



Technical Brief

**NVIDIA nForce MCP
Audio Processing Unit**

*N*VIDIA

I. Importance of PC Audio

The last five years have seen tremendous improvements in the multi-media PC experience. One of the most exciting advances has been in PC audio. PCs of the mid-90s had crude monophonic sound whereas today's advanced PCs have rich, multi-channel, three-dimensional audio capability. However, as is often the case with emerging technologies, compelling 3D audio and the resulting theater-quality sound has come with a price. Those willing to part with the approximately \$200 to \$300, were given a front-row seat to the 3D audio festivities. But others, who either didn't want to purchase additional hardware, or were convinced that their integrated sound chip was somehow 3D-compliant, and already capable of "mimicking" Dolby™ Digital 5.1 audio, were left behind.

With the proliferation of first-person games, such as *Quake 3*, the advent of 3D positional audio is now much more important to developers and users alike than ever before. This, combined with the improved audio feature set in Microsoft®'s DirectX® 8.0, including realistic 3D, sub-mixed audio streams, and occlusion/obstruction effects, have set the stage for much more realistic gaming environments. NVIDIA®'s Audio Processing Unit (APU), already found in Microsoft's Xbox™, and an integral part of the NVIDIA nForce™ Media and Communications Processor (MCP) is the only one of its kind. Primed to bridge the gap between games and reality, the nForce APU addresses everyone's increasing appetite for full-featured, high-quality, real-time encoded Dolby Digital 3D audio, and greatly surpasses the audio capabilities of other sound cards and sound solutions, integrated or otherwise.

II. What makes an APU?

With gate counts far exceeding those of CPUs, 3D graphics architectures and GPUs are pushing the envelope every six months. Unfortunately, "leading-edge" 3D audio solutions have yet to make that quantum leap in technological innovation. Enter the nForce Audio Processing Unit (APU). Like its graphics counterpart, the GPU, the APU's main function is to help increase overall system performance by offloading complex audio algorithms and effects processing from the CPU. But, unlike traditional sound cards with an extremely limited function set, an APU is defined by the following three parameters:

- It must have an advanced architecture featuring, at a minimum, support for hardware acceleration of 256 2D voices and 64 3D voices, and advanced 3D positional audio
- It must fully support all DirectX 8.0 features and capabilities
- It must integrate support for Dolby Digital 5.1 and possess a Dolby Digital encoder in hardware to deliver an uncompromised, cinematic-quality audio experience

III. APU Architecture: An Introduction

At its core, the Audio Processing Unit is a multi-processor audio rendering engine (see Fig. 1). The APU is responsible for providing hardware audio acceleration for both output streams (playback) and input streams (record). The APU renders completely to system memory. This decoupling allows the resulting stream to be transferred to any transducer including an AC97 CODEC or a USB speaker system.

The APU is divided into four main sections:

- Setup Engine – This unit is responsible for performing all data and parameter setup for the other processors. All memory management, mapping and DMA resources are controlled in this unit.
- Voice Processor – This unit contains several fixed function digital signal processing (DSP) units responsible for processing voices and mixing the results in the mixer buffers.
- Global Processor – This unit is built around a programmable DSP. The DSP is responsible for adding varied effects to the data in the mixer buffers and producing the final output stream to the OS.

Dolby Digital Interactive Content Encoder – This unit is built around a programmable DSP, which is responsible for encoding Dolby Digital (AC-3) data that'll be sent over the SPDIF to an external consumer decoder. This allows 5.1 speakers (left front, right front, center, right rear, left rear, sub-woofer) to be transmitted over a digital interface.

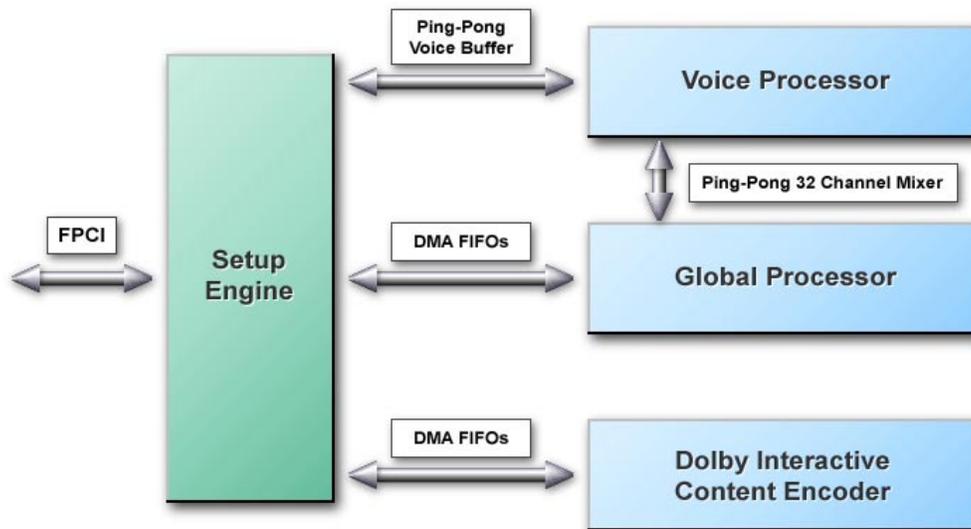


Figure 1: A high-level block diagram of the APU

Setup Engine

The Setup Engine is where a great deal of the complexity lies. The Setup Engine performs the following functions:

- DMA channel management and prioritization.
- Management and updating of the voice structures in system memory.
- Processing of the voice execution lists—2D and 3D.
- Parameter setup for the voice processor.
- Data de-interleaving for >2 channel voices.
- Data formatting—all data is signed 24-bit going into the voice processor.
- Data gathering—the voice processor is presented the sample data it needs in a sample buffers.
- Taking care of alignment and loop conditions, if needed.
- Down sampling averaging, if needed.
- Gathering output data, interleaving in the correct format, and then DMAing to system memory.

Voice Processor

The Voice Processor (VP) will render all the 2D and 3D voices. The HRTF will run in parallel with the rest of the 2D processing. All samples are mixed into one of 32 mixer bins. The input data and parameters are ping-ponged on a voice basis.

Global Processor

The Global Processor is a programmable DSP. It will perform the following functions on each frame:

- Global Effects (reverb, chorus, flanger, etc.)
- Equalization
- 3D Cross-talk cancellation
- I3DL2 Reverb, Occlusion & Obstruction
- Post mix to Setup Engine

Dolby Interactive Content Encoder

The nForce APU integrates the Dolby Digital Interactive Content Encoder into a programmable DSP with a fix-to-float format engine. It's used to take the output of the Global Processor and encode it into a Dolby Digital (AC-3) stream. Now, users will be able to experience true theater-quality, multi-channel surround sound, *rendered in real-time*, from their Dolby Digital-equipped PCs and home-theater set-ups.

IV. Advanced Features

With five DSPs (three fixed, two programmable) dedicated to audio processing, the NVIDIA nForce APU surpasses the capabilities of any sound card on the market today. The APU supports all of the latest features in the DirectX 8 audio API, DirectSound[®] 3D (DS3D), with up to 256 hardware-processed voices (audio streams) or 64 hardware voices in 3D. Support for multi-speaker 3D audio is available for up to six speakers. Furthermore, the APU can also process and output a Dolby[™] Digital audio stream directly to a home theater system via a SPDIF connection (see fig. 2). With such an advanced feature set, the APU brings audio quality and performance to the same level as state-of-the-art graphics performance, completing the 3D experience.

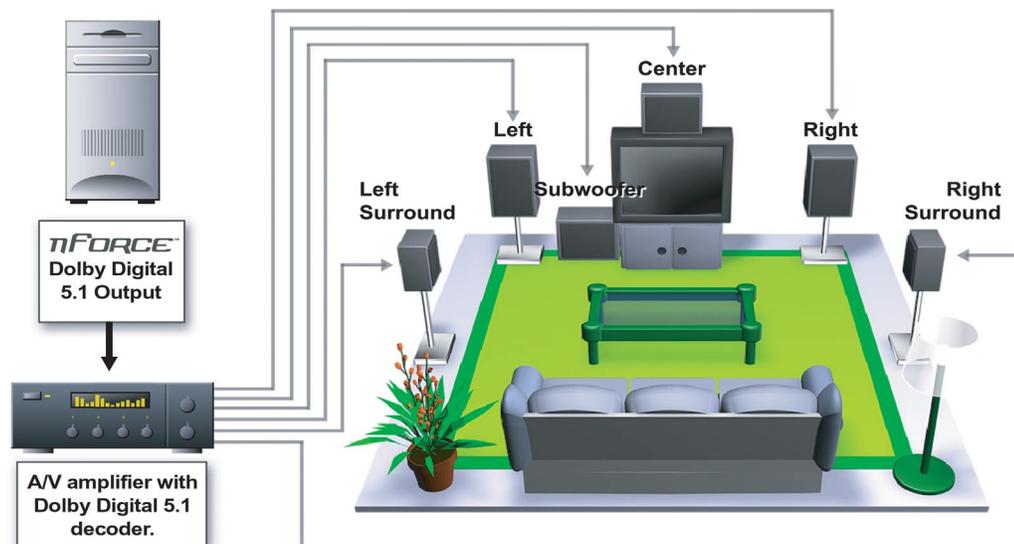


Figure 2: Support for up to six-speaker output and Dolby[™] Digital output via S/PDIF.

DirectX 8.0 Features and Benefits

DirectX 8 introduces a whole new set of audio features the CPU must process if there is no dedicated audio hardware to support it. The APU was designed as a DX8 audio processor and has the ability to process advanced features in hardware such as:

- Downloadable Sounds Version 2 (DLS2) – The APU can perform layering and filtering using stereo waves as the source.
- Effects processing on DirectSound buffers – More sophisticated sound effects as well as higher quality music and sound effects with greater production values and real-time interactive control.
- Interactive 3D Audio Level 2 (I3DL2) - Reverberation and occlusion/obstruction support.

- Sub-mixed audio streams - Complex, rich sounds can be created by sub-mixing two or more audio streams.

Furthermore, the APU is the highest speed audio processor available. In addition to the above DX8 features, the APU also processes the following in hardware:

- Head Reference Transfer Functions (HRTF) algorithms with cross talk cancellation.
- Near field effects.
- Macro effects.
- 7-band graphic equalizer.

V. Features and Benefits

In comparison to other integrated sound solutions or add-in sound cards, the APU's high-level of performance and advanced feature set clearly places it in a class all by itself. Here's a quick rundown on its major features:

Feature Comparison	NVIDIA nForce APU	Current \$99-\$149 Add-in Sound Card
DSP or HW-accelerated 2D (Stereo) Voices	256 voices	64 voices
DSP or HW-accelerated 3D Voices	64 3D voices	32 3D voices
DX8 HW-Submixer	32 HW submixers	No
DLS2 Acceleration	✓	No
Per Voice Parametric EQ	✓	No
Occlusion and Obstruction	✓	No
Near Field Effects	✓	?
EAX2 and I3DL2 Reverb	✓	✓
Globla Effects (Reverb, Chrous etc)	✓	✓
Speakers	2, 4, 6 speakers	2, 4 speakers
S/PDIF Interface	✓	✓
H/W Dolby Digital Encode	✓	No

VI. Conclusion

With features such as the most advanced 3D positional audio, a real-time Dolby Digital encoder, and full support for DirectX 8.0 audio, the NVIDIA nForce APU exceeds the features and performance of today's top-of-the-line sound cards and other integrated sound solutions. When combined into a PC distributed processing design, such as the NVIDIA nForce Platform Processing Architecture, or game console such as Microsoft's Xbox, the end result is an uncompromised 3D audio experience on all levels.

Appendix A - APU Feature List

- **256 Total Voices**
- **Input Format Support**
 - 1-18 samples per block
 - De-interleaving
 - 8-, 16- and 32-bit containers
- **Output Format Support**
 - 1, 2, 4, 6 samples per block
 - 16- or 32-bit containers
- **DirectX 8 Capable**
 - Sub-mix bins
 - Second pass processing
- **64 3D Voices**
 - HRTF with cross-talk for speakers
 - I3DL2 reverb
 - Occlusion and obstruction
 - Near field effects
 - Full cross-fade per voice
- **DLS2 Acceleration**
 - Two envelope generators per voice
 - Two LFO generators per voice
 - Loop and release segments
 - DLS filtering
 - Pan, pitch, vibrato, tremolo
 - Reverb and chorus send
- **32 Bin Mixer**
 - 8 per voice volumes are mapped to each bin
 - DX8 loop back processing
- **7-Band Graphic Equalizer**
- **I3DL2 Reverb and Occlusion**
- **Hardware/Software Interface**
 - Complete system memory based structure and control
 - Fire-and-forget interface – minimum software control

Appendix B - Glossary

Block: The size of container for an audio sample. The container is sized in bytes. For example, APU supports 1, 2 and 4 byte containers.

Buffer: A pre-allocated memory location that contains static or dynamic audio sample data. Dynamic data can be rendered into a buffer, but the buffer itself is not dynamic.

Channel: The number of samples per period that represent separate transducers. For example, a stereo audio file has two channels and a quad-speaker audio voice has four channels.

DLS: Downloadable Sample format. Describes a wavetable sample set and synthesis methods for rendering notes.

Envelope: A value that contains varies segments that can be used to alter the volume, pitch or filter cut-off frequency of a note.

Frame: The number of audio samples rendered per voice list traversal. Currently the voice processor will generate 32 samples for each voice it renders, therefore, the frame size is 32.

HRTF: Head-Related Transfer Function can be thought of as a set of audio filters that contain all the listening cues for each ear. The cues include environment, shoulder reflections and the ear itself.

I3DL2: Interactive Audio 3D Level 2 - part of the IASIG working on 3D audio

IASIG: Interactive Audio Special Interest Group.

IID: Inter-aural intensity difference – the sound is louder at the ear closest to the sound.

ITD: Inter-aural time difference – the sound arrives first at the ear closest to the sound.

LFO: Low Frequency Oscillator – this is used with note rendering for such affects as tremelo and vibrato.

Note: A playback voice used for wavetable music playback.

Sample: An audio sample is a data value taken at some periodic time from a continuous analog wave. Samples are typically 8-bit unsigned, 16-bit signed or 20/24-bit signed.

Sample Rate: The rate at which the periodic data is sampled. Common rates are 8KHz, 11KHz, 16KHz, 22.05KHz, 24KHz, 32KHz, 44.1KHz and 48KHz.

SPDIF: Short for Sony/Philips Digital Interface Format, an audio transfer protocol that can use either optical or electrical connectors. It's used to transfer either 16-bit stereo audio data between various components (CD players, Digital Audio Tape decks, sound cards), or to transfer Dolby Digital (formerly known as AC-3) 5.1 channel audio for movies.

Stream: A real-time stream of audio data typically being created on the fly from a decoder or other type of CODEC. Stream memory is allocated on the fly.

Voice: A general audio stream (input or output).

References

<http://www.microsoft.com/hwdev/meltdown/meltdown2k/DirectX8AudioOverview.zip>

<http://www.microsoft.com/hwdev/meltdown/meltdown2k/DirectX8Audio-EssentialInformation.zip>

© 2001 NVIDIA Corporation

NVIDIA, the NVIDIA logo, and nForce are registered trademarks or trademarks of NVIDIA Corporation. Other company and product names may be trademarks or registered trademarks of the respective companies with which they are associated.