Dolby® E and Master Control

WHITEPAPER

Introduction

Master control with traditional broadcast signals has presented relatively few difficulties for programme playout. Video and Audio in SDI and AES formats are linear signals. Once a small amount of timing adjustment has been applied, each signal is processed sample by sample. This introduces little delay in the system, perhaps a line or so for the video and a few samples for the audio. Any relative timing changes between video and audio are minimal.

With the advent of surround sound, Dolby E has become the de-facto standard for encoding, storing and routing audio. This paper discusses some of the implications of Dolby E, and some possible solutions.

So what is Dolby E?

Dolby E is a proprietary format for audio compression, developed by Dolby Laboratories. It allows for the compression of several linear PCM (pulse code modulation) audio streams into one AES3 stream. There are up to eight available channels in the Dolby encoding scheme, which allows for some flexibility, but typically it is used to generate a surround sound channel consisting of 5 'mono' signals and a low frequency (LFE) channel. This is commonly referred to as 5.1 audio. The other two channels may be used for a stereo mix of the same audio.

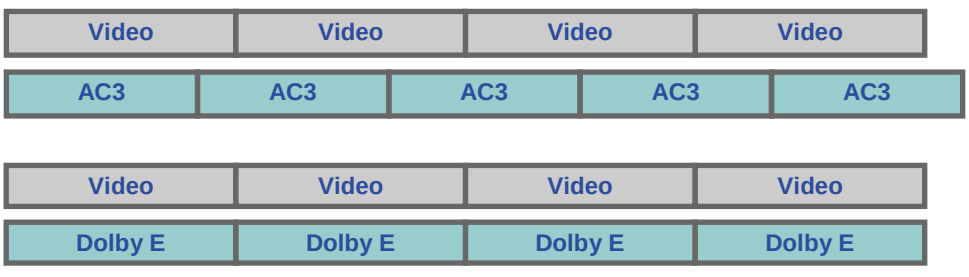
An alternative is to generate additional surround sound channels, and encode 7 surround channels plus an LFE channel. This is known as 7.1.

Once encoded, the Dolby E data is placed in the audio bits of the AES3 bit-stream, and the stream can be stored and routed as a normal AES signal.

Dolby E and Dolby AC3

What's the difference? In terms of how the audio is processed, not a lot. The audio is sampled and broken into time blocks. Each block is then processed to determine the appropriate compression. The encoded audio then has metadata attached, and is formatted as an AES stream.

The difference is in the timing relationship with video. AC3 packets are not related to video, but Dolby E packets are aligned to video frames, set by the reference signal fed into the encoder. A small part of the Dolby E stream (around the frame blanking area of video) has no Dolby E data, therefore the signal can be switched at this point without corrupting the Data in the Dolby E packet.



Dolby encoders will produce Dolby E packets aligned to the video reference. However, there is no guarantee that the Dolby packet will remain aligned to video in a system. Processing delays for video and audio are different, and it is possible for the two signals to become misaligned.

Dolby In Master Control

In a master control environment, the main concern is that the normal process of audio mixing, voiceovers etc. cannot be applied to a Dolby E stream. Any change to the content of the Dolby packet will corrupt it, resulting in a complete loss of output (Dolby decoders will normally mute the output if the data is corrupt). Leaving your master control mixer at unity gain and cutting the audio may also not be possible. Most DSP chips used in audio mixing will have a gain of 0.99999999, not 1.0000000. i.e. they multiply the signal content by a value, rather than passing it through the mixer untouched. The tiny gain change is not noticeable for PCM audio, but will result in corrupted Dolby data.

It is also possible that a programme being transmitted with Dolby E sound may have adverts which have a PCM stereo soundtrack. In this case, a normal master control mix poses some problems, as it's not possible to mix directly from Dolby to PCM (and vice versa).

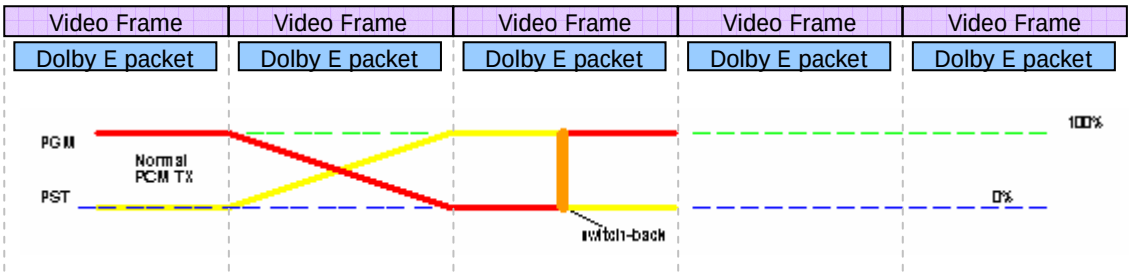
One solution is to have an audio mixer with the ability to detect AES3 PCM audio or Dolby E on its input, and process the two signals differently, depending on the intended mix. As no mixing or gain adjustment is possible with Dolby signals, all commands of this type (V fades, crossfades, etc) are interpreted as cuts. Although the transitions are not ideal, the user does not have to change the system configuration or make any changes due to the presence of Dolby signals. It also means the delay in the audio path is minimised. The mixer delay is typically 8 samples, which will not result in lip-sync problems.

The following examples show the effect of some different transition types when processed in this way. It is assumed that the Dolby E packets are aligned to the video frames at the input of the mixer in all cases:

KEY: **Red = PGM PCM** **Yellow = PST PCM** **Blue = any Dolby-E signal**

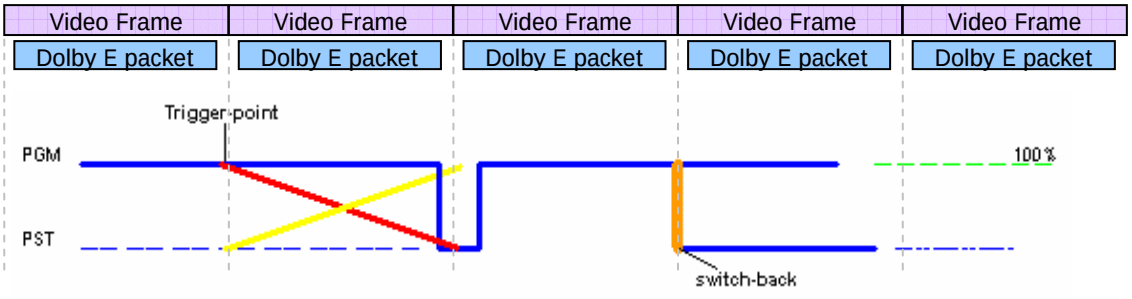
PCM to PCM Crossfade

A PCM to PCM crossfade transition will be implemented as normal.

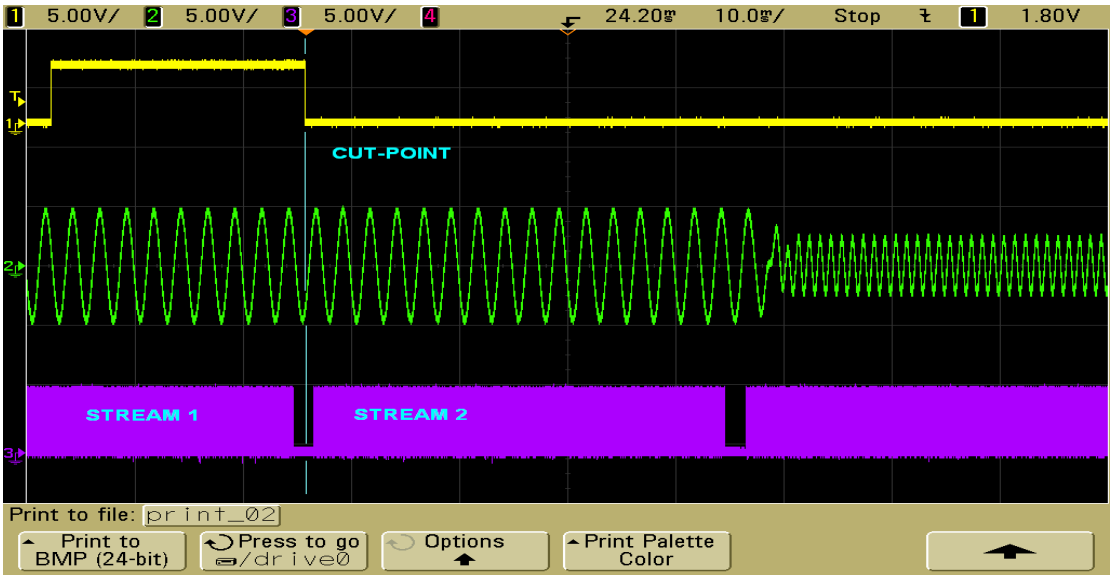


Dolby to Dolby Crossfade

A Dolby E – Dolby E crossfade will be treated as a Cut. The aim is to switch streams at the end of the transition.



The result of this when seen on the output of a Dolby Decoder is shown below:



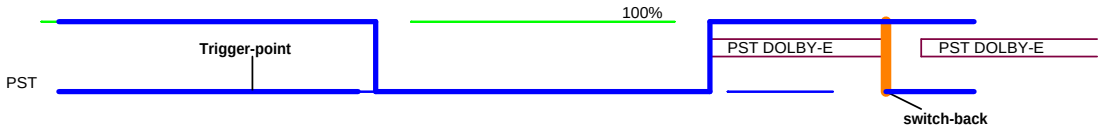
The falling edge of yellow trace indicates where the Program co-efficient reaches zero *& the end of packet has passed*. The purple shows the decoded AES containing Dolby-E data out of the mixer, the gaps being the guard bands, in sync with the Vertical interval of the video reference used to time the system.

The green is the decoded audio from a Dolby-decoder on the mixer output.

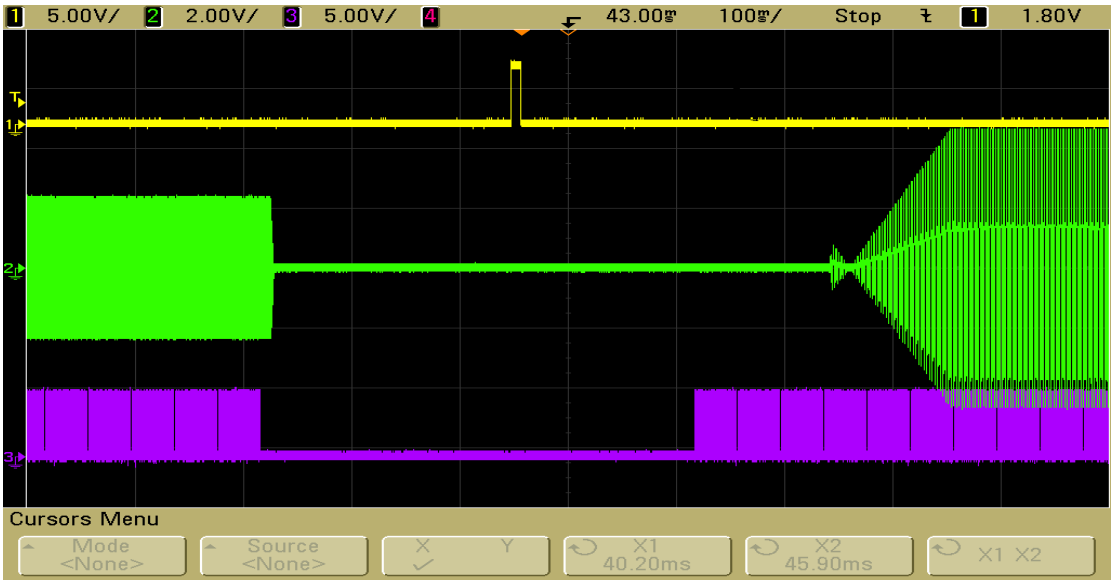
Note the audio changes a frame later than the mixer-switch, as the decoder buffers a whole video-frames worth of audio before decoding it. The Dolby decoder actually implements a fast cross-fade between the two signals encoded in the different packets either side of the switch. It generally takes about 5 milliseconds to accomplish this process. This is also the same for Cut-operations on co-timed signals.

Dolby to Dolby Transitions

A Dolby E – Dolby E ‘V fade’ will be treated as a cut. When a command to V-fade is received, the mixer will cut the Program input, and the mixer output will remain in this state until the intended transition is complete (i.e. the Preset bus should be at 100%). This means the audio output will be silent for the duration of the V-fade.



Again, the final output after Dolby decoding is modified slightly by the Decoder processing:

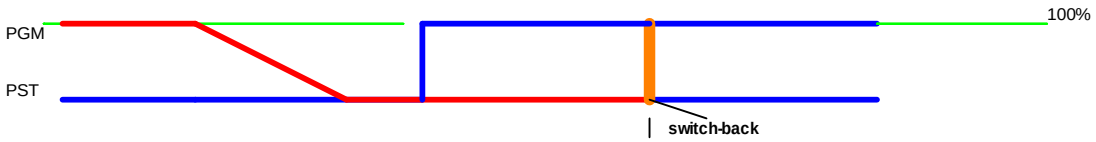


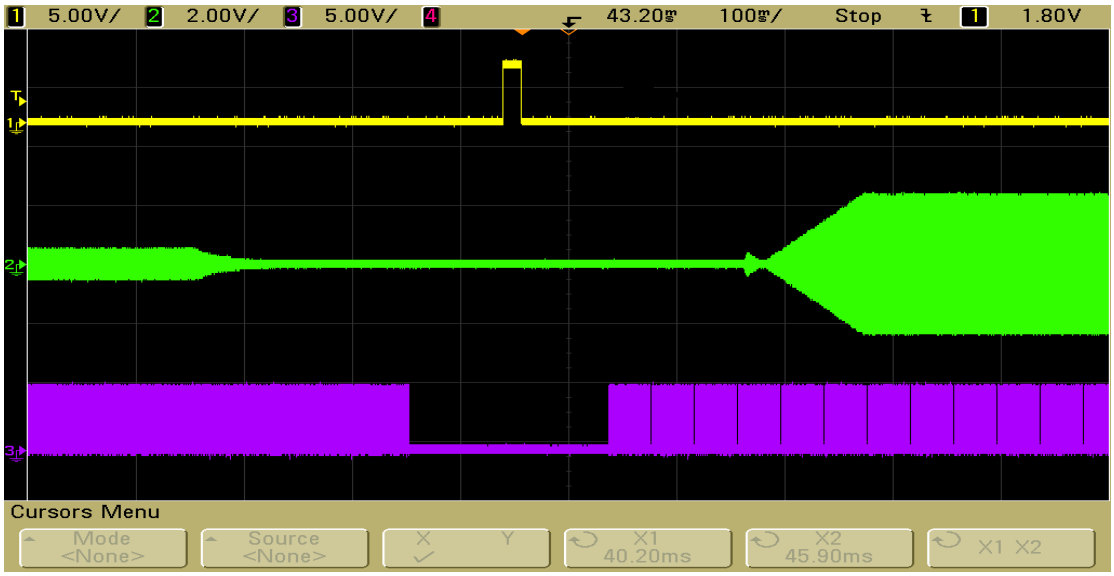
Note here that there is a section in the middle with all audio data muted.

This is effectively a fade to silence followed by a second fade from Silence to program-content. Because the new stream appears as a new “unknown” format to the decoder, the output of the decoder is muted, and it then performs a ramp up to full amplitude over two frames after the first full new frame. It is not possible to do a V-fade with less than at least 4 frames of silence. (2 for the decoder-recovery plus n-frames for the fade rate plus up to 1 extra frame for the end of the current packet).

PCM to Dolby transitions

For PCM to Dolby E transitions, the process is a little more complex. The fade down of the program bus will take place, and the mixer will wait until the Preset bus would be full amplitude. The Dolby E Preset bus is then cut in at the next packet boundary. Dolby decoders will recognise this as a new Dolby E stream and will perform a 2-frame fade to full amplitude. This will sound very similar to a fade & take transition, or fast V-fade, depending on the fade rate setting.

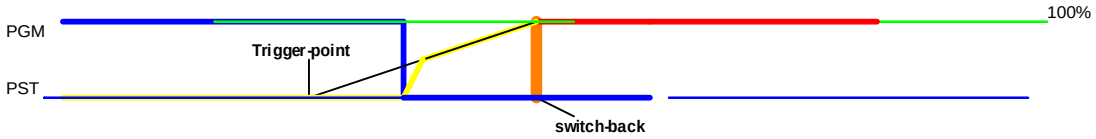




This shows the PCM being faded down as normal, followed by the section of silence. Normally the Preset would be the predominant signal, but as the mixer knows it is DE, it inserts silence until the fade has reached full amplitude before cutting up the DE signal. (Followed by the two-frame recovery by the decoder).

Dolby E to PCM transitions

Dolby E to PCM transitions take a similar approach. The incoming command (e.g. crossfade) triggers a cut at the next packet boundary on the Dolby E Program input. The PCM Preset input will then fade up to follow the gain that would have applied for a PCM-PCM transition. In practice, Dolby Decoder downstream will detect a change from a Dolby encoded input to a PCM input, and the fade up by the mixer is turned into a 2 video frame fade-up by the downstream decoder.



The advantage of this process is the short time delay involved, which minimises lip-sync issues. The audio delay is a mere 8 samples (plus any delay needed to align two Dolby E packets). This compares to the alternative of a full frame of delay in a Dolby decoder, and another full frame in an encoder.

Overs

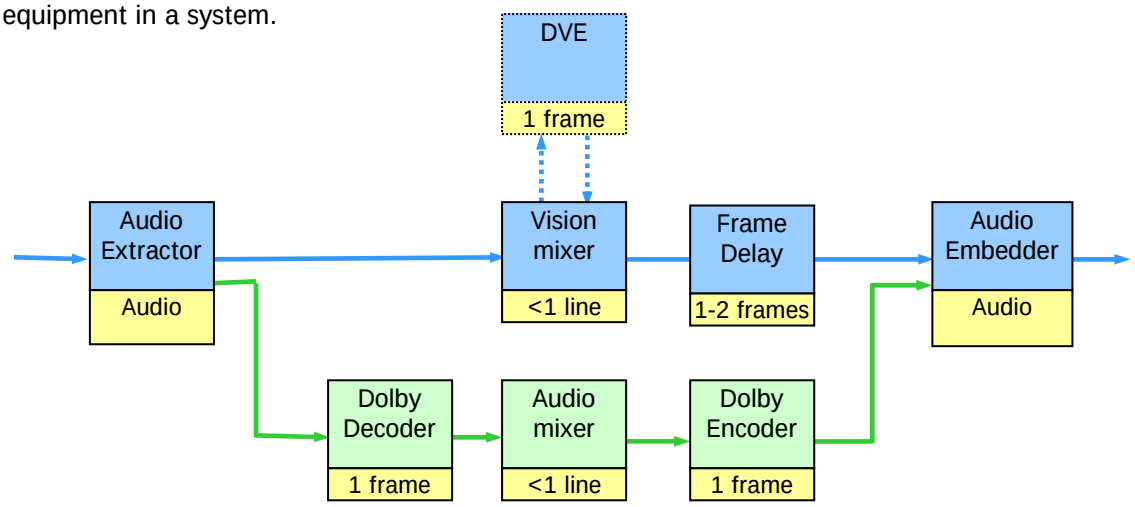
As overs are a PCM signal, it's also not possible to mix overs into a Dolby E stream. To prevent corruption of the output, overs are inhibited for as long as the output signal is Dolby E.

Metadata

All the above transitions will cleanly switch Dolby E stream, and not result in corrupted data. They will not process the Dolby data in any way, and will also not alter metadata content in the Dolby E stream. The metadata content may have parameters which affect the audio output. In particular, the dialogue level metadata can be used to change the level of the audio when it is decoded. The process described above will not eliminate this possibility, and the user should be aware that input material needs to be controlled to prevent large changes in perceived loudness in the home.

Full Mixing Capability

To fully mix audio originated as 5.1, the only option is to decode before the mixer, mix in a traditional AES environment, and encode after the mixer. A typical arrangement of equipment is shown below, with typical delays shown at each piece of equipment for HD (SD generally has slightly longer delays in the audio extracting/embedding path). As mentioned previously, the Dolby decode process takes one video frame. The encode process also takes one frame. Obviously this creates lip-sync problems if the video is passing through a standard vision mixer, the delay could be as little as a line. Traditionally, this would also result in a system with interconnections between different pieces of equipment in a system.



Clearly the video and audio path can have a significant difference in delay, up to 2 frames if a DVE is not present, and one frame if the DVE is present. This causes serious lip-sync issues. The solution would normally be to insert a frame delay into the video path as shown, incurring additional expense and using valuable rack space.

The New Pro-Bel Masterpiece mixer does not require external delay, as it incorporates the required delay within the mixer. This ensures that the Video/Audio timing relationship is preserved, whatever configuration of video mixer, DVE, and audio mixer is chosen. The decoder will also detect Dolby or PCM, and bypass the decoder as required, The encoder can be selected as in or out of circuit, and the video path delay adjusted to suit.

Incorporating all the required functions in one unit allows the user to migrate from a traditional SDV/AES environment to an SD or HD environment, with a mixture of Dolby and PCM audio. Once the system timing has been configured for the delay present in the mixer, there is no further requirement to adjust timing, change signal paths to compensate for differing delays due to different signal types.

As the Program and Preset signals are fully decoded, it is also possible to mix in overs.

Dolby E and HD Progressive Standards

Dolby Encoders and Decoders are designed to operate with interlaced scan video. Dolby has stated that there are no plans to introduce equipment with progressive frame rate capability (e.g. 720p). Therefore audio generated for use with progressive scan video will have audio packets generated using an interlaced reference. These packets will straddle two video frames.

This does not mean the signals cannot be used together, but it does add an extra level of complexity. As Dolby equipment will require an interlaced reference, it is essential that any equipment downstream does not switch or process on the field boundaries of this reference, as this will switch in the middle of a Dolby E packet. The most obvious course is to use SD B&B wherever possible, and configure any routers for frame rate switching. Where equipment uses HD tri-level sync, the two references must be suitably co-timed.

Surround Sound - Stereo Audio Perception

One question which cannot be covered with a technical answer is 'what will it sound like?'

- Should a mix from a full 5.1 programme to a stereo program still have some stereo content on the rear channels? Or the centre channel? If it is, at what level?

- Should the stereo be mono mixed and fed to the centre speaker?

- Should mono overs be on the centre channel? Maybe being able to 'position' the over to one side will add another differentiating factor to your channel?

These are not questions this paper will attempt to answer. However, we are discussing the implications of these questions with several broadcasters, and ensuring that the requirements for the future of Dolby 5.1 mixing can be met.

WWW.PRO-BEL.COM

UK
+44 (0) 1189 866 123

USA
+1 631 549 5159

France
+33 (0) 1 45 18 39 80