Dolby E and Dolby AC-4 over Professional IP Media Networks

15 April 2021

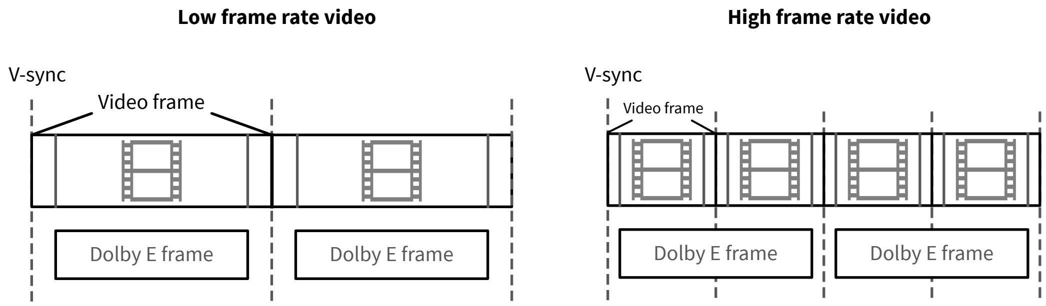
Dolby E and Dolby AC-4 over professional IP media networks

This documentation describes key aspects required to support Dolby E and AC-4 over an IP-based transport, including system timing, audio and video phase relationship, and line position (the point at which the Dolby audio frame starts in relation to the video frame).

As professional media networks transition from serial digital interface (SDI) based systems to IP, media distribution methods are also evolving. Fast IP distribution circuits are replacing dedicated media circuits, providing more flexibility when selecting media formats. Traditionally, encoded audio formats have been popular in distribution to comply with equipment channel limits and bandwidth constraints. With the emerging IP distribution systems, we expect most audio to be carried in an uncompressed format. However, compressed audio formats will not disappear immediately, and methods are required to support legacy compressed audio formats distributed over IP-based systems. As Dolby E has been popular for audio distribution, solutions to carry it over IP infrastructures are required.

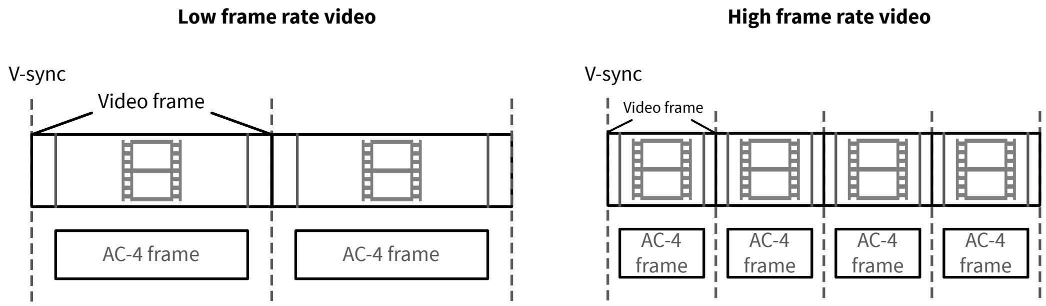
When distributing Dolby E (for example, in an SDI environment), audio frames must always be phase locked to video frames, as shown in this figure.

Figure 1: Phase-locked Dolby E and video frames



When distributing AC-4 (for example, in an SDI environment), and the AC-4 frame rate equals the video frame rate, audio frames must be phase locked to video frames as shown in this figure.

Figure 2: Phase-locked AC-4 and video frames



Phase-locked audio and video frames enable lossless video frame aligned switching and editing. The time difference between the start of the audio frame and the start of the video frame must be strictly controlled in accordance with existing specifications. See SMPTE RDD 33 for Dolby E guidance and SMPTE ST 2101 for AC-4 guidance. This is true for IP environments as well so that gateways can convert IP streams containing Dolby E or AC-4 back to a legacy transport and maintain the correct phase relationship between the audio and video frames.

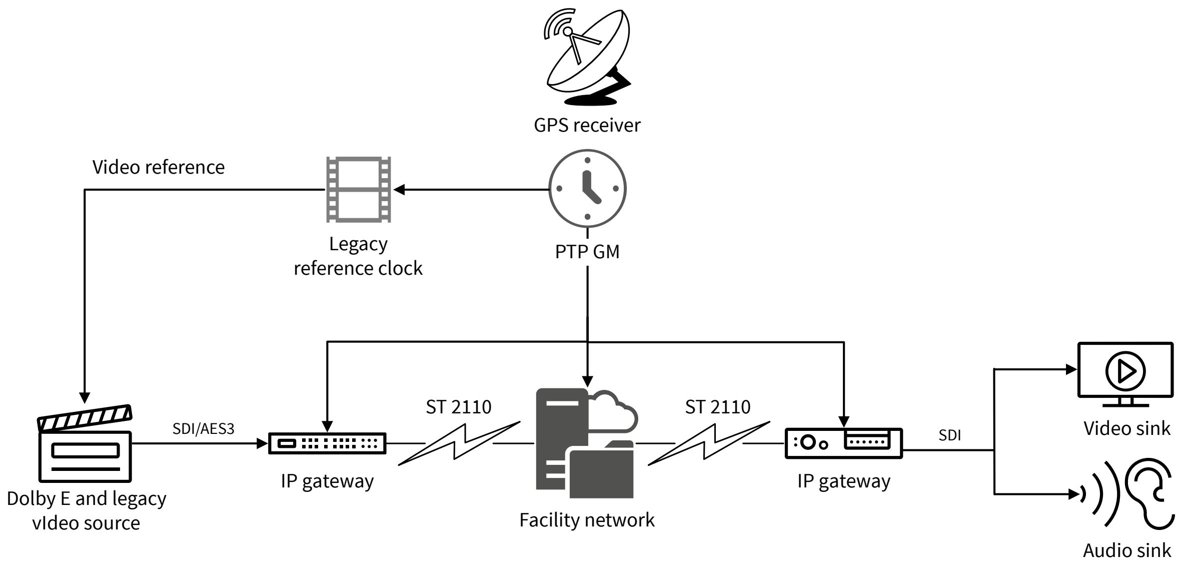
Designers of systems and equipment that support Dolby E or AC-4 over IP must be aware of certain system requirements to ensure that these properties are maintained in the IP infrastructure. The timing requirements for Dolby E and AC-4 across an IP network are the same as the requirements necessary to maintain general audio and video synchronization, so equipment does not necessarily need to be specifically designed for these formats. Any product that modifies the timing of either the video or audio streams must be aware of and maintain audio and video synchronization to be compatible with Dolby E or AC-4. Although such products will maintain correct audio and video frame alignment, to rectify incorrectly aligned Dolby E or AC-4 frames, specific support is required.

Synchronous input

With synchronous input, where the incoming stream is using the same reference clock as the facility, the accurate audio and video synchronization mechanisms of an IP gateway are sufficient to maintain the correct audio and video phase relationship.

A common scenario is for Dolby E to be encoded using legacy SDI/AES3 based equipment, carried across the IP infrastructure and then decoded or processed by legacy equipment. In this scenario, network edge gateways are required to convert between the transport mechanisms.

Figure 3: Dolby E over IP network



A key requirement of the IP gateways is maintaining the phase relationship between audio and video frames. Accurate audio and video synchronization mechanisms are sufficient to maintain the correct audio and video phase relationship. If the IP gateway maintains the timing relationship between PCM audio and video streams required for audio and video synchronization, then this gateway should be suitable for transport of synchronous Dolby E or AC-4. In this case, the video would be carried across the network as per SMPTE ST 2110-20, and the Dolby E or AC-4 stream would be carried over IP as per SMPTE ST 2110-31.

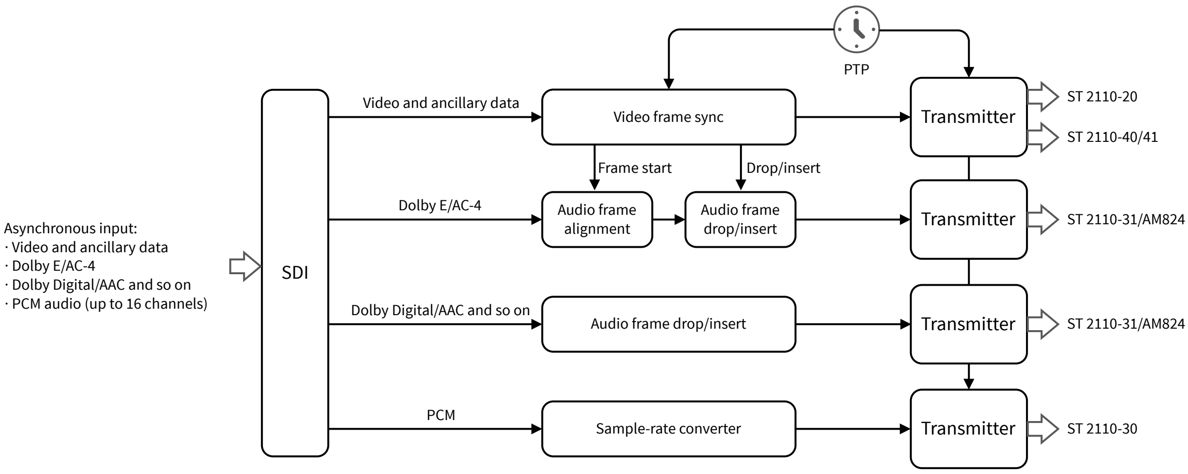
If the correct audio and video frame phase relationship is not maintained, then equipment downstream (such as an SDI router) may corrupt audio frames. An audio frame corruption caused by a bad switch will be concealed by a downstream Dolby decoder, but audio quality will be degraded.

Asynchronous input

With asynchronous input, where the incoming stream is not using the same reference clock as the facility, the incoming stream must be carefully synchronized to that local reference. As there is a possibility that the input stream has misaligned Dolby E or AC-4, we recommend realigning audio frames correctly with the video frames before synchronizing the video to the reference.

Asynchronous input is common for ingress of a contribution or distribution link into a facility. To illustrate how the requirement to maintain line position affects product design in this case, the internals of an IP gateway converting the legacy transport to the IP stream is shown in this figure.

Figure 4: Frame synchronization for asynchronous input



As with a traditional gateway, a video frame synchronizer is required to bridge the two clock domains. This involves the insertion of repeated video frames to maintain consistent latency. For encoded audio, including Dolby E or AC-4, a similar approach must be taken, where frames are inserted or dropped.

It is good practice for Dolby E and AC-4 frame alignment to be checked and any necessary realignment performed before a frame drop or insertion. This ensures that even if the line position of the incoming audio is incorrect, audio frame corruption will not occur at the switch point.

Audio and video frames must be inserted and dropped simultaneously to maintain correct audio and video synchronization. For AC-4, drop or insert only a single frame (rather than multiple frames in a row). In the case of high frame rate video, it should be noted that, unlike AC-4, Dolby E products support only lower frame rates (such as 25 or 29.97 fps). To maintain audio and video frame phase lock, the audio frame rate chosen must be half or quarter the video frame rate. As a result, not every video frame boundary is available as a switch point. See the following figures.

Figure 5: Dolby E safe switching point

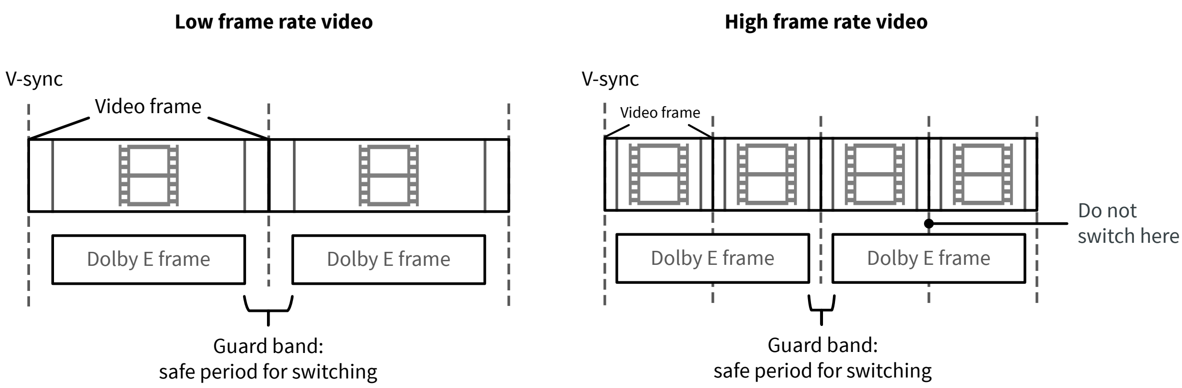
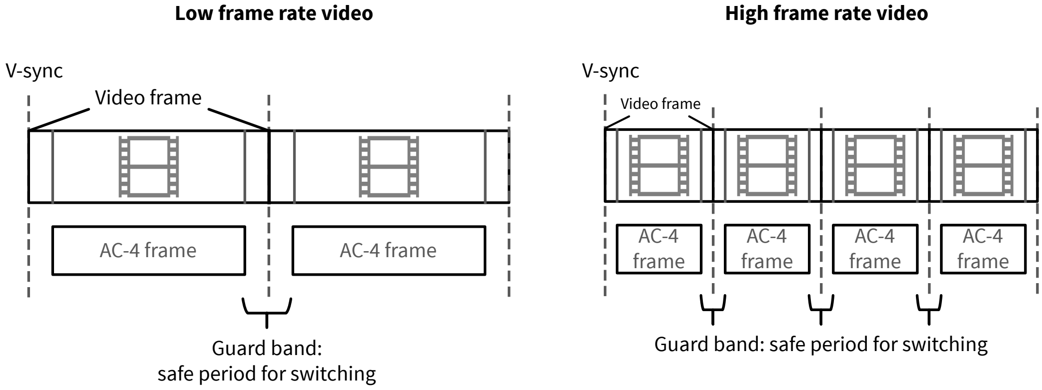


Figure 6: AC-4 safe switching point



When a Dolby AC-4 stream is switched or frame drops and insertions occur, the Dolby AC-4 encoder must be in a mode where the Dolby AC-4 encoder creates independently decodable audio frames only.

Note: This mode may be called Professional Distribution mode, but the terminology used is product specific. Additionally, this control may not be exposed to the user.

For more information on Dolby E working in a high frame rate video system, see *Manufacturer Guidelines: Handling of Dolby E in Environments Using Progressive Frame Rates (50/60p)*. For frame-rate guidance about AC-4, see the section regarding A/V alignment and segmentation in *Dolby AC-4 in MPEG-DASH for Broadcast Services*.

For audio formats with frames that are not phase locked to video (for example, Dolby Digital or Advanced Audio Coding (AAC)), a different approach must be taken. In this case, the audio frame is a different duration than the video frame, so audio and video frames must be dropped or inserted at different times. Unfortunately, this causes variations in audio latency and audio/video synchronization by up to an audio frame. For this reason, we recommend using a video frame phase-aligned format such as Dolby E or AC-4 in this scenario. It should be noted that, for Dolby AC-4, selection of an audio frame duration unmatched to video is possible. Where frame synchronization or live switching occurs, use of an audio format with phase-aligned video and audio frames provides multiple benefits, including consistent latency, accurate audio/video synchronization, and simplicity of implementation.

Line position testing

We recommend checking correct line position of Dolby E or AC-4 both in the SDI domain and the IP domain when carried over SMPTE ST 2110-31.

It is good practice to check for correct line position during the commissioning of new equipment or systems. If the test equipment to check Dolby E or AC-4 line position directly within a SMPTE ST 2110-31 stream is not available, an SDI test input will have to be used. To convert the IP stream under test to SDI, use a reference IP gateway that has known reference audio/video synchronization characteristics. Test and measurement products that support line position measurement in SDI are readily available from multiple manufacturers.

Transport of Dolby E and AC-4 according to SMPTE ST 2110-31

SMPTE ST 2110-31 is the correct transport standard for Dolby E and AC-4, which allows bit transparency and signaling of audio that is not linear PCM.

Although Dolby E and other encoded audio formats can be made to work using SMPTE ST 2110-30 or AES67, this is strongly discouraged. These standards state that only carriage of linear PCM is supported, which allows the use of mixers and sample-rate converters in the audio path. Any such audio processing will corrupt encoded audio; therefore, an alternative bit-transparent AES3 transport mechanism is required.

SMPTE ST 2110-30 and AES67 standards do not support carriage of the AES3 channel status bits required to avoid corrupted audio that is not linear PCM being misidentified as PCM. When SMPTE ST 2110-31 carries a Dolby encoded format, the AES3 channel status bits must be used to signal audio that is not linear PCM.

References

Documentation is available to further describe Dolby E or AC-4.

1. SMPTE RDD 33:2015, *Format for Non-PCM Audio and Data in AES3—Dolby E Data Type*, available from http://www.smpte.org.
2. SMPTE ST 2101:2015, *Format for Non-PCM Audio and Data in AES3—AC-4 Data Type*, available from http://www.smpte.org.
3. *Manufacturer Guidelines: Handling of Dolby E in Environments Using Progressive Frame Rates (50/60p)*, available from https://developer.dolby.com.
4. *Dolby AC-4 in MPEG-DASH for Broadcast Services*, available from https://media.developer.dolby.com/AC4/AC4_DASH_for_BROADCAST_SPEC.pdf.
5. SMPTE ST 2110: Professional Media Over Managed IP Networks suite of standards, available from http://www.smpte.org.
6. AES67-2018, *AES Atandard for Audio Applications of Networks—High-Performance Streaming Audio-over-IP Interoperability*, available from https://www.aes.org.
7. AES3-1-2009, *AES Standard for Digital Audio—Serial Transmission Format for Two-Channel Linearly Represented Digital Audio Data—Part 1: Audio Content*, available from https://www.aes.org.
8. SMPTE ST 337:2015, *Format for Non-PCM Audio and Data in an AES3 Serial Digital Audio Interface*, available from http://www.smpte.org.
9. Additional Dolby E technical documentation is available from https://developer.dolby.com.