Digital Audio Plant Integration: A Tutorial for Designing Digital Audio/Video Broadcast and Production Facilities

Foreword

This is an updated version of the white paper originally published by NVISION in 1994. It served as the basis for all the application notes in the widely acclaimed NVISION "The BOOK" and "The BOOK II". These two publications have been used by hundreds of design engineers around the world to build affordable digital facilities that are easy to maintain and operate, and provide the ability to grow as technology advances. "The BOOK" and the "The BOOK II" are both available in PDF format on the NVISION web site at www.NVISION1.com, or on the NVISION CD ROM Product Catalog which you may order at the same web site. New information on TDM routing, Dobly Digital and Dolby E routing, as well as integrated A/D and D/A conversion is provided in this paper, and in other white papers on the NVISION web site.

This paper takes a very practical approach to solving real problems associated with interfacing analog and digital equipment together into a single working unit. When possible, multiple techniques are discussed as likely solutions for a given problem and when appropriate the practical differences between digital islands, suites, layers and facilities are discussed. There is no single right answer. Sometimes a digital suite is the solution. Sometimes digital and analog must co-exist as part of a transition. More commonly, a new, entirely digital facility is designed from the ground up. However, in all cases, it is possible to capitalize on the benefits of digital audio.

DIGITAL AUDIO MACHINE INTERCONNECT

Providing a reliable interconnection for transferring analog signals between two pieces of equipment requires more than a length of copper and two mating connectors. Digital audio connections are not any different. To maximize the benefits offered by digital technology, an understanding of new procedures and formats is required. In addition, there are two inter connect choices available to the user and the criteria which need to be considered in deciding which to use are defined and explained. Signal impedance matching and careful cable selection are two key areas which must be understood. Digital audio data formats also will have an impact on the design process. Finally, details for grounding, shielding, and distribution of common signals form the basis for actually wiring up a facility.

A Brief History of Digital Formats

A large number of digital audio formats have been used to some degree over the past decade. The three primary signal formats likely to be encountered in the video environment are IEC 958, or S/PDIF, (the Sony Philips Digital Interface developed for serial transmission of digital audio information between consumer products), AES-10, or MADI, (the Multi-Channel Audio Digital Interconnect standardized by the Audio Engineering Society, for the interconnect of 56 channel digital audio between consoles and multi-track recorders), and AES3. AES3 has become the dominant standard for the interconnection of digital audio signals between equipment - audio and video. The AES3 standard is jointly supported by both the AES and EBU societies, and is often referred to as the AES/EBU standard for digital audio. The standard defines the baseband data format for two channels of audio and respective overhead information, the transmission data format, and the electrical interface for the signal. The AES3 standard was most recently revised in 1997 and amended in 1999. In digital television facilities, embedding audio in the serial digital video data has become very popular indeed. SMPTE 272 is the standard describing how this process is to be carried out. While problematic in its earliest implementations, it is now quite robust, and widely installed.

Historically, SDIF-2 was the first widely used digital audio format. An SDIF-2 interface consists of 3 coaxial cables: one each for left channel data, right channel data and a timing signal. Multi-track recorders feature a balanced SDIF-2 interface using twisted pair ribbon cable and 50 pin D-type connector. The timing signal is often referred to as the SDIF-2 word clock, or word clock for short. It is a square wave signal oscillating at the digital audio sample rate - the rate at which analog signals are sampled for conversion into a digital format. Word clock is still very popular for timing in audio only facilities. AES3 is the most popular format in video facilities. Digital audio program distribution is simplified with the AES3 format. First, both left and right channel audio data are placed into one serial data stream, left data first then right. Second, data is coded for transmission into a bi-phase signal, a self-clocking data format. With these enhancements, the AES3 signal transmits 2 channels of audio with timing, as one balanced signal, over one twisted pair cable. Digital audio program production typically requires independent channels of audio. For this reason, the AES3 format allows the 2 channels of audio data to be monaural. Digital audio production mixers then break the AES3 signal apart into two separate channels of audio before the mix. After production, the material is formatted into the desired number of AES3 signals with the correct channel assignments for distribution. Most equipment offers analog audio or AES3 digital audio connections.

It is important to note that the AES3 format is intended to be independent of the audio conversion sample rate. However, the net data rate is exactly sixty-four times the sample rate. Since 48 kHz is the most frequently used sample rate in the video environment, the most frequently encountered bit rate for AES3 data is 3.072 Mb/s, (mega-bits per second). This is clearly a much larger bandwidth signal than traditional analog audio and failure to recognize this fact can lead to problems which are discussed later in this paper, or in The Books. Video engineers recognized that the AES signal had a similar bandwidth to analog video and pushed for standardization of a low level voltage signal format for coaxial AES3 data transmission. AES-3id offers one approach to using coax for AES3 signals. SMPTE 276 offers a slightly different approach. In short, AES-3id assumes that balun devices are used to convert from balance to single ended signals. SMPTE 276 is based on generating a single ended signal directly. Equipment that conforms to either AES-3id, or SMPTE 276 may be interconnected in most cases. However, in the presence of long cable runs, or incorrect usage of balun devices, there could be issues.

AES3 for Twisted Pair

The AES3 twisted pair interconnect is well defined in the AES3 standard document. The signal is bi-phase coded, transformer coupled and transmitted in a balanced format, on twisted pair copper. The output level may be 2 to 7 volts peak to peak with a source impedance of 110 ohms. The receiver should have a sensitivity of at least 200 mV and an input impedance of 110 ohms. The inter connecting cable should also exhibit a nominal 110 ohm characteristic impedance. The standard connector is specified to be the XLR type. A typical AES3 interconnect and electrical path are shown in Figure 1. Please note the shield bypass capacitor at the receiver. This is recommended by NVISION, and others, for increased suppression of high frequency emissions and is not part of the AES3 specification.

AES-3id for Coaxial Cable

The AES3 and SMPTE committees have established electrical interface guidelines for the trans- mission of AES3 data on coaxial cable. AES-3id was developed by the AES while SMPTE developed the SMPTE 276 standard. The AES-3id interface is single ended. The signal level is 1.0 volt, +/- 20 percent, peak to peak, when terminated with 75 ohms. The source impedance is also 75 ohms. It is not required that the signal be transformer coupled, but most implementations are adaptations of existing AES3 circuits so the transformer remains. This interface format is perceived by many video engineers as offering greater compatibility in their operational environment. The AES3 –id and SMPTE guidelines specify the BNC connector as standard. A typical AES-3id interconnect and electrical path are shown in Figure 2. Please note the shield bypass capacitor at the receiver. This is recommended for increased suppression of high frequency emissions and is not part of the AES-3id specification.

As mentioned, these guidelines were developed at the request of users who wished to use video equipment to route and distribute digital audio signals. Unfortunately, the bandwidth of the AES signal can exceed that of many analog video DAs and routers. In particular, the high frequency energy generated by the fast digital edges of the AES signal can lead to ringing, oscillation and crosstalk in analog video equipment. These fast edges are often slew rate limited by slow analog DAs or other video

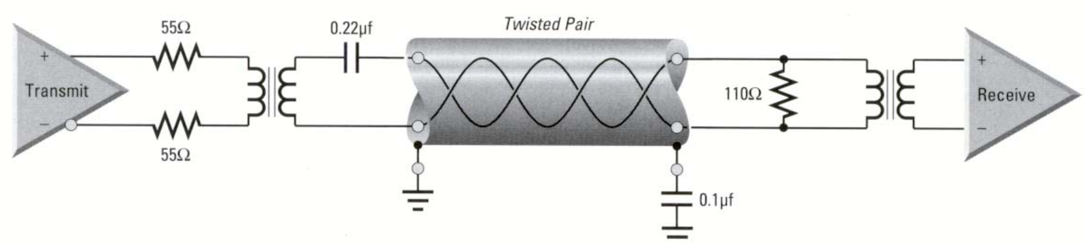

Figure 1: Recommended AES3 Interconnect Circuitry

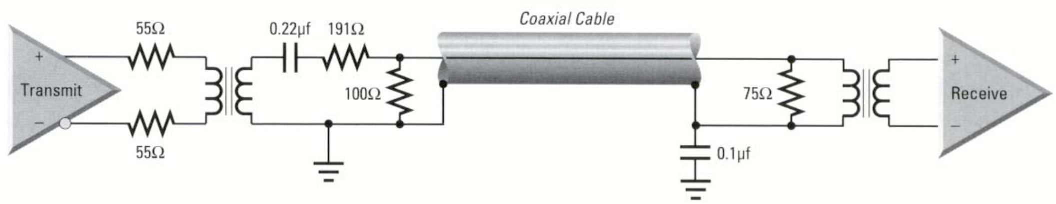

Figure 2: Recommended AES3 — ID Interconnect Circuitry

distribution equipment. These distortion effects may lead to jitter generation at later stages in the signal path. Careful consideration should be given to the use of video equipment for AES signal distribution. Equipment designed for the AES3 standard as well as the AES 3id and SMPTE guidelines should provide a clean, artifact-free signal. For more details see *The Book* and *The Book II.*

Suggested Considerations for Choice of AES Interconnect

An easy procedure to evaluate which interface type to use follows.

First, consider cable costs. In new facilities, this is easy. Determine the number of feet of each type required and get competitive bids. In existing facilities, twisted pair and coaxial cable may already be installed in floors, ducts, and ceilings. Analog audio cable is not recommended for digital applications, but coaxial cable is and may provide a great savings. The costs of cable removal, replacement and downtime must all be considered in addition to the price of the cable itself. High quality digital audio twisted pair cable varies greatly in price, but the typical range is from one-fifth to three-fourths the cost of high quality video coax.

Second, include cable termination costs. BNC connections are arguably easier to make and provide higher density packaging than XLR connections. Router and terminal equipment with BNC connector options provide attractive cost benefits for a facility. Be sure to look for this option in product data sheets.

Third, the cost of supporting existing equipment must be considered. Since XLR connectors are standard, virtually all professional audio and video equipment uses them. If the BNC option is considered, include the costs of converting all existing XLR and twisted pair equivalent connections to the BNC format. Quotes for this equipment, which is usually supplied on a per connector basis, will be required. Be sure to specify the 1 volt operation level per AES 3id or SMPTE 276 guidelines since more than one voltage option is available for these converters.

Fourth, evaluate your equipment needs and the ability of existing analog equipment to truly meet the specifications required for accurate transmission of digital signals. Compare operational and maintenance costs also. As an example, an inexpensive coaxial interface digital AES3 router designed specifically for digital audio data, requires considerably less space and power than an existing analog video router, it will “drop in” to the existing pre-cabled facility and provide both reduced operational costs and peace of mind.

Electrical Properties of the AES3 Signal

It is important to treat digital audio as a high frequency signal. Its 3.072 Mb/s data rate requires a bandwidth similar to that of analog video. Digital audio signals should be treated with the same care in cabling and installation. To this end, a well executed interconnect in either interface format will have matched source, destination, and cable impedances. Unfortunately, equipment manufactured to the AES3 standard prior to 1992 violated this rule. The standard specified a 110 ohm source and a 250 ohm load resistance. Fortunately, this can easily be corrected. Equally fortunate is the fact that the 1992 AES3 Standard replaces the mismatched 250 ohms with a matched 110 ohm load. An application exists in The Book, should you experience this problem.

Cable Types

Cable selection is important for AES3 applications. For coaxial use, select a good, 75ohm characteristic impedance cable. Any cable which provides acceptable analog video transmission performance should work. Many twisted pair options are available which meet the exact 110 ohm impedance requirement. If one of these is not available, or does not meet your budget, consider a good quality data cable. A cable exhibiting an impedance value within ten percent of 110 ohms and a low capacitance per foot rating, such as 12 or 13 pF/ft, will work quite well. More detail can be found in “The Book” Using matched impedances and a good quality cable; transmission distances of up to 1000 feet can be achieved for either format without equalization.

Signal Distribution

A common signal often needs to be distributed throughout a plant. Typical examples are synchronization, digital test tone and digital quiet. These signals are discussed later in this paper and in “The Book” and “The Book II”. This process requires strict attention to detail since this is a high frequency signal. Typically, one or more distribution amplifiers are used in a tree topology to insure network reliability and constant signal phase between various pieces of equipment. A conceptual picture is shown in figure 3. “The Book” describes AES3 signal distribution with loop through techniques and helps to explain when phase is important. It also discusses network reliability. While not defined as part of AES3 standard practice, loop-thru topologies work if good transmission line techniques are practiced.

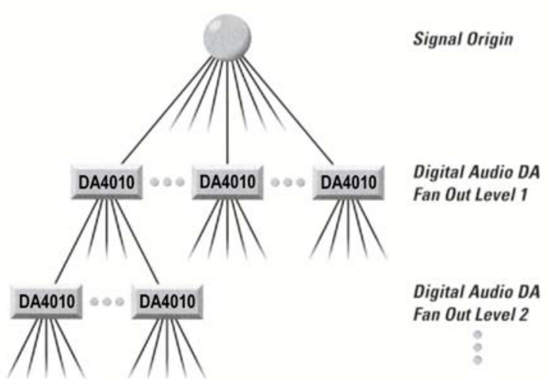

Figure 3: Recommended Distribution Topology for AES3 Signals

Details for Shield and Ground Connections

Shield and ground connections are still important for data signals. Figures 1 and 2 demonstrate a preferred technique for connecting the receive cable shield in order to minimize signal to noise, and RF emissions without creating low frequency or DC ground loops which cause various ills ranging from HUM to fires.

Other Standards to Be Aware of and to Anticipate

AES 10, more popularly known as MADI (the Multi-Channel Audio Digital Interface), and AES 11 (Digital Audio Synchronization), are two standards increasing in importance. AES 10 defines a serial format for the transmission of 56 channels of digital audio data. MADI is intended for machine to machine connection of two pieces of equipment with either fiber or coax. It is not a telecommunications compatible standard, nor is it an Ethernet compatible standard. AES11 is the synchronization standard. It defines jitter and long term stability for a digital audio reference signal. Proposed guidelines recommending a fixed phase relationship for audio and video reference signals may be included as part of AES 11. PAL and AES3 reference signals are to be phased so that AES3 block is aligned to the start of PAL line 1 within +/- 5% of an AES3 frame. An absolute, operationally convenient AES3 relationship to NTSC is much more difficult to define because of the 1000/1001 ratio in the NTSC frame rate. Finally, the popularity of low cost 8 channel digital audio recorders, as well as the requirements for Dolby E transport and production, have led to a new proposed SMPTE Standard, SMPTE 324, which provides for up to 16 channels of audio in a single coaxial interconnect. Also, there is a standard for encapsulating AES3 data in ATM data transport streams.

A DESIGN STRATEGY FOR PLANT INTERCONNECT

The above information can be put to use immediately. Whether designing a new facility, or refurbishing an old one, an interconnect strategy is crucial to the success of the final design.

- Choose a common plant standard for interconnect. AES3 and embedded audio have become the dominant formats. Remember that any original analog material not originated in the standard format must be converted.
- Choose an electrical interface, BNC or Twisted Pair. This is thoroughly discussed above.
- Select cable. Consider cost per foot, cross section diameter, ease of termination, existing cable and associated labor costs.
- Consider the need for multi-channel audio interconnect between devices. Embedded audio may provide a good solution for broadcast facilities. Be sure to consider how much audio production you do. The less production you do, the more attractive embedding becomes.

SYNCHRONIZATION OF DIGITAL AUDIO FACILITIES

Virtually all pops and clicks which plagued the early adopters of digital audio production are eliminated when equipment is synchronized. In fact, the only way to achieve direct digital audio transfers between machines without pops and clicks or the expense of sample rate converters, is to synchronize them to a common reference. No digital audio transfer can be completed if the source and destination equipment are not locked. Any digital plant design must include a strategy for synchronization of all equipment. Other benefits are derived from this approach. A common reference removes the effects and idiosyncrasies of poorly designed data and clock recovery circuits. Configuration time is decreased, and machine dependent operational considerations are removed, reducing the number of red herrings and the amount of wasted maintenance time. If considered early in the design process, the costs of synchronization are small. Virtually all broadcast quality video equipment will be correctly synchronized when locked to house color black. This helps reduce the costs of synchronization in any facility.

A digital audio signal such as an AES3 data stream may serve as a reference. NTSC or PAL video signals work equally well. These sync signals can be used in a local island, a suite or across an entire facility. Large facilities will most likely benefit from an integrated audio video synchronization network. Including audio in this network is a new concept previously not implemented for analog audio. This section describes various techniques for synchronizing islands, suites or layers and plants. It will define and discuss frequency only synchronization as well as combined phase and frequency synchronization, and when to be concerned about each. Some equipment does not allow for external synchronization. This section describes two methods for integrating this equipment into a synchronized facility. The importance of phasing audio and video signals is also discussed.

Frequency Synchronization

All 48 kHz oscillators are not created equal. Each is specified to oscillate at 48 kHz within some tolerance, typically +/- 25 parts per million or +/- 1.2 Hz for professional audio equipment. Now consider NTSC and PAL digital video tape recorders. When digital audio material is transferred from one machine to another, pops and clicks occur regularly even though both recorders nominally produce 48 kHz equivalent sample rates. The PAL and NTSC video timing references are not locked to each other. If the PAL and NTSC signals could be locked to a common clock, pops and clicks would disappear. This is the answer. Synchronize all equipment to a common reference. PAL and NTSC video plants have been synchronized within common video formats for years. It is required for color accurate editing, among other things. Audio has been without this burden until now. The emergence of digital audio as a primary format requires that all equipment must be locked to a common clock. NTSC, PAL and digital audio equipment must all be synchronized together.

Fortunately, nearly all professional video equipment generates a 48 kHz digital audio sample rate when locked to a standard video reference. NTSC equipment operating at 59.94 fields per second and PAL machines operating at 50 fields per second both output 48kHz audio. The chart shown in figure 4 illustrates that PAL and NTSC are not the only signals which can be locked to a common clock. Film and CD players can be locked together as well. In fact almost any piece of equipment can be locked to this timing chain. Once locked, digital audio material can be transferred between equipment of nearly any video standard and format transparently. This is how digital technology is intended to work. This is why a common time base is required.

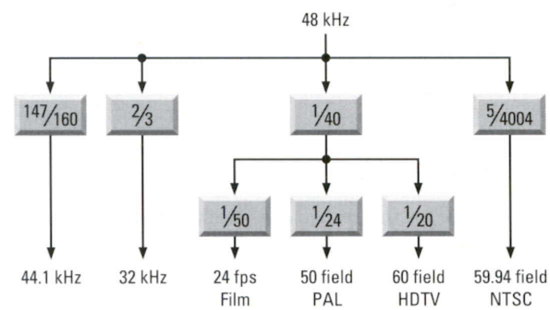

Figure 4: Integer Relationships For Audio and Video Fequencies

Phase Synchronization

Digital audio reference signals must provide phase and frequency information. While a video signal can provide frequency synchronization, it cannot always provide for phase accurate audio referencing. Phase alignment is extremely important for audio, particularly when the analog signal is converted to digital. Any audio processing and recording equipment will force the alignment of all AES3 inputs to a common AES3 frame phase. Any difference in the frame phase of analog to digital converters will generate a proportional phase error between the audio signals when an AES3 frame phase alignment is executed. A quick calculation reveals that if a 48 kHz sample rate is used, a one half sample time error in phase is equivalent to 75 degrees of phase shift for a 20 kHz analog tone, a serious degradation in the audio image. The only conclusion which can be made, is that digital audio phase must be preserved when converting multiple channels of audio between the analog and digital domains. This is almost impossible to achieve with a typical NTSC video reference. Figure 5 shows two phase locked loops which, when used in cascade, lock to horizontal line rate and generate an audio reference frequency. Unfortunately, these two phase locked loops include four digital dividers which start at random states on power up. Hence the output phase of the audio reference is random, incoherent. The only way to provide an absolute phase reference is to use an SDIF-2 word clock or an AES3 signal as the common master clock for conversion.

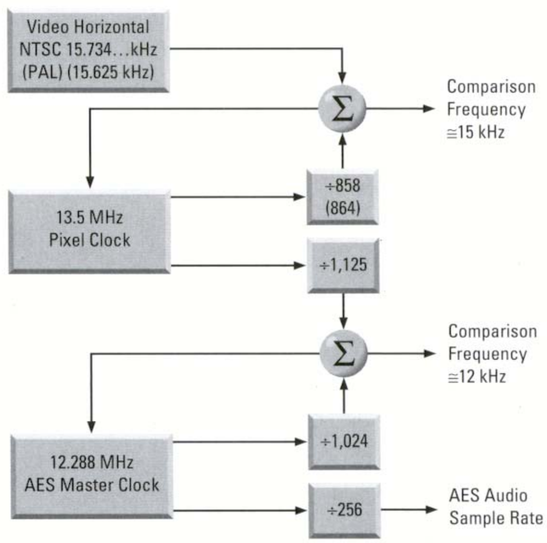

Figure 5: Typical Topology for Locking Audio to Video Hoizontal Line Rate

STRATEGIES FOR SYNCHRONIZATION

The Local Island

One simple and effective technique for locking a piece of audio equipment to a piece of video equipment, a DVTR for example, is to use an extra AES output of the DVTR. This output is a standard AES signal operating at the correct video referred sample rate. The down side is that this connection will probably need to be made uniquely for every session via a patch bay or jumper cables. It is important to note that it is virtually impossible to lock the DVTR to the audio equipment. See *The Book* for more examples and detailed diagrams.

A Small Suite

In the past, for a suite of audio equipment including analog to digital converters (ADCs), which were going to be used for video applications, a local digital audio reference locked to video was used. All audio and video equipment were frequency locked and the conversion equipment could operate with complete phase coherence, provided the digital inputs of the video equipment were used. With this technique, the audio image is completely preserved and the transfer of digital audio between equipment in this suite is assured. Phase coherence generates savings when a production mixer is part of the suite. Since all the inputs are in phase, no input by input delay adjustment is required to correct the audio image.

In today's broadcast facilities, the same type of situation exists. Late model analog VTR's must be integated into the digital plant. NVISION provides low cost multi-channel audio converters and SDI embedders on a single module. This greatly simplifies synchronization and audio channel phasing since up to 8 audio channels are simultaneously converted and embedded.

Video Facilities, Existing and New

The biggest synchronization gain is achieved by locking an entire video facility to a common reference for both NTSC and PAL formats. All audio is running at exactly the same 48 kHz rate. Audio program material may be transferred universally and transparently throughout the facility. The signal feels exactly like analog. The exciting part is that both new and existing facilities can be configured this way with little effort. Simply generate a common timing signal which feeds both the PAL and NTSC master sync generators. This will lock all video equipment to a common time base. Downstream video timing is preserved. For configurations where audio phase is crucial, the suite approach described above is useful. For facilities which already operate in either PAL or NTSC and are expanding into the opposite format, this same topology may be used, or the timing network for the new format may be locked downstream of the existing video network. In either case, all digital audio sample rates can be locked to a common reference. "The Book" offers some excellent examples of how this may be implanted in a few easy steps.

The Fine Points of Plant Synchronization

The most stable oscillator available should exist at the top of the timing chain. Typically, this is a 5 MHz ovenized crystal oscillator, but some facilities may choose to use a 5 MHz rubidium or cesium reference. These three options provide stability which exceeds the NTSC and PAL standards. If a PAL facility is to be slaved to an NTSC facility, be aware that NTSC stability requirements are less stringent than those for PAL. When either video format is slaved to the other, more phase locked loops are introduced into the timing chain than in the case where two video formats are generated from a common starting point. This may cause some increase in jitter in video equipment referenced to the slaved generator.

Audio Sources without Sync Inputs

Converting asynchronous audio to a standard audio sample rate will simplify plant operations. This is particularly true in larger facilities. Asynchronous material exists. It is inescapable. Digital audio equipment without synchronization inputs has been manufactured and used for years. The biggest offender is equipment which offers a vari-speed option. This allows the sample rate to vary by +/- 12.5%, slightly more than one musical step. For example, a 48 kHz nominal sample rate can vary between 54 kHz and 42 kHz. When confronted with this material, there is only one course of action: digital audio sample rate conversion.

A digital audio sample rate converter operates analogously to a video standards converter. The signal to be converted is fed to the input of a dynamic low pass filter which continuously adjusts its output phase, producing interpolated sample values which occur at a rate determined by an external timing reference. Sometimes, the input and output sample rates may be locked together through an integer relationship; for example, 48 kHz and 44.1 kHz are related by the ratio of 160 to 147. This type of conversion is called a synchronous sample rate conversion. Other times, there is no integer relationship between the two rates. This is called an asynchronous rate conversion and is often required when conversions between video formats or video film transfers are made. It is always required for vari-speed applications.

Sample rate converters are widely available, and units which are specifically tailored to the 4 channel architecture of digital video are available. Converters can be installed as shared devices, accessed by a router, or as stand alone devices which are patched in when and where required. Sample rate converter outputs may be synchronized to a reference using all three techniques described above. A number of sample rate converter ICs are available on the market offering exceptional conversion quality. Signal to noise and distortion are both well below 20 bits of resolution.

Broadcast facilities often receive signals from digital telecommunications lines, such as DS-3, or E-3, and satellite feeds. These input signals are not locked to the house reference. A frame synchronizer with tracking audio delay is used to provide synchronized audio and video signals to the rest of the facility.

Signal Path Timing

Digital audio delay lines are essential for preserving the phase relationship between audio and video. Loss of Lip Sync provides easy indication that phase has slipped. The relative timing of audio and video signals varies as video processing equipment is inserted into the signal path. Equipment such as DVEs and color correctors insert delay into the video path which much be compensated for to maintain the original phase relationship between audio and video signals. An adjustable digital audio delay provides this functionality.

A DESIGN STRATEGY FOR SYNCHRONIZATION

No plant should be without this strategy. What follows are some key points to bear in mind as the plant design develops.

- All digital audio transfers require at least two pieces of equipment to be synchronized.
- Almost all digital video equipment provides accurately locked digital audio outputs. Professional audio equipment locks to AES3 or SDIF-2 signals, if not video. Only some consumer gear will not lock to an external reference input.
- Video timing signals can all be locked to one common reference.
- Existing video facilities can be synchronized without disturbing any downstream video timing alignment.
- A phase accurate digital audio reference should be used for conversion between analog and digital formats. This reference should be frequency locked to video.
- Sample Rate Converters should be needed only rarely, and are typically needed only for specific machines.

ROUTING DIGITAL AUDIO SIGNALS

AES/EBU digital audio routers may be sorted into three classes: embedded, synchronous, and asynchronous. Each has its benefits and weaknesses. All are intended to provide automated interconnect of AES/EBU digital audio signals within a facility, or between suites. Routing digital audio data is not as simple as one might think. Design engineers must understand the serial digital signal format being transmitted to the router, the change a particular type of router has on the format of signal outputs it provides, and the effect this change may have for any piece of equipment receiving this output. To make an educated decision between the three routing types, it is important to first understand the AES/EBU serial data format.

The AES/EBU format defines a sub-frame (one audio channel), a frame (an ordered pair of sub-frames), and a block (192 frames). Two audio channel sub-frames make up an AES3 frame as shown in figure 7. If the channels are stereo, Left comes first and is called channel A or 1, and Right is channel B or 2. The C, V, U and P bits indicate channel status, sample validity, user data and sample parity, respectively, for each channel. Validity and Parity refer to sub-frame channel data - the immediate audio sample. Channel status bits accumulate on a block basis to form a 192 bit, or 24 byte, data header which contains unique information about the audio data in each channel. Useful information contained in the channel status bits indicates sample rate, emphasis coding, stereo format, and professional mode. A CRC is also included for the 24 bytes of channel status data. The channel status data headers in a given AES3 data stream are independent. As an example, A channel audio may be emphasized and B channel audio not. Remember, the two channels are not required to be stereo. The User bit does not have a defined application and is not often used. Four Auxiliary bits are reserved for additional audio services. By default, they can provide room for 24 bit audio data, however, these 4 bits can also be used to encode a third audio channel into the data stream. The channel status bits indicate exactly what information these 4 bits contain. Finally, 4 bits are reserved for framing. In fact, these four bits contain three unique bi-phase code violations. These violations indicate if the subsequent data is channel A, B or the start of block - a special case of channel A. Figures 6 and 7 show the AES3 format in detail.

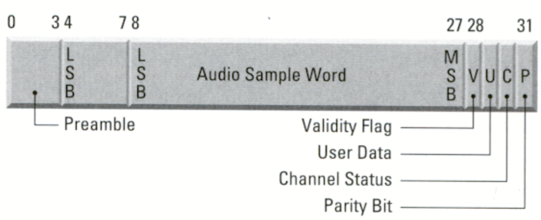

Figure 6: AES3 Sub-Frame Format

Embedded Audio Routing

Embedded audio routing refers to the practice of multiplexing AES3 audio data into a serial digital video bit stream. Embedded routing can be viewed as the most complex technique or the simplest technique for routing audio in a video environment. If audio breakaway is not required and the machines have integrated embedders, an embedded audio router may be the right solution. Current embedding techniques treat digital audio signals as asynchronous inputs. Hot switching of serial digital embedded signals requires significant signal processing to eliminate the effects of clipped or dropped audio frames. This will introduce uncertainty into the actual take point for an audio breakaway relative to the video signal. If audio processing is required, or audio pre-selection is needed, the complexity and cost of stand alone embedders and disembedders will usually point to an unembedded AES pre-select router as the preferable solution. Operationally, a synchronous AES router can execute the take at a point which is exactly phased with the video take. This topology also ensures that multi-channel audio phase is preserved. If an embedded output is desired, the output audio can be embedded into the resultant output video data steam. This signal can then proceed to its destination.

Embedded routers are implemented as space matrices rather than time matrices. The high data rates associated with digital video make time switch architectures impractical. A single embedded router is typically the most expensive option. The number of machines with built-in embedders and disembedders is a prime factor to be considered in any comparative cost analysis. For larger plants where a significant amount of material originates from embedded sources and breakaway matrix dimensions are small, embedded routing may become attractive.

Early embedders and dis-embedders were not designed for clean switch routing, NVISION developed the first embedders and disembedders that avoided putting audio samples in ancillary data

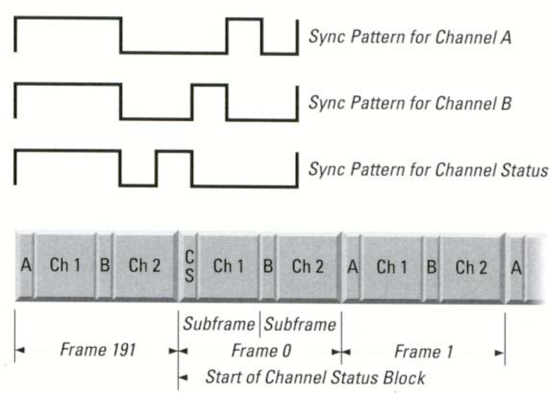

Figure 7: AES3 Frame Format

packets of video lines that could be affected by the video switch. For this reason, the combination of NVISION video routers and embedders and disembeeders provides identical results to that of using an NVISION video router and separate synchronous AES router.

Asynchronous Routing

Asynchronous routing is the simplest in form and the most versatile. This type of router is truly like an electronic patch bay. Because this signal is not processed, an asynchronous router will accept digital audio inputs at any sample rate. The destination equipment must of course be able to lock to the sample rate of the data routed to it. Most equipment will mute if the sample rate is not compatible for its selected mode of operation. The sample rate of asynchronously routed digital audio signals is allowed to change dynamically. This feature allows full vari-speed operation, an important feature for audio production. Asynchronous routers can also switch signals which are synchronized to a common reference. There is one potentially significant drawback to asynchronous routing: when a switch is made between inputs, corruption of the output AES3 serial data format is virtually guaranteed. An AES3 frame will be truncated. Remember that this can also occur when switching embedded audio if the originating embedder was not carefully designed.

When and AES frame is truncated, destination equipment may have to reacquire lock. Equipment must be realigned to the AES3 frame and block boundaries in order to resume normal operation. This process takes a perceptible amount of time, sometimes as long as one second. Some equipment behaves gracefully during this time, that is, it mutes. Not so well behaved equipment may continue to generate audible pops and clicks.

An asynchronous router is effective, easily integrated into any facility, and efficiently operated. Asynchronous routers work well in environments where circuits are pegged for the current job and not changed for its duration. They are the only choice for environments where many different audio sample rates are used. Telecine applications, or professional audio production are good examples. They do not work well as the final stage for ON-AIR switching or where video frame accurate audio editing is required.

Asynchronous routers are always implemented as space matrices. They are also the least expensive option. Price and flexibility make this option very popular.

Synchronous Routing

The synchronous router is more complex. It is designed to provide an AES3 frame accurate switch, predictably timed to the start of a video field if desired. Any transition between two inputs is made at an AES3 defined frame boundary. This insures that the destination equipment sees no discontinuity in AES3 frame or block data formats. The synchronous router requires each input be frequency locked to a common digital audio reference. Synchronous routing is similar to field accurate video routing. A properly timed video switch is visible only as a cut. If not correctly timed, color shift, H-shift, and even vertical roll can result. A correctly timed AES switch provides a clean cut. The alternative is pops, clicks, and mutes.

A synchronous router must receive each AES3 input and align its frame phase to that of a local reference, usually the house AES3 digital audio reference. When a take occurs, audio data is removed from the serial bit stream of the new source and striped into the correct position in the serial bit stream of the output. Each AES3 input is frame aligned and the switch transition also occurs at an AES3 frame boundary. The resulting switch preserves stereo phase and channel assignment. The destination machine receives a synchronous router output data stream which contains no framing interruptions. The result is a clean, instant take - no pops, clicks, or pregnant pauses. An analog type audio break- away is achieved. This transition can be aligned to video. It is also possible to achieve sub-frame routing with similar synchronous techniques.

The synchronous router is the perfect answer for clean “HOT” takes and “ON-AIR “ applications. This makes it an excellent choice for video facilities. It provides a guaranteed video timing relationship for audio transitions. Synchronous routers are manufactured as both space and time matrices. Space matrices offer unbounded expansion. Time matrices are limited by internal bus bandwidth considerations. Synchronous routing is typically more expensive than asynchronous, but less expensive than embedders and disembedders.

TDM Routing

During the 2000 NAB show, NVISION introduced the first Time Division Multiplexed router that provided a robust, reliable method to route mono AES3 signals; the NV7256. It provides glitch free routing of AES3 sub-frames with complete channel status bit preservation. In order to route any mono sub-frame input to any mono sub-frame output, the router core must be synchronous. Certainly, it is possible to reorder signals in a given input signal; for example, swap the left and right inputs of an AES3 input signal. This functionality is now carried out on an input by input, or output by output basis, only. It is not true mono routing. Dolby digital audio requires a 6 channel full bandwidth master prior to compression to either Dolby Digital, AC3, or Dolby E. So, this means working with two Dolby master signals requires 6 AES3 inputs, or 2 SDI Embedded inputs. In either case, the originating machines will be synchronized, and therefore the AES3 sub-frames may be routed separately with the NV7256 router. In addition, the NV7256 is designed to operate with Dolby E frame only switching. And, since all channels status information is preserved any compressed audio material will be recoverable by the output device, provided of course, that device is compatible with the compression format used.

Analog Routing

Analog audio signals still exist. The best approach today, is routers that allow both analog and digital cards in the same frame. These routers, such as the NV7256 and the NV5128, provide A/D conversion, followed by either ASYNC, or SYNC routing. In the case of SYNC routing, Mono routing is also available. Then, at the output, the router can be either digital, or have D/A conversion. The cost difference for analog audio is much less then additional terminal equipment converters, and the router becomes more flexible since it provides a bridge from any input format to any output format. Router control is also simplified with this approach because external tie-line management is not required for the router control system.

Integrated Fades in Routers

It is technically possible to provide a DSP engine in the output of a digital router to generate a fade-fade effect at ever switch. The argument for such capability is that switching instantly from a high signal level to a low signal level can create an audible pop, even thought the AES3 signal frame is perfect. This pop has been true of analog routers for years. The argument against such circuitry is that it adds additional cost for no significant benefit. Audio that is going to air will typically go through master control, where the fade will occur, and then to a compressor prior to FM modulation. Or, in a digital plant, the audio will feed the Dolbly Digital compressor, and the energy of the step, if it even occurs, will be greatly attenuated.

A DESIGN STRATEGY FOR ROUTING

The following steps should help clarify what type of routing is best for a given need and what variables should be analyzed as part of a thorough design analysis.

- Embedded routing is clearly best if the audio is never broken out of the data stream. If this is not the case, another routing option may be indicated. Determine the size of the audio break-away matrix needed. Include the cost of embedders and disembedders to support the matrix. Then compare to a separate audio matrix feeding embedders only.
- “ON-AIR” or clean “HOT” switching requires synchronized AES signals. If this feature is required or desirable, synchronous routing is necessary.
- Asynchronous routing is the most flexible, and can always be used unless a clean transition must be guaranteed.
- Asynchronous routing is the least expensive option. Synchronous routing is typically priced thirty to forty percent higher.
- Embedded routing must be priced as a combination of both audio and video routers. Do not forget to include the price of embedders and disembedders for audio only inputs and outputs.
- Think hard before buying a new router that does not provide integrated A/D and D/A conversion. This is much less expensive than separate external converters and rack frames.
- When synchronous routing is used, all digital audio sources must be synchronized. This is recommended for any routing option.

ANALOG AUDIO CONSIDERATIONS

Analog equipment such as human ears, human vocal tracts, microphones and speakers have not been eliminated. Analog machines which cannot be disposed of for economic or archival reasons still need to be supported. For these applications, ADC and DAC equipment, standalone, or in the router, is required as an interface between the analog and digital facilities. ADC and DAC equipment must be properly aligned to prevent unwanted gain or attenuation. Two factors need to be considered; Full Scale Digital level and input/ output impedance settings.

A Full Scale Digital (FSD) input is that analog signal level which, when converted to digital, results in the largest possible digital code output. This is analogous to the clip level of an analog signal. Full Scale Digital level is usually set to the analog clip level, or maybe 1 dB above. It is important to standardize this level on all equipment which provides conversion between the analog and digital domains. It is possible to find machines which are calibrated to different absolute analog levels. Transfers between these pieces of equipment may generate unwanted gain or attenuation. Avoid transfers of digital material in the analog domain if at all possible. If unavoidable, *The Book* provides some possible solutions for Full Scale Digital level problems.

Input and output impedance selection is also important. Audio signal distribution has traditionally been implemented with a matched 600 ohm source and destination impedance. Advances in equipment design have made it possible to use low source impedance, nearly zero ohms, and high destination impedance, twenty to forty kilo-ohms. This has become the standard practice in the majority of facilities because it does not introduce 6 dB of attenuation at every interconnect. “The Book” provides more details on how to accommodate matched impedance environments, or manage old equipment with 600 ohm terminations.

FACILITY DESIGN STRATEGY

The following list of steps can serve as an initial guideline for designing a digital audio facility. It certainly prompts the main questions which need to be asked at the earliest stages of the design. Hopefully, enough information is provided in this paper, “The BOOK, and other NVISION White Papers, to answer many of the questions which arise during the design stage. Figure 8 is a conceptual representation of an ideal plant which incorporates all the key features and techniques discussed in this paper. Many other approaches are possible, but this one has proven to be adaptable to many different requirements and functional in most any operational environment.

In fact, since this white paper was published, the largest digital audio facilities in the world have been designed using these techniques and NVISION routers. It is safe to say that after nearly 10 years, and tens of thousands of synchronous digital audio router ports installed in the world, that these techniques are genuinely tried and true.

SUMMARY

Digital Audio technology may be readily integrated into any facility, existing or new. Attention to detail and good design practice will result in a reliable, efficient facility. Synchronization and routing are the foundation upon which the plant will be built. An investment of design energy in these areas will yield a substantial return over many years. Plant interconnect requires more detail than with analog. HUM is gone, but jitter and reflections have taken its place. Careful selection of cable and transmission line wiring practice eliminate virtually any opportunity for jitter or signal reflections to disrupt the transparent exchange of digital audio signals within the plant. Digital technology is ready for use.

- Consider the advantages of synchronizing the facility. Direct digital transfers across dissimilar format machines will eliminate unnecessary sample rate conversion, save equipment costs and configuration time. The cost of synchronization is small compared to the benefits which are realized even in asynchronous routing environments.
- Consider Routing Options. Synchronous for ON-AIR or video field accurate editing. Asynchronous for pegged circuits and mixed sample rate audio production. A small asynchronous router as a pre-selector to a synchronous router will allow sharing resources to process asynchronous inputs, or use a router with integrated SRC inputs.
- For synchronous routing, lock all audio and video to a common reference.
- For synchronous routing, sample rate convert asynchronous inputs to the plant master time base.

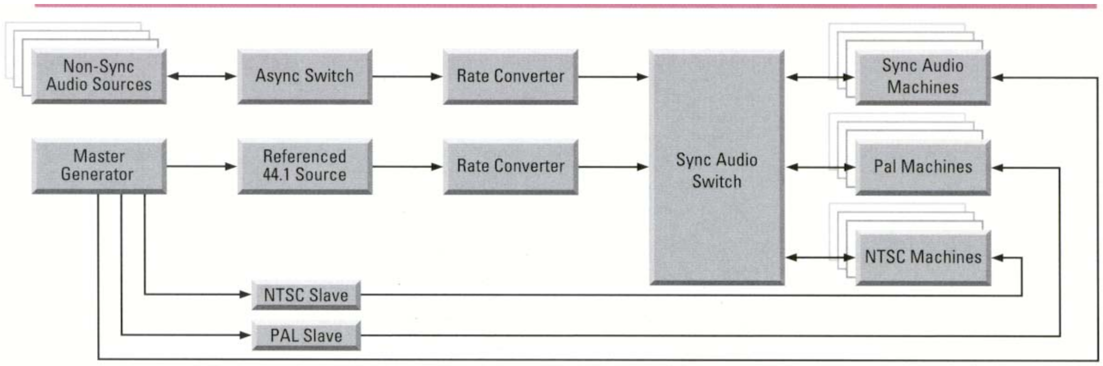

Figure 8: A Conceptual Plant Topology

- Consider an all digital router with ADCs and DACs integrated into the modules. Digital routers are less expensive, take up less room and cost less to operate than their analog counterparts.
- Analyze video signal processing paths to determine if audio delay compensation is required.
- Delay and sample rate conversion units may be used as shared resources with a router.
- Define an operational Full Scale Digital audio level and stick to it. Verify that this is consistent with digital equipment having analog inputs and outputs.