AES3 Status Information – What is it For?

MIKE LAW of BCD Audio explains some of the mysteries of the AES interface and shares his thoughts concerning its development.

The hooking together of audio equipment has always required standards, so that one bit of equipment can be connected to another in a reliable way. The professional audio industry borrowed the balanced audio signal at 0dBm from the telephone industry, and adopted and developed it with great success. Consumer equipment used a simpler format at a lower level of –10dBV over a single wire that relied on the cable screen to complete the circuit.

The AES-EBU standard for the transmission of digital audio was formalised in 1985 to create a similar standard for connection of digital audio equipment. The same familiar balanced pair terminated with XLRs was used, and two channels of audio were specified as most circuits required stereo. In fact, the system was based on the S/PDIF consumer format that had already been introduced for the connection of CD players, enabling stereo linear signals to be connected with a single wire. This wire carries all the information to reconstruct the audio signal, including information about the signal itself, track number and duration, and copy protection.

In keeping with the prevailing analogue interface philosophy, the professional signal is normally used with high-voltage balanced signals and the consumer signal with lower level unbalanced signals. The signal is also readily connected with fibre optics, and as the same system underpins both standards they may, with care, be connected together.

Signal Format

A signal format was chosen to allow for two channels of linear PCM samples at a constant sample rate, plus additional control information, to be transmitted or received with one connection. The maximum sample word length was chosen to be 24 bits, but signals of lower resolution can be accommodated: CD signals are 16 bit, and other resolutions can be used where appropriate. If the word length was 20 bits, for example, the spare four bits could be used for another 'auxiliary' purpose – which made sense when converters offered 18 or 20 bit maximum resolution.

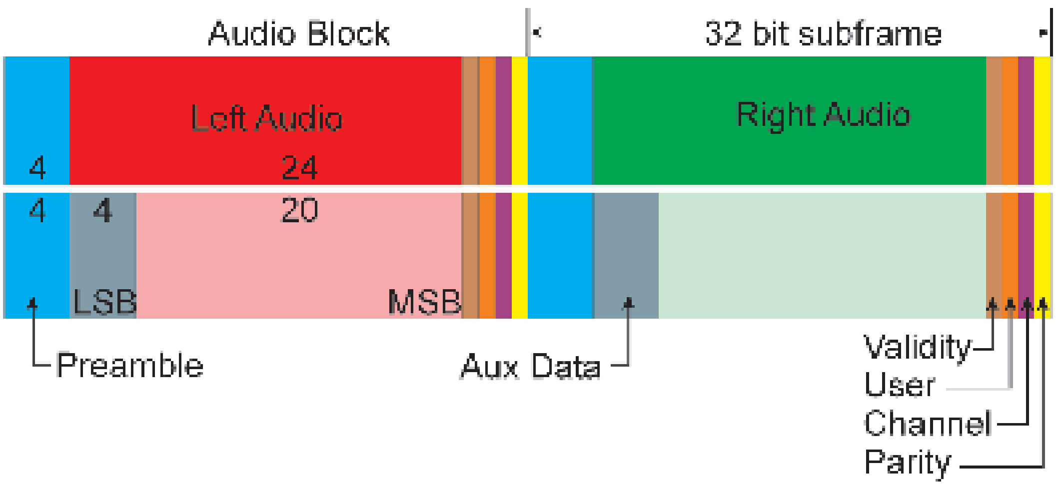
Eight additional bits were added to each audio sample – bringing the total number up to 32 bits – mainly to keep the circuitry simple. Four of these bits are used to keep the receiver synchronised with the receiver, a Validity bit flags whether the audio sample is valid, a Channel Status bit pipes across control information, a User bit conveys track information, and a Parity bit ensures that the whole lot are accurate. If the parity bit is wrong a receiver can mute the output to ensure that only proper signals are heard – but there is no error correction facility. The main purpose of the parity error checking is to ensure spurious signals are not heard when reconnecting and switching signals.

Channel Status Information

The main purpose of the Channel Status data is to convey information about the audio signal so that the receiver knows how to deal with it. The control information associated with the source is not very complex and can be transmitted at a low rate, so it was decided to send only one bit with every audio sample to build a block of useful data, and to recycle this information at a low rate. A block of 192 bits was chosen, so that the whole block is completed after 4ms at 48kHz. Surprisingly, separate blocks are provided for the left and right samples, and the 192 bits contained in each block are often treated as 24 eight-bit bytes.

Although the consumer S/PDIF and professional AES formats send the control information in the same way, they use it in totally different ways. Furthermore, since the information can be different for the left and right channels, the confusion really sets in! Additionally, the structure is dependent on the

The two possible forms of AES data. The top half shows the usual format with 24 bits of audio data. The lower half shows the alternate form with 20 bits of audio per sample, plus 4 auxiliary data bits



semiconductor manufacturers producing chips that allow for the manipulation of these bytes – some are better than others but none implement the standard fully. More confusion!

It is easy to become overwhelmed with this confusion, and ignore it all – after all, if we can hear the signal it must be okay, mustn't it? However, it is worthwhile coming to terms with the format as the reward is the reliable interconnection of equipment. I have listed below my interpretation, and how the various data bits might affect your audio.

Channel Status Fields – A Review

Byte 0, bit 0: Consumer/Professional mode.
The confusion starts with the very first bit, as this decides whether the following bits should be interpreted in one way or the other! As a receiver could be presented with either type of signal it needs to switch between two totally different subroutines or logic to interpret the remaining data based on the status of this bit.
Confusion: 100% Importance: 100%
Usefulness: 100%

Byte 0, bit 1: Linear/Non-linear samples.
This bit is the same in both the consumer and professional modes. If the bit is asserted the audio samples represent digital data (eg. from a data CD) or need special decoding to restore a normal audio signal (DTS or Dolby Digital, for example). A digital-to-analogue converter should mute if this flag is found.
Confusion: 0% Importance: 100%
Usefulness: 100%

Byte 0, bits 2 to 4: Audio signal emphasis.
(See also Byte 0, bits 4, 5 in Consumer mode.)
In the earliest days of digital audio, converters were 14 bit and recordings were optionally made with emphasis so that de-emphasis could reduce converter hiss. Emphasised signals were marked with a data flag so that D-A could switch in the appropriate de-emphasis filter. A similar technique is employed on FM radio to great effect.

Rumour has it that on some early mastering machines the emphasis selection was random (!) and not many CDs were mastered with pre-emphasis. As converters improved, the idea quickly became obsolete. It is also conceivable that there are CDs that are emphasised but don't flag it, and vice-versa. I don't know of any professional system that still uses emphasis, but as it is written into the interface standards it ought still to be implemented. It is worth pointing out that in order for the de-emphasis to work correctly the sample rate needs to be known (see frequency flags).

Most D-As include facilities to correctly de-emphasise a signal, and some AES receivers can perform the task in the digital domain. In digital mixers if the AES receiver can't apply the filter internally (and most don't) the DSP should ideally switch in a filter. I wonder how many do, however.
Confusion: 50% Importance: 100%
Usefulness: -1000000%!

Byte 0, bit 5: Locked/Unlocked.
This little bit could be the most important flag if implemented in a certain way. A big problem with digital audio is to decide whether to switch in sample-rate conversion or not within a signal chain, and this single bit could help.

If the equipment is running from a local, unsynchronised clock this flag should be set to 1. If the equipment is externally synchronised this flag should be zero. Therefore sample rate conversion should be asserted if this flag is set, or if the incoming sample frequency is found to be different to the wanted rate.
Confusion: 50% Importance: 100%
Usefulness: 100%

Byte 0, bits 6, 7: Frequency information.
(See also bits 0-3, Byte 3 in Consumer mode.)
These two bits indicate whether the sample rate is a standard 32, 44.1, 48kHz, or something else. The AES3 standard has introduced further flags in Byte 4 to cover some other sample rates. Sample rate indication is necessary to implement de-emphasis correctly, and useful in deciding if sample rate conversion is required. Unfortunately, the flags are optional in the AES3 standard and are often incorrectly implemented, so are difficult to rely on in practice. Certain AES receivers are appearing now, however, that understand the flags properly, so a transmitter ought to transmit them correctly – and this should become mandatory. Ideally, any receiver should measure the sample rate itself, but if that's not possible it should obey these flags.
Confusion: 10% Importance: 50%
Usefulness: 50%.

Byte 1, bits 0 to 3: Channel mode.
These bits indicate whether the two audio samples should be considered as stereo, two independent mono signals, a primary and secondary signal, or one double rate mono signal. The ability to detect a double rate mono signal is important, and three separate operating modes have been defined – although at first glance they all look identical! On close inspection, the first mode gets the left/right information from Byte 3, the next mode is coded to always be a left signal, and the last mode is coded to always be a right signal.

Equipment implementing double-rate, double-wire connections use this information to select the left and right signals automatically.
Confusion: 25% Importance: 100%
Usefulness: 75%.

Byte 1, bits 4 to 7: User bit management.
These bits indicate the usage of the User bit data block. CDs use these bits for track information and the AES18 standard allocates them for data transfer. User bits are normally not used in most systems, so it is useful to know when they are used.
Confusion: 0% Importance: 25%
Usefulness: 75%

Byte 2, bits 0 to 5: Use of Aux bits and sample length.
Bits 0, 1, 2 indicate whether the audio samples occupy the 24 bit maximum length, or if 20 bit samples with a

separate four bits of aux data are being used. Bits 3, 4, 5 select the sample length in great detail from 16 bits though to 24 bits, with all choices possible. These bit allocations are somewhat historic, in that earlier converters were often 18 or 20 bit resolution. If an auxiliary signal were used, these flags would be used to trap the aux signal from appearing within the main signal and creating noise problems. Aux signals are not used these days [*if they ever were! – **Ed***] and as the lower resolution sample bits are normally zeroed if unused, the effect of incorrectly setting the bits is not devastating.

However, if the transmitter gives 24 bit resolution samples but doesn't set the flags correctly, a receiver might incorrectly truncate your signal to 20 bits! Some sample rate converters obey these flags, and DSP's could correctly dither low-resolution signals if the flags were used properly.

Confusion: 25% Importance: 100%
Usefulness: 100%

Byte 2, bits 5 to 7: Indication of alignment signal.
Digital audio signals are sometimes configured with no headroom (eg, in mastered consumer CDs) while uncontrolled signals and normal broadcast signals are generally configured with some degree of headroom. We all know that it is catastrophic to allow a digital signal to clip, as that introduces all kinds of odd frequencies aliased back from the sample rate.

There is current no defined user practice for implementing these signal alignment bits, but they could solve all kinds of calibration and metering problems in the industry. At present, there is a defined standard to indicate the use of -18.06 and -20dBfs alignment standards, but there are also some unused and unspecified alignment settings available. These could be implemented at another, lower setting for uncontrolled levels – I vote for -30.1dBfs (5 bits of headroom.)

Confusion: 10% Importance: 100%
Usefulness: lots (if used!)

Byte 3: Multichannel modes.
These bits can indicate a channel number, usefully identifying which signal is which. Obviously, the flags would only really be of significance to the outputs of multichannel recorders or mixers, identifying the channel numbers for installation and routing purposes.

Confusion: 75% Importance: not sure
Usefulness: don't know.

Byte 4, bits 0,1: Digital audio reference signal.
These bits should be set if the transmitted signal is to be used as an audio clock reference. There is an argument that reference signals to minimise jitter should have NO channel status information and NO audio signal present – so in this instance your very best master signal might not be indicated as such!

In order to decide downstream whether to apply sample rate conversion, it would be possible to copy these flags from the mixer reference port, to all output ports. However, there are no guidelines to this idea.

Confusion: 100% Importance: 0%
Usefulness: 0%.

Byte 4, bits 3 to 7: Extended sample frequency information.
These bits indicate the new double and quadruple sample rates, as well adding in two half rates. They are useful to complete the picture as the original frequency bits were full. Bit 7 has been added to indicate the video drop frame rate of 1/1.001.

Confusion: 10% Importance: 50%
Usefulness: 50%.

Byte 5: Currently spare.
This byte is currently spare, and is not used.

All AES receiver devices implement Bytes 0-5, but there is a trend not to support the further bytes. The exception is the last byte which is automatically generated and received and contains error protection data. Therefore, the industry should use these bytes or lose them forever! The later bytes are not defined for consumer equipment, which explains why much of the available hardware does not support them.

Bytes 6 to 9: Alphanumeric channel origin.
Bytes 10 to 13: Alphanumeric channel destination.
These are useful information bytes, with four characters available in each of the two fields for left and right. Originally intended to signify the internal routing of large matrices – as an installation and maintenance aid – by monitoring the information at each output the routing can be discovered. The Bytes are not often used in practice, and if they are used it is entirely up to the manufacturer to decide what to send.

Confusion:10% Importance: 50%
Usefulness: 50%

Bytes 14 to 17: Local sample address code.
Bytes 18 to 21: Time of day sample address code.
These bytes can be used to time stamp the signal, although as the channel status information is only updated every 192 wordclock frames (4ms at 48kHz) it is not possible to identify individual samples. Time of day coding comprises two digit hours, two digit seconds and two digit frames, implying a maximum resolution of 10ms. To my knowledge these bytes are not used in broadcast systems, although multichannel recorders might use them. The bytes could alternatively be used for additional alphanumeric information.

Confusion: 10% Importance: 10%
Usefulness: 25%

Byte 22: Reliability flags.
These bits were intended to indicate whether the preceding bytes were reliable or not. Zeros mean the fields are reliable, and ones mean the fields are unreliable. Unfortunately the default setting is zero, which means that in practice the byte is never used, and is ignored!

Confusion: 50% Importance: 0%
Usefulness: 0%

Byte 23: Checksum (CRC).
An AES transmitter adds up all the transmitted bytes in fields 0 to 22 using a defined method, and sends this data at the end of each block as Byte 23. (The hardware

normally does all this automatically, which is handy!) An AES receiver performs the same calculation and compares its answer with the value in Byte 23 – this has to be done separately for the left and right fields. If the answer is different an error is signalled, although it is not really clear what should be done if this occurs.

If the channel status information is in Consumer mode the checksum is not transmitted at all and the receiver does not check the information byte. These errors can also appear if hardware is not producing block-synchronisation signals correctly, which can happen with SDI embedded signals in my experience.

The Future – a Personal View

Quite frankly, as I have tried to show above, the whole Channel Status information block is a mess, and does not really help the industry in obtaining reliable connections between equipment, or aid the maintenance department in diagnosing signal routes and problems. There are three areas where it could evolve to be more useful however, given a will by the industry to implement these ideas.

The first involves sample rate conversion switching and clock reference origins. Information should be added to indicate what clock reference applies to the incoming signal, and this information should be relayed by all devices downstream of that reference. A receiver would then know whether to automatically switch in sample rate conversion or not. Subsequently, a sample rate converter would insert the appropriate information concerning its own reference into the Channel Status data, maintaining accurate clock reference information for the rest of the signal chain.

The second area is in the indication of signal headroom. If signal headroom flags were implemented properly, every piece of equipment would be aware of the correct (nominal) signal level, and be able to set its local metering and signal processing accordingly. Consumer equipment could be assumed to have zero headroom (corresponding to current CD mastering practice!).

Finally, the interface could convey signal path flow information. The alphanumeric text fields could be enhanced so that every piece of equipment in a signal chain adds its own text to the existing text field, in a periodic way, perhaps leaving around two seconds for each message. It would then be possible to identify the signal flow through a broadcast installation in a very simple but useful way. Existing equipment, provided it transfers these bytes from input to output would be fully compatible with this idea, although not contributing additional information. This concept should also be used to indicate the clock reference path.