A NATIVE STEREO EDITING SYSTEM FOR DIRECT-STREAM DIGITAL

James A. Moorer, Sonic Solutions
Ayataka Nishio, Sony Corporation
Yasuhiro Ogura, Sony Corporation

The editing in native form of 2.8224 Mhz 1-bit audio (Direct-Stream Digital) is problematic since the required computation rate to achieve level control, pan, and crossfade in real time is beyond the capabilities of general-purpose digital signal processing chips. A system has been developed which incorporates a special-purpose chip designed specifically for computations on 1-bit audio. This is embedded into the framework of a general-purpose audio editing and premastering system. It behaves, looks and feels exactly like a stereo, PCM editing system with full flexibility of sample-accurate cross-fades, level control, metering, and non-linear random-access editing, except that internally it is operating directly on the 1-bit signal. This paper describes the data flow and architecture of the system.

Introduction:

One of the barriers to wide use of the 1-bit 2.8224 MHz Direct Stream Digital audio ([1], [2]) is the lack of editing systems that are capable of dealing with the audio in its native form. Although we have more than 20 years of experience editing, mixing and producing audio in PCM form, we are just beginning to deal with the difficulties of the high-rate 1-bit audio format. The sampling rate of Direct Stream Digital (or DSD®) is high enough that general-purpose digital signal processors are not capable of doing any but the most trivial of operations on DSD directly (in real-time). Any kind of non-trivial processing requires custom silicon at this time.

A custom chip has been developed expressly for the mixing and editing of DSD in native format [3]. This chip has been incorporated into a stereo editing system so that level control, pan, and editing can be done directly on the DSD signal in real-time. When no processing is applied, the signal is copied in a bit-for-bit accurate manner.

A 24-bit PCM signal accompanies the DSD signal at every stage. This is useful for waveform display and for metering. As the DSD signal is processed, the PCM waveform is derived in real-time so that accurate metering and waveform display is possible at each stage. At no time is the PCM signal substituted for the DSD signal in the audio chain.

This paper will describe the architecture of the system, the data flow, and possibilities for future expansion.

Goals and Philosophy

The design of the editing system was driven by a number of separate desires. The first and most important was to present a standard, familiar interface to the editor. We wanted the editing of DSD to look exactly like editing of PCM. All the handling of DSD is done automatically in a transparent manner.

Another goal was to use standard hardware as much as possible. To this end, our standard USP[1] boards were used. The two pieces of new hardware are the DSD I/O box, which interfaces to DSD A/D and D/A units, and the DSSP board, which embodies the E-chip for DSD processing. The DSSP communicates with the USP using its standard audio I/O interface. This carries up to 4 channels of DSD data and up to 4 channels of 24-bit PCM.

The result of this decision is that the system can alternately be used for PCM editing and DSD editing by just reloading the DSP code on the USP boards. In the same fashion, standard NoNOISE® processing is available for PCM data. It also makes it possible, in the future, to write software-based (non-real-time) algorithms for DSD itself. We will be faced with releasing 16-bit, 44.1 kHz recordings for some time to come, so it seems prudent to be able to edit and master standard CDs on the same system.

Note also that this automatically provides us with real-time networking through MediaNet®. No extra work need be performed to allow a number of editors to work on

[1] The Ultra-Sonic Processor is an audio processing board for the PCI bus. It has four 56002 DSPs. There is a serial audio I/O port capable of sending and receiving up to 60 bytes of data on each sample period. There is a SCSI interface directly on the board with DMA access into 12-Mbytes (24-bits x 4Mwords) of DRAM. It has board-to-board connections for routing data to other USPs in the system.

common projects. Similarly, standard background operations, such as tape backup and restore, can proceed simultaneously with editing for efficient workflow.

Although the current system architecture is somewhat bulky, it can be smoothly extended to 4 or 6 channels of DSD editing should that be desired.

The E-Chip

Since the E-chip has been described elsewhere [3], we will simply review the capabilities here. The device takes 4 DSD signals as input and produces 2 DSD signals as output. There are 8 24-bit coefficients that are sent to the chip every 22.6757 microseconds (once per sample at 44.1 kHz, or once every 64 DSD samples). Internally, the coefficients are linearly interpolated for each DSD sample. The 8 coefficients form all possible connections from the 4 input streams to the 2 output streams. There is a switching matrix after each channel of processing that performs the following functions:

- When the coefficients for one output channel are all set to zero, the switching matrix sources a dither pattern that results in a zero level signal in the audio band.
- When the coefficients are set so that one and only one input is directed to the output with a gain of unity, then the switching matrix simply selects that signal and forwards it to the output without change.

There is a special set of circuitry that controls the switching matrix so that after an edit, the matrix waits for the noise shaper to "settle" and begins reproducing the input sequence, then it switches to the input sequence, bypassing the processing entirely. Obviously, this is only possible if the gains after the edit are exactly unity. Otherwise, the noise-shaper must continue operating.

The E-chip also has a stereo PCM input. This allows us to other processing (*e.g.,* equalization) on the PCM signal, then modifying the DSD signal accordingly.

The DSSP Board:

To use the E-chip in the system, we need to be able to send and receive DSD data from the USP board. We also need to be able to calculate fades in real-time, and forward 24-bit coefficients to the E-chip in a sample-accurate manner. Since we only have 8 multiply operations for the two output channels, the trim gain, pan gain, master gain, and fade coefficient must be combined into a single coefficient for each data path. For real-time controls, such as sliders and pan knobs, smoothing must be applied on a sample-by-sample basis to eliminate all "zippering" effects.

The architecture of the DSSP board is shown in Figure 2. The data from the USP is separated into four serial DSD streams and two PCM streams. These are delivered to the E-chip audio data inputs. There is a DSP56301 on the board which calculates all the coefficients and sends them to the E-chip. The DSP does all the calculations noted above for each coefficient, including smoothing and forwarding to the E-chip in a sample-accurate manner. There is an interrupt from the audio clock circuitry to the DSP at the beginning of each sample period for lock-step synchronization.

The DSP deposits the 10 coefficients (8 for the DSD data and 2 for the PCM data) into an FPGA which then serializes and multiplexes them for the E-chip.

The data from the E-chip is multiplexed and packed for transmission back to the USP. It is also converted to 24-bit PCM that is also sent to the USP.

.

System Architecture

Figure 3 shows the architecture of the DSD editing system (the host computer is not shown). There are 3 PCI boards that are used: two USP boards and one DSSP board. Optionally, a MediaNet board may be installed for networked operation. The first USP board is connected through a bit-serial link to the DSD I/O box, which is then connected to the DSD AD and DA converter. The second USP communicates to the DSSP board using two bit-serial links.

The division of tasks between the boards is relatively straightforward. The first USP is responsible for DSD I/O, audio clocking, and for direct recording to the hard disk. On each sample, we record two channels with 8 bytes of DSD data and 3 bytes of PCM data. This is a combined data rate of 970,200 bytes/second, which is easily within the scope of the SCSI bandwidth of the USP. A standard 9-gigabyte hard disk can hold over two hours of stereo DSD (and PCM) recording. Up to 6 drives can be connected to the system. The DSPs on the USP are 24-bit devices. The DSD data is packed into the high-order 16-bits of four consecutive 24-bit words. When the data is transferred to the disk, a hardware byte sequencer automatically extracts the high-order two bytes of the 24-bit word. The low-order byte of each word is used by the DSD I/O box for status and control information (*e.g.*, the lock state of the PLL in the DSD I/O box). Figure 4 shows the data packing into the bit-serial stream that the USP uses to communicate both with the DSD I/O box and the DSSP board. The USP can handle two such streams. For the DSSP, both streams are used to send 4 channels of DSD and PCM data to the DSSP board, and one stream is used to receive 2 channels of DSD and PCM data from the DSSP.

The same SCSI bus is attached to both USP boards, so both boards have equal access to the hard disk. The second USP is responsible for digital I/O to and from the DSSP board. For direct processing, the DSD data goes across the board-to-board interface from USP #1 to USP #2 and then to the DSSP. Data coming back from the DSSP may be written to the disk directly from USP #2, or can be sent back across the board-to-board interface to USP #1 and then back to the DSD I/O box for auditioning.

On the hard disk, each channel of DSD data and PCM data occupy a different file. The data are not normally interleaved into a single file. This allows straightforward generalization to more than 2 channels. There are 4 files associated with each stereo recording (2 PCM and 2 DSD).

The audio clock source for the entire system comes from the DSD I/O box. In the case that the DSD A/D is in use, then the clock comes from the A/D. The DSD I/O box has an internal crystal for self-clocking, and a word-sync input for external reference. There is a PLL in the DSD I/O box that generates a 256*Fs clock (11.2896 MHz) that the USP uses for digital I/O. This clock is routed from the receiving USP to the other boards in the system.

Future Directions:

Now that we have some experience with the DSD editing system, we can look to a more integrated system. The functions of the two USP boards and the DSSP board can be combined into a single board so that a 1-board stereo DSD editing system is possible. Although equalization directly in the DSD domain requires custom silicon that does not exist at this time, it is possible to realize some amount of equalization by processing the

PCM data on the DSP, then feeding the difference (before and after equalization) to the E-chip so that it may be added into the DSD stream and requantized to 1-bit DSD. A 1-board solution also provides a straightforward extension to multi-channel DSD recording and processing, even if the final product is to be stereo.

Conclusions:

An editing system for stereo DSD data is designed to look and feel exactly like the editing of normal PCM audio. Metering and waveform display is done using a 24-bit PCM signal that is carried along with the DSD at each stage. The I/O can be synchronized to a common clock for smooth, if bulky extension to multi-channel use.

For normal editing functions (splicing, mixing, panning, gain adjust), *all* processing is done in the DSD domain. The companion PCM data is never auditioned or mixed with the DSD data.

References:

[1] A. Nishio, G. Ichimura, Y. Inazawa , N. Horikawa, and T. Suzuki "Direct Stream Digital Audio System" Proceedings of the AES 100th Convention, Copenhagen, Denmark, May 1996

[2] A. Nishio, M. Akune, G. Ichimura, Y. Ogura, and Y. Tsuchida, "A New CD Mastering Processing Using Direct Stream Digital" Proceedings of the AES 101st Convention, 8-11 November 1993, Los Angeles

[3] Masayoshi Noguchi, Gen Ichimura, Ayataka Nishio, and Shigeo Tagami “Digital Signal Processing in Direct Stream Digital Editing System,” Proceedings of the AES 102nd Convention, Munich, Germany, March 1997

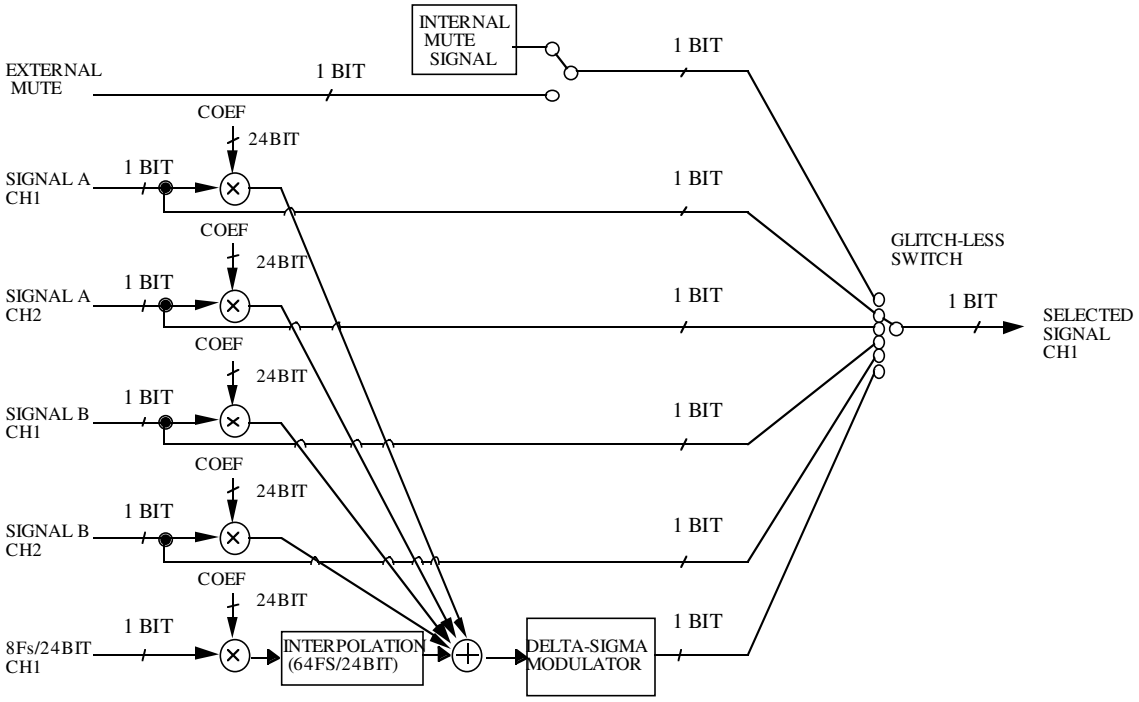

Figure 1: Simplified block diagram of one channel of the E-chip processing. The coefficients are 24-bits. The output of the summation is 26 bits. The coefficient smoothing is not shown, nor is the steering logic for the glitch-less switch.

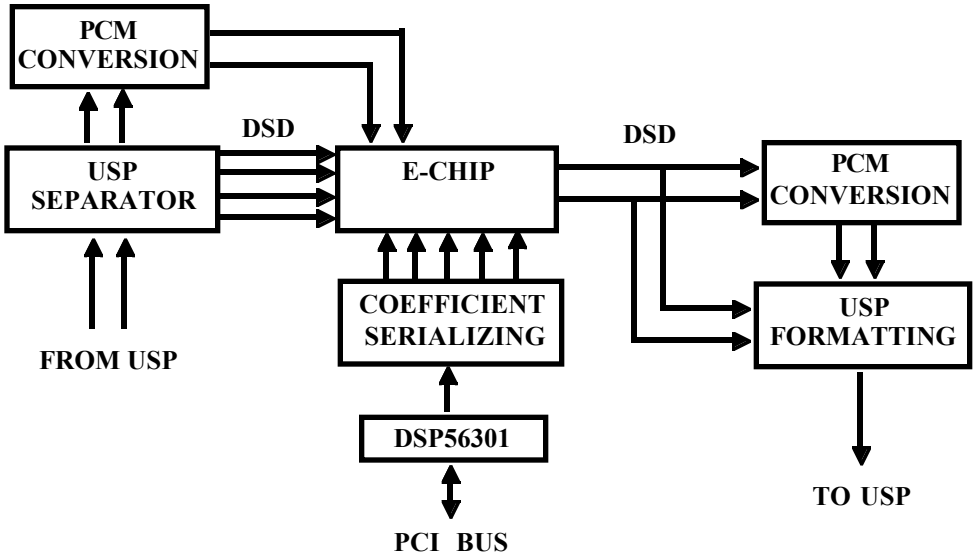

Figure 2: Simplified block diagram of DSSP board. Data comes from the USP in a packed, bit-serial form. The DSD and PCM data are separated into individual serial streams, then sent to the E-chip. The output of the E-chip, plus 24-bit PCM data, are packed into bit-serial form and sent back to the USP. The DSP56301 is responsible for synthesizing fades and forwarding coefficients to the E-chip in a sample-accurate manner.

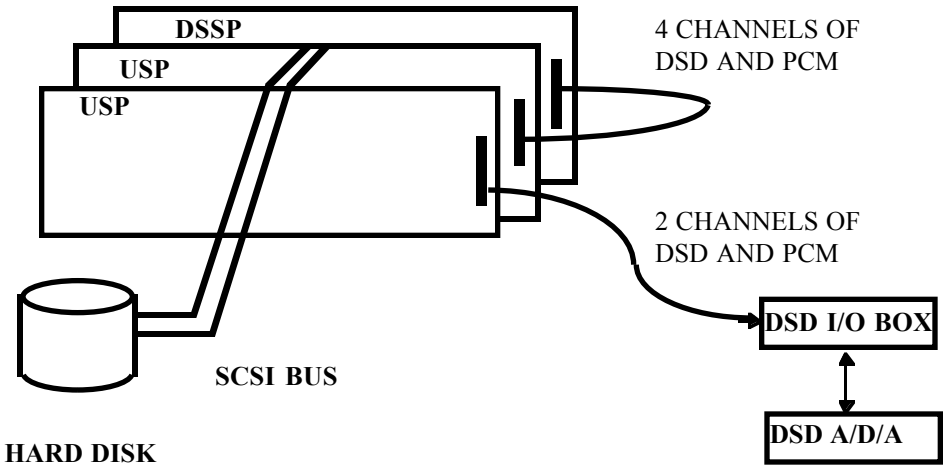

Figure 3: System diagram of DSD editing system. Two USP boards are used for data handling, data formatting, disk control, and control of the PCM data. There is a direct connection between the two USP boards that is not shown here.

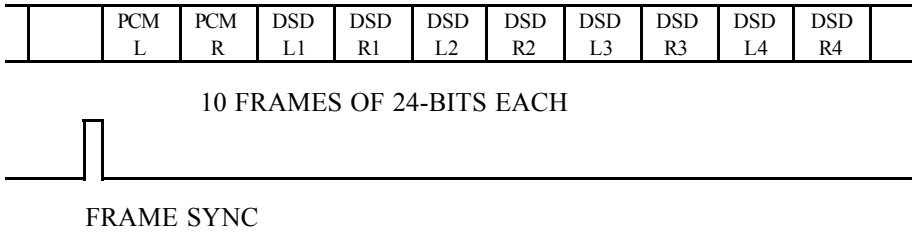

Figure 4: Data packing in the bit-serial stream used by the DSD I/O box and the DSSP to exchange stereo PCM and DSD data. The DSD data occupy the high-order 16 bits of each 24-bit frame. The low-order 8 bits of the DSD frames are used for status and control.