

EBU – Recommendation R 123



EBU Audio Track Allocation for File Exchange

Recommendation

Source: P/AGA & EHDF Audio Groups

Geneva
July 2009

Contents

1. Introduction	5
2. Recommendation	5
3. Metadata	6
4. Technology.....	6
5. Bibliography	9
Annex: Clarification of audio terms used in this document.....	11
Table 1: Audio Track Allocations for International File Exchange	7

Page intentionally left blank. This document is paginated for recto-verso printing

EBU Audio Track Allocation for International File Exchange

<i>EBU Committee</i>	<i>First Issued</i>	<i>Revised</i>	<i>Re-issued</i>
PMC	2009		

Keywords: Track Allocation, Multichannel Audio, File Exchange, Metadata

1. Introduction

In the past, audiovisual programme material was exchanged by the physical transfer of tapes and tape cassettes containing video and audio tracks. As tape formats developed, the number of tracks that could be accommodated on the tape increased, and the situation became further complicated when compressed, multichannel audio formats were also recorded onto tape alongside linear audio tracks. The EBU addressed this environment with its Recommendation R 48 [1], which went through several revisions over the course of a few years. Note that R 48 is not replaced by the present recommendation.

If programmes are to be exchanged with North American broadcasters, note should be taken of, SMPTE 2035 [2], which specifies track allocations up to eight tracks for VTR formats.

This recommendation addresses current production practice, where audiovisual material is commonly exchanged as files. These are stored on suitable IT-based data carriers (servers, removeable hard drives, optical disks etc.). Files are exchanged either by the physical transfer of the data carriers or they are transferred over broadband connections using suitable file transfer (or streaming) protocols.

The crucial point is that audiovisual files have none of the physical restrictions imposed by tape formats, so the potential number of audio track arrangements in a file is huge. Obviously, to avoid operational confusion during the international exchange of programmes, the allocation of audio channels must be unequivocal.

The main focus of this recommendation is the delivery and exchange of audiovisual programme material in a file based environment. In addition precautions have been made that this document is applicable for the compatible use of both non-multichannel programmes and multichannel programmes. This reflects the operational requirements of the EBU members in an increasingly IT-based production environment.

2. Recommendation

The EBU recommends that for the international exchange of programmes, the audio channels within all digital audiovisual files be allocated as shown in Table 1 for up to 16 channels. This table will be extended as necessary when multichannel requirements exceed 16 channels.

This recommended allocation does not preclude other, unlisted track allocations for in-house applications of the receiving organisation, but these ad-hoc arrangements must be specified by prior agreement between the two parties concerned.

3. Metadata

It is to be expected that audiovisual files will have sufficient Metadata embedded in them and/or associated with them to describe their Content to the production and transmission systems with which they will interact. Standards such as the MXF suite of standards, AES52, EBU Tech 3285 (BWF) and EBU Tech 3306 (RF64) fully document this aspect of the use of files.

4. Technology

This is a listing of representative technologies used to bear multi-channel audio streams and associated metadata. This is not a comprehensive list.

Technology	Audio Channels
Hdcam VTR	4
270 Mbit/s ASI-mux Link	4/8
Dolby E (16 bit)	6
Dolby E (20 bit)	8
HdcamSR VTR	12
E ² (Linear Acoustics)	16
SD-SDI	16
HD-SDI (1.5 Gbit/s)	16
HD-SDI (1.5 Gbit/s) Dual link	16 x 2
<i>HD-SDI (3 Gbit/s)</i>	<i>32(Proposed)</i>

Table 1: Audio Track Allocations for International File Exchange

R 123 Ref.	No. of Tracks	Programme Type	Notes	Origin (used by)	AUDIO TRACK NUMBER															
					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2a	2	Encoded MCA	4	[1] 10a	encoded MCA															
4a	4	Ⓢ with encoded MCA	1, 4, 6	[1] 11a	Ⓢ L	Ⓢ R	encoded MCA													
4b	4	Ⓢ with IT or EFX		[1] 2b (ARD, ZDF, ORF)	Ⓢ Mix L	Ⓢ Mix R	IT or EFX L	IT or EFX R												
4c	4	Ⓢ with AD or FL		(ARD, ZDF, ORF)	Ⓢ L	Ⓢ R	AD/FL L	AD/FL R												
8a	8	Ⓢ with discrete MCA	1, 2	[1] 11b	Ⓢ L	Ⓢ R	MCA L	MCA R	MCA C	MCA LFE	MCA L Sur	MCA R Sur								
8b	8	Discrete MCA with Ⓢ	1, 3, 5	[1] 11c	MCA L	MCA R	MCA C	MCA LFE	MCA L Sur	MCA R Sur	Ⓢ L	Ⓢ R								
8c	8	Discrete MCA with 2 x Ⓞ	1, 3, 5, 6	[1] 11c*	MCA L	MCA R	MCA C	MCA LFE	MCA L Sur	MCA R Sur	Ⓞ	Ⓞ								
8d	8	2 x Ⓢ with 2 x encoded MCA	2, 4, 7	(SKY)	Ⓢ Mix L	Ⓢ Mix R	Ⓢ EFX L	Ⓢ EFX R	encoded MCA Mix		encoded MCA EFX									
8e	8		4, 7		Ⓢ Mix L	Ⓢ Mix R	encoded MCA Mix		Ⓢ EFX L	Ⓢ EFX R	encoded MCA EFX									
8f	8	Ⓢ with M/S and/or B-format Ambisonics	8		Ⓢ L	Ⓢ R	ⓄM	ⓄS	W	X	Y	Z								
8g	8	Ⓢ with AD or FL and IT and encoded MCA		(ARD, ZDF, ORF)	Ⓢ Mix L	Ⓢ Mix R	AD/FL L	AD/FL R	IT L	IT R	encoded MCA									
8h	8	Ⓢ with encoded MCA and IT and AD/FL		(ARD, ZDF, ORF)	Ⓢ Mix L	Ⓢ Mix R	encoded MCA Mix		IT L	IT R	AD/FL L	AD/FL R								
8i	8	Ⓢ with IT/ME and IT and encoded MCA		(ARD, ZDF, ORF)	Ⓢ Mix L	Ⓢ Mix R	IT/ME L	IT/ME R	IT L	IT R	encoded MCA Mix									
12a	12	Ⓢ with discrete MCA + encoded MCA + monos/Ⓢ	1, 2		Ⓢ L	Ⓢ R	MCA L	MCA R	MCA C	MCA LFE	MCA L Sur	MCA R Sur								
12b	12		1, 3, 5	(SVT)	MCA L	MCA R	MCA C	MCA LFE	MCA L Sur	MCA R Sur	Ⓢ L	Ⓢ R								
12c	12	2x Ⓢ with 2x encoded MCA + extra Ⓞs/Ⓢs	2, 4, 7		Ⓢ Mix L	Ⓢ Mix R	Ⓢ EFX L	Ⓢ EFX R	encoded MCA Mix		encoded MCA EFX									
12d	12		4, 7		Ⓢ Mix L	Ⓢ Mix R	Encoded MCA Mix		Ⓢ EFX L	Ⓢ EFX R	encoded MCA EFX									

R 123 Ref.	No. of Tracks	Programme Type	Notes	Origin (used by)	AUDIO TRACK NUMBER															
					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
12e	12	Discrete MCA + extra \textcircled{M} s/ \textcircled{S} + encoded MCA & encoded M&E	2, 3, 5	(SVT)	MCA L	MCA R	MCA C	MCA LFE	MCA L Sur	MCA R Sur	\textcircled{S} L or \textcircled{M}	\textcircled{S} R or \textcircled{M}	encoded MCA incl. audio from ch 1 - 8 & Dolby Metadata	encoded MCA M&E sound & Dolby Metadata						
12f	12	\textcircled{S} with encoded MCA and IT and discrete MCA		(ARD, ZDF, ORF)	\textcircled{S} Mix L	\textcircled{S} Mix R	encoded MCA	IT L	IT R	MCA L	MCA R	MCA C	MCA LFE	MCA L Sur	MCA R Sur					
12g	12	\textcircled{S} with IT + discrete MCA + encoded MCA		(ARD, ZDF, ORF)	\textcircled{S} Mix L	\textcircled{S} Mix R	IT L	IT R	MCA L	MCA R	MCA C	MCA LFE	MCA L Sur	MCA R Sur	encoded MCA					
12h	12			(France TV)	\textcircled{S} Mix L Lang 1	\textcircled{S} Mix R Lang 1	encoded MCA Language 1	\textcircled{S} Mix L Lang 2	\textcircled{S} Mix R Lang 2	encoded MCA Language 2	\textcircled{S} Mix L Lang 3	\textcircled{S} Mix R Lang 3	encoded MCA Language 3							
16a	16	2x \textcircled{S} with 2x discrete MCA	2, 7		\textcircled{S} L Mix	\textcircled{S} R Mix	MCA L Mix	MCA R Mix	MCA C Mix	MCA LFE Mix	MCA L Sur Mix	MCA R Sur Mix	\textcircled{S} EFX L	\textcircled{S} EFX R	MCA L EFX	MCA R EFX	MCA C EFX	MCA LFE EFX	MCA L Sur EFX	MCA R Sur EFX
16b	16		3, 7		MCA L Mix	MCA R Mix	MCA C Mix	MCA LFE Mix	MCA L Sur Mix	MCA R Sur Mix	\textcircled{S} Mix L	\textcircled{S} Mix R	MCA L EFX	MCA R EFX	MCA C EFX	MCA LFE EFX	MCA L Sur EFX	MCA R Sur EFX	\textcircled{S} EFX L	\textcircled{S} EFX R
16c	16		2, 7	(SKY)	\textcircled{S} Mix L	\textcircled{S} Mix R	\textcircled{S} EFX L	\textcircled{S} EFX R	MCA L Mix	MCA R Mix	MCA C Mix	MCA LFE Mix	MCA L Sur Mix	MCA R Sur Mix	MCA L EFX	MCA R EFX	MCA C EFX	MCA LFE EFX	MCA L Sur EFX	MCA R Sur EFX
16d	16	Dual Language: 2x discrete MCA + 2x \textcircled{M} / \textcircled{S} or AD (dual language)		(France TV)	Language 1						Additional \textcircled{M} or \textcircled{S} , e.g. AD - Lang 1	Language 2						Additional \textcircled{M} or \textcircled{S} , e.g. AD - Lang 2		
16e	16			(France TV)	MCA L	MCA R	MCA C	MCA LFE	MCA L Sur	MCA R Sur		MCA L	MCA R	MCA C	MCA LFE	MCA L Sur	MCA R Sur		Additional \textcircled{M} or \textcircled{S} , e.g. AD - Lang 1	Additional \textcircled{M} or \textcircled{S} , e.g. AD - Lang 2
16f	16	Triple Language: 3x \textcircled{S} s + 3x MCA + 2x \textcircled{M} / \textcircled{S} or AD		(France TV)	\textcircled{S} Mix L Lang 1	\textcircled{S} Mix R Lang 1	Encoded MCA Lang 1	\textcircled{S} Mix L Lang 2	\textcircled{S} Mix R Lang 2	Encoded MCA Lang 2	\textcircled{S} Mix L Lang 3	\textcircled{S} Mix R Lang 3	Encoded MCA Lang 3	Additional \textcircled{M} or \textcircled{S} , e.g. AD	Additional \textcircled{M} or \textcircled{S} , e.g. AD					
	(32+)				Work to be done when more experience is gained of >16-channel systems.															

*It is felt that this version of EBU R 48-2005 variant 11c should be a discrete and different variant.

Key to table abbreviations:

\textcircled{M} = Mono, \textcircled{S} = Stereo, MCA = Multichannel Audio, EFX = Effects, Sur = Surround, IT = International Sound, AD = Audio Description, FL = Foreign Language, M&E = Music and Effects (no dialogue), WXYZ = Ambisonic B-format channels (the W channel is the non-directional mono component of the signal, the X, Y and Z channels are the directional components in three dimensions).

General notes:

Audio should be 48 kHz, 24 bit, and time-aligned with picture on tape, file, connection or link [3].

Encoded MCA shall also be time-aligned with picture on tape, file, connection or link and not advanced or retarded to account for down-stream processing delays.

Whichever format is used should be clearly indicated in any recording report or other associated metadata.

NOTES:

- 1 *The stereo should be related to the multichannel mix, and can be either a downmix from the multichannel sound or a separate stereo balance.*
- 2 *This format supports compatibility with legacy stereo environments by placing the stereo version(s) first.*
- 3 *The main use of this format is to support compatibility with encoded transport streams (such as 20 bit Dolby E) audio track allocation for multi-track recordings, to that recommended for sound-only programmes, EBU Technical Recommendation R91-1998 [4] where the first 6 channels are the multichannel mix and 7 & 8 are the stereo mix.*
- 4 *Some forms of compression have different options for the audio coding (e.g. 16 versus 20 bit coding). It is essential that this choice be noted on the labels and recording report.*
- 5 *In some circumstances neither a specific stereo mix nor an automated downmix will exist during the early stages of the production process. Under these circumstances, tracks 7 and 8 would be unused.*
- 6 *Tracks 7 & 8 could be used for alternative commentary(s), different language tracks or audio description, in mono or stereo*
- 7 *This track allocation provides a main mix and a clean effects version, and is based on common operating practice, particularly in sport.*
- 8 *This format is to support B-format Ambisonics. It is expected to be useful for archiving, clean effects and as a production resource.*
- 9 *Tracks 11 & 12 could be used for alternative commentary(s), different language tracks or audio description, in mono or stereo.*

5. Bibliography

- | | |
|----------------------------------|---|
| [1] EBU Recommendation R 48-2005 | Allocation of audio tracks on digital television recorders |
| [2] SMPTE 2035 | Audio Channel Assignments for Digital Television Recorders (DTRs) |
| [3] EBU Recommendation R 37-2007 | The relative timing of the sound and vision components of a TV signal |
| [4] EBU Recommendation R 91-1998 | Track allocations and recording levels for the exchange of multichannel audio signals |

Annex: Clarification of audio terms used in this document

1. Full Mix / Complete Mix

SMPTE Definition:

A mix consisting of all the elements (e.g. music, effects and dialogue) required to form a standalone audio programme.

Comment: This is as used in the ARD, ZDF, ORF and SRG.

2. International Sound

This term has become very difficult to define, as television programme production has increased in sophistication and complexity. In addition, the term 'International Sound' is often used interchangeably with 'Clean FX', particularly in the context of live sports OBs where the name may have originated, but the two terms (used to) mean quite different things.

2.1 Conventional Definition:

All the audio elements of a programme that appear in-vision.

The elements that are included will depend on what is in the accompanying vision (this is an essential difference to other forms of Music&Effect tracks). E.g. in a simple programme, the video transmission could consist of exactly what was captured on camera, in which case the 'International Sound' includes everything in the programme except the commentary (since that's 'out-of-vision').

2.2 SMPTE Definition:

A mix consisting of all the elements required to form a standalone audio programme, except for the dialogue element. The main idea is that a local commentary could easily be added to an 'anchor' programme in this way.

2.3 Revised Definitions:

In current practice, any audio mix that is not the main programme audio could consist of a combination of various elements, and so 'International Sound' has become a more loosely defined term. Frequently, 'International Sound' is a term that must be accompanied by a qualifying explanation, often genre-dependent. So, for example, there are different versions of 'International Sound' encountered.

2.3.1 International Sound - Documentary M&E:

All the audio elements except commentary (narration) or voice-over.

A documentary M&E would include sync (anyone talking in-vision). This would continue if that person went out of vision. These are "un-dipped" i.e. the level of the mix is constant so if the new narration is a different length the mix can be redone properly. This explanation fits the meaning of ARD, ZDF, ORF and SRG for the term International Sound. Even the SMPTE definition may fit into it (A mix consisting of all the elements required to form a standalone audio programme, except for the dialogue element).

2.3.2 International Sound - Fiction or Drama M&E (aka "footsteps"):

Audio containing no dialogue whatsoever. Often termed "footsteps" because new effects are applied when actors are walking and talking off-mic, so Foley artists are needed to generate them.

Comment: This explanation fits the meaning of ARD, ZDF, ORF and SRG for the term Music&Effects.

2.3.3 International Sound - Clean FX:

Full event effects coverage, including continuous crowd atmosphere, typically used when the host-broadcaster has cut away from the event venue to its presentation studio (or advertising breaks etc.).

This audio would generally be accompanied by a vision feed (often confusingly termed the 'World Feed') which shows the host-broadcaster's coverage of the event during the action, but which cuts to a wide-shot of the venue when the host-broadcaster does other things such as VT playbacks, interviews and studio discussion.

2.3.4 International Sound - World Feed:

In these days of international TV OBs the 'World Feed' vision is generally assembled by a second vision mixer (both animal and mineral) and a separate sound mixer supplies the audio to match the pictures.

This 'World Feed' audio may include some of the host-broadcaster's VT play-ins and some of their interviews. Therefore all the necessary elements need to be made available as separate 'stems' (see definition below) to the mixers.

3. Terms used in an international environment

Various versions of the following elements for all internationally exchanged programmes may be delivered. Because of the demands of surround sound, the BWF/RF64 may be used for virtually all but the Full Mix.

3.1 E track:

An audio track that just consists of effects (two versions as above Documentary E track or Fiction/Drama E track).

3.2 M track:

An audio track that just consists of music as used in the final programme pre mixed but un-dipped.

3.3 Stems:

Clean sections of audio that allow sections of a programme to be re-cut.

Stems are almost completely un-mixed, so that music that dips under effects in the Full Mix would be full level on the music stem track. Similarly, if two pieces of music overlap in the Full Mix, there would be stem tracks for each piece with no fade between them and they would be extended, allowing new transitions to be added.