SMPTE REGISTERED DISCLOSURE DOCUMENT

SMPTE RDD 40:2016

Essence-independent IP Live Networked Media Transport

The attached document is a Registered Disclosure Document prepared by the proponent identified below. It has been examined by the appropriate SMPTE Technology Committee and is believed to contain adequate information to satisfy the objectives defined in the Scope, and to be technically consistent.

This document is NOT a Standard, Recommended Practice or Engineering Guideline, and does NOT imply a finding or representation of the Society.

Errors in this document should be reported to the proponent identified below, with a copy to eng@smpte.org.

All other inquiries in respect of this document, including inquiries as to intellectual property requirements that may be attached to use of the disclosed technology, should be addressed to the proponent identified below.

Proponent contact information:

Toshiaki Kojima
Sony Corporation
4-14-1 Asahi-cho, Atsugi
Kanagawa, 243-0014
Japan

Email: Toshiaki.Kojima@sony.com

Approved
October 11, 2016

Table of Contents

Page

Introduction

Although network systems have become common in file-based applications, until recently it has been difficult to apply networking technologies to live production applications in which there is a strong requirement for real-time processing.

However, with the continuing rapid evolution of networking technologies, there is now a trend towards the use of networked systems for live production applications.

This RDD describes an IP-based transport mechanism (referred to hereafter as ‘IP mapping') intended mainly for live production applications.

1 Scope

This RDD defines the IP mapping which includes the following main characteristics:

1. Essence-independent mapping

   Each essence (video, audio and ancillary data) can be dealt with independently.

2. Packet protection using FEC and/or hitless failover

   Appropriate protection can be chosen according to application requirements.

3. Frame-boundary-aware mapping for both essence and FEC

   Frame accurate processing can be performed in both essence and FEC.

4. Support for up to 4K uncompressed and compressed video

   Either uncompressed video or compressed video can be chosen up to 4K video seamlessly according to application requirements.

2 Normative References

IETF RFC 3550, RTP: A Transport Protocol for Real-Time Applications

IETF RFC 3551, RTP Profile for Audio and Video Conferences with Minimal Control

SMPTE RDD 34:2015, LLVC — Low Latency Video Codec for Network Transfer

SMPTE ST 125:2013, SDTV Component Video Signal Coding 4:4:4 and 4:2:2 for 13.5 MHz and 18 MHz Systems

SMPTE ST 272:2004, Television — Formatting AES Audio and Auxiliary Data into Digital Video Ancillary Data Space

SMPTE ST 274:2008, Television — 1920 × 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates

SMPTE ST 291-1:2011, Ancillary Data Packet and Space Formatting

SMPTE ST 296:2012, 1280 × 720 Progressive Image 4:2:2 and 4:4:4 Sample Structure — Analog and Digital Representation and Analog Interface

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

SMPTE ST 424:2012, 3 Gb/s Signal/Data Serial Interface

SMPTE ST 425-1:2014, Source Image Format and Ancillary Data Mapping for the 3 Gb/s Serial Interface

SMPTE ST 425-3:2015, Image Format and Ancillary Data Mapping for the Dual Link 3 Gb/s Serial Interface

SMPTE ST 425-5:2015, Image Format and Ancillary Data Mapping for the Quad Link 3 Gb/s Serial Interface

SMPTE ST 2022-5:2013, Forward Error Correction for High Bit Rate Media Transport Over IP Networks (HBRMT)

SMPTE ST 2036-1:2014, Ultra High Definition Television — Image Parameter Values for Program Production

SMPTE ST 2048-1:2011, 2048 × 1080 and 4096 × 2160 Digital Cinematography Production Image Formats FS/709

SMPTE ST 2059-1:2015, Generation and Alignment of Interface Signals to the SMPTE Epoch

3 Overview of the IP Mapping

Figure 1 gives an overview of the IP mapping which consists mainly of the following four processes:

1. De-multiplexing

In the case of an SDI input signal it is first de-serialized and de-multiplexed in order to extract each essence (video, audio and ancillary data). This process is out of scope of this document.

2. Essence RTP Packing

Each essence is independently mapped as an Essence payload in conjunction with an Essence header, followed by Transport headers added to complete the Essence datagrams. The Transport headers include Common and RTP (Real Time Protocol, RFC 3550) headers. In order to generate Essence and Common headers, information obtained through the FEC creation process is required (see Section 5).

If compressed video is preferred, the input video signal shall be compressed before entering the Essence Payload Packing module. LLVC described in SMPTE RDD 34 is defined as the primary video codec.

3. FEC RTP Packing

FEC shall be used in this mapping. FEC creation is performed based on the data from the Essence payloads. Common and RTP headers are added to complete the FEC RTP datagrams.

4. Network Transmission Processing

Network (UDP, IP and Ethernet) headers are added to all Essence and FEC RTP datagrams and output as network datagrams. This process is out of scope of this document.

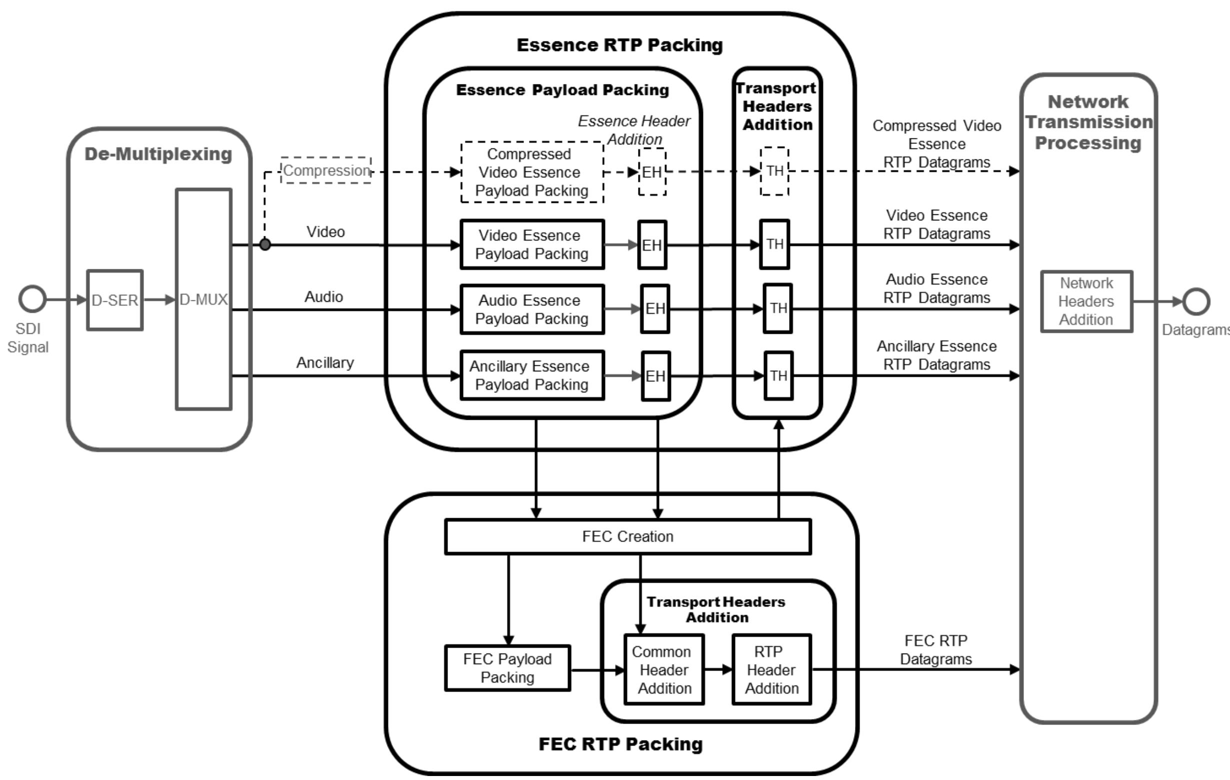

Figure 1 – Overview of the IP Mapping

4 Structure of the IP Mapping

Figure 2 and Figure 3 show the structure of the Essence Datagram and FEC Datagram respectively.

Because all required information for dealing with Essence and FEC datagrams is included in the Essence header and the Common header, all datagrams have the same destination IP address and UDP port number. In other words, all datagrams are transported as a single stream.

The FEC Payload shall be 1,382 bytes as the FEC creation is applied to the Essence Payload which consists of Essence Header (4 bytes) and Essence (1378 bytes).

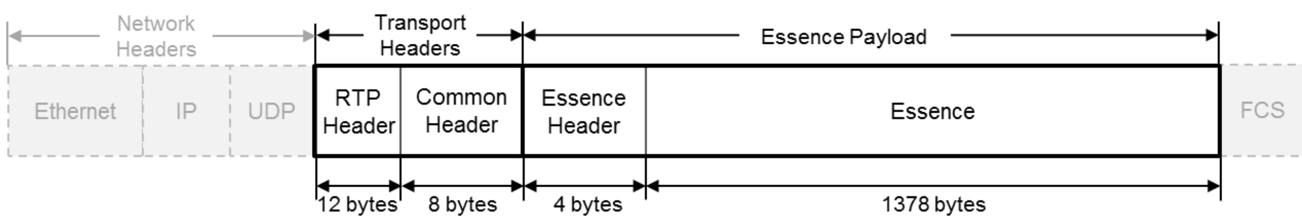

Figure 2 – Structure of Essence Datagram

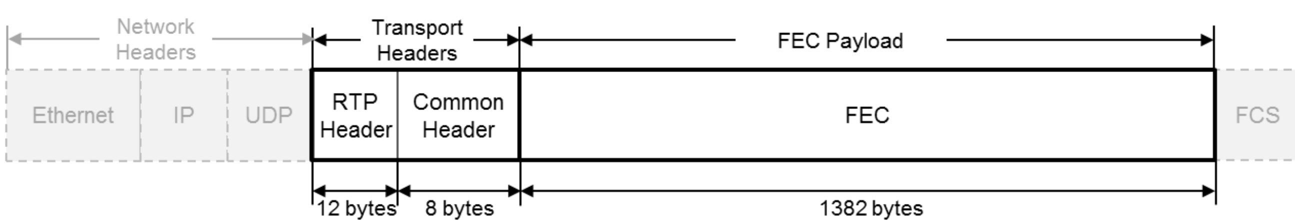

Figure 3 – Structure of FEC Datagram

4.1 Essence Header

Figure 4 defines the Essence header. Detailed information on each field is specified in Table 1.

<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td><td>29</td><td>30</td><td>31</td></tr>
<tr><td colspan="2">PT</td><td colspan="14">Payload Length</td><td>S</td><td>E</td><td colspan="7">Frame Count</td><td>F</td><td>C</td><td>G</td><td colspan="4">Reserved</td></tr>
</table>

Figure 4 – Definition of Essence Header

Table 1 – Essence Header Fields Description

Field Name		Size (bit)	Description
PT	Payload Type	2	Shall be set according to the essence type of the payload. 0: Video, 1: Audio, 2: Ancillary data, 3: Reserved
	Payload Length	14	Active essence size of the media payload in octets.
S	V Start	1	For the first datagram of a video frame (field in the case of interlace) of each essence type: Shall be set to 1. For other Essence datagrams: Shall be set to 0.
E	V End	1	For the last datagram of a video frame (field in the case of interlace) of each essence type: Shall be set to 1. For other Essence datagrams: Shall be set to 0.
FC	Frame Count	7	All input A/V signals shall be synchronized so that if Frame Count is added to the mapped input A/V signals (Essence datagrams) at a sender, it can be used for synchronizing datagrams at a receiver. The initial value shall be set to the specific value such that Frame Count is 0 at the Epoch defined in SMPTE ST 2059-1. The value shall be incremented by one when a new video frame starts. When a video stream is not present, Frame Count shall be incremented at intervals corresponding to the system frame rate chosen by the user. The same value shall be set for all related Essence datagrams. Range of value is from 0 to 127, and when the value reaches its maximum, the count shall rollover and start again at 0.
F	Field Flag	1	Shall be set as follows: 0: Field 0 1: Field 1 Shall be set to 0 for all datagrams of progressive format.
C	Compression	1	Shall be set as follows: 0: Uncompressed video 1: Compressed video
G	Forced End FlaG	1	For the datagrams which contain invalid Essence data e.g. adding zero padding: Shall be set to 1 For all other datagrams: Shall be 0.
RSV	Reserved	4	Reserved for future use. Shall be set to 0. This field shall be ignored by receivers.

4.2 Common Header

Figure 5 defines the Common header. Detailed information on each field is specified in Table 2.

<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td><td>29</td><td>30</td><td>31</td></tr>
<tr><td colspan="7">Frame Count</td><td>F</td><td colspan="2">FT</td><td colspan="2">ST</td><td colspan="2">DT</td><td>B</td><td>M</td><td colspan="16">Sequence Number</td></tr>
<tr><td>T</td><td colspan="7">Reserved</td><td colspan="4">L Max</td><td colspan="4">D Max</td><td colspan="4">L Count</td><td colspan="4">D Count</td><td colspan="8">BLK_ID</td></tr>
</table>

Figure 5 – Definition of the Common Header

Table 2 – Common Header Fields Description

Field Name		Size (bit)	Description
FC	Frame Count	7	Shall be set to the same value as that of the corresponding Essence header. Because the Common header is outside of the FEC process, this value can be used before FEC decoding. This enables pre-processes such as clean switching at RTP datagram level. Validity can be checked against the same value from the Essence header with FEC. This value is unique in a single FEC Block. For FEC datagrams, it shall also be set to the same value as that of the corresponding Essence datagrams belonging to the same FEC Block.
F	Field Flag	1	Shall be set to the same value as that of the corresponding Essence header. Because the Common header is outside of the FEC process, this value can be used before FEC decoding. This enables pre-processes such as clean switching at RTP datagram level. Validity can be checked against the same value from the Essence header with FEC. This value is unique in a single Block. FEC datagrams shall also be set to the same value as that of the corresponding Essence datagrams belonging to the same FEC Block.
FT	FEC Type	2	Shall be set to 0. This field indicates the FEC type being used. 0: XOR based 1: Reed-Solomon based 2: Reserved 3: Reserved
ST	Stream Type	2	Reserved for future use. Shall be set to 0. This field shall be ignored by receivers.
DT	Datagram Type	2	Type of datagram. 0: Essence datagram 1: Row FEC datagram 2: Column FEC datagram 3: Reserved.
B	FEC Block Last	1	Shall be set to 1 for the last Essence datagram, the last Column FEC datagram and the last Row FEC datagram of every FEC Block transmission. For other datagrams, shall be set to 0.
M	Transfer Mode	1	Reserved for future use. Shall be set to 0. This field shall be ignored by receivers.

SN	Sequence Number	16	Shall be set in each category separately, namely video Essence, audio Essence, ancillary data Essence, Column FEC and Row FEC datagrams. Shall be incremented by one in each transmission of the same category and shall be sequential across FEC Blocks. Initial value should be set randomly. When this value reaches its maximum, the count shall rollover and start again at 0.
T	FEC Block Top Flag	1	Shall be set to 1 for all Essence and FEC datagram of the first block of a video frame (field in the case of interlace). All other FEC blocks of the video frame (field in the case of interlace) shall be set to 0.
RSV	Reserved	7	Reserved for future use. Shall be set to 0. This field shall be ignored by receivers.
-	L Max	4	L value of the LxD Basic FEC Block Size for XOR. It shall never be changed during the RTP session regardless of the actual size.
-	D Max	4	D value of the LxD Basic FEC Block Size for XOR. It shall never be changed during the RTP session regardless of the actual size.
-	L Count	4	Row number of the Essence or FEC datagrams (from 0 to L).
-	D Count	4	Column number of the Essence or FEC datagram (from 0 to D).
BLK_ID	FEC Block ID	8	Shall be set to a unique number for each FEC Block. Same number shall be set for all Essence and FEC datagrams in the same FEC Block. Shall be incremented by one in each transmission. Initial value should be set arbitrarily. When this value reaches its maximum, the count shall rollover and start again at 0.

4.3 RTP header

The IP Mapping adopts the RTP header specified in RFC 3550. The definition of the RTP header is shown in Figure 6. All fields in the RTP header shall be set according to RFC 3550. Table 3 clarifies the value of each field.

<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td><td>29</td><td>30</td><td>31</td></tr>
<tr><td colspan="2">V=2</td><td>P</td><td>X</td><td colspan="4">CC</td><td>M</td><td colspan="7">PT</td><td colspan="16">Sequence Number</td></tr>
<tr><td colspan="32">Time Stamp</td></tr>
<tr><td colspan="32">Synchronization Source (SSRC) Identifier</td></tr>
</table>

Figure 6 – Definition of RTP Header

Table 3 – RTP Header Fields Description

Field Name		Size (bit)	Description
V	Version	2	Shall be set to 2
P	Padding	1	Shall be set to 0.
X	Extension	1	Shall be set to 0
CC	CSRC Count	4	Shall be set to 0. Note: There are no CSRC lists in Essence and FEC Datagrams.
M	Marker	1	For the last Essence datagram of video frame (field in the case of interlace): Shall be set to 1. For other Essence and all FEC datagrams: Shall be set to 0. As this field is not used for FEC datagrams, in the case of FEC datagrams transmitters are recommended to set to 0 and receivers shall ignore the field. This field is used to detect the frame (field in the case of interlace) boundary of only Essence datagrams. Frame Count in both Common and Essence headers should be used for all datagrams rather than referring to this value.
PT	Payload Type	7	Shall be set to 110 at the start of transmission. Alternatively, this field may be set by other means in accordance with RFC 3550/3551. This field shall not be changed during the RTP session.
SN	Sequence Number	16	Shall be increment by one in each transmission. Initial value should be set randomly as stated in RFC 3550. When the value reaches its maximum, the count shall rollover and start again at 0.
TS	Timestamp	32	The value of Timestamp shall be derived from a clock that increments monotonically and linearly in time to allow synchronization and jitter calculations. This value is not used in the IP mapping.
SSRC	Synchronization Source	32	Should be set randomly as stated in RFC 3550. The value shall not be changed during the RTP session.

5 FEC Creation

As described in Section 3, using FEC is mandatory in this mapping. The FEC creation process is described first because information obtained through the FEC creation is required to generate both the Essence header and Common header.

One of two FEC mechanisms can be chosen in the IP Mapping, XOR or Reed-Solomon (RS). The RS based FEC shall be used for low bit rate Essence RTP datagrams (less than or equal to 500Mbps) while the XOR based FEC shall be used for high bit rate Essence RTP datagrams (greater than 500Mbps).

5.1 XOR Based FEC

This FEC scheme is similar to what is specified in SMPTE ST 2022-5:2012 with the following differences:

1. Only Level B (using both Column FEC (L) and Row FEC (D)) scheme shall be used.
2. Only Block-Aligned FEC operation shall be used.
3. A truncated FEC block shall be allowed on specified conditions.

The default FEC block size shall be 12(L) x 12(D). However, the FEC block shall be truncated when the frame (field in the case of interlace) boundary is detected.

Figure 7 shows an example of a truncated 4x4 FEC block caused by reaching the frame boundary. Each ED represents Essence Datagram and each FD represents FEC Datagram. No dummy datagrams or header only datagrams will be sent for the remaining Essence and FEC datagrams.

This truncated FEC avoids the situation in which a FEC block contains information related to two video frames (or fields). As a result, frame-accurate video processes in the datagram domain with FEC can be realized.

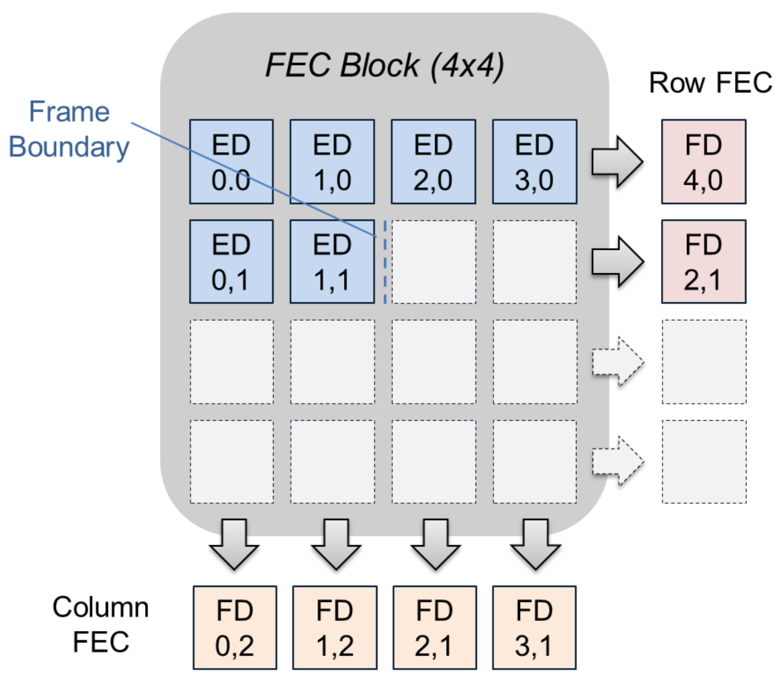

Figure 7 – Example of a truncated FEC Block

5.2 RS Based FEC

There are several differences between XOR based FEC and RS based FEC.

- RS based FEC is a one-dimensional FEC Block arrangement.
- RS code is capable of creating multiple FEC datagrams from a single group of Essence datagrams.

An RS code is specified by RS(n, k). This notation indicates that the FEC block is composed of k Essence datagrams. With these k Essence datagrams, n-k FEC datagrams (in other words, redundancy datagrams) are created using RS code. This composes a FEC block of total n Essence and FEC datagrams as shown in Figure 8.

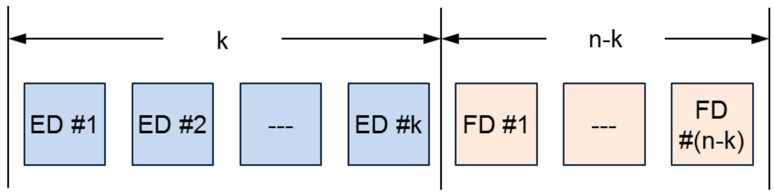

Figure 8 – Structure of RS based FEC block

The default FEC block size shall be RS(16, 14). However, the FEC block is truncated when the frame (field in the case of interlace) boundary is detected.

The truncated FEC block size shall be one of the following:

RS(15, 13), RS(14, 12), RS(13, 11), … RS(5, 3), RS(4, 2), RS(3, 1).

The RS specification used to construct the RS FEC payload is shown in Table 4.

Table 4 – RS Specification

Galois Field	$GF(2^8)$
Primitive Polynomial	$p(x) = x^8 + x^4 + x^3 + x^2 + 1$
Primitive Root	*MSB* *LSB* $\alpha = [0\ 0\ 0\ 0\ 0\ 0\ 1\ 0]$
Generator Polynomial	$g(x) = \prod_{i=0}^{r-1}(x - \alpha^i)$

The value r, the Generator Polynomial, is the number of redundant symbols.

(Note that r redundant symbols can generate r FEC Diagrams)

6 Essence Payload Packing

Each essence (video, audio and ancillary data) shall be packed in the separate payloads of the respective Essence datagrams. As described in Figure 2, the payload size of the Essence datagram shall be 1378 bytes. When the frame (or field) end is detected, zero-paddings shall be added so that the payload size is kept to 1378 bytes. The successive frame (or field) data shall be packed into a payload of a new Essence datagram.

6.1 Video Payload Packing

According to the video format, the following four types of video packing are defined in the IP mapping.

6.1.1 $Y'C'_BC'_R$ 4:2:2 10-bit

Only active video data per 4-pixel unit starting from the data next to SAV shall be packed in the video payload. The order within the 4-pixel structure shall be Y Data(Y0Y1Y2Y3) first followed by the C Data($C_{B0}C_{R0}C_{B1}C_{R1}$).

Figure 9 shows an example of video payload packing. Video data of a new frame (or field) shall be packed from the start of a video payload. Adding an offset or performing zero-padding before the video data shall not be allowed. Video data are packed in the video payload consecutively within a frame (or field). Adding spaces or zero-padding shall not be allowed.

0 1 2 3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1

Y0[9:0]	Y1[9:0]	Y2[9:0]	Y3
Y3[7:0]	C_B0[9:0]	C_R0[9:0]	C_B1[9:6]
C_B1[5:0]	C_R1[9:0]	Y4[9:0]	Y5[9:4]
Y5[3:0]	Y6[9:0]	Y7[9:0]	C_B2[9:2]
C_B2	C_R2[9:0]	C_B3[9:0]	C_R3[9:0]

Figure 9 – Example of Video Packing (YC_BC_R 4:2:2 10-bit)

6.1.2 RGB 4:4:4 10-bit

Only active video data per 4-pixel unit starting from the data next to SAV shall be packed in the video payload. The order within the 2-pixel structure shall be as follows: G Data(G0G1) first, B Data(B0B1) next, and finally R Data(R0R1).

Figure 10 shows an example of video payload packing. Video data of a new frame (or field) shall be packed from the start of a video payload. Adding an offset or performing zero-padding before the video data shall not be allowed. Video data are packed in the video payload consecutively within a frame (or field). Adding spaces or zero-padding shall not be allowed.

0 1 2 3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1

G0[9:0]	G1[9:0]	B0[9:0]	B1
B1[7:0]	R0[9:0]	R1[9:0]	G2[9:6]
G2[5:0]	G3[9:0]	B2[9:0]	B3[9:4]
B3[3:0]	R2[9:0]	R3[9:0]	G4[9:2]
G4	G5[9:0]	B4[9:0]	B5[9:0]

Figure 10 – Example of Video Packing (RGB 4:4:4 10-bit)

6.1.3 RGB 4:4:4 12-bit

Only active video data per 2-pixel unit starting from the data next to SAV shall be packed in the video payload. The order within the 2-pixel structure shall be as follows: G Data(G0G1) first, B Data(B0B1) next, and finally R Data(R0R1).

Figure 11 shows an example of video payload packing. Video data of a new frame (or field) shall be packed from the start of a video payload. Adding an offset or performing zero-padding before the video data shall not be allowed. Video data are packed in the video payload consecutively within a frame (or field). Adding spaces or zero-padding shall not be allowed.

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
```

G0[11:0]	G1[11:0]	B0[11:4]	
B0[3:0]	B1[11:0]	R0[11:0]	R1[11:8]
R1[7:0]	G2[11:0]	G3[11:0]	
B2[11:0]	B3[11:0]	R2[11:4]	
R2[3:0]	R3[11:0]	G4[11:0]	G5[11:8]

Figure 11 – Example of Video Packing (RGB 4:4:4 12-bit)

6.1.4 Compressed Video

The mapping of LLVC (Low Latency Video Codec) described in SMPTE RDD 34 is defined in the IP mapping as the default compressed video scheme.

Figure 12 shows the structure of the LLVC compressed video payload. Data within one frame is contained in multiple TUs (Transmission Unit) from TU 0 to TU n preceded by Picture Info. This dataset is called the Compressed Codestream.

The codestream shall be packed from the beginning of a video payload. Adding an offset or performing zero-padding before the video data shall not be allowed.

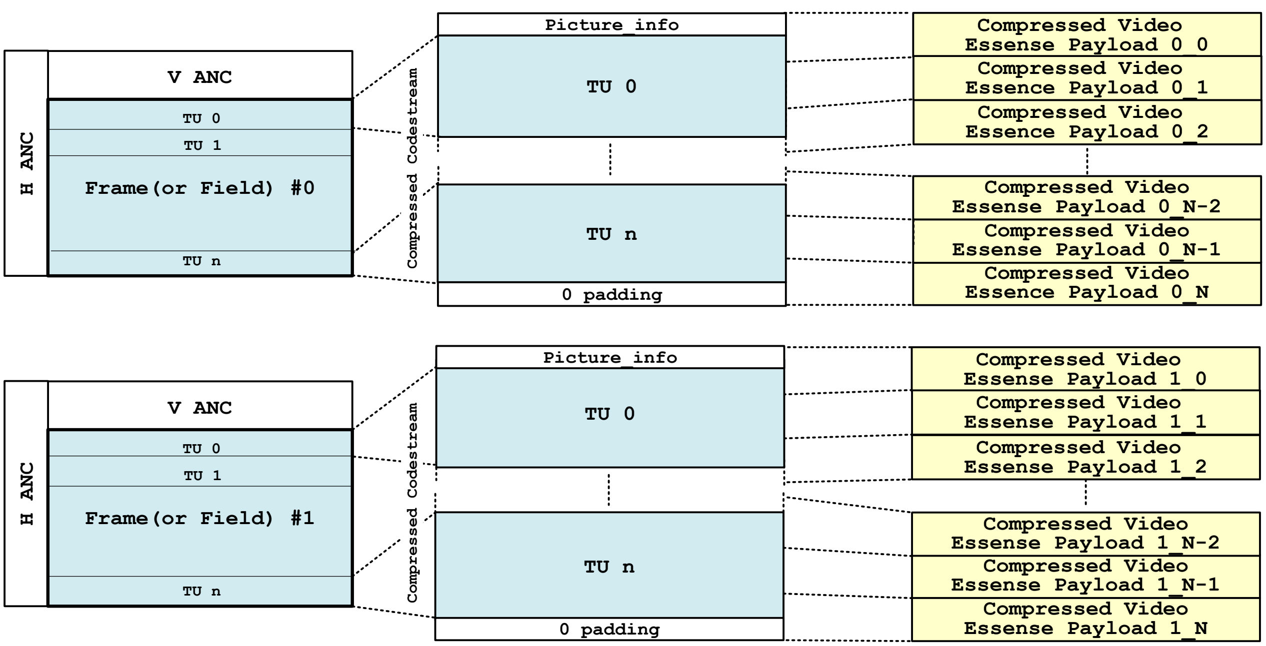

Figure 12 – Structure of LLVC Compressed Video Payload

6.1.5 Supported Video Formats

Table 5 and Table 6 show the supported video formats for 4:2:2 $Y'C'_BC'_R$ and 4:4:4 RGB respectively. All the video formats except SD (Standard Definition) formats are supported by LLVC.

Table 5 – Supported Video Formats for 4:2:2 $Y'C'_BC'_R$

Resolution	bit	Standard	Typical SDI Bit Rate	Frame Rate
720x480	10	SMPTE ST 125	270 Mb/s	59.94i
720x576	10	SMPTE ST 125	270 Mb/s	50i
1280x720	10	SMPTE ST 296	1.5 Gb/s	50p, 59.94p, 60p
1920x1080	10	SMPTE ST 274	1.5 Gb/s	50i, 59.94i, 60i
		SMPTE ST 274	1.5 Gb/s (P[*1] and PsF)	23.98p, 24p, 25p, 29.97p, 30p
		SMPTE ST 274	1.5 Gb/s Dual-Link	50p, 59.94p, 60p
		SMPTE ST 274	3 Gb/s	
2048x1080	10	SMPTE ST 2048-1	1.5 Gb/s Dual-Link (P and PsF)	23.98p, 24p, 25p, 29.97p, 30p
		SMPTE ST 2048-1	1.5 Gb/s Dual-Link	47.95p, 48p, 50p, 59.94p, 60p
		SMPTE ST 2048-1	3 Gb/s	
3840x2160	10	SMPTE ST 2036-1	Dual 3 Gb/s	23.98p, 24p, 25p, 29.97p, 30p
		SMPTE ST 2036-1	Quad 3 Gb/s	50p, 59.94p, 60p
4096x2160	10	SMPTE ST 2048-1	Dual 3 Gb/s	23.98p, 24p, 25p, 29.97p, 30p
		SMPTE ST 2048-1	Quad 3 Gb/s	47.95p, 48p, 50p, 59.95p, 60p

**1) Progressive*

Table 6 – Supported Video Formats for 4:4:4 RGB

Resolution	bit	Standard	Typical SDI Bit Rate	Frame Rate
1920x1080	10	SMPTE ST 274	3 Gb/s	50i, 59.94i, 60i
		SMPTE ST 274	3 Gb/s (P and PsF)	23.98p, 24p, 25p, 29.97p, 30p
		SMPTE ST 274	Dual 3 Gb/s	50p, 59.94p, 60p
1920x1080	12	SMPTE ST 274	3 Gb/s	50i, 59.94i, 60i
		SMPTE ST 274	3 Gb/s (P and PsF)	23.98p, 24p, 25p, 29.97p, 30p
		SMPTE ST 274	Dual 3 Gb/s	50p, 59.94p, 60p
2048x1080	10	SMPTE ST 2048-1	3 Gb/s (P and PsF)	23.98p, 24p, 25p, 29.97p, 30p
2048x1080	12	SMPTE ST 2048-1	3 Gb/s (P and PsF)	23.98p, 24p, 25p, 29.97p, 30p
3840x2160	10	SMPTE ST 2036-1	Quad 3 Gb/s	23.98p, 24p, 25p, 29.97p, 30p
3840x2160	12	SMPTE ST 2036-1	Quad 3 Gb/s	23.98p, 24p, 25p, 29.97p, 30p
4096x2160	10	SMPTE ST 2048-1	Quad 3 Gb/s	23.98p, 24p, 25p, 29.97p, 30p
4096x2160	12	SMPTE ST 2048-1	Quad 3 Gb/s	23.98p, 24p, 25p, 29.97p, 30p

6.2 Audio Payload Packing

The following items shall be contained in the Audio payload of the Essence datagram.

- Audio Group Information.
- Audio data portion in the Audio Data Packet specified in SMPTE 299 (24-bit audio) and SMPTE 272 (20-bit audio).

6.2.1 Audio Group Information

Figure 13 shows the structure of Audio Group Information which indicates whether an audio group is valid or invalid. When the specific audio group is valid, the correspondent field shall be set to 1, e.g. if audio group1 is valid, Gp1 field is set to 1. If invalid it shall be set to 0.

In the IP mapping, only audio groups which are set to valid shall be packed to the audio payload.

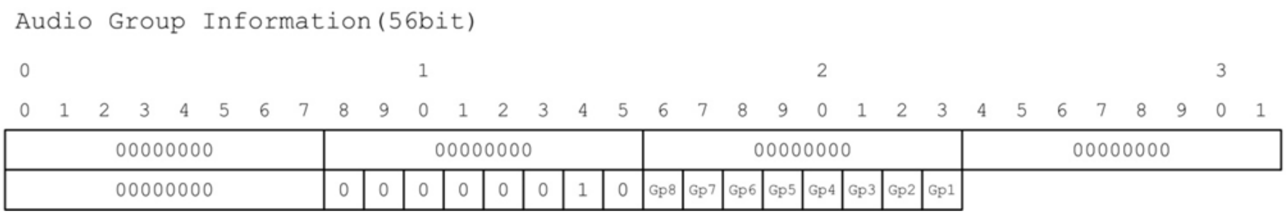

Figure 13 – Audio Group Information

6.2.2 24-bit Audio Packing

In the case of 24-bit audio packing for 4 channels (1 Audio Group), Audio Group Information shall be packed from the start of an audio payload followed by CH2 - CH1 - CH4 - CH3 audio data repeatedly. Adding zero-padding before Audio Group Information, between Audio Group Information and audio data, and between audio data shall not be allowed. Figure 14 shows an example.

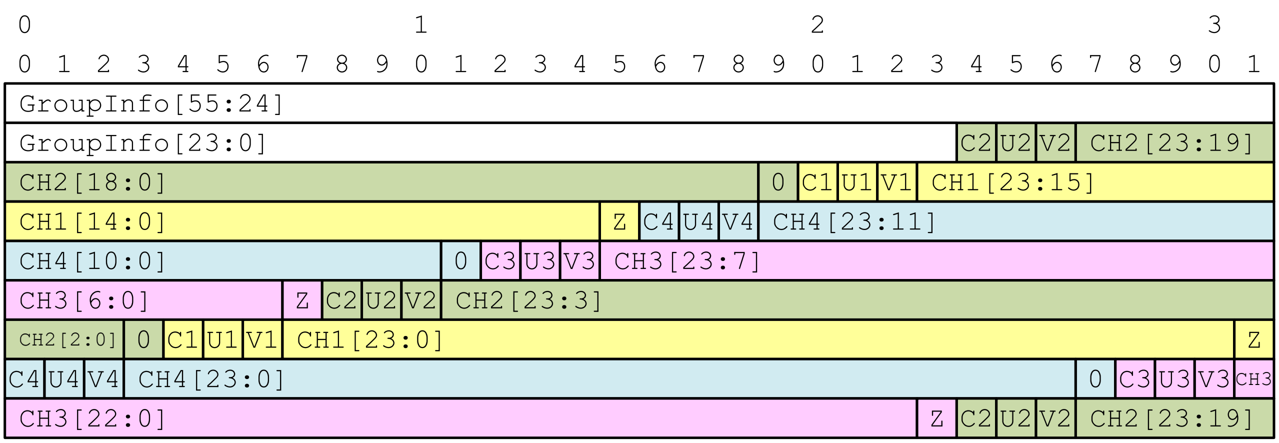

Figure 14 – Example of Audio Packing (24-bit/4ch)

Z = AES block sync.

Cn = AES channel status bit of CHn.

Un = AES user bit of CHn.

Vn = AES sample validity bit of CHn.

In the case of 24-bit audio packing for 8 channels (2 Audio Groups), Audio Group Information shall be packed from the start of an audio payload followed by CH2 - CH1 - CH4 - CH3 - CH6 - CH5 - CH8 - CH7 audio data repeatedly. Adding zero-padding before Audio Group Information, between Audio Group Information and audio data, and between audio data shall not be allowed. Figure 15 shows an example.

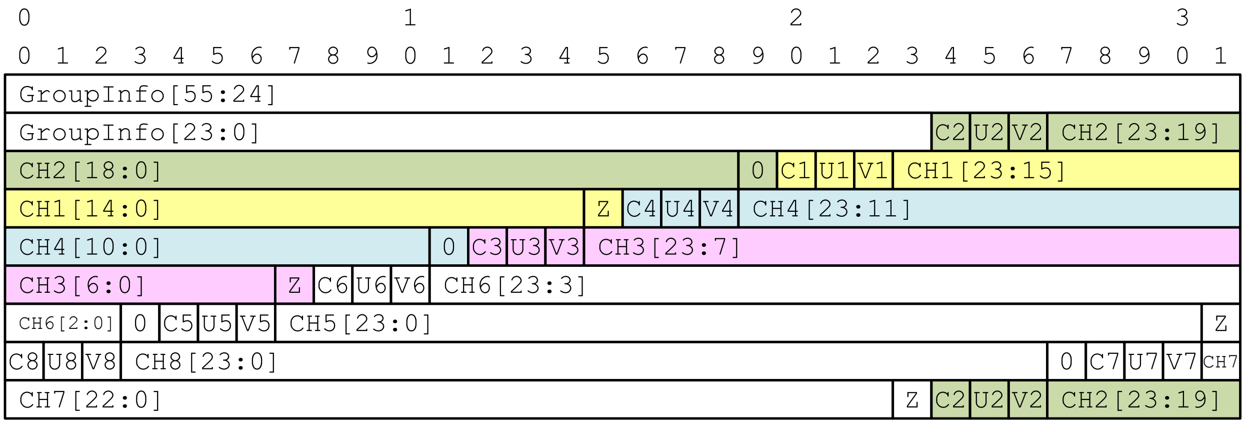

Figure 15 – Example of Audio Packing (24-bit/8ch)

In the case of 24-bit audio packing for 16 channels (4 Audio Groups), Audio Group Information shall be packed from the start of an audio payload followed by CH2 - CH1 - CH4 - CH3 - CH6 - CH5 - CH8 - CH7 - CH10 - CH9 - CH12 - CH11 - CH14 - CH13 - CH16 - CH15 audio data repeatedly. Adding zero-padding before Audio Group Information, between Audio Group Information and audio data, and between audio data shall not be allowed. Figure 16 shows an example.

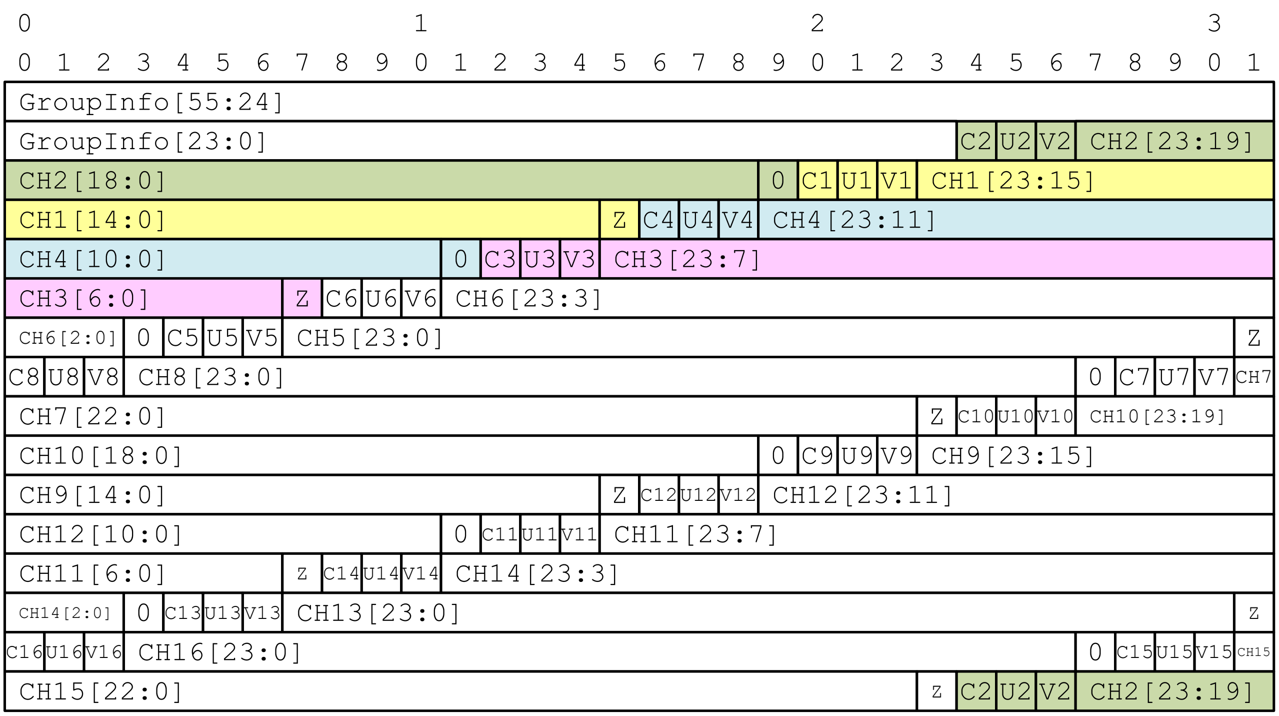

Figure 16 – Example of Audio Packing (24-bit/16ch)

6.2.3 20-bit Audio Packing

In the case of 20-bit audio packing for 4 channels (1 Audio Group), Audio Group Information shall be packed from the start of an audio payload followed by CH2 - CH1 - CH4 - CH3 audio data repeatedly. Adding zero-padding before Audio Group Information, between Audio Group Information and audio data, and between audio data shall not be allowed. Figure 17 shows an example.

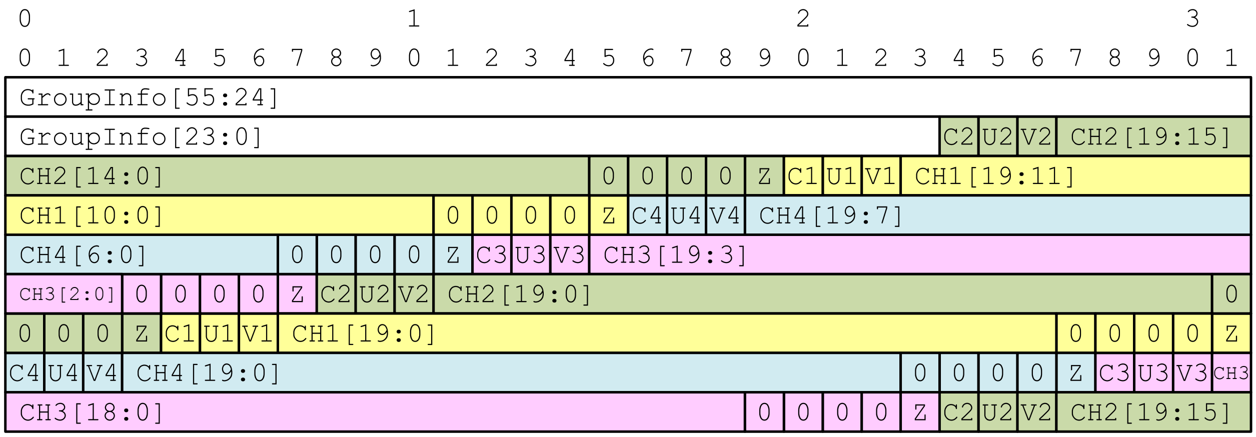

Figure 17 – Example of Audio Packing (20-bit/4ch)

In the case of 20-bit audio packing for 8 channels (2 Audio Groups), Audio Group Information shall be packed from the start of an audio payload followed by CH2 - CH1 - CH4 - CH3 - CH6 - CH5 - CH8 - CH7 audio data repeatedly. Adding zero-padding before Audio Group Information, between Audio Group Information and audio data, and between audio data shall not be allowed. Figure 18 shows an example.

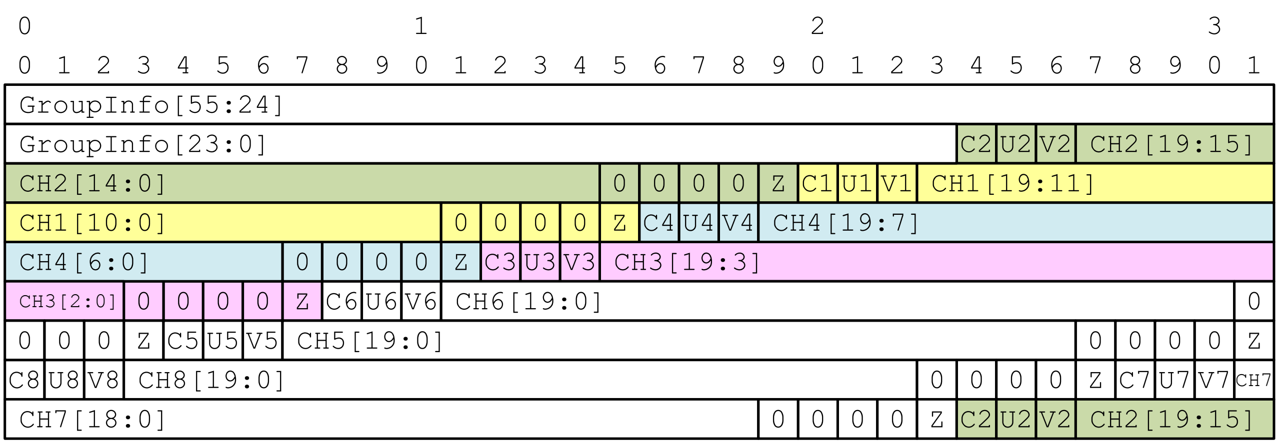

Figure 18 – Example of Audio Packing (20-bit/8ch)

In the case of 20-bit audio packing for 16 channels (4 Audio Groups), Audio Group Information shall be packed from the start of an audio payload followed by CH2 - CH1 - CH4 - CH3 - CH6 - CH5 - CH8 - CH7 - CH10 - CH9 - CH12 - CH11 - CH14 - CH13 - CH16 - CH15 audio data repeatedly. Adding zero-padding before Audio Group Information, between Audio Group Information and audio data, and between audio data shall not be allowed. Figure 19 shows an example.

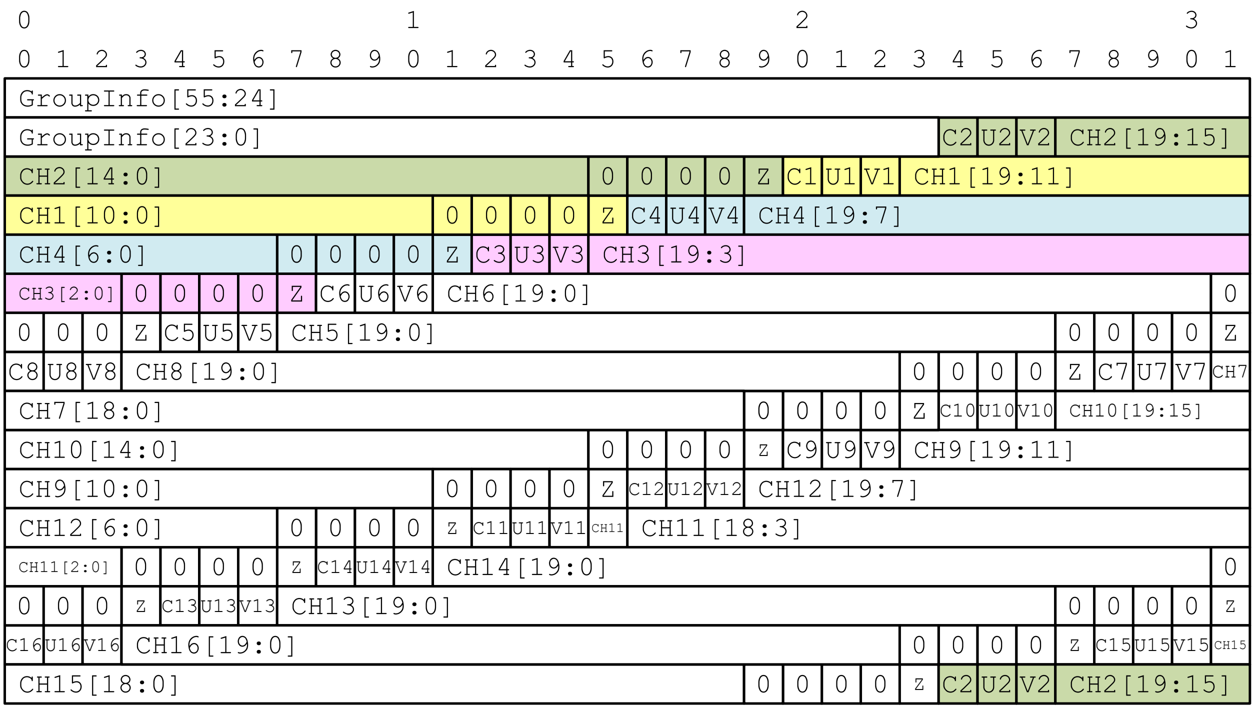

Figure 19 – Example of Audio Packing (20-bit/16ch)

6.3 Ancillary Payload Packing

Ancillary Data Packet type 2 specified in SMPTE ST 291-1 shall be used for packing to an Ancillary payload in the Essence datagram with the following exception:

The first 3-word Ancillary Data Flags (000h 3FFh 3FFh) shall be replaced with the following items respectively:

- 3FFh
- PIW0 (Position Information Word 0)
- PIW1 (Position Information Word 1)

PIW0 and PIW1 indicate the place where a specific Ancillary data is obtained such as:

- H Ancillary or V Ancillary
- 3G-SDI Link number
- Line number

Figure 20 defines the PIW0 and PIW1. The detailed information is specified in Table 7. The relationship between Link Number and Data Stream (see SMPTE ST 425-1, SMPTE ST 425-3 and SMPTE ST 425-5) is shown in Figure 21.

<table>
<tr><td>9</td><td>8</td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td><td>0</td><td>9</td><td>8</td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td><td>0</td></tr>
<tr><td colspan="10">PIW0</td><td colspan="10">PIW1</td></tr>
<tr><td>0</td><td>V/H</td><td colspan="3">Link Number</td><td colspan="5">Line Number [13-9]</td><td>0</td><td colspan="9">Line Number [8-0]</td></tr>
</table>

Figure 20 – Definition of PIW0 and PIW1

Table 7 – PIW0 and PIW1 Description

Field Name	Size (bit)	Description		
V/H	1	1: V Ancillary, 0: H Ancillary		
Link Number	3	0	3G-SDI Link 1	Data Stream One
		1		Data Stream Two
		2	3G-SDI Link 2	Data Stream Three
		3		Data Stream Four
		4	3G-SDI Link 3	Data Stream Five
		5		Data Stream Six
		6	3G-SDI Link 4	Data Stream Seven
		7		Data Stream Eight
Line Number	14	Line number		

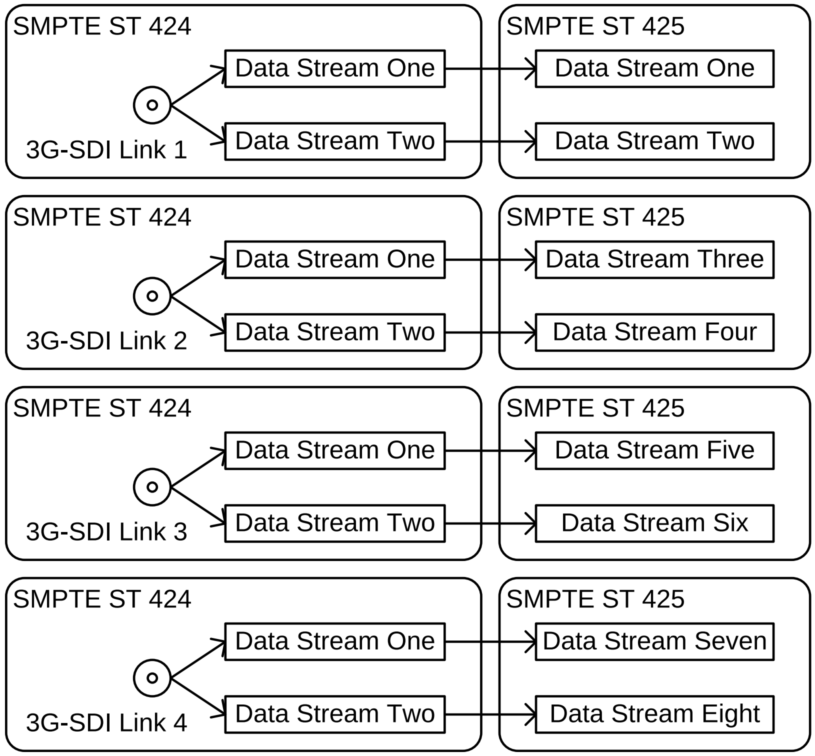

Figure 21 – Relationship between Link Number and Data Stream

Figure 22 shows an example of Ancillary data packing.

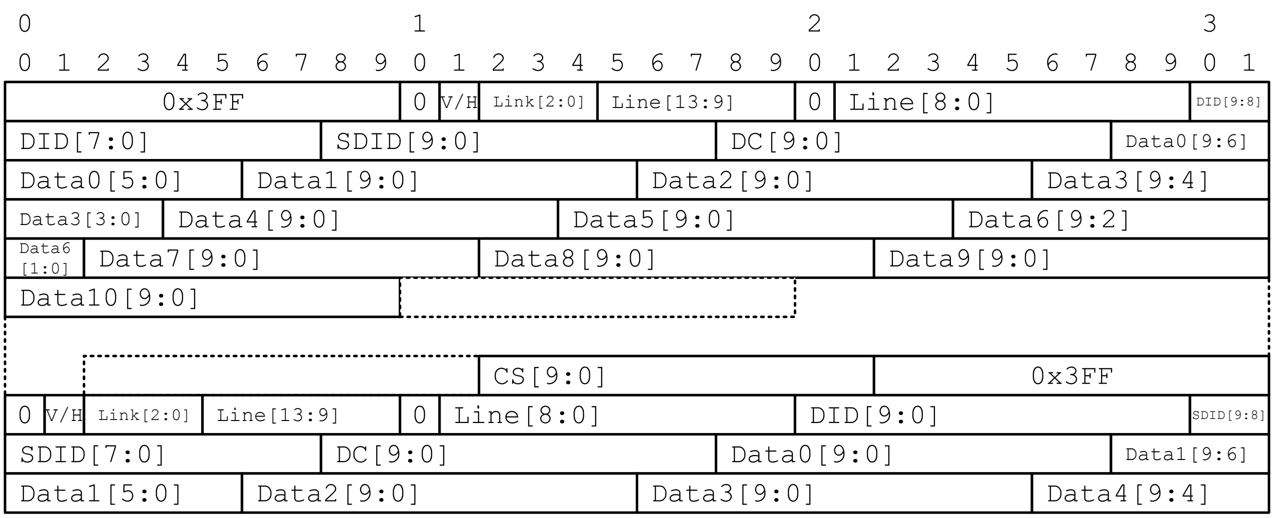

Figure 22 – Example of Ancillary Data Packing

DID = Data Identification

SSDI = Secondary Data Identification Word

DC = Data Count Number

CS = Checksum

7 Essence Synchronization

In the IP mapping, Frame Count in the Essence Header shall be used for synchronization between essences. As described in Section 4.1, all input A/V signals shall be synchronized so that if Frame Count is added to the mapped input A/V signals (Essence datagrams) at a sender, it can be used for synchronizing datagrams at a receiver. Frame Count is used whether or not a video stream is present, as defined in Table 1.

As the initial value shall be set to the specific value such that Frame Count is 0 at the Epoch defined in SMPTE ST 2059-1, the alignment of Frame Number is absolute. As a result, the phases of any essences from various sources can be uniquely identified for differences of up to 127 frames (or fields).

In addition to synchronize essences, Frame Count can be used for detecting frame (or field) boundaries so that clean video switching between two video streams can be easily realized.

8 Hitless Failover

The IP mapping supports a hitless failover function. In order to perform hitless failover, a stream sender shall transmit a pair of redundant streams which consist of a pair of redundant datagrams. RTP Datagram, consisting of RTP Header and RTP Payload, shall be the same for each pair of redundant datagrams.

In addition, Identification field in the IP header shall be used. Figure 23 shows the definition of Identification field. Packet Sequence ID is a 13-bit number which is incremented in successive datagrams of the same stream. Packet Sequence ID runs from zero to 8191(dec) and returns to zero when exceeding 8191(dec). Only bit1 is used in Flag. When hitless failover is intended to be performed, the bit1 shall be set to 1.

A stream receiver performs the hitless failover process only when the bit1 is set to 1. As the stream receiver does not use memory and passes a datagram to the downstream module immediately on receipt of one of the redundant streams, there is no delay.

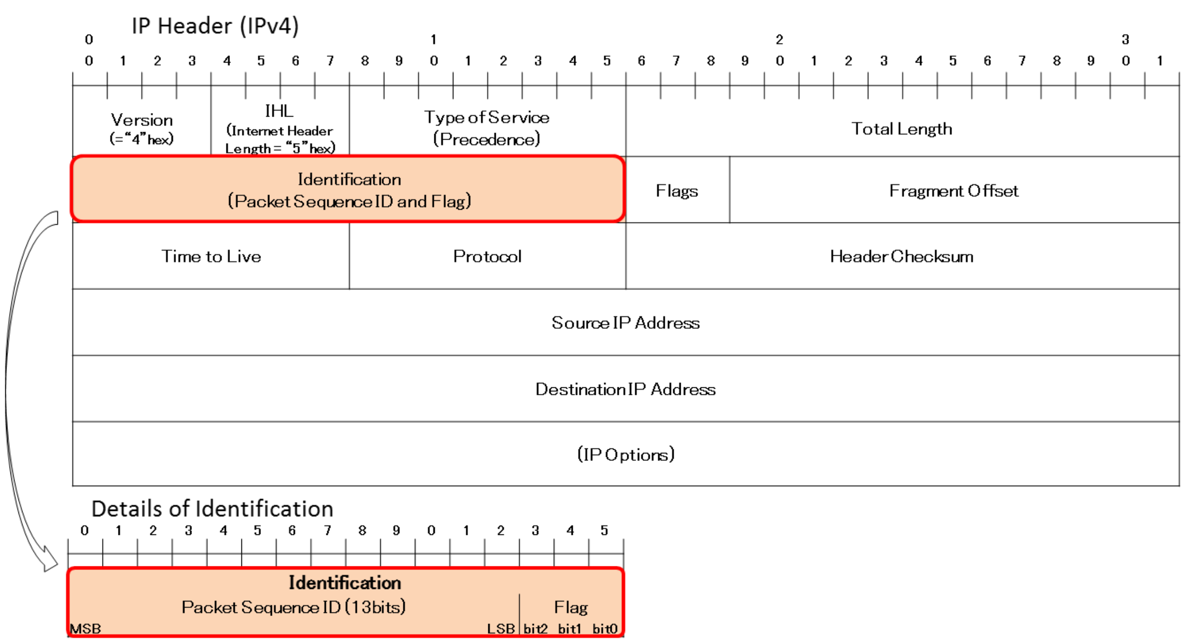

Figure 23 – Definition of Identification