Is Your Plant Infrastructure Up to Handling Multi-channel Digital Audio?

by David L. Bytheway

Principal Engineer
Philips Broadcast
Salt Lake City, Utah
Introduction

- Digital audio systems are a bigger challenge than digital video for some users.

- Multi-channel systems only compound problems.

- What are some of the problems?

- What are some ways to solve them?
Background Philosophy

- Any routing or distribution system should be as transparent as possible.
- Never create or cause any performance degradation to the signals passing through the system.
- Rarely achieved in analog systems.
- Potential exists in digital systems, but not guaranteed.
Synchronization and Phasing

- All sources must be synchronized to a common reference.
- Any device that fades, mixes or switches signals must align or time signals.
AES11-1997 Quick Summary

- All sources are synchronized to a common audio reference signal.

- AES frames (sample boundaries) are defined as the reference point.

- All outputs must be transmitted within ± 5% of the reference signal.

- All inputs must accept signals within ± 25% of the reference signal.
All devices must have a deterministic amount of delay for signals meeting the previous criteria. Delay is measured and stated in an integer number of audio samples.

When using video reference signals, audio is aligned with a specific edge of the video reference.

For PAL type signals, falling edge of line #1 sync is used.

For NTSC type signals, a 5-frame flag is required along with falling edge of line #1 sync.
Automatic Audio Frame Alignment

- Phase diagram showing AES11 specified Receive and Transmit windows.
Example of Audio Frame Alignment

- Scope capture of the reference input and an output of a routing switcher that meets the requirements of the AES11-1997 synchronization standard.
Example of Audio Frame Alignment

- Scope capture showing two different input timing relationships to the reference signals. All output will be aligned with the reference.
Example of Audio Frame Alignment

- Scope capture of a system with the maximum permissible delay of 25% due to combined effect of a transmitter that is 5% late combined with cable delay.
Using PAL Video References

- Scope capture of the required relationship between PAL type video references and AES reference or output signals of a system meeting AES11.
Using NTSC Video Reference

- Phase diagram showing AES11 specified Receive and Transmit windows when using NTSC type reference.
Using NTSC Video Reference

- Scope capture of the required relationship between NTSC type video reference and AES reference or output signals of a system meeting AES11.
Using NTSC Video Reference

- Scope capture to illustrate the 5-frame sequence in relationship to a stable AES output stream. Note that the 5 video frames are one audio frame in length.
What does the 5-frame problem mean?

- All devices that deal with audio and video must address the problem.
- Most devices determine the 5-frame sequence on power-up.
- Most devices keep all AES pairs within them aligned with each other (but many don’t - beware!).
- You must avoid splitting any surround sound signal that uses multiple pairs over multiple devices when using video reference.
What does the 5-frame problem mean?

An example of a hookup that will cause misalignment of AES pairs due to the 5-frame problem.
Advantages

- Simplest system interconnect.
- No external I/O devices to interface VTRs, disk servers, etc.
- Most transparent to the AES signals, up to 24 bit.
- Preferred for production work, because of ease of use.
- Can provide near-click-less or completely click-less switching with the right routing hardware.
- Can support both synchronous or asynchronous with the right routing hardware.
- Can split AES pairs, routing separate audio channels, and other stereo/mono effects with right routing hardware.
Disadvantages

- Suffers from the 5-frame problem for all devices locked to NTSC video references (Not all PAL equipment is immune from problems either).
- Not enough channels available on typical storage devices such as VTRs and Disk Servers.
- Baseband systems need many more paths for complete surround-sound support or other multi-channel production support.
Compressed (Dolby-E) Audio Systems

Advantages

- Full surround can be contained on a single AES pair, or a pair of audio tracks on a digital VTR or Disk Server
- No chance to experience channel misalignment due to differing path lengths or the 5-frame problem
- Can be switched silently if switched on video frame boundaries.
- Can be edited on video frame boundaries.
- Can save routing switcher and wiring resources.
- Can be used for both surround-sound and multiple languages if needed.
- Automatically detects Dolby E signals and switches between a standard AES pair and the surround-sound version when detected.
Compressed (Dolby-E) Audio Systems

- **Disadvantages**
  - Audio signals are compressed, some minimal quality loss is present.
  - Cannot manipulate or mix signal without decompression.
  - Cost of encoders and decoders
  - Required decoders for monitoring or metering
  - Not all AES devices properly pass the AES stream.
  - Not all AES devices properly pass the AES status bits.
  - Creates a one-frame delay for encode and an additional one-frame delay for decode. This must be accounted for in any distribution system. Video delays may be needed to make correct the differences.
  - All signals must be synchronous.
  - Will lock only to analog video reference, no tri-level or AES reference at this time.
Embedded Audio Systems

Advantages

- Can significantly reduce routing resources and interconnect cabling, patch bays, etc.
- Can carry up to 16 channels.
- Most devices will pass a Dolby E stream.
- Many VTRs and Disk Servers already support a limited number of embedded audio channels.
Embedded Audio Systems

- Disadvantages

  - Support for all 16 channels of audio is missing in most VTRs and Disk Servers. Most only support first four channels.
  - Cannot be switched click-free (or is nearly impossible).
  - Cannot perform audio breakaways or manipulate separate audio channels.
  - Not all video-processing devices pass the audio packets properly.
  - Many devices do not align the multiple AES pairs that are embedded.
  - Cannot pass asynchronous signals.
  - Audio must be extracted to do any processing such as delay, mixing, crossfading, or compression.
  - Great care must be taken in system design (it is hard!).

David L. Bytheway, Principal Engineer - Philips Broadcast
Summary and Recommendations

- If you have absolutely no need to do any audio manipulation then a solely embedded audio plant may work for you (you will have limitations).

- Many users are finding a blend of baseband and uncompressed AES with Dolby E streams the most flexible.

- Great care must be taken to avoid so that the 5-frame problem does not bite you.
Summary and Recommendations

- The use of Dolby E will largely overcome the 5-frame problem.
- When using Dolby E you must account for the one-frame of additional audio delay for encode and decode.
- Pay attention to your reference system.
- It is possible to make this all work!
Thank You!