Dolby Laboratories delivers state-of-the-art, reliable entertainment technologies that allow television broadcasters to create and deliver a premium audio experience to the widest audience with the genuine experience that the content creators intend. Dolby technologies provide consistent and reliable playback quality and loudness levels to drive the consumer experience over broadcast and over-the-top (OTT) networks.

With Dolby technologies, incorporated in Imagine Communications products, broadcasters can position their service for emerging and future trends and consumer demands like 7.1 surround sound, audio description service, adaptive bit rate streaming, dialogue enhancement, mixed system sounds, and multiple language support.

This paper will discuss the following subjects:

- An overview of the various Dolby technologies supported by Imagine Communications, including Dolby® E, Dolby Digital and the latest technology, Dolby Digital Plus, and the differences between these technologies
- The best ways to broadcast PCM (Pulse Code Modulated) audio and Dolby compressed audio over AES embedded into SDI (Serial Digital Interface) and over TS (Transport Streams)
- How to handle incorrectly identified PCM and non-PCM audio in AES and embedded SDI signals
- Eliminating audio disturbances when switching between PCM and non-PCM audio
- Adding audio up-mixing, down-mixing and loudness control to Dolby audio processing
- An overview of audio metadata and the importance of maintaining this metadata throughout the workflow

Included at the end of the paper are references to Imagine Communications products that support Dolby audio technologies.

**Dolby Technologies**

**Dolby E: Eight Channels of Audio for Production Facilities**

While consumers have become accustomed to enjoying multichannel audio, TV broadcast infrastructures were originally geared almost exclusively to two-channel audio, making the distribution of multichannel audio prior to final transmission difficult. (Without compression, one AES signal can carry only two channels of audio.) To meet this challenge, Dolby E allows the distribution of multichannel digital audio over existing two-channel infrastructures. By using Dolby E technology, broadcasters can deliver up to eight channels of audio with metadata over existing stereo infrastructures in a variety of configurations, such as a 5.1 surround sound mix with a 2.0 stereo signal or four paired 2.0 stereo signals.

Dolby E offers a cost-effective, practical solution to the multichannel distribution dilemma. For example, Dolby E audio can be decoded and re-encoded as many as 10 times, as is typical during the contribution phase of DTV programs, without developing any audio artifacts, thereby providing significant flexibility in the audio production workflow.

Another important feature of Dolby E technology is audio/video synchronization. When the audio is encoded, it is locked to a video reference and has the same frame boundaries as the video. As long as the video and audio frame boundaries are lined up properly, a user can switch between video signals with a quiet audio switch. When Dolby E is sent over a TS (Transport Stream) using SMPTE ST 337:2008, it typically occupies a bandwidth of 2.3 Mb/s.

**Dolby Digital and Dolby Digital Plus: Sophisticated Audio Coding Technologies**

For distribution into the home environment, Dolby Digital is a de facto standard worldwide and is part of many relevant standards such as DVB and DLNA. Also known as AC-3, Dolby Digital supports anything from 1.0 channel mono sound to a 5.1 surround sound audio mix. Dolby Digital can be carried over one AES link or embedded into SDI. The bit rate for Dolby Digital ranges from as little as 56 kbps for 1.0 mono to 448 kbps for 5.1 surround sound over a transport stream.
Dolby’s latest audio technology — Dolby Digital Plus — provides data-rate improvements over traditional Dolby Digital coding for a given quality level. It also adds two more channels. With Dolby Digital Plus, broadcasters can provide for 7.1 surround sound mixes or the ability to support additional channels to support multi-language support or video description services. Video Descriptive Services, also known as Audio Description (a dialogue-only audio signal that describes the action on the screen for those with vision impairments), are increasingly becoming a regulatory requirement for broadcasters, and Dolby Digital Plus allows broadcasters to meet these new regulations.

Dolby Digital Plus is fully backward compatible with Dolby Digital. Dolby Digital Plus meets the demands and capabilities of new content delivery and storage media, while simultaneously maintaining backward compatibility with the wealth of existing content that is encoded with Dolby Digital when consumed on one of the more than 1 billion devices (such as set-top boxes, PCs or tablets) that support Dolby Digital Plus. Dolby Digital Plus also allows content encoded in Dolby Digital Plus to be played back on existing Dolby Digital home theater systems when they are used with set-top boxes or decoders that support Dolby Digital Plus.

Dolby Digital Plus can provide higher quality audio at equivalent bit rates when compared to Dolby Digital. For example, Dolby Digital Plus can deliver 5.1 surround sound at data rates as low as 128 kbps. Dolby Digital Plus delivers the flexibility and quality demanded by content providers, broadcasters and home theater enthusiasts. It delivers superb multichannel audio without impacting the data needed for high-quality video or added-feature content, while also enabling multiple surround audio streams when required.

Dolby Digital Plus excels not only in home theaters, but also in playback on smartphones and tablets and in streaming media, making it the ideal format for today’s multi-screen world. With Dolby Digital and Dolby Digital Plus, broadcasters can deliver a solution that supports seamless ad insertion, channel configuration and audio adaptive bit-rate switching for high-quality and reliable second screen applications, as illustrated in the following diagram:
This diagram shows how broadcasters offering second screen applications can scale the audio bit rate and channel configuration provided to a customer as the available bandwidth changes. With Dolby Digital and Dolby Digital Plus, this can be done seamlessly without audible clicks and pops as the bit rate changes or as ads are inserted in the video stream.

The Best Ways to Process Dolby E Audio
During television production, Dolby E signals typically are carried over an AES digital audio interface (either unbalanced 75 ohm or balanced 110 ohm). The final mix (5.1 surround sound and 2.0 stereo) is encoded into Dolby E and embedded into an SDI coaxial serial digital video interface. This simplifies transmitting video and audio signals. But to maintain synchronization between video and audio, broadcasters must take into account any additional audio processing that will take place before the content is finally displayed.

As mentioned previously, the Dolby E signal is locked during the encoding process so that video and Dolby E frame boundaries can be lined up for switching purposes. The allowable variation between video and audio is called the guardband. In some cases, when the Dolby E signal is embedded, the video delay relative to the audio signal is set to two frames so that when the Dolby E signal is decoded and monitored (or more audio processing is done), it is lined up with the video signal. When audio is de-embedded, it causes a small delay, which is not noticeable. However, if the Dolby E signal is de-embedded, processed and re-embedded, the Dolby E signal will require a one-frame delay in the Dolby E signal to maintain proper alignment of frame boundaries and a one-frame delay in the video path to prevent any lip synch problems.

Decoding Dolby Digital and Dolby Digital Plus Audio
In the audio workflow, either Dolby E content or PCM non-compressed audio is carried through to the emission point, where the decoded Dolby E signal or PCM non-compressed audio content is encoded into Dolby Digital or Dolby Digital Plus. Ideally, Dolby Digital and Dolby Digital Plus should be decoded once in the home environment. If Dolby Digital and Dolby Digital Plus are decoded and re-encoded, audible artifacts can result. However, there are cases where Dolby Digital and Dolby Digital Plus may need to be decoded, processed and re-encoded. While this is possible, the audio quality may be noticeably affected.

Mixing PCM and non-PCM Audio Processing
When converting video and passing through embedded Dolby E, Dolby Digital and Dolby Digital Plus audio, de-embedding, decoding into PCM, passing through the SRCs, re-encoding and re-embedding are necessary because the ancillary data space and de-embedding and embedding methods vary between video format standards.

Attention must be paid to avoid issues with the audio when embedding a mixture of non-compressed PCM audio and Dolby E. Dolby Digital and/or Dolby Digital Plus compressed (non-PCM) audio over AES. Dolby E, Dolby Digital and Dolby Digital Plus cannot be processed in any way (including adjusting gain, invert, equalization, etc.) or the data stream will be disrupted. Some devices use a video frame synchronizing processing device that de-embeds PCM audio to be passed.
through an SRC (Sample Rate Converter) to be synchronized and timed with video. The use of the SRC, though, will cause problems with audio encoded in Dolby E, Dolby Digital or Dolby Digital Plus; the sample rate converter should be disabled. It is important to understand that if the Dolby E, Dolby Digital or Dolby Digital Plus signals are re-embedded into the output of a video frame synchronizer, they will not pass into the new clock domain at the output of the video frame synchronizer, and the audio will eventually drift away from the video, causing lip sync issues.

Dolby E can pass through a video frame synchronizer so that it enters the new clock domain; however, it must be aligned properly with the video signal so that there are no audio disturbances downstream when the frame synchronizer repeats or drops frames. If the Dolby E is not aligned, it can be passed through the video frame synchronizer, de-embedded and realigned, and re-embedded after the video frame synchronizer, taking into account the repeated or dropped video frames.

Dolby Digital and Dolby Digital Plus can pass through a video frame synchronizer; however, neither Dolby Digital nor Dolby Digital Plus are locked or synchronized to video when encoded, so frame boundaries in Dolby Digital and Dolby Digital Plus will not line up with video frame boundaries. When the video frame synchronizer repeats or drops frames, there may be audio disturbances. The only way to prevent this is to de-embed the audio signal, decode it into PCM, pass it through the SRC, re-encode and re-embed it. Care must be taken so that the audio quality is not significantly degraded, as the Dolby Digital and Dolby Digital Plus signals may degrade because of the decode/re-encode process.

Dolby decoders can be set to automatically decode Dolby non-PCM audio and pass through PCM audio. This allows an upstream device such as a video/audio server to switch audio content types. Some devices do not properly align Dolby E audio with video when it’s embedded in SDI. As a result, when a switch takes place in the middle of a Dolby E frame at the input of the decoder, it will cause a downstream audio disturbance. To eliminate this audio disturbance, the partial Dolby E frame must be muted. When switching between Dolby Digital or Dolby Digital Plus audio streams and PCM audio streams, there may be a disturbance since the frame boundaries may not line up.

Dolby technologies use compatible metadata, and by identifying PCM and Non-PCM Audio

In the past, the AES standard has used a status flag (C-bit) to distinguish between PCM and non-PCM audio. However, it is common for this status flag to be incorrect or missing. This being the case, a video processing device with embedded audio capability must identify whether the audio content is PCM or non-PCM and then set the status flag correctly at the output for the next device downstream. This complicates matters, but it is the reality for how devices operate in today’s digital audio and video systems.

The Importance of Maintaining Audio Metadata

When an audio mix is done in production, audio metadata will be generated that matches the audio content. Metadata, carried in the encoded bitstream, describes the encoded audio and conveys information that precisely controls downstream encoders and decoders. Metadata enables the correct use of the audio material based on the configuration of the user setup and on the user’s preferences, and gives content providers control over how original program material is reproduced in the home.

Key audio metadata parameters include the following:

- **DOWNMIX** - controls stereo/mono downmix, so a stereo TV can play back a 5.1 show
- **DYNAMIC RANGE CONTROL** - controls dynamic range for stereo/mono and “late night” outputs
- **DIALOGUE NORMALIZATION** - aligns program loudness to reduce differences between ads and program loudness levels
- **CHANNEL CONFIGURATION** - flags channel configuration, such as mono, 2.0 stereo and 5.1 surround sound

The audio metadata should be passed through the entire signal chain and decoded in the home to ensure that the consumer gets an experience in their home or on the go that is as close as possible to what the content creator intended.

All Dolby technologies use compatible metadata, and by using Dolby technologies from end-to-end, broadcasters can ensure that the benefits of using metadata to reproduce the audio signal at the home are realized. For example, all Dolby Digital Plus bit streams require embedded metadata for program loudness and dynamic range control, leading to:

- Consistent program loudness on all decoders
- Predictable loudness when downmixing
- Leverage of existing CALM Act compliance workflow in broadcast

The audio metadata should be passed through the entire signal chain and decoded in the home. When Dolby technologies are used in the workflow from end-to-end, the metadata can be delivered to the end device and used to deliver the content as it was meant to be consumed. This ideal scenario doesn’t always play out in practice though. In situations where Dolby content is decoded in the workflow, the audio metadata may be dropped. Even if the audio metadata is passed through, the audio content may be processed and the audio metadata may not be updated. In addition, some audio sources that go into the emission channel may not have audio metadata, and default audio metadata settings at the Dolby encoder may not match the audio content.

To account for this, when Dolby audio content must be decoded for loudness control and up-mixing (from 2.0 to 5.1 or 7.1) or down-mixing (from 5.1 or 7.1 to 2.0), the audio metadata should be updated when additional audio processing occurs. For example, if 2.0 is decoded from Dolby E and up-mixed and re-encoded, the coding mode audio metadata parameter needs to be updated. Another example is if the audio is decoded and loudness controlled, the audio metadata for dialnorm has to match the target loudness setting and be updated before re-encoding. And finally, if the audio metadata is incorrect or missing, it should be updated or generated as needed.
Conclusions
Dolby’s deeply instilled audio expertise and close partnerships with key industry players from content creation to playback, like Imagine Communications, enable broadcasters to deliver high quality premium entertainment services onto current and future-generation consumer entertainment devices. Dolby technologies enable broadcasters to deliver audio in the way that content creators intended and provide consumers with the quality experience that they have come to expect. These Dolby technologies integrated with video/audio synchronization, conversion, video encoding/decoding, audio monitoring/logging and audio processing (such as loudness control) in the audio workflow allow broadcasters to realize the benefits of multichannel audio while saving space and cost. For people who care about HD entertainment, Dolby Digital Plus is the audio solution that delivers genuine, high-quality sound everywhere.

References for Imagine Communications Products That Support Dolby Technologies
In the following workflow diagram, Imagine Communications product platforms that support Dolby (and DTS Neural technologies) are highlighted:

Reference Materials:
- Dolby Metadata Guide
- Dolby® Digital Plus Audio Coding Technical Paper
- Post It With Dolby E – A Post Production Primer
<table>
<thead>
<tr>
<th>Model</th>
<th>SEL-1ENC1</th>
<th>SEL-1ENC2</th>
<th>SEL-1DEC1</th>
<th>SEL-1DEC2</th>
<th>SEL-2FS1</th>
<th>SEL-2XD1</th>
<th>AMD6803+T</th>
<th>DAPM6802+L</th>
<th>MF6800+T</th>
</tr>
</thead>
</table>

- **Dolby E Decoder**: X X X X X X X X
- **Dolby E Encoder**: X X X X X X
- **Dolby Digital Decoder**: X X X X X X X
- **Dolby Digital Encoder**: X X X X X X
- **Dolby Digital Plus Decoder**: X X
- **Dolby Digital Plus Encoder**: X X
- **Dolby Down Mix**: X X X X
- **AES Interface Support**: X X X X X X X
- **Embedded SDI Support**: X X X X X X X X
- **Video and Audio Synchronization**: X X X X
- **Video Conversion**: X X X X
- **Dolby E (Guardband) Alignment**: X X X X
- **Audio Metadata Pass Through**: X X X X X X X X
- **Audio Metadata Generator**: X X X X X X
- **DTS Neural Surround Up Mix**: X X X X X X X X
- **DTS Neural Surround Down Mix**: X X X X X X
- **DTS Neural Surround Multisegment**: X X X X X X
- **DTS Neural Surround Loudness Control**: X X X X X X X X

---

<table>
<thead>
<tr>
<th>Model</th>
<th>Selenio X50™</th>
<th>Selenio X85™</th>
<th>Videotek Test and Measurement Products: TVM9150PKG</th>
<th>TVM9100PKG</th>
<th>VTM4150PKG</th>
<th>VTM4100PKG</th>
<th>CMN-LA</th>
<th>LLM-1770-D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Intelligent Frame Synchronizer and Converter</td>
<td>Multiple Application Video and Audio Platform</td>
<td>Multiformat 3G/HD/SD Video Waveform Monitor and Audio Signal Analyzer</td>
<td>Multiformat HD/SD-SDI Video Waveform Monitor and Audio Signal Analyzer</td>
<td>Multiformat Modular Rasterizer Monitor</td>
<td>Multiformat Modular OnScreen Monitor</td>
<td>Loudness Analyzer</td>
<td>Loudness Logger</td>
</tr>
<tr>
<td>Dolby E Decoder</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dolby E Encoder</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dolby Digital Decoder</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dolby Digital Encoder</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dolby Digital Plus Decoder</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dolby Digital Plus Encoder</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dolby Down Mix</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>AES Interface Support</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Embedded SDI Support</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Video and Audio Synchronization</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Video Conversion</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dolby E (Guardband) Alignment</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Audio Metadata Pass Through</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Audio Metadata Generator</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>DTS Neural Surround Up Mix</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>DTS Neural Surround Down Mix</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>DTS Neural Surround Multisegment</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>DTS Neural Surround Loudness Control</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>