Illustration by Neil Lucas.

RAISING THE STANDARDS

SEARCHING FOR THE DIGITAL AUDIO DISC

RICHARD ELEN wonders about the future of high-quality audio distribution media.

The next format for the distribution of audio material is a big topic of discussion in the US as well as in the UK. Of course, we are not talking seriously about high-quality Internet music distribution for a while yet, but we are wondering about the successor to CD. Obviously, it's DVD, isn't it? Or is it DVD-Audio? Or Super Audio CD? And what are they, anyway?

DVD-Video offers, in theory, up to 96kHz sampling and 24-bit word-length, although most existing players won't do better than 20/48. The draft DVD-Audio spec, on the other hand, is alleged to contain provisions for more or less anything you like, up to and including Sony and Philips' Super Audio CD, which uses one-bit Direct Stream Digital (DSD) sampling at 2.82MHz.

Standards

There is a big problem here. A standard which is so wide you can drive anything through its doors isn't a standard: it's an excuse. And what actually happens is that manufacturers simply don't implement all of a standard that's too wide. Even so, DVD-Video, which is fairly tightly defined, allows you to release completely legal discs which nobody can listen to — MPEG-2 multi-channel audio, for example. It's an option in the spec, but nobody is going to make players that support it outside Europe, and there are precious few even there. Something like this even happened with CD which, on the face of it, has been a wonderful standard. But the original spec allowed for four-channel surround audio, it's just that nobody made any players, so no-one produced any discs.

I'm a great believer in standards: when well-written, they ensure that everybody can make devices that do the same thing and they will all talk to each other; discs will play on any machine, which will only need to do one or two things. In a perfect digital audio disc world, there would be one standard sample rate and word length, you could output stereo or surround, and it would be playable in stereo on a CD player — or something like that. A well-written standard can even be somewhat future-proof, capable of extension as technological developments permit. Look at AES/EBU for example — or the ADAT spec for that matter. They had room for 24-bit recording even though, at the time, you could only do 16. Look at MIDI.

But the much-rumoured DVD-Audio spec (yes of course you can keep it secret — I'm sure nobody who actually makes records or anything would like to be consulted or find out about it) is much wider than that, not necessarily for the right reasons, being defined at least partly for political purposes (include everyone's suggestions so that they won't leave the group). What do you suppose manufacturers will actually make? Your guess is as good as mine — so here's mine.

Digital audio is operating in an increasingly rarefied atmosphere. We rightfully want to make records with the highest quality gear available — higher quality than consumer equipment, for example, would be nice (this has not necessarily been true for a while, but never mind), and thus we can justify indulging ourselves in ever-higher sampling rates and longer word lengths. In my opinion, improvements in audio quality due to increased sampling rates for conventional digital audio (ie. not DSD) tail off above about 60kHz — but 88.2 and 96 are nice multiples of current practice, so why not? And going from 16-bit to 20-bit is certainly an improvement, while going to 24-bit noticeably gives you a little more.

Not that most people use all that dynamic range, of course. We're still squashing things to make our CDs sounds as loud as possible, just like we did in the old days of black plastic — although at least we hear the benefit of longer digital words on fades and reverb tails.

48 vs 96

Can you and I hear the difference between 24/48 and 24/96, in our comfortable, quiet studios? Probably. How many members of the record-buying public can tell the difference? Not very many.

Even if they have home theatre systems with surround capability, there are plenty of people around who will tell you that they're already happy with 20/48 — and not just members of the public. Noted producer Robert Margouleff (Stevie Wonder, Boyz II Men) here in Los Angeles, for example, is of the opinion that DTS 5.1 surround 20/48 CDs and DVDs already satisfy the vast majority of listeners quality-wise, even discerning ones. Bob argues that digital audio has already surmounted the quality requirements of the content, and that the missing link is surround sound, which we can also provide — and release — now: no future DVD-Audio player required. I would say that we can do surround a good deal better, but that's not to do with DTS or AC-3 or sampling rates, it's to do with the whole concept of 5.1, which is a story I've already told (*Audio Media,* May 1998).

Another person who should know, leading audiophile label owner and producer Gene Pope, of Pope Music over on the East Coast of America, who produces some of the best-sounding discs I've ever heard, actively dislikes 96kHz, because he can hear the jitter.

So, quite a few people here are happy with the currently-available distribution formats (ie. CD and DVD-Video) and don't think that most listeners need an improvement, even if they may use more advanced production equipment in the studio, now or in the future.

Indeed, there is an argument that says that we can't deliver the promise of 24/96 anyway, let alone more esoteric specifications. You may be able to get 24 or so bits to wiggle 96,000 times per second, but that doesn't mean that the data itself carries any more real information than more traditional specifications. Jitter in the clock can be a real problem, and you may need to look at temperature-controlled oscillators, while who knows what will happen to your nice stable clock signal by the time it gets from the clock circuitry to the actual converter section. Meanwhile, the noise level may be limited by Brownian motion in the ➤

THE HDSP PARTNERS

Sonic Solutions announced at NAB an exclusive HDSP Partnership with the leading developers of audio processing technology to deliver a full range of special-purpose processing applications for 24-bit, 88.2/96kHz high-density audio and multi-channel surround sound. Working with Sonic Solutions, these partners will design specialised software for the newly-announced Sonic HDSP plug-in processor, thereby extending its capabilities to a wide range of audio applications. Sonic's new HDSP plug-in processor has been specifically designed to deliver multi-channel 24-bit, 88.2/96kHz and 192kHz high-density audio and to provide audio engineers in CD and DVD premastering with tools that allow them to master for the next generation of digital audio delivery. Through the HDSP partnership, Sonic is seeking to deliver the widest range of tools and plug-in processing capabilities for mastering these new formats.

The HDSP Partners are a group of the industry's most highly-regarded audio technology firms, who are working with Sonic Solutions on an exclusive basis to develop plug-in technology for high-density audio and surround sound. The HDSP Partners include:

George Massenburg Labs — Founded in 1982 by multiple Grammy and TEC Award-winning producer and engineer George Massenburg, GML designs and manufactures high-quality audio products for professional applications. GML is developing a double-precision mastering equaliser and other 96kHz/24-bit mastering tools for the HDSP plug-in processor.

Metric Halo Labs — Known for their sophisticated, real-time feedback and analysis tools for audio mastering, Metric Halo is developing specialised tools for SonicStudio and high-density audio and surround, including on-screen instrumentation and metering based on their award-winning SpectraFoo technology.

Microsonics — A Berkeley, California company, well known for their patented HDCD process for delivering on conventional compact discs the full richness and detail of high-resolution master recordings. PMI will produce a suite of their 24-bit HDCD high-density audio process tools for the HDSP plug-in processor.

POW-r Consortium — The Pow-r Consortium is an organisation of some of the world's finest digital audio designers who have joined to create technology for reducing digital word lengths; they will deliver advanced, high-resolution de-correlated dithering technology based on their algorithm called Psychoacoustically Optimized Wordlength Reduction.

Spatializer Audio Laboratories, Inc. — Founded in 1993, Spatializer's patented 3-D audio processing technology has provided the company with considerable exposure in motion. Spatializer is planning on developing standard and high-density audio surroundising tools for the plug-in processor.

Weiss Engineering, Ltd. — Established by Daniel Weiss in 1984, Weiss manufactures digital audio signal processors for the professional studio, including the Gambit Series; the 102 Series, a modular signal processing system; IBIS digital mixing consoles and the Penguin workstation. Weiss will develop their highly sought-after 96kHz 24-bit EQ for the HDSP plug-in p processor.

Z Systems Audio Engineering — Z-Systems, founded by Glen Zelniker, manufactures digital audio interface and signal processing equipment for the recording, mastering, and high-end audiophile playback markets, and will develop a six-channel equaliser for high-density audio with surround sound.

Some of the partners commented on the aims of the group: "24-bit 96kHz high-density audio is clearly the next step in professional audio recording," said George Massenburg, President of GML. "Sonic's new HDSP plug-in processor gives me the power, precision and control I need to bring our renowned processing to the new audio formats; we plan to make extensive use of its capabilities to deliver new, double-sampling tools for professional mastering.

"We believe that our high density audio technology is among the finest available and we have been looking for ways to do even more," said Daniel Weiss, President of Weiss Engineering, Ltd. "Sonic's new HDSP technology gives us a platform to take our technology in exciting new directions, and we will be designing new high-density audio and surround tools based on its capabilities."

Glen Zelnicker, President of Z Systems Audio Engineering, said: "24-bit 96kHz high-density audio gives us the capability we need to produce the highest-quality DSP and we are enthusiastic about our partnership with Sonic in bringing these new technologies to market."

The new Sonic HDSP plug-in processor utilises an advanced multi-processing DSP-based architecture to deliver the fastest and most flexible hardware platform available. With four parallel-patched 80MHz 24-bit 56301 DSP processors, the plug-in processor can provide up to quad-precision accuracy for ultra-high-fidelity audio processing and stereo 192kHz audio. 1.5MB of ultra-fast SRAM on-board provides significant bandwidth for high-precision processing and customisable plug-in loads. Effects processing can take the form of inserts, sends/returns in user-selectable serial or parallel modes, or controllers, effectively replacing sections of the mixing desk. Multiple boards can be daisy-chained, delivering simultaneous effects processing. Sonic also offers a Hardware Developers Kit (HDK), which provides third-party developers with the information and resources they need to develop applications and plug-ins for Sonic's HDSP plug-in processor.

Sonic Solutions, 101 Rowland Way, Suite 110, Novato, CA 94945, USA.
Ⓣ +1 415 893 8000.
Ⓕ +1 415 893 8008.

analogue circuitry, which is why you can't get real 24-bit conversion systems that even approach the theoretical 144dB dynamic range (although you can actually hear through that noise floor, of course, so there is some benefit there). This is what author John Watkinson has suggested, anyway — but compare converters and see what you think.

The fact that it is unlikely that anybody can make 24/96 converters that even approach the theoretical audio possibilities suggested by the word length and sample rate doesn't mean that the converters can't sound better, of course, and there are two reasons for that in my view: the first reason is a commonly known one and the second I have never seen in print before (so it may be rubbish, but here goes).

Steep Filters

Shannon-Nyquist tells us that the bandwidth of a communications system is limited to half the carrier frequency. In digital audio, that means half the sample rate, minus a little because you have to place filters at the top end to get rid of nasty effects caused by trying to capture frequencies beyond the 'Nyquist limit', or by trying to play them back. These filters are traditionally very steep indeed and, as a result, created one of the big early problems with digital audio: it sounded horrible because of enormous phase shifts in the HF end of the signal, caused by those filters. Errors of several thousand degrees at 10kHz were not unheard of in early digital systems, and this represents one reason why people held on to black plastic for so long (the other big problem was jitter, of course, to which we are far more sensitive than is generally admitted). We found out how to deal with the filter problem to an extent, partly by designing them better and partly by discovering that, if you went for a more gentle and phase-safer slope, the results were much less ghastly than theory predicted.

There is no doubt, however, that higher sample rates allow you to move those filters so far out of the way that they are never going to influence phase relationships (or anything else) in the audio band at all, so there alone is a good reason for doing it. But there is another angle, too.

Ultrasonics

In the article referred to earlier, Mr. Watkinson notes that there is no point being able to record up to 48kHz because, unless we are very young, we can only resolve signals up to 18kHz (and even then we hardly get to 25kHz). This is true, but it does not mean that there isn't anything up there that we need to record — and here's my second reason for higher sampling rates. Plenty of acoustic instruments produce usable output up to around the 30kHz mark — something that would be picked up in some form by a decent 30in/s half-inch analogue recording. A string section, for example, could well produce some significant ultrasonic energy. And what happens is that the ultrasonic content of all those instruments blends together to produce audible beat frequencies which contribute, to a greater or lesser extent, to the overall timbre of the sound. Well, that's what's supposed to happen, anyway. And if you record your string section at a safe distance with a coincident pair, or a Soundfield mic or something, all those interactions will have taken place in the air before your microphones ever capture them. So you can happily record such a signal at 44.1kHz sampling and never worry about losing anything — as long as your filters are decent and you have enough bits.

If, however, you are a total control freak, and your idea of recording a string section is with a couple of 48-track digital machines in the back room, a mic on each desk feeding its own track so that you can mix it all later, you are doomed. Your close-mic technique does not pick up any interactions, so the only place they can happen is when you mix it — by which time the ultrasonic stuff has all been knocked off by your 48kHz multitrack machines, so that will never happen.

So if I was to be uncharitable which, of course, I am not, I could say that high sampling rates allow you to use bad mic technique with better results.

Pick A Number

Having established that higher sampling rates are a good idea — though not whether we can actually do them or not — there is a question as to what the sample rate should actually be in a studio environment. On the face of it, if 96kHz takes care of capturing any audio that might ever happen, and 24 bits offers so many quantisation steps that your fade is as smooth as a… whatever, then that will do fine, won't it?

Well, yes, in a way. But there are some potential problems, real or imaginary, to having a production environment which has the same operating parameters as the consumer distribution format. One way of putting it is a kind of 'headroom'. We need to work at higher resolutions than consumers so we can start with a higher level of quality in case some gets lost on the way, which might well happen. Another way of looking at it is to think about what happens when you modify a digital signal in the digital domain, say by EQ'ing it: you create more bits. You ought to have spare bits (and other things) so you have room to work. You can always lose resolution, but you can't easily get it back again.

There is also a less technical and more, um, philosophical way of considering the relationship of the production environment to the consumer environment. Imagine that you made a traditional vinyl record in the Seventies, on 16-track, in a good studio of the time. Your disc was cut by a true master of the art, and it was pressed on virgin vinyl by an amazing little record company in Wales or somewhere. Of course, the white labels you got back from the record company still sounded horrible, but we were all used to that.

Now along comes some audiophile with a system whose turntable alone cost more than the royalties from the last two albums you produced and, on his system, not only can you hear the comforting hum of the guitar amps, but you can also hear the fluorescent tube up in the ceiling of Studio Two — you know, the one that used to buzz, and we obviously forgot to turn it off that day, it must've been the psychopharmacopeia… and you can hear the lead guitarist tapping his foot in the quiet bit.

Oh dear. You could never hear that in the control room when you were actually making the album, of course, because the band's manager was arguing with the by-then-former lead vocalist ('musical differences'), and the air conditioning was going, and there was the population of a small village in Rutland in the control room (complete with beer, of course) and their girlfriends, and then there was the Dandruff of the Gods of course, that didn't help…

The point is that, for several years after the introduction of quite decent-sounding reasonably-priced CD players, absolutely everyone except you could hear the '80s and '90s equivalent of that guitarist tapping his foot.

We might complain about audiophiles being able to extract more information from a recording than we had ever knowingly put in (actually I have never heard anyone complain about this apart from me, on a long-forgotten edition of BBC Radio London's *Sounds Good*), but I was only slightly joking at the time when I remarked that, quite frankly, listeners at home should degrade the replay quality of their gear to match the industrial audio setting of the studio — with its air conditioners and fans and too many people in the control room and coffee in the faders, and where there were pieces of cooker wire and 500 solder joints between the amp and the speakers, which only had 32SWG wire connecting their terminals to the voice coils anyway and, if you asked anyone if the electrons were flowing in the right direction, they would hit you — so that they could hear our records the way that we heard them when we all agreed that Take 146 was the master (until we heard it the next morning, at least).

Digital audio is operating in an increasingly rarefied atmosphere. We rightfully want to make records with the highest quality gear available — higher quality than consumer equipment, for example, would be nice (this has not necessarily been true for a while, but never mind), and thus we can justify indulging ourselves in ever-higher sampling rates and longer word lengths.

This is why you need to have a production environment that has a higher intrinsic resolution than the consumer distribution medium. They are a whole lot less likely to get more information out of your recording than you knowingly put in. We simply can't afford to have people recovering undefined sonic experiences from your albums, enjoying things that you never knew were there and would have removed if you had.

And so, as a result, if the consumer distribution format of the future is going to be 24 bits, 96kHz sampling (that is the upper limit in the present DVD-Audio spec, as far as I understand it, and as far as regular PCM is concerned), then we will ultimately have to make our recordings at 192kHz — and please don't ask me how many bits are required to provide the appropriate metaphysical headroom because I really couldn't tell you. I hesitate to say '32-bit' because there are so many 'marketing bits' in there by then that I really couldn't say if anyone will ever notice the difference.

Spanners — DSD And Super Audio CD

This, you might say, is quite difficult — and expensive — enough already, thank you, but regrettably there is a spanner already winging its way into the works as we speak, which has been thrown by the very people who brought us 'pure, perfect sound forever' — Philips and Sony.

The latest effort from these two giants of audio innovation is Direct Stream Digital and the Super Audio CD, so cogently covered by Benedick Grant in *Audio Media,* June 1998. Basically, these techniques utilise one-bit converters running at a sampling rate of 2.82MHz and, by all accounts, the technology sounds excellent. In broad, oversimplistic terms, this is equivalent to a data rate similar to 16-bit PCM at 192kHz — four times that of a conventional CD (though that equivalence should not be interpreted as conveying much information about the quality). It features lossless compression, which is an idea whose time has definitely come, and one that groups like Acoustic Renaissance in Audio (the people who have consistently come up with the best ideas for high-quality digital audio discs) have been pushing for years. People who should know, like Michael Bishop at Telarc, say that the initial recordings sound absolutely stunning, and Telarc have recently been out on location recording quite a lot of material with DSD — they are obviously taking it seriously.

These technologies represent a couple of dark horses here, however. First, they are scalable technologies, in the sense that you can do simple, cheap, OK-sounding converters relatively easily (for an SACD 'WalkPerson' for example) while, with a good deal of effort and expense, truly high levels of quality are available. So the same recording — the same physical disc — could satisfy the teenage jogger, the in-car player and the audiophile. This is a very encouraging prospect. Converter manufacturers would not have too much trouble developing conversion systems for this technology in some ways, as the sigma-delta style of conversion used in the vast majority of outboard converters today is still employed — you simply record the one-bit stream directly instead of decimating it.

Secondly, assuming you don't simply want to convert your recording to DSD at the mastering stage — which many might see as rather missing the point of using the technology at all — then you need to replace virtually every piece of digital equipment in your studio. Oops.

DSD Converters

First to appear on the DSD front, as one might expect, have been stereo converters and

Michael Bishop, of Telarc, says, of DSD, that initial recordings sound absolutely stunning.

recorders. Sony have been developing this technology for a while, and now third parties are beginning to get in on the act — dcs have probably released their DSD converter by now, for example. The next requirement would be an editing system, and Sonic Solutions announced such a system at the Amsterdam AES.

This is about all you need for classical recording. Multi-channel DSD converters to capture a surround signal, a recorder to store the output of the converters, an editing system to put it all together and, while you're at it, put a couple of extra mics up and record a stereo version at 44.1 for the Red Book layer that's also on the disc (or the CD version if you're still doing them).

In fact, I have heard differing reports about whether either PCM-based DVD-Audio or SACD would have a Red Book layer: the arguments for single inventory are strong, and I suspect this is what will happen, but some record companies seem to have become less concerned about single inventory lately.

When it comes to multitrack recording and mixing, however, rather more gear is involved, and this is where DSD gets a bit scary. If we end up with a distribution format with traditional PCM — ie. a 'DVD-Audio' disc offering up to 24-bit at up to 96kHz, then all we do is upgrade our studio gear bit by bit until we at least reach, and possibly exceed (depending on the headroom we would like to have) that level of performance. It's technology we know; at some level we kind of understand it; and it's an evolutionary strategy: comparatively safe. The only obvious problem is that generating a 44.1 version will not be simply a matter of putting up a couple of mics for most people: DVD is based around 48kHz and multiples, while Red Book CD is based around 44.1. Sample-rate conversion from 96 to 44.1 may not sound very nice — it's not a simple 'divide-by-n' — which might mean that we will have to do two mixes; one stereo at 44.1 and one 5.1 at 96. This may be why there are second thoughts about single-inventory.

DSD, on the other hand, is a revolutionary strategy. It may be, after all, that we can use PCM systems in the studio and convert to DSD at the end of the day, in which case we just determine that some level of PCM performance (say 24/192) gives us enough headroom to make a multitrack-derived Super Audio CD, and the classical people are the only people who need to invest in (relatively simple) systems relying on DSD throughout.

It may, alternatively, be that DSD production techniques themselves need to provide headroom above the specification of Super Audio CD, and that, in the studio, it will actually be necessary to use some enormous sample rate, listenable to directly on a short-wave radio, that will down-convert nicely to both 44.1 and SACD. In any event, if we end up using DSD for production, a whole load of gear is required, almost all of which is currently imaginary, although, in quite a few cases, it may simply be a matter of changing the type of converters and recorders that sit on the outputs of the analogue mixing console — if, that is, analogue consoles are in any way sufficient to create either of the next generation of digital audio discs.

Crystal Gazing

I am afraid that, at this point, my crystal ball goes cloudy and my spirit guides retreat back into the quantum ether from which they may, or may not, have come. I simply don't know what subset of the devices specified in the Flying Rolls jealously guarded by the Ancient and Mystical Order of the Digital Audio Disc will actually be manufactured by its members, however wide the allowable specs might be.

My guess is that we will end up with a consumer audio distribution format based either on SACD, or upon PCM discs recorded at 24/96 (or both, which might be the worst of all possible worlds). Both will have a Red Book CD-compatible layer, and 5.1 as well as stereo mixes in most cases.

Plenty of acoustic instruments produce usable output up to around the 30kHz mark — something that would be picked up in some form by a decent 30 in/s half-inch analogue recording. A string section, for example, could well produce some significant ultrasonic energy. And what happens is that the ultrasonic content of all those instruments blends together to produce audible beat frequencies.

We have all the signs of a format war on our hands that, as always, will be expensive and controversial however it falls out in the end. The worst-case scenario — both formats co-existing — could have the same result as that caused in the past by having two incompatible open-reel digital systems, which arguably held back the introduction of serious digital recorders into the studio for years and resulted in both formats being eclipsed in many minds by the MDM.

The question we get asked most frequently at my place of work, is 'Does it do 96kHz?' Everybody thinks that's what they want. What they actually need — or will need — may be quite a different matter. ❑

Richard Elen is a frequent writer on professional audio can be contacted on relen@igc.org. He is VP of Marketing at Apogee Electronics in California. The views expressed in this article are the author's own and are not necessarily those of his employers.