 non-seamless splice n.: A splice which is not a seamless splice: it is either not **visually seamless** or not **syntactically seamless**. May cause a momentary freeze, blank screen, or motion judder. Many non-seamless splices may appear to be seamless to some viewers; whether it is good enough is subjective. The appearance of the splice usually depends on the relationship between the two spliced streams at the time of the splice.

I.6.1.4.1.3 seamless splice n.: A splice which is both **syntactically seamless** and **visually seamless**.

I.6.1.4.1.4 splice (1) n.: The process of leaving one bitstream and joining another. (2) v. The act of such joining. (3) n. The place in the resulting bitstream where the joint has been made.

I.6.1.4.1.5 syntactically seamless splice n.: A splice which results in a bitstream which meets the syntactic and semantic requirements of MPEG-2 Systems spec [2] and Video spec [3]. Not necessarily a **visually seamless splice** (possibly due to insertion of **transition frames**).

I.6.1.4.1.6 transition sequence, transition frames n.: A short bitstream sequence of frames synthesized by a **splicer** to interpose between the end of the **old stream** and the beginning of the **new stream**, usually to control buffer levels.

I.6.1.4.1.7 visually seamless splice n.: A splice which results in an unbroken sequence of decoded frames such that the last frame of the old stream is followed by the first frame of the new stream without intervening.

I.6.1.4.1.8 transition frames: This kind of splice may or may not be a **syntactically seamless splice**. Visual performance may depend on decoder characteristics that are not defined by MPEG-2.

The environment within which a splicer is operating can have a significant effect on how seamless a splice it is able to make. Even a TS splicer can perform a seamless splice on a SMPTE 312M [5] conditioned stream when commanded to splice exactly at the conditioned splice point. However, many types of splicers may have a problem producing a visually seamless splice between programs

that contain no I-frames. Some splicers can take advantage of variable bit rate coding or delay to produce visually seamless splices. Therefore, it is important to note under what circumstances a given splicer will produce a splice with a given level of artifacts.

I.6.1.4.2 Frame accuracy

Frame accuracy describes whether the transition between sequences occurs at exactly the frame specified. Assuming that a given splicer can be commanded to splice at an exact frame, the issue is whether the splicer can perform the splice at exactly the desired point.

As discussed in prior clauses, splicers will typically place some restrictions on the types of frames that can be used for in-points and out-points. Hence, when discussing the frame accuracy of a splicer, it is important that the condition under which the accuracy is to be determined is specified. For example, almost any splicer can be frame accurate when commanded to splice at pre-conditioned splice points whereas, only splicers that can produce I-frames on demand can be frame accurate when commanded to splice in a situation where neither the in-point or out-point are anchor frames. Other splicers can be frame accurate when commanded to splice where the out-point is an anchor frame and the in-point is an I-frame.

Unfortunately, there are different views among the manufacturers regarding what constitutes "frame accurate". Therefore, it is important to articulate what the range of possible circumstances is and how the splice point might be adjusted in these circumstances.

I.6.1.4.3 Delay

While minimum fixed delay is important to the design of retransmission facilities, certain splicers may be able to take advantage of variable delay to improve the quality of their splices in situations where the appropriate in-points and out-points do not appear at opportune times.

It would be useful to identify the upper limit of acceptable delay for a given type of facility, as well as whether variable delay is acceptable and, if so, how much. Facilities at different points in the broadcast chain may have very different requirements; for example, an origination facility may be very sensitive to delay variations, whereas the last retransmission facility before reaching the home may be quite insensitive to delay variations.

I.6.1.4.4 MPEG-2 compliance of output stream from splicer

Just as the input network stream should comply with the MPEG-2 Transport Specification [2], the output transport stream from the splicer/headend (after the processing of cue messages and appropriate ad-insertion) that is processed by the subsequent splicer or the set-top box should comply with the MPEG-2 Conformance Specification [4], as well as the SCTE Network Interface Specification SCTE 40 2001 [6]. Even though the network streams may not have many discontinuities (such as sequence_end_codes or time base changes), the output streams from the splicers may include one or more discontinuities that are allowed by MPEG. These may include termination of a video sequence using seq_end_code followed by a new sequence with different coded frame size or bit rate, or changes to and from film mode in addition to time base discontinuities signalled by the PCR discontinuity flag. Splicers and set-top boxes that comply with the MPEG-2 Conformance Standard are expected to process the discontinuities in the transport streams described below. The following italicized text related to handling of discontinuities by the set-top box is reproduced from ISO/IEC 13818-4 [4] and should serve as a guideline for splicers on compliance of output stream after the splicing operation:

Handling of decoder discontinuities (Sequence concatenation; decoding discontinuities; splicing; format changes).

In compliant Transport Streams, at any audio access unit boundary or any video sequence boundary, the following discontinuities in the decoding process parameters can occur:

- *For video, any parameter set in the sequence header or lower layer headers, such as profile/level, frame rate, bit rate, GOP parameters, picture format, etc.;*
- *For audio, any parameter such as audio layer, bit rate, sample rate, etc.;*
- *For both video and audio, the decoding time of the first access unit after the boundary can be larger than would have been expected had the boundary not been present. This can happen independently for all, some, or one of the elementary streams of a program. It may or may not be indicated by the presence of extra information referring to a seamless or non-seamless splice point.*

Assuming any combination of change(s) in decoding process parameter(s) which lead(s) to parameter values that are supported by the decoder under test, the decoder under test shall:

- *Maintain correct presentation synchronization between the different elementary streams of the program;*
- *Not produce unacceptable audio or video artifacts, such as chirps, blocking etc. However, when a decoding discontinuity occurs, there may not be any data to present during some time interval. At such instants, audio decoders are recommended to mute and video decoders to freeze frame/field.*
- *"In addition, when a phase shift in display timing of video (after the discontinuity) is indicated by the PTS (i.e, the difference between the current PTS and the previous one is not an exact integer number of frame periods) decoders can continue the display process without any discontinuity in the vertical timing. This involves re-mapping the decoded video on the display process, which may require some additional memory (than what is specified in the T-STD model). On the other hand, decoders can also resynchronize (genlock) by directly adjusting the vertical display timing to the decode timing, thereby allowing for a discontinuity in the phase of the vertical display timing. This may result in a visible resynchronization effect on display devices. Both these implementations are allowed in compliant decoders."*

Here are some examples of set top behaviour under different splicing conditions:

1) If the input stream (output from the headend or the splicer to the set top) fully complies with video syntax, the transport T-STD and time base continuity (i.e., no PCR discontinuities and PTS continues in phase) before and after the point where insertion occurs, the set top should change over to the new sequence in a seamless fashion. This includes sequence header changes at the insertion points where coded frame size or bit rate or quantizer matrices change. Frame rate changes will result in resynchronization and possibly a reset of some set-top boxes.
2) If a PCR discontinuity appears before the start of the inserted sequence (and there is still compliance with the T-STD), the set top will acquire the new sequence with some display artifact which accounts for the phase shift of the time base.
3) If the input stream breaks the T-STD at the point where the sequence changes (i.e., the inserted sequence overflows the T-STD buffer), the set top may reset as the input is no longer 'compliant'.
4) The case where no time base discontinuity is signalled and the inserted sequence starts off with a new time base is also non-compliant, and the set-top box may reset.

The cue message standards require the transport stream going to the next splicer or a set-top box to fully comply with the MPEG-2 transport and video specifications and in this case there should be no error in the output of the set-top box. Even though set-top boxes handle errors, they will not likely be able to handle 'non-compliant' streams gracefully. All MPEG compliant set-top boxes must handle changes of the video resolution after a seq_end_code and changes to parameters in the seq_header that follows (such as quantization matrices, bit rate, frame size, aspect ratio, low delay etc). Processing PMT changes might incur a larger delay compared to the changes in video stream only, since the PMT processing is often done in firmware, whilst the video changes are typically done in an ASIC.