SMPTE STANDARD

Extension of the 24-Bit Digital Audio Format to 32 Channels for 3 Gb/s Bit-Serial Interfaces

Table of Contents

Page

Foreward 2
Intellectual Property 2
Introduction 2
1 Scope 3
2 Conformance Notation 3
3 Normative References 4
4 Definition of Terms 4
5 Extended Audio Data Packet 4
5.1 DID Values 4
5.2 Packet/Group Relationships 4
6 Extended Audio Control Packet 4
6.1 DID Values 4
6.2 Packet/Group Rrelationships 5
Annex A Bibliography (Informative) 6
Annex B Audio Multiplexing in S425-A Image Format Mappings (Informative) 7
Annex C Worst-Case Scenarios for Audio Multiplexing in S425-A Image Format Mappings (Informative) 10

Approved
July 21, 2010

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 ST 299-2 was prepared by Technology Committee 32NF.

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential to the implementation of this Standard. However, attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. SMPTE shall not be held responsible for identifying any or all such patent rights.

Introduction

This section is entirely informative and does not form an integral part of this Engineering Document.

SMPTE ST 299-1 defines the 24-bit audio format for up to 16 audio channels at 32, 44.1, or 48-kHz sample rate, or 8 audio channels at 96-kHz sample rate, in the SMPTE 292 Serial Data Interface at 1.5Gb/s. SMPTE ST 299-2, extends the audio format to 32 audio channels at 32, 44.1, or 48-kHz sample rate, or 16 audio channels at 96-kHz sample rate. Specifically, this extension defines the 24-bit audio format for channels 17 up to 32 so that up to 32 audio channels may be multiplexed with source image formats mapped to a 3 Gb/s serial interface.

1 Scope

SMPTE ST 299-2 defines an SMPTE 291M Type 1 packet structure for identifying audio channels numbered from 17 to 32, beyond the 16 channels defined in SMPTE ST 299-1. Four extended audio data packets and four extended audio control packets are identified. One of the extended audio control packets and one of the extended audio data packets is assigned to transport each of the four extended audio groups . Each extended audio group has four channels that carry up to four 24-bit audio channels with 32, 44.1, or 48-kHz sample rates, or up to two 24-bit audio channels with 96-kHz sample rate.

The purpose of defining four extended audio groups is to allow for multiplexing up to 32 audio channels into the horizontal ancillary (HANC) data space of the video source image formats mapped to a 3 Gb/s Level A serial data interface. The 3 Gb/s Level B serial data interface is capable of carrying 32 audio channels via application of SMPTE ST 299-1 to each of two 1.5 Gb/s SDI channels.

The format and location of audio data and control packets in the video data stream is identical to that of SMPTE ST 299-1, except for differences required to define extended audio groups.

2 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.

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.

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.

Unless otherwise specified, the order of precedence of the types of normative information in this document shall be as follows: Normative prose shall be the authoritative definition; Tables shall be next; followed by formal languages; then figures; and then any other language forms.

3 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 291M-2006, Television — Ancillary Data Packet and Space Formatting

SMPTE ST 299-1:2009, 24-Bit Digital Audio Format for SMPTE 292 Bit-Serial Interface

4 Definition of Terms

Channel: One half of a channel pair as defined in SMPTE ST 299-1.

Audio channel: One channel of audio program, which occupies one channel at sampling rates of 32, 44.1, or 48 kHz, or two channels at a sampling rate of 96 kHz

Extended audio group: An audio group as defined in SMPTE ST 299-1, but numbered from 5 to 8

Extended audio data packet: An audio data packet with structure as defined in SMPTE ST 299-1, but with identity corresponding to Extended audio group numbers 5 to 8

Extended audio control packet: An audio control packet with structure as defined in SMPTE ST 299-1, but with identity corresponding to Extended audio group numbers 5 to 8

5 Extended Audio Data Packet

The structure and multiplexing rules for extended audio data packets shall be as defined for audio data packets in SMPTE ST 299-1, with the following exceptions.

5.1 DID Values

The DID values for extended audio data packets shall be $1A7_h$ for audio group 5 (channel 17-20), $2A6_h$ for audio group 6 (channel 21-24), $2A5_h$ for audio group 7 (channel 25-28) and $1A4_h$ for audio group 8 (channel 29-32), respectively.

5.2 Packet/Group Relationships

Extended audio groups 5 to 8 shall be transported only using extended audio data packets defined in this document. Audio groups 1 to 4 shall be transported only using audio data packets defined in SMPTE ST 299-1.

6 Extended Audio Control Packet

The structure and multiplexing rules for extended audio control packets shall be as defined for audio control packets in SMPTE ST 299-1, with the following exceptions.

6.1 DID Values

The DID values for extended audio control packets shall be $2A3_h$ for audio group 5 (channel 17-20), $1A2_h$ for audio group 6 (channel 21-24), $1A1_h$ for audio group 7 (channel 25-28) and $2A0_h$ for audio group 8 (channel 29-32), respectively.

6.2 Packet/Group Relationships

Extended audio groups 5 to 8 shall be represented only using extended audio control packets defined in this document. Audio groups 1 to 4 shall be represented only using audio control packets defined in SMPTE ST 299-1.

Annex A Bibliography (Informative)

SMPTE 12M-2-2008, Television — Transmission of Time Code in the Ancillary Data Space

SMPTE ST 352:2010, Video Payload Identification for Digital Interfaces

SMPTE 425-2008, 3 Gb/s Signal/Data Serial Interface — Source Image Format Mapping

SMPTE RP 168-2009, Definition of Vertical Interval Switching Point for Synchronous Video Switching

Annex B Audio Multiplexing in S425-A Image Format Mappings (Informative)

Figure B.1 shows an example of Multiplexing for 32 channels of audio into a 1080p/59.94 video stream. Groups 1, 5, 6, 7, and 8 are synchronous to each other, and Groups 2, 3, and 4 are asynchronous to each other and to the other 5 groups.

Figure B.2 shows an example of Multiplexing for 32 channels of audio into a 1080i/59.94 video stream. Groups 1, 5, 6, 7, and 8 are synchronous to each other, and Groups 2, 3, and 4 are asynchronous to each other and to the other 5 groups. Note in this example that packets for multiple samples for the channels in Groups 1 to 4 are embedded in the HANC data space prior to multiplexing Groups 5 to 8. This particular group ordering is not a requirement. The timing of the nth sample of 32 audio channels on a video line is represented by eight sample instances in eight audio data packets. Since these eight sample instances are independent of each other, the ordering of these eight packets in the HANC data space is arbitrary.

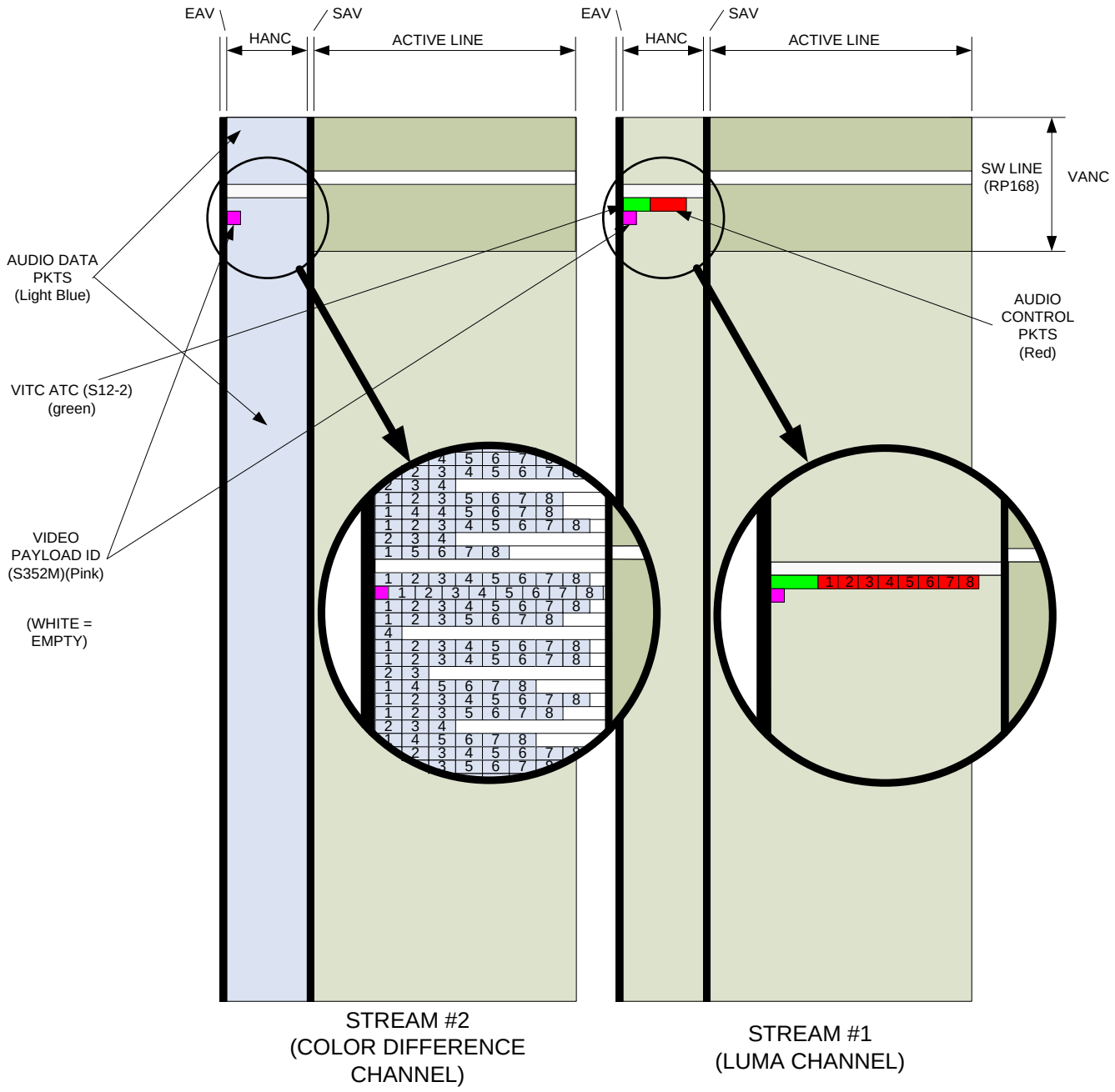

Figure B.1 – Example of Audio Packet Embedding for 1080p/59.94 Level A

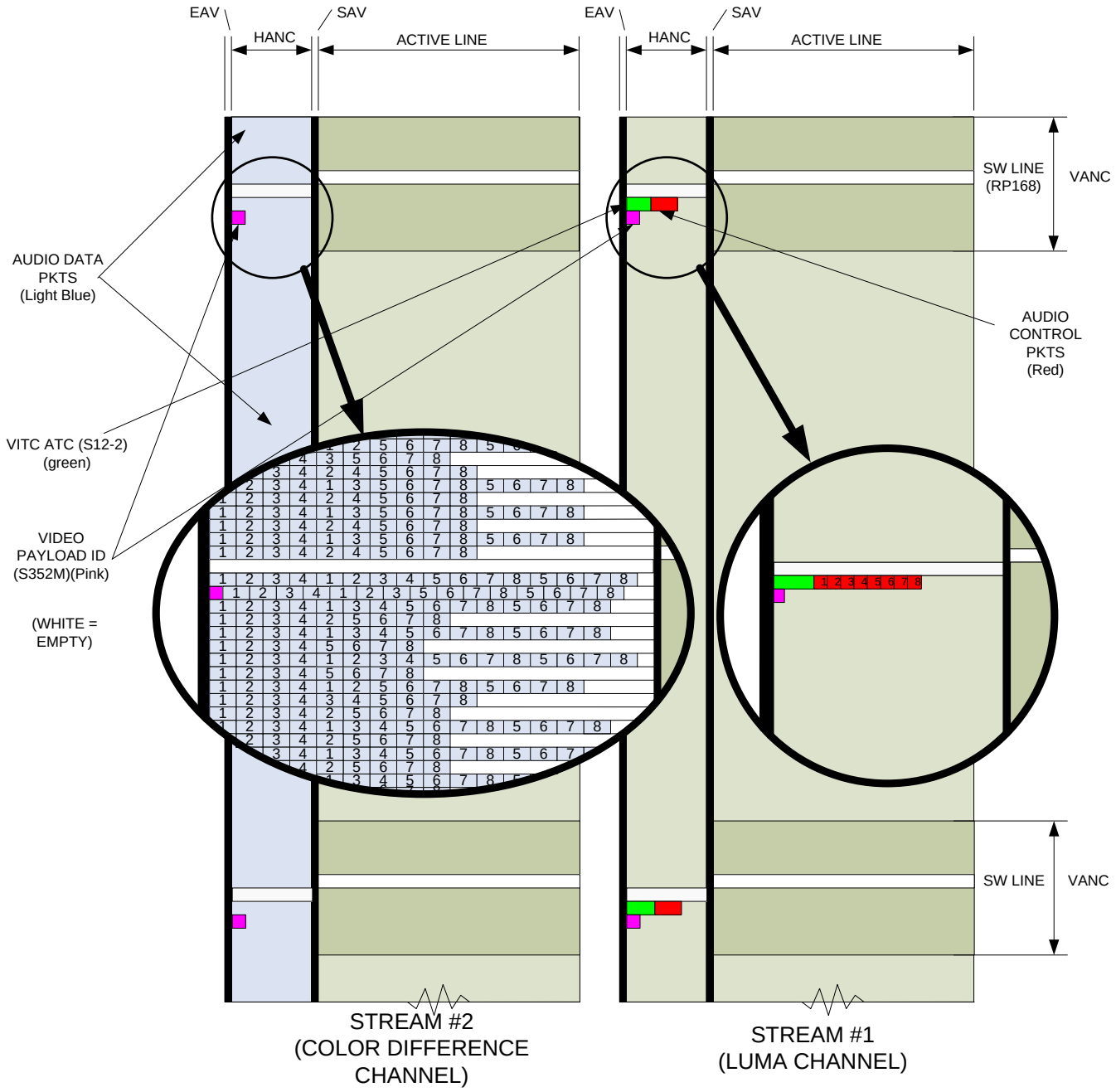

Figure B.2 – Example of Audio Packet Embedding for 1080i/59.94 Level A

Annex C Worst-Case Scenarios for Audio Multiplexing in S425-A Image Format Mappings

(Informative)

1. Worst-case HANC data space availability for audio multiplexing occurs on Stream #2 of Level A Mapping Structure #1 with the 1080p/59.94 or 1080p/60 video format. On Line 10, the 11-word S352M Video Payload ID packet must be embedded immediately after the EAV TRS. The available space for Na samples of 8 audio groups is given in Table C.1.

Table C.1 – 59.94Hz Level A Mappings – Stream #2 HANC Space

S425 Level A Mapping	Stream #2 Words per Line	Stream #2 Video Words	Stream #2 SAV+EAV+CRC+LN+ S352	8 Audio Group Pkts*Na	Words Remaining
1	2200	1920	4+4+2+2+11	31*8*1	9
2	3300	2560	4+4+2+2+11	31*8*2	221
	4400	3840	4+4+2+2+11	31*8*2	41
3	4400	3840	4+4+2+2+11	31*8*2	41
	5500	4096	4+4+2+2+11	31*8*2	885
4	4400	3840	4+4+2+2+11	31*8*2	41

2. Worst-case audio packet multiplexing around the switching line as defined in SMPTE RP 168 occurs on Stream #2 of Level A Mapping Structure #1 with the 1080p/50 video format. The probability of a synchronous audio sample in a given video line period is p(sync). For Level A Mapping #1, 48-kHz audio, and a 50-Hz frame rate, p(sync) = audio sample rate/line frequency = 0.8533, and Na = 1. Over a span of 7 lines, 5.973 audio samples will be embedded. Therefore, the maximum number of consecutive lines with one audio packet from each group is 6. If an audio sample is pushed off from Line 8 (after the switching line), it is placed on Line 9. In the worst-case situation, the Line 9 thru Line 13 packets are pushed off to Lines 10 thru 14, respectively, after which normal multiplexing resumes since no audio packets were to be embedded on Line 15. The use of mpf in flagging delayed audio samples for a given group is shown in Figure C.1.

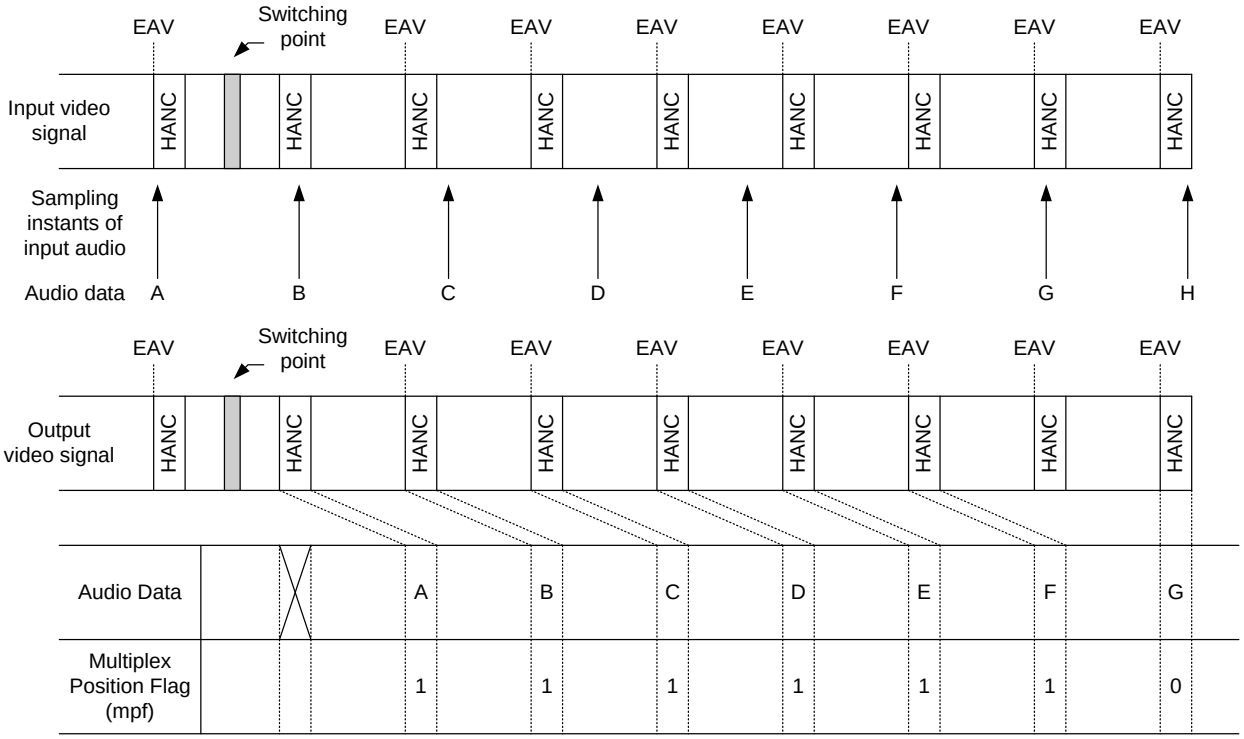

Figure C.1 – Relationship between the multiplex position flag and the multiplex position of audio data packets for Audio Embedding to 1080p/50 Level A video

3. The frequency offset of audio groups asynchronous to the video rate would have to be greater than {6/7 – p(sync)}/p(sync) = 0.45% for there to be 7 audio samples in a 7 line periods.of 1080p/50 Level A video. The worst-case situation of Figure B.2 applies for asynchronous frequency offsets less than this amount. For any audio group with a frequency offset greater than 0.45%, the number of consecutive samples with mpf = 1 increases, but it is not likely that frequency offset will approach even the 7-line limit described above. Since Na = 1, each asynchronous audio group has a reserved position in each video line for one audio data packet. When mpf = 1 for an audio group, the associated packet can only be “pushed off” to the next video line, not to the packet position of another group on the same line.

4. Worst-case Stream #1 availability for audio control packets assumes, according to SMPTE12M-2, that VITC ATC is possibly present in the HANC data space on Line 9. Figures B.1 and B.2 show examples of this ATC packet occurring at the beginning of the HANC space.