Memory Centric Architectures for Video Production

Pinnacle Systems, Inc. • 280 N. Bernardo Avenue
Mountain View, CA 94043 • Phone: (650) 526-1600 • Fax: (650) 526-1601
www.pinnaclesys.com •

PINNACLE® SYSTEMS
The Choice For Digital Video.™

Memory Centric Architectures for Video Production

Carl Calabria, Laurin Herr & Patrick McLean

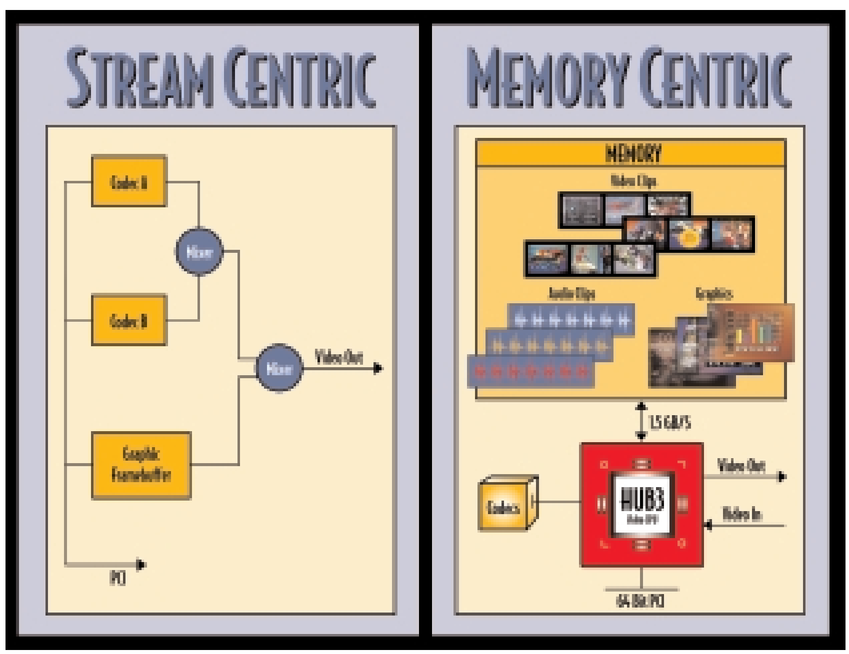
Overview

The accelerating transition to digital media production and distribution is driving an unprecedented fusion of cinema, television, video and the Web. The world's media/entertainment/publishing industries are undergoing consolidation of ownership and systemic integration on a global scale. In parallel, there is a broad decentralization of media production and distribution, a democratization of great creative and commercial energy made possible by the development of powerful personal production systems for digital media.

New content sources, new production workflows, new distribution channels, new audiences and new applications for digital video all come with their own particular requirements in terms of picture resolution, frame rate, aspect ratio, and color sampling. Multiple compression algorithms are needed, each optimized for different applications, and their number is growing. Videotape formats have proliferated in recent years, as have the variety of interface signal formats and connectors used in production.

Pinnacle's challenge, as a leading vendor of production tools for digital video professionals, is to build a unified architecture for digital media authoring at a competitive price that will be flexible enough and powerful enough to attract application software developers and users during this next period of turbulent growth and rapid transition in the digital media business.

In practice, most digital video systems in use today are digital adaptations of traditional analog designs, what we call stream-centric architectures. In a stream-centric architecture, baseband video and audio flow through a sequence of real-time processors in a pre-determined sequence of operations based on the design of the specific hardware used. Input video is processed by a hardware resource of some sort, with the resulting baseband signal sent to the next resource, and so on. Cross point switches are used to interconnect resources. Stream-centric architectures work particularly well on fixed-format signals, such as NTSC/PAL, or when the sequence of operations to be performed are clearly understood in advance of hardware design.



Stream Centric vs. Memory Centric.

In theory, the ideal digital video system should embody the principles of scalability, programmability and extensibility. As it turns out, the heart of such an ideal system should be a large shared memory with extremely fast, flexible access to programmable processing resources. Pinnacle calls this approach a memory-centric architecture. In a memory-centric architecture, any baseband signal input to the system is immediately converted to a pixel array stored in memory. All processing resources then operate on these pixel arrays and deposit their results in new pixel arrays. All operations are to and from a large shared memory. "Routing" is achieved by changing memory address pointers. The final pixel array is returned to a baseband video format for output.

Technical Advantages of Memory Centric Architectures

A memory-centric architecture potentially has several technical advantages over traditional stream-centric approaches, such as:

- Faster than real-time operation.
- Multiple instantiations of a single hardware resource
- Fully asynchronous I/O
- Shared memory efficiency
- Host access to resources
- Resolution independence
- Programmable resource configuration
- Compressed or baseband data streams are handled with equal ease

Over the years, various memory-centric architectures for video and audio processing have been implemented using general-purpose computer CPU coupled with available system RAM. While such systems have demonstrated the flexibility of the memory-centric approach, they have typically been either very costly or very slow or implemented with only limited functionality and quality. To solve these limitations, Pinnacle Systems' engineers have pioneered the use of add-on accelerator cards implementing memory-centric architectures since 1987, perfecting this approach through successive product generations.

	HUB 2	HUB 3
Total Memory Bandwidth	320 MB/s	1500 MB/s
Memory Type	DRAM/VRAM	SDRAM
Pixel Format	RGBK	YUVK and/or RGBK
Bits per Component	8	8, 10 and 16 (internal)
Sampling Structure	4:4:4:4	4:2:2:4 and /or 4:4:4:4
Black Level/White Level	0-255	16-235 and/or 0-255
4x4 Stream Count (8-bit)	8	37.5
4:2:2:0 Stream Count (8-bit)	-	75
Video Parts	1 in, 1 out, 1 i/o	5 in, 5 out
Video Port Clock	up to 27 MB/s	up to 108 MHz
Aggregate Port BW (I or O)	108 MB/s	310 MB/s
Video Port Aspect Ratio	4:3	Programmable(16:9, etc)
Video Port Scanning	525i, 625i	Programmable (i or p)
Video Port Embedded Control	None	16 bytes/vsync
Audio Channels	2 in, 2 out	8 per port (embedded)
Audio Sampling Resolution	48/44.1 KHz @16bps	48KHz @ 20 or 16bps
Local Bus Type	32-bit DVR	32 or 64-bit PCI
Local Bus Bandwidth	40 MB/s	200 MB/s
Maximum Number of Codecs	2	4
Max. Compressed Data Rate	15 MB/s	up to 200 MB/s
Desktop	Integrated	Independent
External 3D Effects Loop	Shared with Codec2	Dedicated
Sub-Pixel Resizer (Up/Down)	ID 8-bit	2D 14-bit (with crop)
Color Space Convert/Correct	None	3x3
LUTs	None	x4 interpolated 14-bit
Chroma Keyer	None	Yes
Luma Keyer	None	Yes
Compositor Type	2 layer	4 layer
Compositor Bandwidth	50 MPixels/s shared	100 MPixels/s
Resizer /LUT/CSC Bandwidth	50 MPixels/s shared	100 MPixels/s
Scatter Gather DMA Channel	None	200 MB/s (2D)
I2C Controller	None	Inernal Master/Slave

Comparison of HUB 2 and HUB 3 Video CPU.

During the late 1990's, Pinnacle's TARGA 2000 was the preeminent memory-centric platform for NLE, supporting such applications as Adobe Premiere, Apple Final Cut Pro, Avid MCXpress and NewsCutter, Discreet edit*, in-sync Speed Razor, Panasonic DVEdit and NewsBytes. Other versions of the TARGA 2000 were used in Avid Media Composer, Sony ClipServer and numerous other professional video systems.

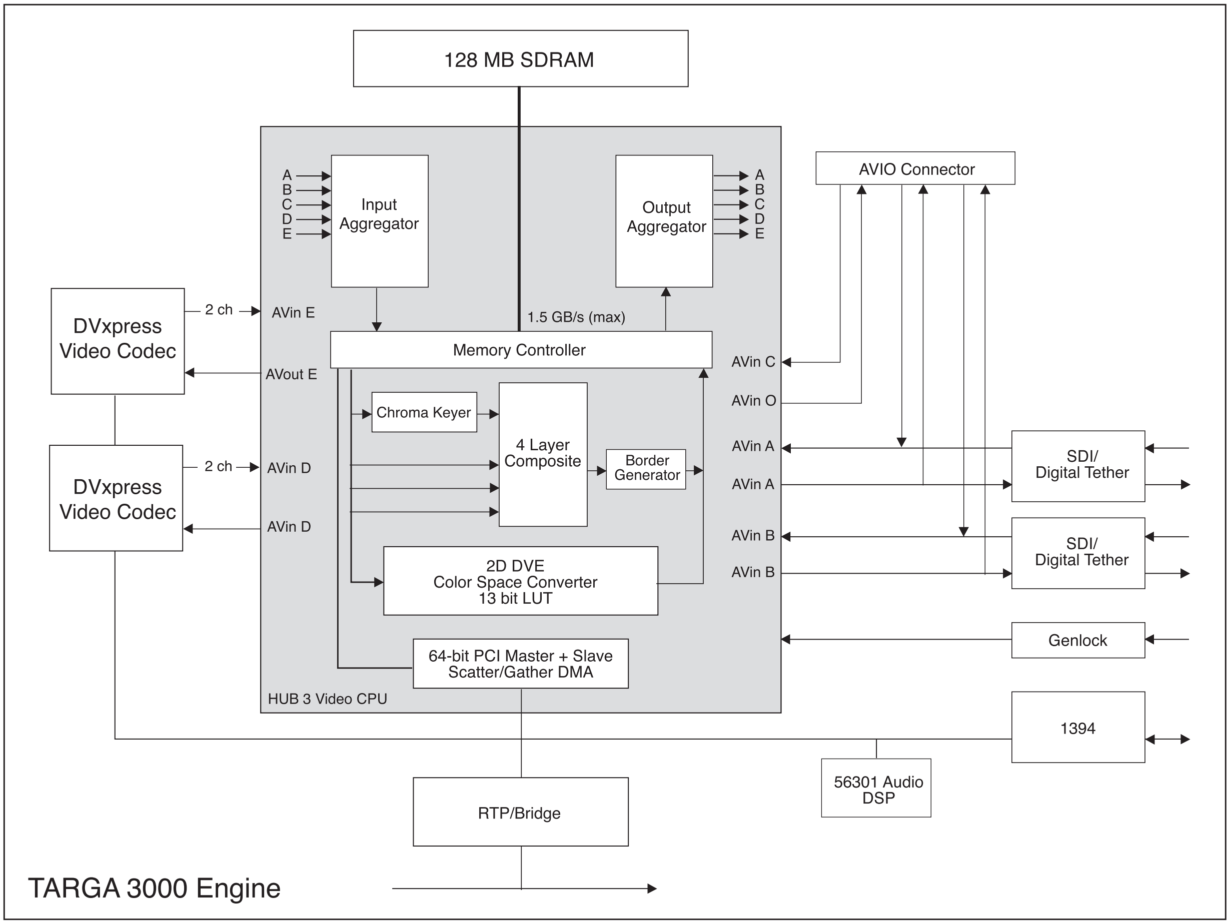
The heart of the TARGA 2000 product family was the HUB2 Video CPU. A new Video CPU called the HUB3, four times more powerful and including many new functional capabilities, has now surpassed HUB2. A comparison of the main characteristics of these two chips shows the many improvements designed into the HUB3.

HUB3 "Video CPU"

All the members of the TARGA 3000 and TARGA Ciné product families are built around the same extremely high performance HUB3 "Video CPU", a 1.5 million gate ASIC designed by Pinnacle and fabricated by IBM Microelectronics using advanced 0.25 micron 5-layer metal technology. It combines in a single chip the functions of a high-speed multi-ported programmable DMA engine with dedicated computational resources capable of executing all the essential image processing operations needed for video post-production. A single HUB3 chip delivers high performance and high quality in an integrated circuit design compact enough to enable breakthrough price/performance.

The HUB3 has five input and five output A/V Ports, each independently controllable, each capable of handling an uncompressed digital video signal with up to eight channels of embedded digital audio. Every A/V Port has its own DMA (direct memory access) to the HUB3 memory interface, itself a 128-bit wide 100 MHz connection with an aggregate bandwidth of 1.5 GBps, enough to carry 75 streams of 4:2:2 uncompressed 601 video in real-time. The HUB3 also incorporates its own 64-bit PCI interface operating at 33 MHz and capable of up to 200 MBps transfers across the PCI bus between HUB3 memory and system memory or disk storage. Video data in the form of pixel arrays is stored in HUB3 memory with 8, 10 or 16 bits per component. Sampling structure can be either 4:2:2:4 or 4:4:4:4. Maximum image size is 2048 x 2048.

The processes that can access the TARGA's central memory store include video encoders and decoders, audio encoders and decoders, codecs, a blitter/resizer, compositor, keyer, audio mixer, vertical blanking encoder and decoder and FX co-processors, such



as a 3D DVE. Within the HUB3 itself, there are two groups of image processing resources, independently programmable and accessible in parallel. Each of these two resource blocks can process 100 Mpixels/sec, roughly eight times real-time for NTSC/PAL resolution video.

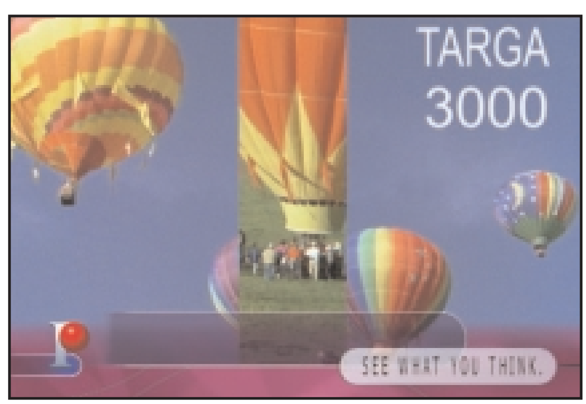
The combination of all these capabilities in a memory-centric architecture using the HUB3 makes the TARGA 3000 and TARGA Ciné powerful, high quality digital video engines with enormous flexibility. Processing resources can be utilized in any sequence. Routing between processes is very fast using memory address pointers. And there is no architectural limit to the number of "passes" through any given process. For example, it is possible to execute a 32-layer composite using standard definition video in real-time by passing through the 4-layer compositor 8 times within a single frame time. Bits per pixel, color space and sampling structure can all be varied as needed, for example using 8-bit YUV for preview and 16-bit RGBA for final rendering. The rendering power of the HUB3 is itself scaleable, capable of operating at faster than real-time speeds or gracefully degrading to slower than real-time speeds when the image size or processing complexity grows. Standard definition, high definition and low definition video can all be handled in the same architecture, with easy conversion between the different formats.

Creative Advantages of Memory Centric Architectures

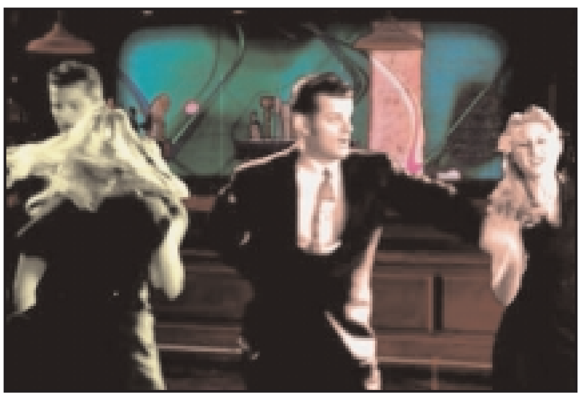
A memory-centric architecture combines the flexibility of a software based application with the speed of hardware acceleration. In previous hardware based applications, the editor has been pushed towards creating images that are driven by the capabilities and routing of the hardware. The memory-centric approach removes this barrier, and lets the editors choose their own compositions.

Below are examples of the types of compositions that are possible with the TARGA 3000 when used with Adobe Premiere and in-sync Speed Razor:

- **"Classic" A/B rolls with titles:** 2 transitioning videos with title overlays such as rolls and crawls. Transitions can be chosen from a large range of pushes, slides and other 2D effects as well as "alpha wipes" (wipes defined by any arbitrary shaped bitmap). In addition, with TARGA 3000, up to 6 real-time graphics layers can be added on top of the basic A/B video edit, each one of them can be animated with a motion path.



–Three resized moving video layers: For instance, multiple videos are being scrolled across the screen to represent different segments of a piece. Below this runs a text crawl complementing the images.



–Cascadable effects: Take a bluescreen image, key it, duplicate it multiple times, color correct each instance independently and composite together as in the example below. Only the TARGA 3000 can do this.

–Deeply layered pieces: This is where the TARGA 3000 excels with Instant Compositing and accelerated rendering of any number of layers. Compositions such as the one below with multiple keys, wipes, animated keys and 2D resizing can be worked on interactively.

New Generation of TARGA Engines

Pinnacle's new HUB3 chip is the foundation for a whole new generation of TARGA boards called the TARGA 3000 and TARGA Ciné engines. Both boards offer pristine quality for multi-channel uncompressed effects coupled with scaleable rendering that operates at faster than real-time for standard definition (NTSC/PAL), yet gracefully degrades as image resolution and/or processing complexity increases. They work in both YUV and RGB at very high mathematical precision, enabling multi-generational transparency for 10-bit video processing, all under programmable control.

The TARGA memory-centric architecture is fundamentally codec independent and handles both compressed and uncompressed data streams. There can be multiple DV and MPEG2 hardware codecs on board. The primary difference between single stream and multiple stream configurations is the number of codecs and audio/video encoders and decoders present in that configuration. In addition, data can be freely moved between TARGA memory and host system memory, enabling the convenient use of software (host-based) codecs.

Picture resolutions, frame rates and aspect ratios are programmable in this architecture. Both interlaced and progressively scanned formats are supported. The HUB3 establishes an extensible foundation for all HDTV resolutions, including 1080p/24, even while providing an excellent platform for producing low-definition video for Web distribution. A family of scaleable I/O peripherals for the new TARGA engines snap-on externally using Pinnacle's new Digital Tether port.

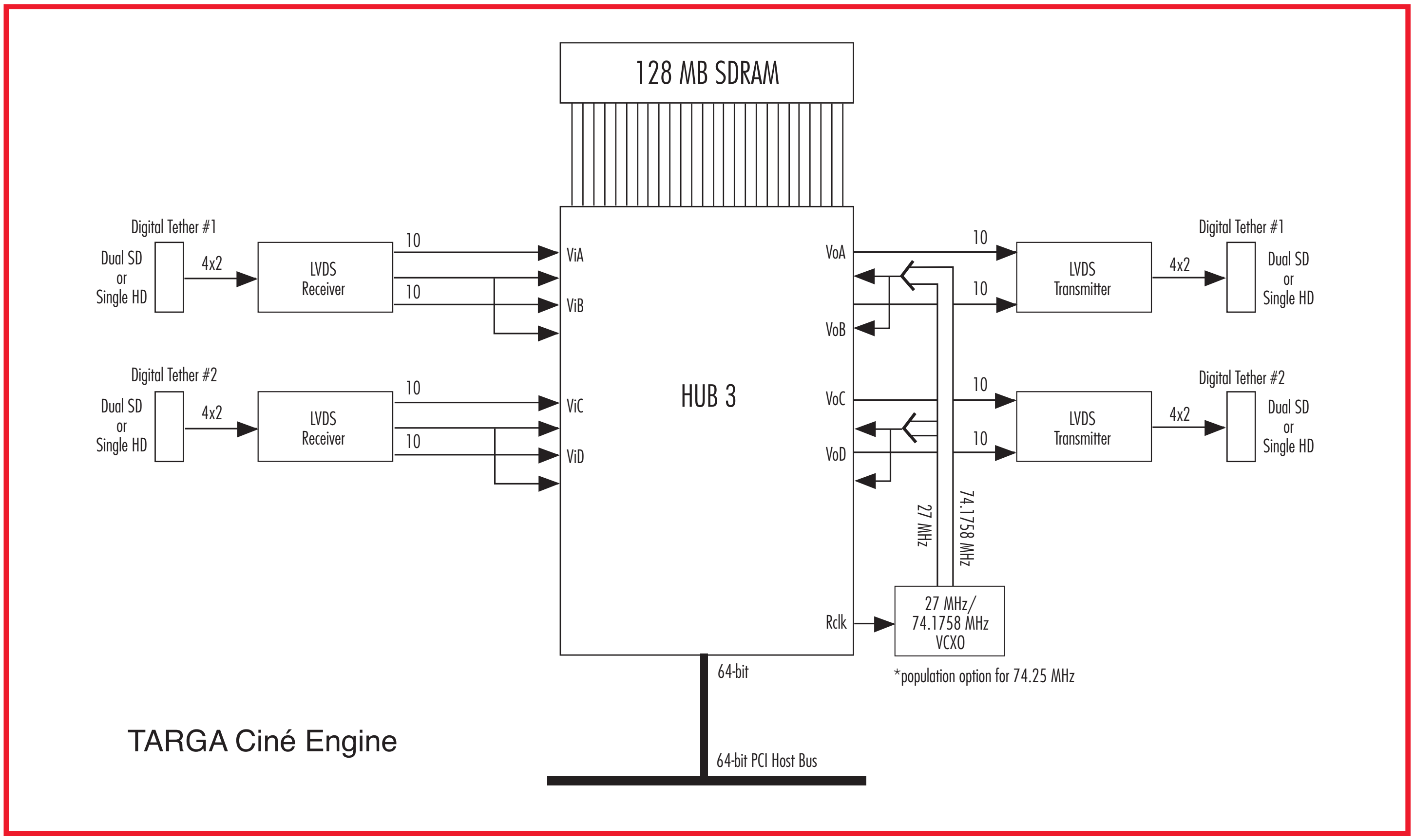
TARGA 3000 Engine

The TARGA 3000 configuration for Windows NT and Windows 2000 platforms comes with 128 MB of 100 MHz SDRAM on a single-slot full-length PCI card ready for faster 64 bit PCI busses. It includes two C-Cube DVXpert codecs capable of up to four channels of decompression or two channels of compression in parallel, supporting both DV 25 and MPEG2, with an optional software upgrade to support DV 50. An internal AVIO expansion connector is built-in for future daughter-card add-ons, such as the Pinnacle Infinite 3D DVE. IEEE 1394 devices can plug directly into the back fence of the board. Or, using Pinnacle's new Digital Tether port, a variety of analog and/or digital active breakout boxes can been attached externally.

The advantages of a memory-centric architecture can also be applied to digital audio. The TARGA 3000 integrates a powerful memory-centric audio subsystem using dedicated Motorola DSP working in conjunction with the HUB3. The DSP implement the audio input and output ports (for external audio I/O), as well as the record and playback transports (for audio to/from disk). Audio data is stored as 20 bit/48 KHz samples, with internal processing done at 24-bit precision. Record, playback, pause and scrub operations are supported, as is 3-band parametric equalization, mixing and multi-point metering.

TARGA Ciné Engine

The TARGA Ciné configuration for the MacOS G4 platform uses the same HUB3 Video CPU but is optimized for uncompressed video operations at scaleable resolutions. TARGA Ciné comes with 128 MB of 100 MHz SDRAM, but does not include either the C-Cube codec hardware or the Motorola audio DSP, opting to

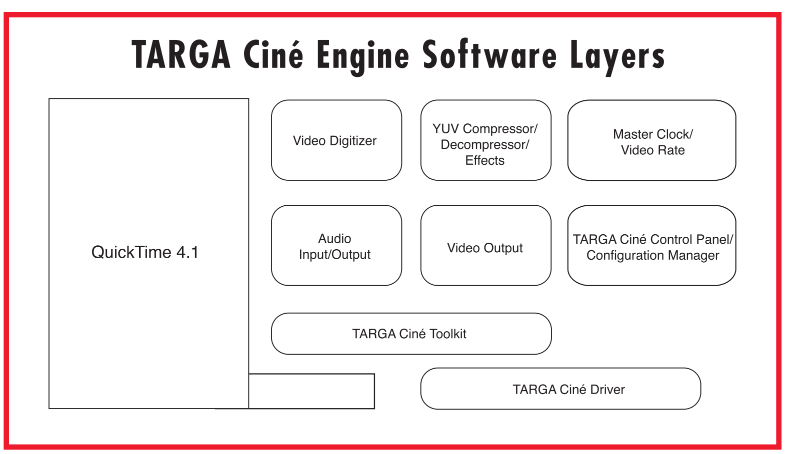


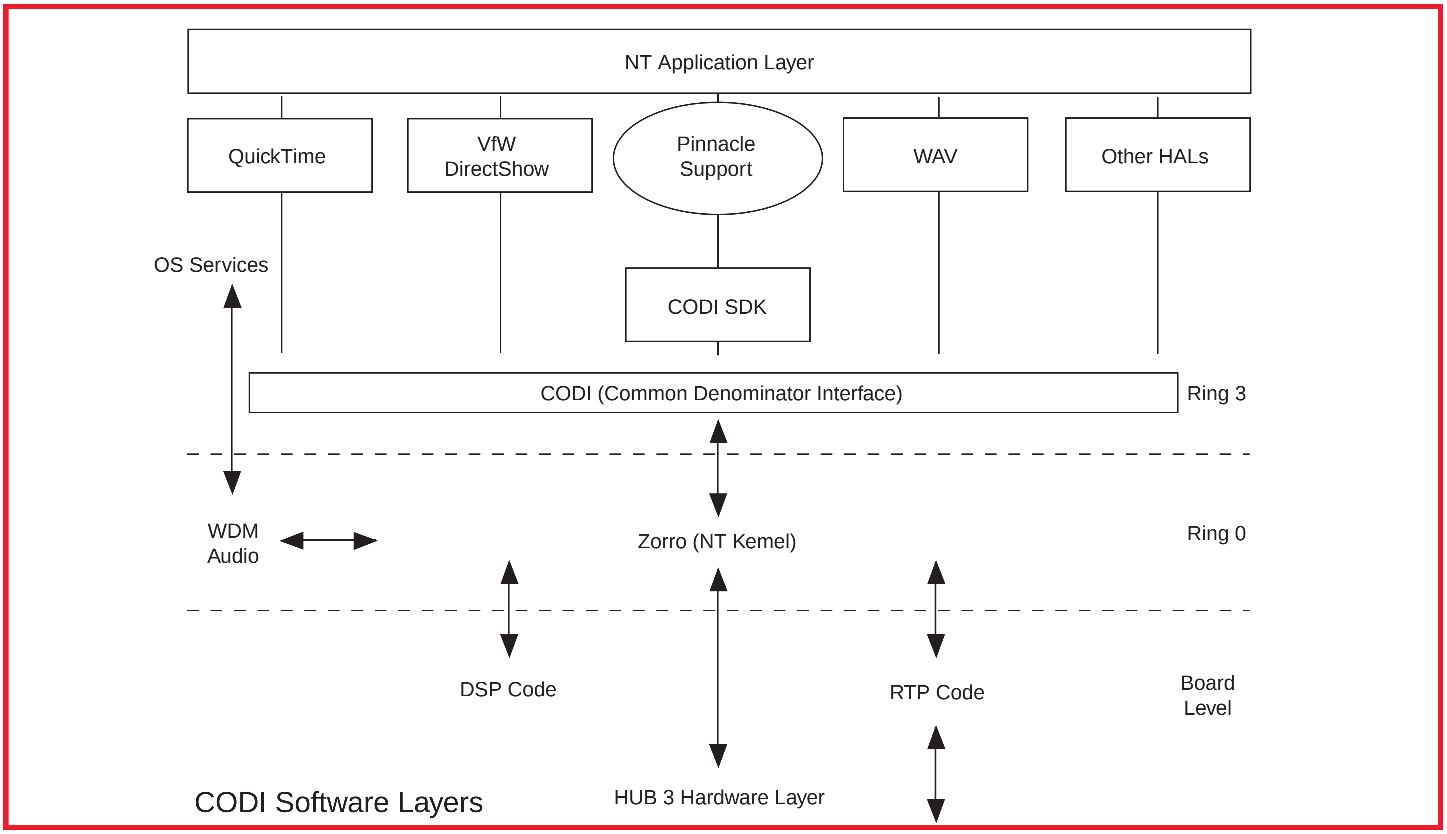
use host-based software codecs and audio processing instead. As a result, TARGA Ciné can be implemented in a half-length 64-bit PCI card. The TARGA Ciné design assumes that IEEE 1394 connectivity for DV video I/0 will use the Firewire built into all G4 systems. In a novel configuration, the TARGA Ciné engine places two independent Digital Tether ports on the back fence of the card, allowing simultaneous input/output of multiple video channels through the connection of multiple breakout boxes.

TARGA Software Layers

One of the most powerful aspects of the new TARGA family is the software programmability and extensibility inherent in a memory-centric architecture. Pinnacle can not only vary the hardware configuration but can also build different software API atop the different configurations, allowing designs to be optimized for different platforms. For example, the TARGA Ciné implements native QuickTime drivers tuned for the Apple G4 platform running the MacOS. This means that all QuickTime applications can run on the TARGA Ciné, including Apple's Final Cut Pro, Adobe Premiere, Adobe After Effects, Pinnacle Commotion Pro, and others. Real time effects are implemented using the QuickTime FX architecture. And independent software development is possible using the Apple QuickTime API.

On the other hand, for Windows NT and Windows 2000 platforms using Intel CPU, Pinnacle has developed its own low-level API called CODI (Common Denominator Interface) that will be exposed for OEM and developer use through a fully documented SDK. This full-featured API is implemented as various pieces of software running on the host CPU, the on-board real-time processor and the on-board audio DSP. CODI represents the lowest level access available to the TARGA and allows 3rd party developers to control virtually every aspect of the TARGA's processing. Applications can choose to control each of the TARGA's ports individually, managing the full complex array of processes involved in TARGA operation. Or they can configure the TARGA for streaming operation whereby arbitrarily complex processes can be preconfigured and then managed by the TARGA's onboard real-time processor. The CODI API is implemented as a collection of COM (Component Object Model) objects. The CODI object model is a hierarchy of objects that represent the various logical and physical components that make up a TARGA 3000 system. The object model allows for multiple boards in a system, as well as multiple users sharing resources on a single board. The CODI API abstracts the functionality of these





CODI Software Layers

components into easy-to-understand and easy-to use software objects.

The CODI processing model centers on the allocation of the TARGA memory to application-specific uses and the management of the various TARGA processes as they simultaneously read and write data to and from the TARGA's shared memory store. Mechanisms are available to gate these TARGA processes so they perform in a particular sequence, and to synchronize them to perform in parallel. Furthermore, these processes are programmatically re-configurable.

CODI exposes virtually all of the functionality necessary to manage TARGA memory and the processes that access it through three basic Interface types: Buffer Interfaces, Port Interfaces and Control Interfaces. Buffer Interfaces provide the means by which an application allocates and organizes the contents of the TARGA's central memory store. Port Interfaces provide the mechanism by which "work" (data processing) is performed on the TARGA, with separate Port Interfaces provided in the API for each different process that reads/writes TARGA memory. The Control Interfaces allow access to the configuration details of specific ports, but do not affect data processing.

Pinnacle is already working with key Windows software vendors and system integrators such as Adobe, Discreet, and in-sync, to port applications that will exploit the full power and flexibility of the TARGA 3000 architecture. In addition, Pinnacle will publish a CODI Software Development Kit (SDK) that rests above the CODI API, providing information and examples regarding TARGA usage. The SDK describes and demonstrates the various TARGA resources, and documents the proper management of the TARGA architecture for application developers.

Conclusion

The TARGA 3000 and TARGA Ciné represent the state of the art today in memory-centric architectures implemented using custom chips. They deliver high performance video processing in a fully programmable environment. By unifying digital content creation around memory and providing high speed access to that memory in conjunction with dedicated processing power, Pinnacle is offering scaleable and extensible capabilities with unprecedented price performance. Memory-centric architectures are the future of video and audio production in the digital age. Pinnacle will continue to advance the state of the art in production technology using this approach leveraging further refinements in processing speed, increasing densities of memory devices and faster bus bandwidths.

For further information regarding the TARGA 3000 CODI SDK, please check the OEM section of the Pinnacle Systems website at http://www.pinnaclesys.com.

For more information on the TARGA 3000, please check the Pinnacle Systems website at http://www.pinnaclesys.com/t3k

For more information on the TARGA Ciné, please check the Pinnacle Systems website at http://www.pinnaclesys.com/targacine