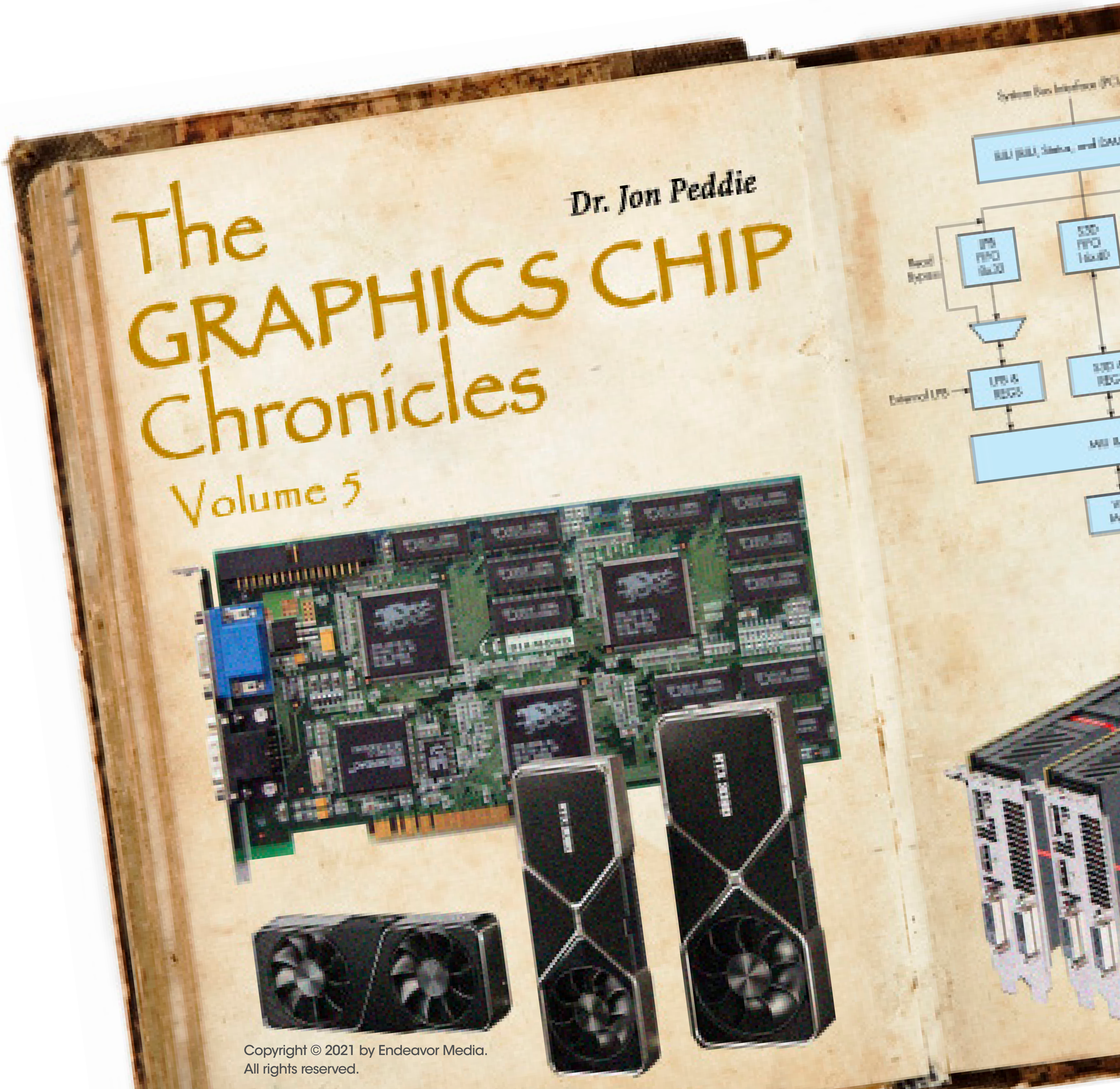
Electronic Design® LIBRARY

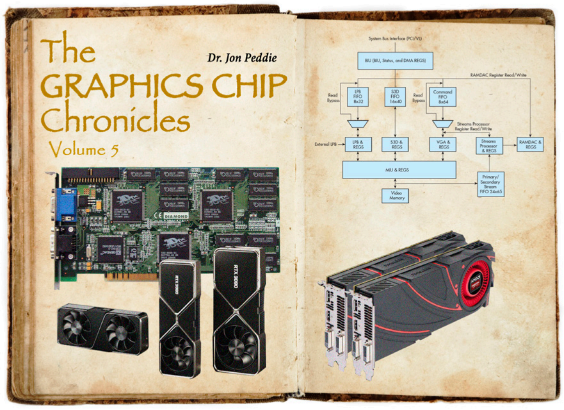
A compendium of articles from *Electronic Design*



Dr. Jon Peddie

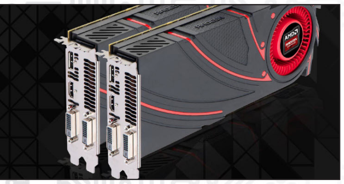
The GRAPHICS CHIP Chronicles

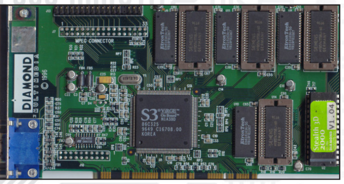
Volume 5

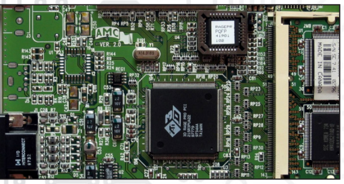


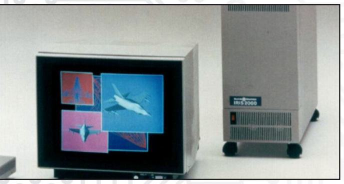
CONTENTS

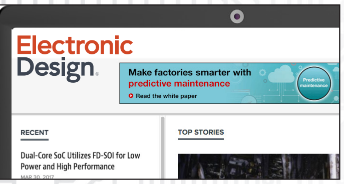

02 John Peddie biography


CHAPTER 1
03 Multi GPUs: A Story of Promise and Potential Failure


CHAPTER 2
06 S3 ViRGE: Trying For Virtual Reality Too Early


CHAPTER 3
11 ATI 3D Rage: One of the First "Free-D" Chips


CHAPTER 4
17 Geometry Engine: The Legendary Chip That Launched SGI



21 More Resources from *Electronic Design*

Bill Wong
Editor,
Senior Content
Director

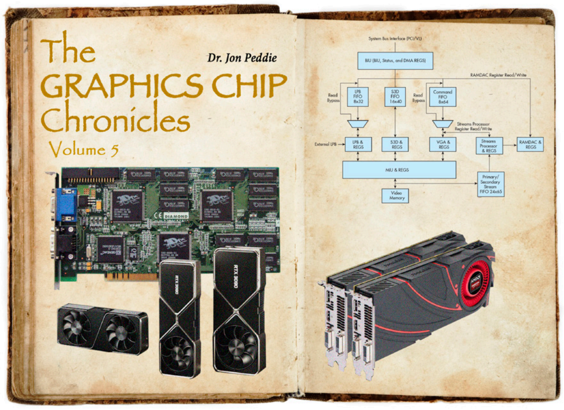
GRAPHICS CONTROLLERS have been an important part of computer systems almost since their inception and they have steadily progressed from providing limited support for low resolution displays to platforms that provide real time ray tracing support. Early display controllers provided features like bit blitting and sprites. Many of these features started out on PCs to support gaming and other applications. These technologies migrated into cell phones that turned into the smartphones of today that are essentially supercomputers-in-a-pocket.

Staying abreast of the latest and greatest technology, including computer graphics, is what we do at Electronic Design but it is often interesting and useful to look back at what was and how we got to where we are at now. Plus, old technologies and methodologies are often useful now. For example, the 8-bit game programmers moved from PCs to cell phones as the PCs moved onto higher end graphics controllers but the cell phones were more limited at the time. I worked with many of these early graphics chips even before I was the first Lab Director at PC Magazine's famous PC Labs many years ago.

One name that stands out in the graphics world is Jon Peddie who has been following and reporting on this technology area as long as I have. I have known Jon Peddie for many years. He has always been my go-to person when it comes to graphics technology.

Jon has been writing short articles about the graphics chips from a historical perspective and the articles can be found online but we have collected them here so you can easily find them. In this volume we start by examining IBM's XGA.

We hope you enjoy this series of ebooks that delve into the history of graphics.



The GRAPHICS CHIP Chronicles
Volume 5
Dr. Jon Peddie

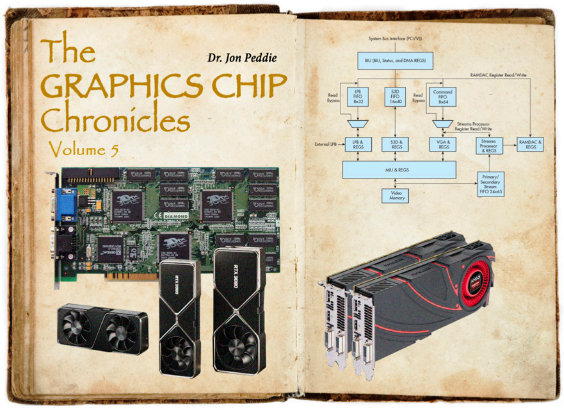
DR. JON PEDDIE is one of the pioneers of the graphics industry and formed *Jon Peddie Research* (JPR) to provide customer intimate consulting and market forecasting services where he explores the developments in computer graphics technology to advance economic inclusion and improve resource efficiency.

Recently named one of the most influential analysts, Peddie regularly advises investors in the technology sector. He is an advisor to the U.N., several companies in the computer graphics industry, an advisor to the Siggraph Executive Committee, and in 2018 he was accepted as an ACM Distinguished Speaker. Peddie is a senior and lifetime member of IEEE, and a former chair of the IEEE Super Computer Committee, and the former president of The Siggraph Pioneers. In 2015 he was given the Life Time Achievement award from the CAAD society.

Peddie lectures at numerous conferences and universities world-wide on topics pertaining to graphics technology and the emerging trends in digital media technology, as well as appearing on CNN, TechTV, and Future Talk TV, and is frequently quoted in trade and business publications.

Dr. Peddie has published hundreds of papers, has authored and contributed to no less than thirteen books in his career, his most recent, Augmented Reality, where we all will live, and is a contributor to TechWatch, for which he writes a series of weekly articles on AR, VR, AI, GPUs, and computer gaming. He is a regular contributor to IEEE, Computer Graphics World, and several other leading publications.

☞ BACK TO TABLE OF CONTENTS



CHAPTER 1:

Multi GPUs: A Story of Promise and Potential Failure

DR. JON PEDDIE

The Graphics Chip Chronicles Vol.5 No.1 - The concept of combining multiple graphics cards to scale up performance emerged at 3Dfx more than two decades ago. Since then, Nvidia and AMD have struggled to sell the idea to consumers and gaming enthusiasts.

When 3D graphics controllers were emerging in the late 1990s, 3Dfx was experimenting with ways to scale up performance and accelerate 3D gameplay. One technology it developed was called scan-line interleave (SLI), which was introduced in 1998 as part of its second-generation graphics processor, the Voodoo2.

In SLI mode, two Voodoo2 add-in-boards (AIBs) could run in parallel, with each one drawing every other line of the display. The original Voodoo Graphics also had SLI capabilities, but the feature was generally used only in the arcade and pro graphics markets.

In addition to reducing scan time, SLI also promised to increase the available frame

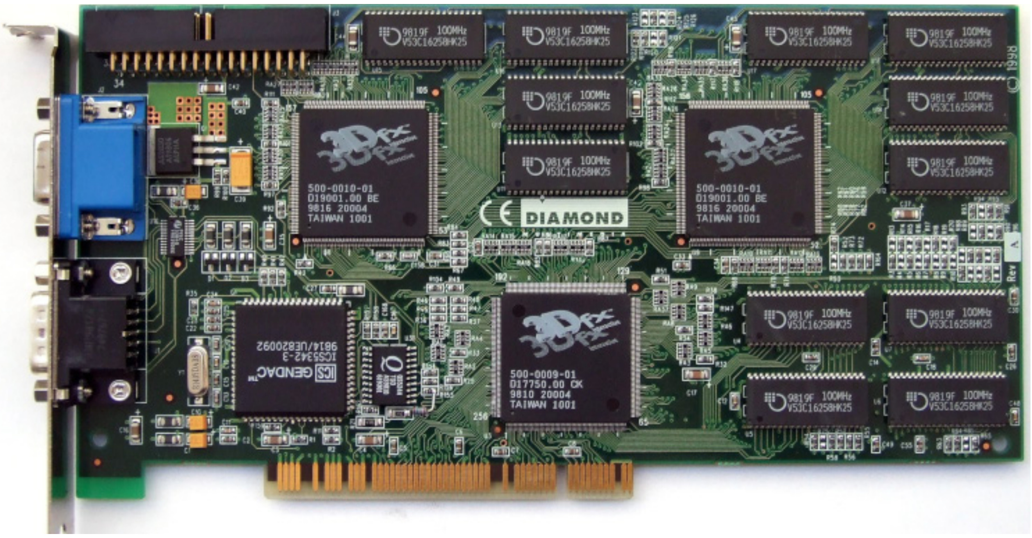


Fig 1. The Voodoo2 AIB.

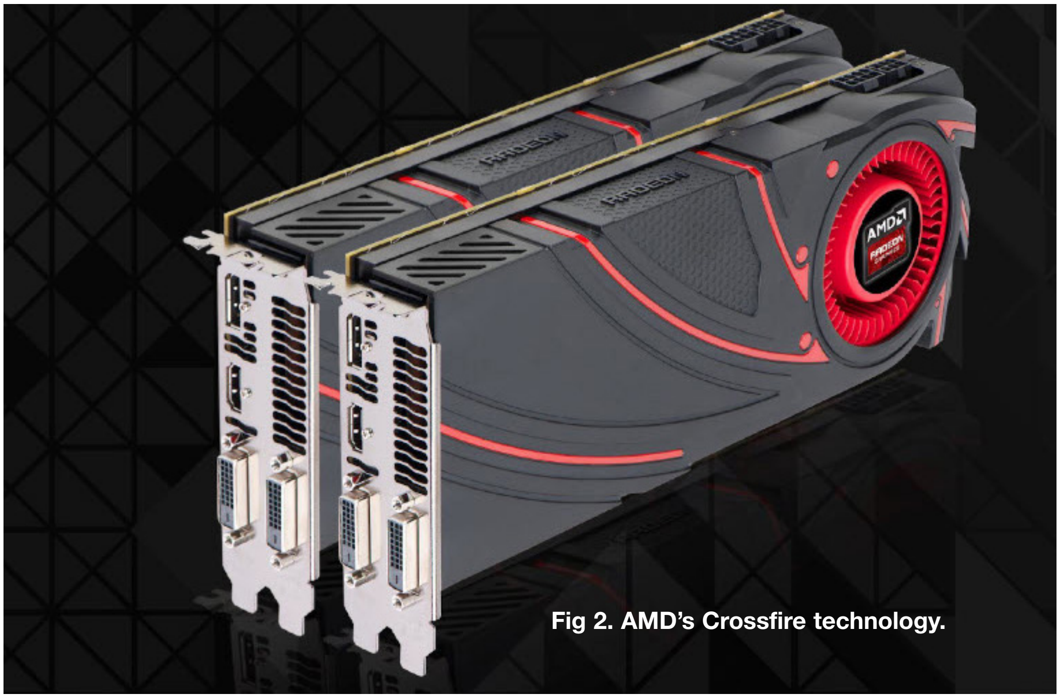


Fig 2. AMD's Crossfire technology.

buffer's memory. That would allow larger models to be loaded and it would also increase the maximum resolution of the screen. Unfortunately, the texture memory remained the same because each AIB needed to duplicate the scene data. That, combined with other overhead issues, dragged on the theoretical performance boost. As 3D models and screen resolutions continued to grow, so did the size and number of texture maps, further cutting into the promised benefits.

3Dfx tried to overcome this problem by adding another chip: the texture mapping unit (TMU). The TMU allowed a second texture to be drawn during the same graphics engine pass with no performance penalty. When it was introduced, Voodoo2 was the only 3D AIB capable of single-cycle dual-texturing. Using the Voodoo2's second TMU depended on the application software. Two very popular games of the time, Quake II and Unreal, successfully took advantage of dual-texturing. In fact, in 1998, multi-textures were almost the standard.

It took a little longer before the price-performance analysis showed up. An 8MB Voodoo2 AIB sold for $249 in 1998, about $480 today. A pair of Voodoo2 AIBs would be around $500 then. The problem was the average performance improvement was only 60 to 70% depending on the game and the central processing unit (CPU). The payoff was never there, nor could it ever be. However, in the end, the concept never faded out completely.

When Nvidia bought 3Dfx's assets in 2000, included in the IP package was SLI. Nvidia didn't reintroduce it until 2004 due to a lack of motherboards with dual AGP ports. And – Nvidia being Nvidia – they rebranded it to scan-line interface. Nvidia also expanded the concept, making it capable of combining up to four AIBs, which 3dfx had accomplished in the professional space with its Quantum3D products. The company also added several operating modes: Split-frame rendering (half per AIB), alternate frame rendering, and even SLI anti-aliasing as well as the ability to use an integrated GPU, a mode it called Hybrid SLI.

But expansion and rebranding could not solve SLI's fundamental problem: the technol-

Fig 3. Nvidia's RTX 3000 Series GPUs.

ogy never delivered anything more than 170% improvement for 200% of the cost. On top of that, AIBs were increasing in price year after year. In addition, the driver support Nvidia had to provide, amounting to a tweak for almost every game, was adding up with each new generation.

In late 2005, reacting to Nvidia's SLI rebranding, AMD, which had just acquired ATI, introduced its own take on the technology, called CrossFire. Then, in 2013, AMD ushered the concept to the next level and eliminated the over-the-top (OTT) strap. Instead, the company used an extended direct memory access (XDMA) to open a direct channel of communication between multiple GPUs in a system, connected via the PCI Express (PCIe) interface.

AMD's XDMA eliminated the external bridge by opening a direct channel between the multiple GPUs in a system. That channel operated over the same PCIe interface as AMD's AIBs. PCIe is typically used to transfer graphics data between GPUs, main memory, and CPU. When AMD introduced XDMA, the AIBs at the time were not using all the bandwidth PCIe could offer, which was considerably more than an OTT strap. The bandwidth of an external OTT bridge was only 900 Mbps, whereas PCIe Gen 3 with 16 lanes could supply up to 32 Gbps.

AMD's added bandwidth and elimination of the OTT (a perk that later on Nvidia charged extra for) gave it a competitive edge. However, AMD's AIBs at the time struggled to match the performance level of Nvidia's, which hurt it in the marketplace. Ironically, when AMD introduced the RX480 in 2016, the company pushed users to purchase a pair of AIBs that it claimed would outperform a single Nvidia AIB at a lower cost. It was a clever marketing pitch, but it didn't help AMD's sales. It also wasn't true.

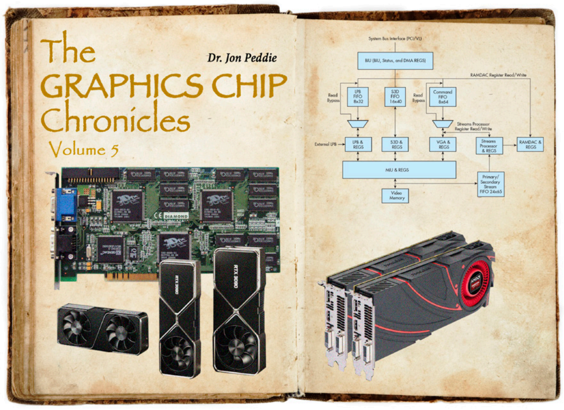
In 2017, as AMD and Nvidia rolled out Dx12 AIBs, AMD dropped support for CrossFire. The company stated, "In DirectX 12, we reference multi-GPU as applications must support mGPU, whereas AMD has to create the profiles for DX11. We've accordingly moved away from using the CrossFire tag for multi-GPU gaming."

Nvidia followed suit in 2019 and made it official in 2020. For its professional graphics AIB line, Quadro, Nvidia introduced a higher-bandwidth scheme it calls NVLink for multi-AIBs. NVLink specifies a point-to-point connection with data rates of 20, 25 and 50 Gbps.

In late 2020, the company introduced a high-end consumer graphics card, the RTX3090, and made NVLink an option for it. The 350-watt RTX 3090 was introduced at $1,499. It is unlikely many gamers will spend $3,000, plus another $90 for the NVLink technology. They may also need to add a larger power supply (PSU) to manage all the extra performance. However, content creators may want to shell out for the added performance.

to view this article online, ☞ *click here*

☞ **BACK TO TABLE OF CONTENTS**



CHAPTER 2:

S3 ViRGE: Trying For Virtual Reality Too Early

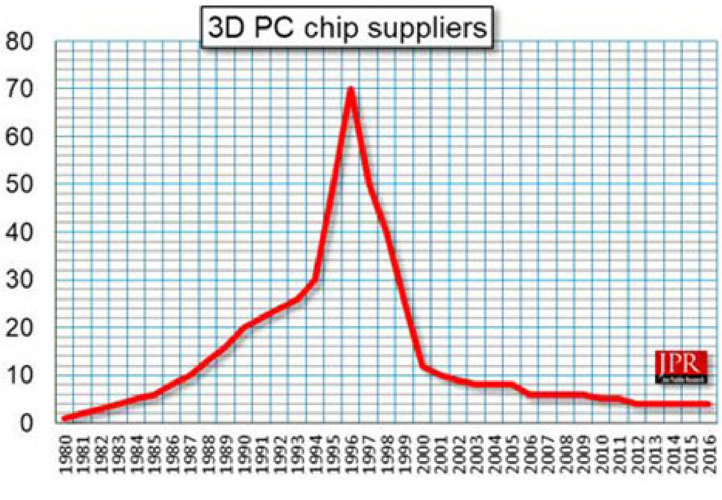
DR. JON PEDDIE

Graphics Chip Chronicles Vol.5 No.2 - Other advanced features of the ViRGE included S3's proprietary compressed texture formats, which the company said resulted in improved performance and reduced memory requirements.

In the mid-1990s, two major developments in the graphics industry were taking place. One was the explosion of 3D graphics chip companies, and the other was an attempt to bring virtual reality from the lab to consumers.

In 1990, twenty companies were making or had declared they would make a 3D graphics chip. By 1996, the number of suppliers exploded to 70. But in 2000, the number of suppliers had dropped to 12. No one had anticipated the complexity of developing such a chip. And the dot-com bubble bursting leveled many of the companies.

Around 1990, the concept of virtual reality began to become popular. Before then, the virtual reality industry mainly provided devices for medical, flight simulation, automobile industry design, and military training purposes. By the end of the century things had gotten quiet as the dream of VR failed to live up to its promises.



In 1989, S3 was founded and began development of a 2D graphics controller. In 1991, the company introduced it S3 911 graphics chip as a Windows (or GUI) accelerator. The company did very well and introduced a series of follow-up 2D controllers.

When the Sony PlayStation, and then the Nintendo 64, were

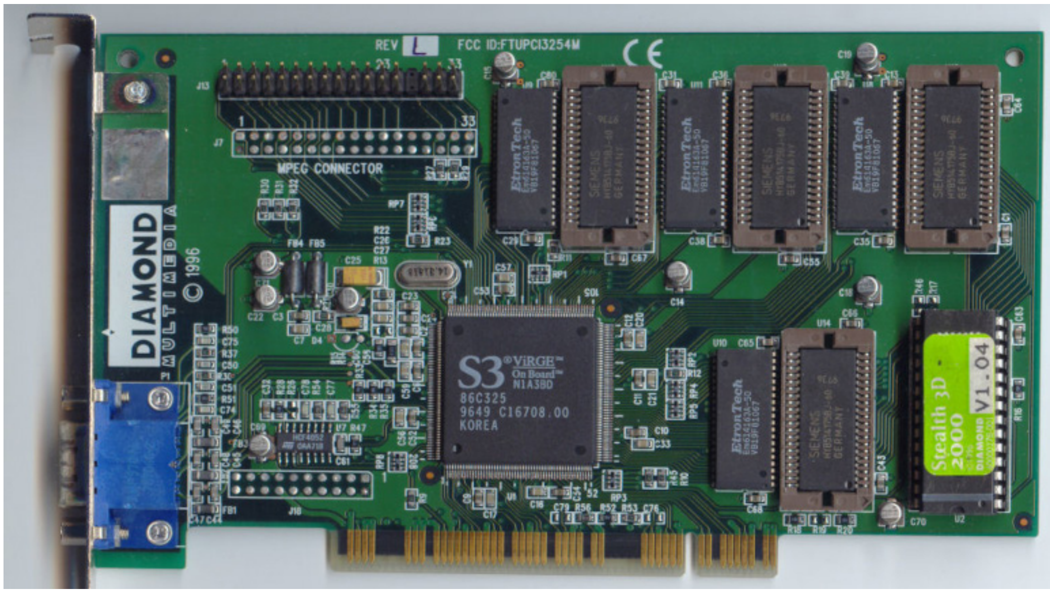


Fig 2. S3's ViRGE on Diamond Multimedia board.

introduced in the mid-1990s with 3D capabilities, consumers and OEMs began to demand 3D functionality from PC graphics cards. S3 responded to the demand and, in 1995, introduced the S3 Virtual Reality Graphics Engine (ViRGE) graphics chipset. It was one of the first 2D/3D accelerators designed for the mass market. S3 sought to capture two hot market opportunities at once—VR and 3D.

The ViRGE, also called the 86C385, was actually a relatively run-of-the-mill 3D controller, but was successful on the market because of its low price and excellent 2D capabilities, which at the time was still the major market in video games.

The S3d Engine provided 2D acceleration for Windows applications performance and a high-performance 3D rendering engine for games and other interactive 3D applications.

It incorporated key Windows accelerator functions of BitBLT, line draw, and polygon fill. 3D features included flat shading, Gouraud shading, and texture mapping support. Advanced texture mapping features spanned from perspective correction, bi-linear and tri-linear filtering, to MIP-Mapping, and Z-buffering. The S3d Engine also had direct support for utilizing video as a texture map. Those features provided the most realistic user experience for interactive 3D applications at the time.

Other advanced features of the S3d Engine included S3's proprietary compressed texture formats, which the company said resulted in improved performance and reduced memory requirements. The engine also provided support for S3's MUX buffering feature, which allowed for Z-buffering support with no additional memory cost.

S3 was also one of the first companies to offer a Streams Processor. The streams processor provided the stretching and YUV color space conversion features required for full-screen video playback with both software codecs and hardware MPEG-1 sources.

The streams processor opened the door for the simultaneous display of graphics and video of different color depths. For example, it was possible to display 24 bpp-equivalent video on top of an 8-bit graphics background. That saved memory bandwidth and storage capacity while permitting higher frame rates.

The chip also offered what S3 called its Scenic Highway. This technology created a low-cost direct connection to S3's Scenic/MX2 MPEG-1 audio and video decoder as well as video digitizers such as Philips' SAA7110/SAA7111.

THE FULL LIST OF FEATURES:

High-Performance Integrated DRAM-based 2D/3D Graphics and Video Accelerator
- High-performance 64-bit 2D/3D graphics engine
- Integrated 135 MHz RAMDAC and clock synthesizer
- S3 Streams Processor for accelerated video
- S3 Scenic Highway for direct interface to live video and MPEG-1 peripherals
- Pin compatible with S3 Trio64V+

S3d Graphics Engine Features
- High performance 2D Windows acceleration
- Flat and Gouraud shading for 3D
- High quality/performance 3D texture mapping
- Perspective correction
- Bi-linear and tri-linear texture filtering
- MIP-Mapping
- Depth cueing and fogging
- Alpha blending
- Video texture mapping
- Z-buffering

S3 Streams Processor Features
- Supports on-the-fly stretching and blending of primary RGB stream and RGB or YUV (video) secondary stream
- Each stream can have a different color depth
- High-quality hardware-assisted video playback with horizontal interpolation Support for Indeo, Cinepak, and software and hardware-accelerated MPEG-1 video

S3 Scenic Highway Interface
- Philips SAA7110/SAA7111 video digitizers
- S3 Scenic/MX2 MPEG-1 audio/video decoder

High Screen Resolution (Non-interlaced) Support
- 1280x1024x256 colors at 75 Hz refresh
- 1024x768x64K colors at 85 Hz refresh
- 800x600x16.7M colors at 85 Hz refresh

High-Performance Memory Support
- 64-bit DRAM memory interface
- 2 MB and 4 MB DRAM video memory
- Single-cycle EDO operation

Non-x86 CPU Support
- Big endian/little endian byte ordering

- Relocatable addressing

Industry-Standard Local Bus Support
- Glueless PCIe Gen 2.1 bus interface
- Glueless VESA VL-Bus interface

PCI Bus Mastering for Display List Processing and Video Capture Support

Multimedia Support Hooks
- S3 Scenic Highway
- VESA advanced feature connector
- 8- and 16-bit bi-directional feature connector

Full Software Support
- Drivers for major operating systems and APIs, including Windows 95, Windows 3.11, Windows NT, OS/2 2.1 and 3.0, ADI 4.2, Direct 3D, BRender, RenderWare, and OpenGL

Green PC/Monitor Plug and Play Support
- Full hardware and BIOS support for VESA Display Power Management Signaling (DPMS) monitor power savings modes
  - DDC monitor communications

Extensive Static/Dynamic Power Management

Industry-Standard 208-pin PQFP package

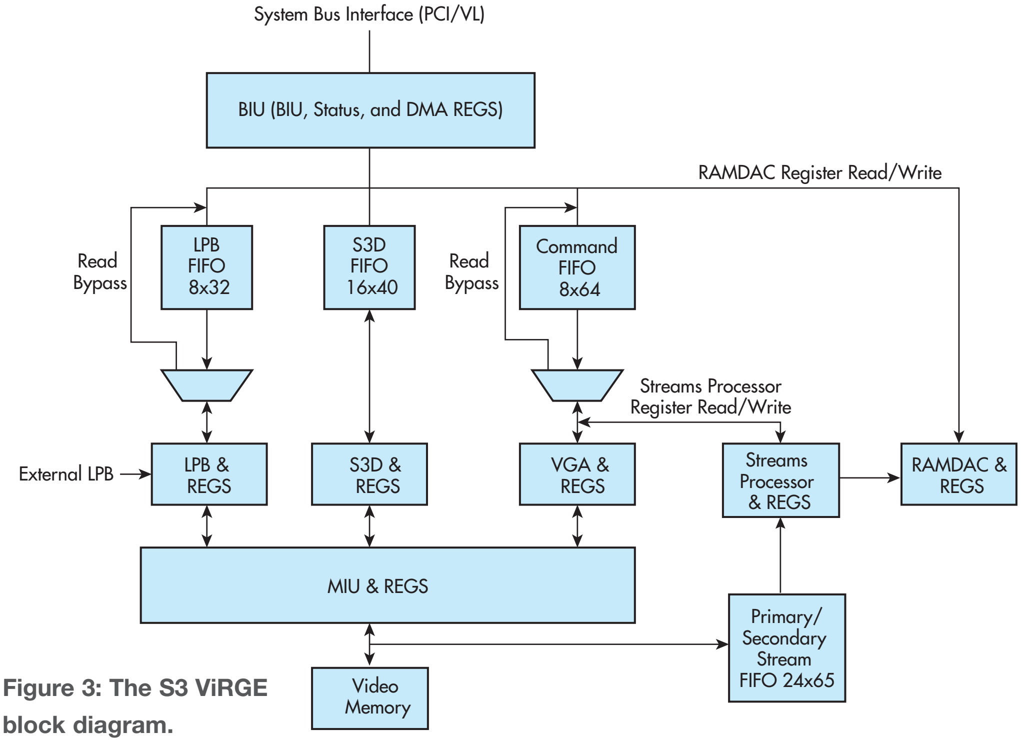


Figure 3: The S3 ViRGE block diagram.

The S3 ViRGE provided linear addressing of up to 4 MB of display memory. Linear addressing of more than 64 KB required that the CPU be operated in protected mode. Linear addressing is useful when software requires direct access to display memory. ViRGE offered two linear addressing schemes. The first method could be used when memory-mapped I/O (MMIO) is disabled or with the existing MMIO method. The second was used in conjunction with the new MMIO method. The newer method for ViRGE offered a 64 MB addressing window.

S3 went public in 1993, but the industry was shifting, and S3 struggled to maintain growth by building additional chips, such as an audio processing chip. By 1999, the company found itself in financial trouble, in part due to the collapse of the dot-com bubble and the emergence of Nvidia and ATI. The number of 3D chip companies was dropping rapidly.

Looking for other markets and financial support, S3 formed a partnership with Taiwan-based VIA Technologies. S3 would incorporate its 3D chip technology in an integrated graphics chip (IGC) made by VIA. The IGC was produced in a 180-nanometer process and would work with AMD and Intel x86 processors, providing North-bridge functions. (Ironically, in 1989, the original business plan of S3 was to build core logic parts. That was partly due to the founders' experience from their C&T days). A 10-year IP swap in 1998 gave S3 access to Intel's buses and processor I/O.

At the time, S3 held a 16% ownership stake in UMC, a semiconductor foundry. UMC was manufacturing the IGC, giving S3 and VIA a pricing advantage.

Still trying to find a path forward, S3 bought one of its OEM add-in board customers, Diamond Multimedia, in late 1999. That just about ended all the company's other AIB partnerships, damaging its cash flow and leading to other losses with PC OEMs.

A month later and running low on cash, S3 announced the formation of S3-VIA, a joint partnership. At the same time, the company was developing a music player it would show at the Comdex conference.

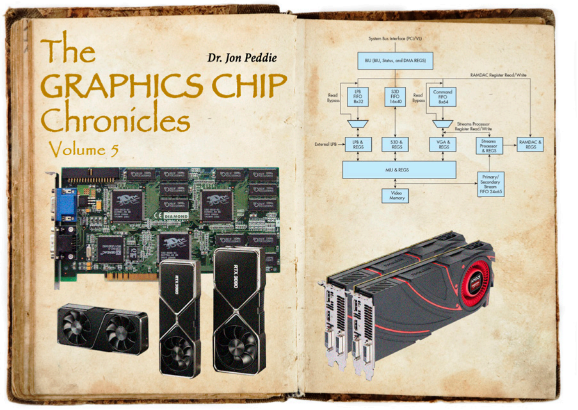
Trying to buy more business in early 2000, S3 considered purchasing what was left of Number Nine in the hopes that it would gain its supply deal with IBM. At the last second, S3 pulled out of the talks, and Number Nine filed for bankruptcy. Rumors circulated that cash-strapped S3 would be acquired by VIA. Months later, VIA acquired S3 for $320 million in a convoluted deal. In 2003, the company filed for bankruptcy.

Epilogue

In July 2011, HTC Corporation announced that it would buy the VIA Technologies stake in S3 Graphics, becoming the majority shareholder of S3 Graphics. That signified the last chapter in the company's history. Even though the S3 ViRGE never used in any VR devices, HTC has transformed into one of the leading VR headset vendors.

to view this article online, ☞ *click here*

☞ **BACK TO TABLE OF CONTENTS**



CHAPTER 3:

ATI 3D Rage: One of the First “Free-D” Chips

DR. JON PEDDIE

Graphics Chip Chronicles Vol.5 No.3 - ATI had great success with the 3D Rage, so much so that IBM agreed to implement it on the motherboard of IBM’s Aptiva multimedia home PCs.

ATI, founded outside of Toronto in 1985 as Array Technology, was a pioneer in the graphics-chip and add-in board (AIB) market. It was acquired by AMD in 2006 and formed what is now the Radeon Technology Group.

In late 1995, ATI announced its first combination 2D, 3D, and MPEG-1 accelerator chip with the name 3D Rage. The 3D Xpression AIB was based on its 3D Rage graphics chip and had elemental 3D acceleration, a year behind the pioneering Matrox Millennium PC 3D chip and around the same time as the S3 ViRGE.

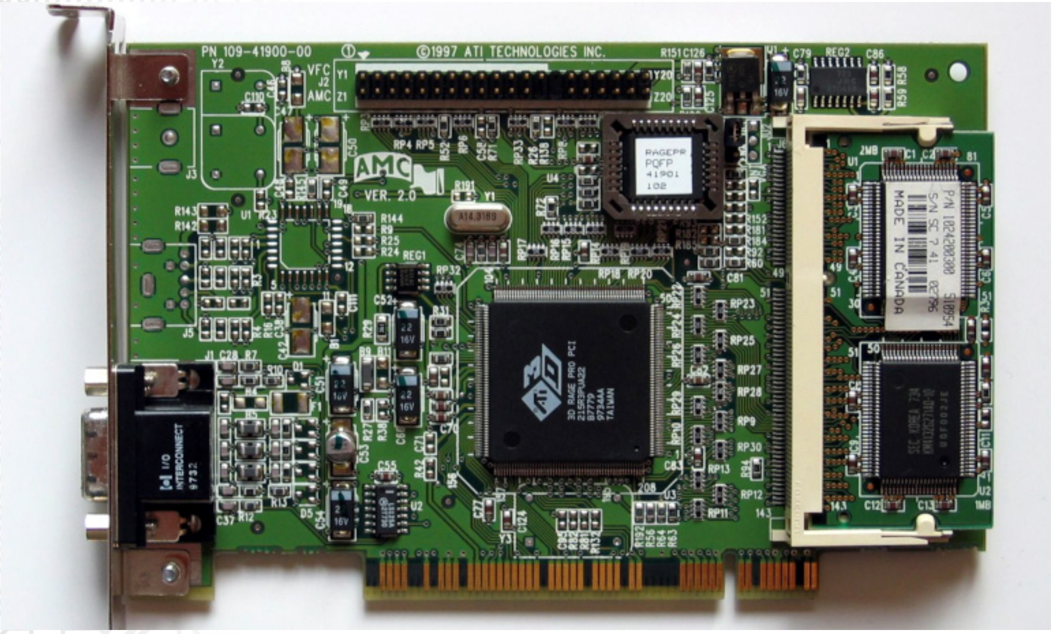
Back then most graphics chip companies used coded numbers to designate a chip because it was just a component and branding wasn’t considered important. (Matrox was the exception).

ATI’s chips were distributed on ATI boards, and it was the 2nd largest graphics company in Canada behind Matrox.

Before 3D Rage, ATI had a chip called the Mach64. It was a 2D GUI, or Windows accelerator, and became popular on the Graphics Pro Turbo board. The chip numbers evolved from there: 264CT – 2nd generation Mach64 2D accelerator, 264VT – was the first ATI chip with video acceleration (V), and the 264GT – the first ATI chip with 3D acceleration (G), which was built with the team from Kubota graphics in Boston that ATI acquired.

As the COMDEX 1995 launch of the 3D Rage approached, ATI learned that S3 was planning to release its ViRGE 3D accelerator at the same time. Phil Eisler, the new chip’s product manager, was not thrilled about launching a product called the ATI 264GT against it. So, he started searching for a name with some verve behind it to compete with the S3 ViRGE.

Back then ATI did have a small run of a 2D board product called the ATI Arcade Rage. Eisler appropriated the “Rage” part and changed the name of the 264GT and announced it at COMDEX 95 as the ATI 3D Rage.



The 3D Rage was a very versatile and scalable controller, and ATI made at least seven versions of it. It was one of the first graphics controllers to integrate the RAMDAC and clock synthesizer. The CT version had an integrated RAMDAC, and all versions had a VGA core.

The controller had support for all the popular busses of the time: ISA, VLB, PCI, and limited VESA VBE support. It had one-pixel shader, no vertex shaders, one texture mapping unit (TMU), and one raster-operations pipeline (ROP). The controller's clock ran at 44 MHz.

The memory clock could run at 57 MHz and be overclocked up to 30% before the memory became unreliable. It had a 32/64-bit memory bus and could provide up to 456 MB/s bandwidth. It could support up to 8 MB of memory (16 MB for 3D Rage Pro).

The RAMDAC could generate up to 16.7 million colors (at 1280 × 1024 resolution) and 65,000 colors (at 1600 × 1200 resolution).

The chip was manufactured at SGS in a 500-nm process, packed 5 million transistors, and was supplied in a 90-mm package.

The 3D Rage represented a departure from the mach32/64 in that it was no longer register-compatible with previous ATI graphics accelerators or the 8514/A. (VGA register compatibility was retained, however.) This departure was necessary to resolve some design limitations that were a legacy of older generation chips. Fortunately, almost all the functionality that was in the mach32/64 was preserved in the 3D Rage design, and some useful additions and enhancements were incorporated.

The many variations of the 3D Rage family, the GX (first four columns), and the CT family are shown in the following table.

Feature	GX-C/D	GX-E*	GX-F	CX	CT	VT	GT (Rage 1, Rage II, Rage II+, IIC, PRO) – LB/GM (Rage LTPRO, Rage XL) – LM (Rage Mobility M/P/M1)
Maximum memory	8 MB	8 MB	8 MB	4 MB	4 MB	4 MB	8 MB, 3D Charger (2 MB EDO DRAM), 3D xpression+ (2 or 4 MB SDRAM)
Minimum memory	512 KB	1 MB	1 MB	512 KB	1 MB	1 MB	1 MB
Supported bus types	ISA, VLB, PCI	PCI	ISA, VLB, PCI	ISA, VLB, PCI	PCI	PCI	PCI, AGP
*Revision E was a short-lived version that was only used in Apple's Power Macintosh-based boards							

From a very rough architectural perspective, the 3D Rage family more closely resembles the mach32 than it does the mach64CT family. However, from a functionality and register level standpoint, the 3D Rage GX is almost identical to the 3D Rage CT. ATI made several versions of the AIB with TV video outputs and some AIBs had TV tuners on them.

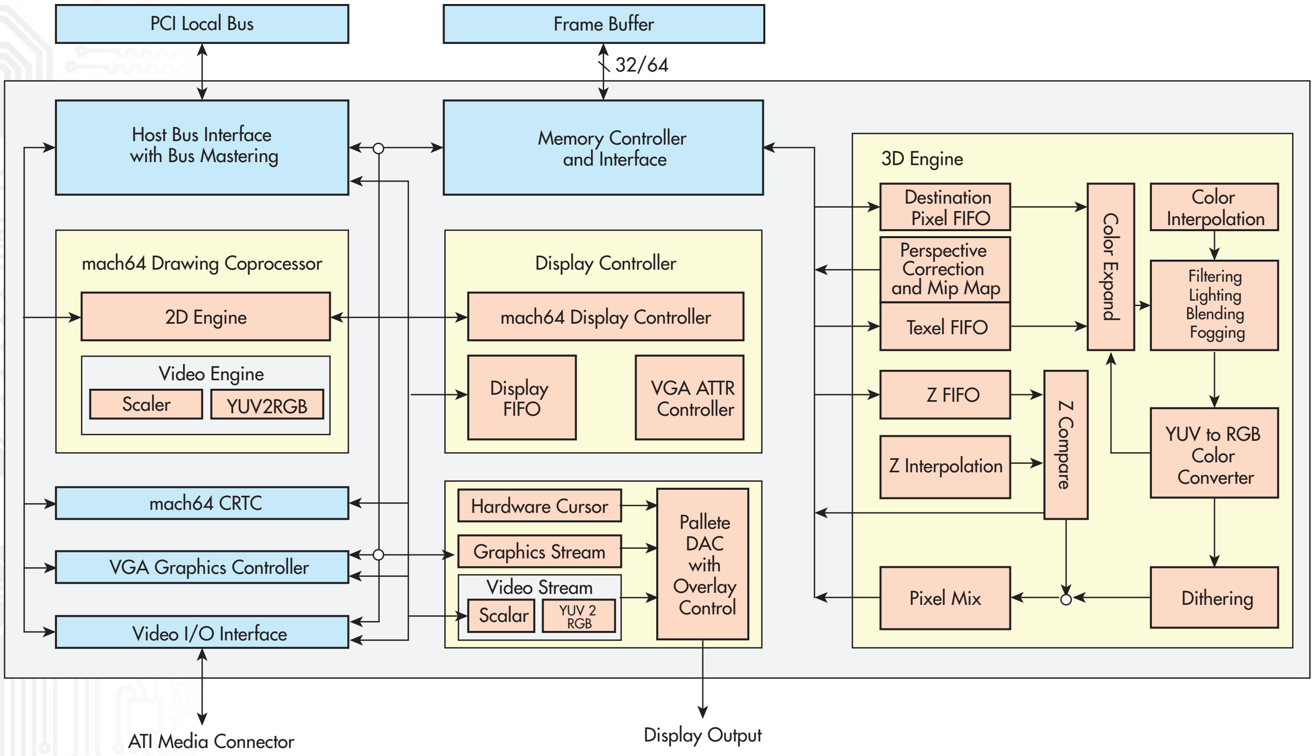
Free-D

The price difference between 2D AIBs and 3D AIBs was so slight that the term “Free-D” became popular when describing them.

In late 1996, ATI started shipping its second-generation 3D Xpression PC 2 TV. The 3D Xpression board featured the company’s latest 3D Rage II chip as well as a new home-grown NTSC/PAL encoder chip called ImpacTV. Like its predecessor, the product was targeted at consumer multimedia applications. With the new TV output support, the board was suited for deployment in the family room where it could drive a big screen TV for playing games or recording digital video or animation to video tape. It was also suited for business settings where it could drive large-screen monitors for displaying presentations.

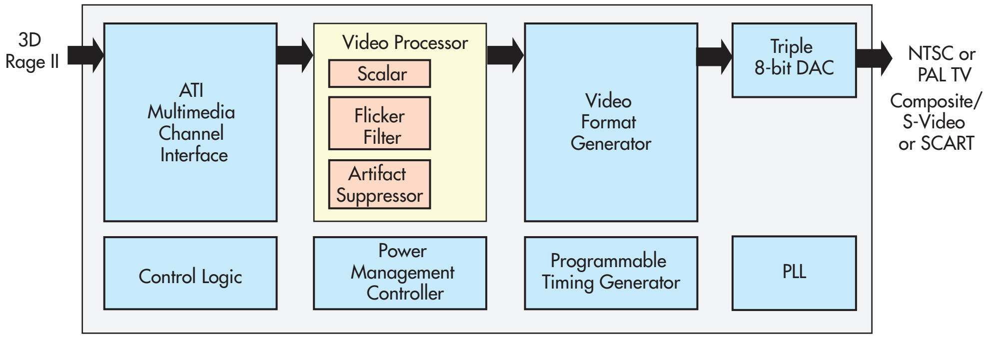
The company said at the time that the new device supplied around 20% better 2D performance, and twice the 3D performance of the original 3D Rage. That was accomplished by increasing the memory clock from 63 MHz to 73 MHz, boosting the size of the on-chip texture cache, and making pipeline improvements to increase concurrency. New features of the 3D Rage II were a 16-bit Z-buffer, and support for 4- and 8-bit palletized textures.

ATI supplied the chip with an integrated 170 MHz RAMDAC standard; however, the company was also able to screen for qualification at 200 MHz which they offered to OEM customers as an option. The 3D Rage II was fabricated by SGS in a 500-nm 5V CMOS process, and it was packaged in a 208-pin PQFP. The following diagram shows the internal organization of the chip:



The ATI board also incorporated its ImpacTV TV encoder with a direct interface to the 3D Rage II chip. The 28-pin PLCC device communicated with the Rage II through the ATI Multimedia Channel (AMC) interface. The chip included a flicker filter and special circuitry to eliminate dot crawl. Both composite and S-Video output was provided, as well as SCART output for European component video systems.

The chip supported all NTSC and PAL formats and featured programmable timing to enable the correct signals to be generated from a variety of computer display modes, including legacy VGA modes and low-resolution DirectX game modes.



The 3D Xpression PC 2 TV had an estimated price of $219 (or around $375 today) in a 2 MB configuration. The ImpacTV chip was available for about $10 in OEM quantities. The 3D Rage II remained at about $30 (or slightly more than $50 today) in OEM volumes. ATI used a variety of silicon suppliers back then: SGS, NEC, UMC and eventually TSMC.

ATI had great success with the 3D Rage, so much so that IBM chose the 3D Rage for implementation on the motherboard of IBM's Aptiva multimedia home PCs. It was the first announced 3D chip motherboard design win for the home entertainment market. Other companies like Sony and NEC followed suit. At the time, the company was averaging a million Rage chip shipments a quarter.

The Saga of ATI

ATI developed plenty of innovative and exciting technology over the years and rose to become the top supplier. However, the market for graphics solutions rapidly fragmented into the low-end and high-end gaming sector as well as the console gaming segment, the professional graphics workstation segment, and the commercial market. Several companies offered solutions for one of those segments, but by themselves the returns were not enough to sustain R&D and manufacturing costs. ATI recognized that the situation called for an economy of scale approach, a strategy similar to what Nvidia has done so successfully.

As a result, ATI acquired several companies over the years and helped write the history of graphics semiconductors.

Its first big acquisition, which caught the industry off-guard, came in 1994 when ATI picked up the design team from high-flying Kubota graphics, a leading-edge workstation supplier that had a long history of its own in acquisitions. And yes, it was the same Kubota that still makes heavy construction equipment. They ventured into computer graphics as part of a diversification program.

Then, in 1997, ATI picked up the assets of a once-leading PC graphics chip and board supplier, Tseng Labs. Tseng, like many other graphics chip suppliers at the time, struggled to make the transition to 3D. In the end, though, it wasn't totally Tseng's fault, though. It wasn't so much that Tseng, or other companies couldn't hire 3D graphics engineers who knew what to do and how to do it. They did. It was the unusually high investment in dollars and time that prevented the transition. ATI, Matrox, and S3 successfully shifted to 3D, and startups like 3Dfx and Nvidia rounded out the 3D graphics chip arena.

Chromatic Research, based in Mountain View, California, was founded in 1993 and announced its first product the same year—the Mpact media engine. Designed as a software-upgradeable multimedia processor, the device combined video, 2D graphics acceleration, 3D graphics acceleration, audio, FAX/modem, telephony, and videophone.

In October 1998, ATI acquired Chromatics Research. ATI had long wanted a Silicon Valley design center, and Chromatic had been looking for a big brother. Chromatics was one of the first highly integrated single-chip SoC suppliers.

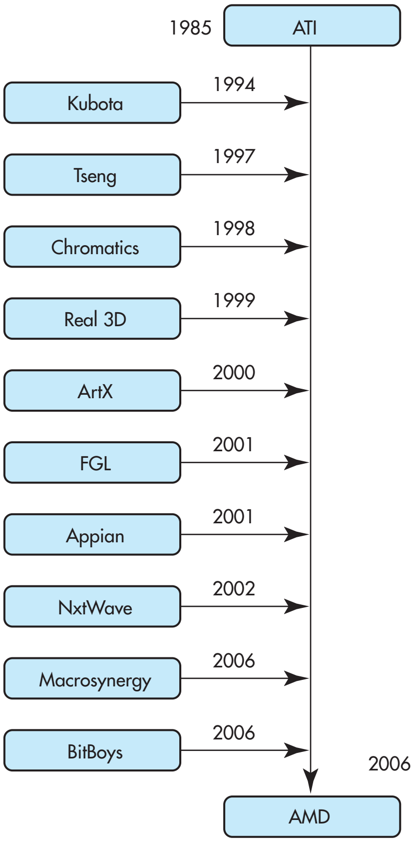
In 1999, ATI again surprised the industry by acquiring part of Lockheed Martin's Real3D team. (Intel acquired the other part). Real3D made ultra-high-end graphics for simulators and tried to parlay that expertise into the consumer market, but the parent company's patience and checkbook waned over time.

The biggest move for ATI was when it acquired ArtX in February 2000, the ex-SGI developers of the Nintendo 64 and subsequent suppliers of Nintendo graphics chips. That propelled ATI into the console business. They had all almost the major market segments in graphics covered by this point.

On top of that, in March 2001, ATI acquired Fire GL (FGL), formerly a division of S3 (a different company than S3 Graphics) that became Sonicblue. FGL had been competing in the workstation market against 3Dlabs, Intense 3D (which was later acquired by 3Dlabs), HP and E&S. How did S3 get into the high-end workstation business? By acquiring a string of graphics companies.

By 1996, multi-monitors had become both practical and affordable, and they were considered essential by everyone from CAD engineers to Wall Street traders. But they were complicated to set up. A former graphics board maker, Appian Graphics, had distinguished itself by developing a robust and easy-to-use software driver that made multi-monitors less complicated. That was a feature ATI thought it should have. And so, in August 2001, ATI acquired Appian's HydraVision desktop management software. ATI, and subsequently AMD, went on to be the leader in the multi-display market.

ATI's next target was the integrated graphics market. The company had developed an integrated graphics controller (IGP), but it was impressed by the graphics chip designs coming out of Taiwan. Instead of taking the time to duplicate that work – a difficult investment in a fast-moving market – ATI



chose to acquire it.

Television was another area where ATI wanted to expand. The company had been developing TV technology for some time, but to speed up the company's capabilities and presence in what was seen at the time as a new multimedia system, it once again opened its checkbook. In February 2002, ATI acquired NxtWave which produced TV demodulators.

Another headline came across the wire in March 2006, reading, "ATI acquires Macrosynergy." Macrosynergy was a division of the chipset firm XGI Technology, and ATI wanted it to expand its presence in China. As part of the deal, ATI absorbed 100 employees from the Shanghai-based company as well as an undisclosed number of engineers based in XGI's Silicon Valley office.

However, ATI did not buy XGI outright, as had been rumored. In 2003, XGI was formed from the graphics division of SIS, the inventor of the IGC. ATI saw Macrosynergy's international presence as a gateway into the burgeoning Chinese market. Trident was another one of the successful 2D graphics chip companies that couldn't make the transition to 3D. XGI acquired Trident and UMC (a semiconductor fab in Taiwan), and then re-acquired SIS, a company it helped get off the ground.

Around 2002, mobile devices were starting to take off and several startups emerged with graphics co-processors for phones, PDAs, and other battery-powered personal devices.

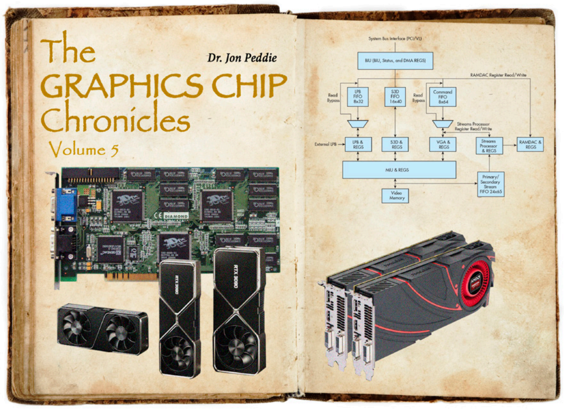
Bitboys Oy was a graphics hardware development company based in Finland and founded in 1991. The company started selling a revolutionary high-end graphics chip that became the TriTech's Pyramid3D. Unfortunately, the chip was a challenge to manufacture, leading to Pyramid3D disappearance soon after it launched. Bitboys, however, gained ground in the mobile phone market by adopting some of TriTech's concepts into a portable device. The company ended up selling the technology to NEC, one of the major early players in the emerging phone market.

Portable devices were another application ATI wanted to conquer and, in May 2006, the company acquired Bitboys.

Shortly after, the story ended. In August 2006, AMD acquired ATI, picking up a company that had developed a graphics solution for every platform imaginable, ranging from handhelds and TVs to workstations and consoles – and everything in between.

to view this article online, ☞ *click here*

☞ **BACK TO TABLE OF CONTENTS**



CHAPTER 4:

Geometry Engine: The Legendary Chip That Launched SGI

DR. JON PEDDIE

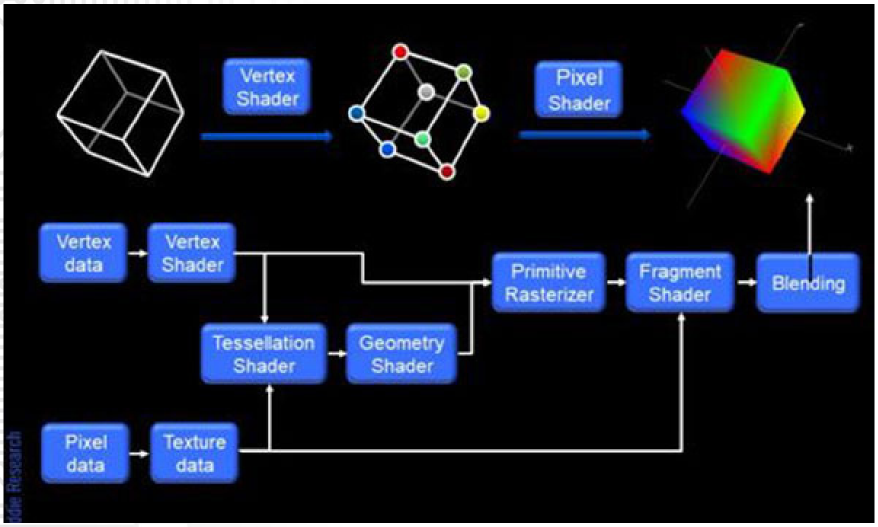
> Graphics Chip Chronicles Vol.5 No. 4 - The Geometry Engine was an engineering marvel, a special-purpose processor able to carry out many of the fundamental computations used in graphics.

Although it could not actually manipulate any pixels, the Geometry Engine introduced in 1981 was a special-purpose processor capable of carrying out many of the most critical mathematical functions used in graphics.

The Geometry Engine at the time was an engineering marvel. The custom-designed chip was based on a four-component vector, floating-point processor for three of the most fundamental operations in computer graphics: matrix transformations, clipping, and mapping to output device coordinates. Developed by Jim Clark and Marc Hannah at Stanford University in 1981, it was the first dedicated vertex processor for what has become a commodity component in modern GPUs. Clark founded Silicon Graphics Inc. (SGI) later the same year, and it became one of the most famous and influential graphics companies ever. SGI was acquired from bankruptcy by Rackable Solutions in 2009, and in 2016, Hewlett Packard Enterprise acquired Rackable/SGI. The deal ended the era of SGI – but not the legend.

While at Sanford University in early 1980, Clark came out with a paper on the nMOS semiconductor capability and its use in very-large scale integrated (VLSI) designs. He said, "The system is designed to perform three of the very common geometric functions of computer graphics. A single chip type is used in 12 slightly different configurations to accomplish 4 × 4 matrix multiplications; line, character, and polygon clipping; and scaling of the clipped results to display device coordinates. This chip is referred to as the Geometry Engine."

The Geometry Engine was a four-component vector function unit that allowed simple operations on floating-point numbers. The instruction set architecture (ISA) is composed of four identical function units, each of which had an 8-bit characteristic and a 20-bit mantissa. The engine accomplished the operations with a very simple structure that consisted of an arithmetic logic unit (ALU), three registers, and a stack. This basic unit could do parallel



adds, subtracts, and other similar two-variable operations on either the mantissa or the characteristic. Because one of the registers could shift down and the other could shift up, the engine could also handle multiplies and divides at the rate of one multiply or divide step per micro-cycle. The 12-chip system consisted of 1,344 copies of a single bit-slice layout that was composed of these five elements. Four pins on the chip were wired to tell the microcode which of the 12 functions it was supposed to do, according to its position in the subsystem organization.

Key to the design was the use of design techniques advocated by Carver Mead and Lynn Conway in the seminal Introduction to VLSI Systems. The book had a significant influence on the methodology used to arrive at the Geometry Engine's architecture.

There are three geometric operations that almost every graphics system must be able to perform: transformation, clipping, and scaling. Figure 1 illustrates these three functions.

The architecture of the Geometry Engine is illustrated in Figures 2 and 3, reproduced from Clark's published works. Each of the four function units at the bottom two-thirds of Figure 3 consists of two copies of a computing unit, a mantissa, and characteristic. The chip's internal clock generator is in the top left corner, while a microprogram counter with push-down subroutine stack is at the top right. The upper third of the chip included the control store, which held the equivalent of 40 kB and contained all of the microcode to implement the instructions and floating-point computations.

Immediately after releasing his landmark paper in *Lambda*, a small group of semiconductor engineers –Clark, Charles Kuta, Kurt Akeley, David J. Brown, and Abbey Silverstone – founded Silicon Graphics Inc. They employed the Geometry System board in a workstation called the Integrated Raster Imaging System, IRIS, which com-

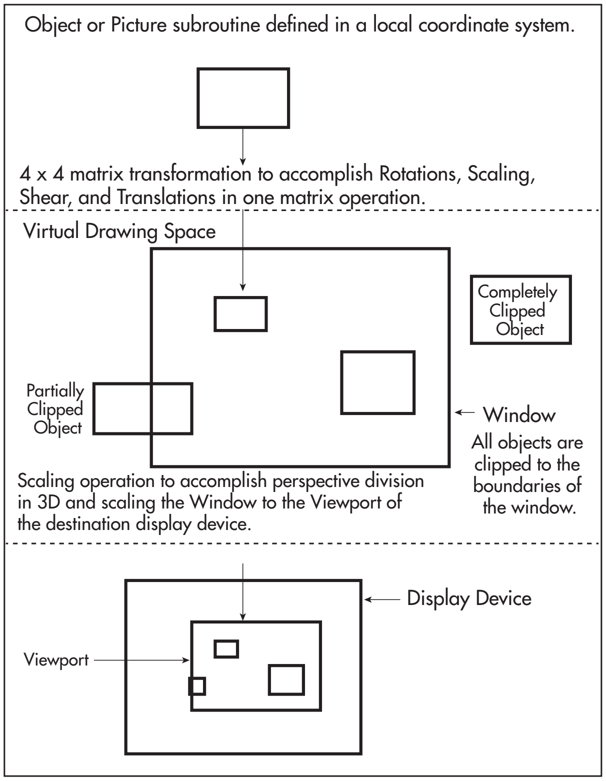


Figure 1. The three core operations of the Geometry Engine: transformations, clipping, and scaling.

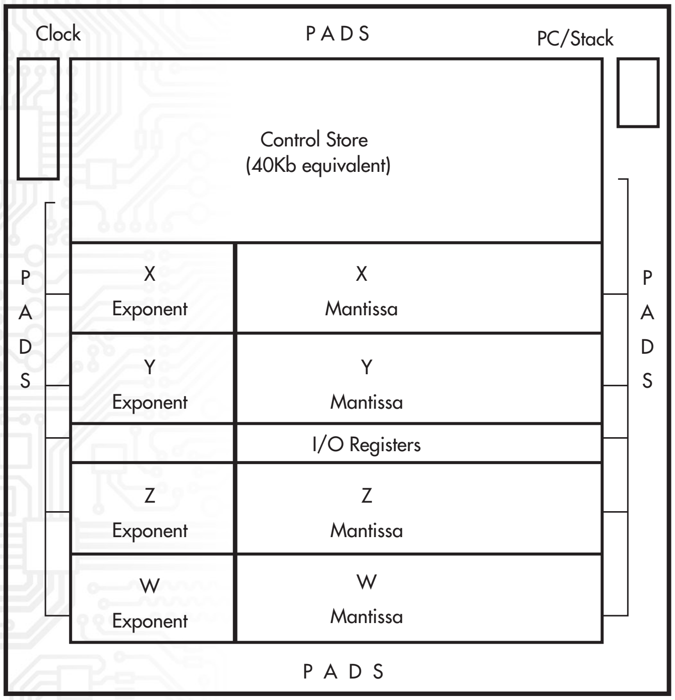
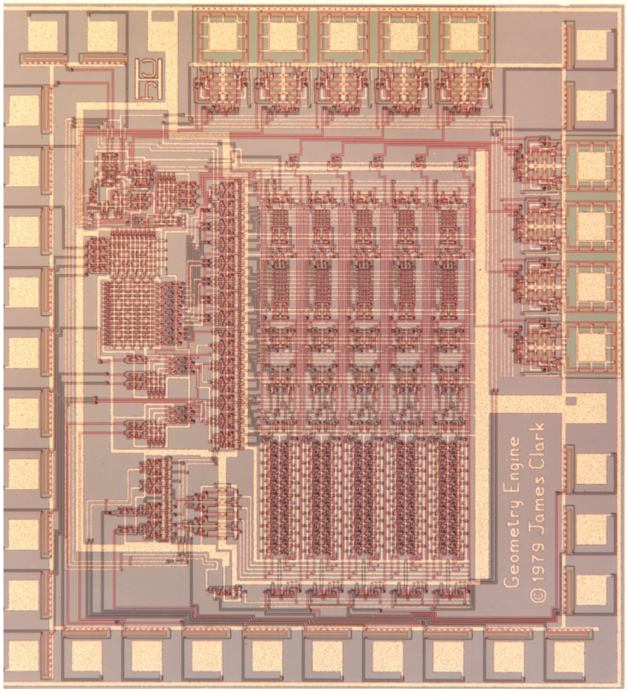


Figure 2. High level block diagram of Geometry Engine.

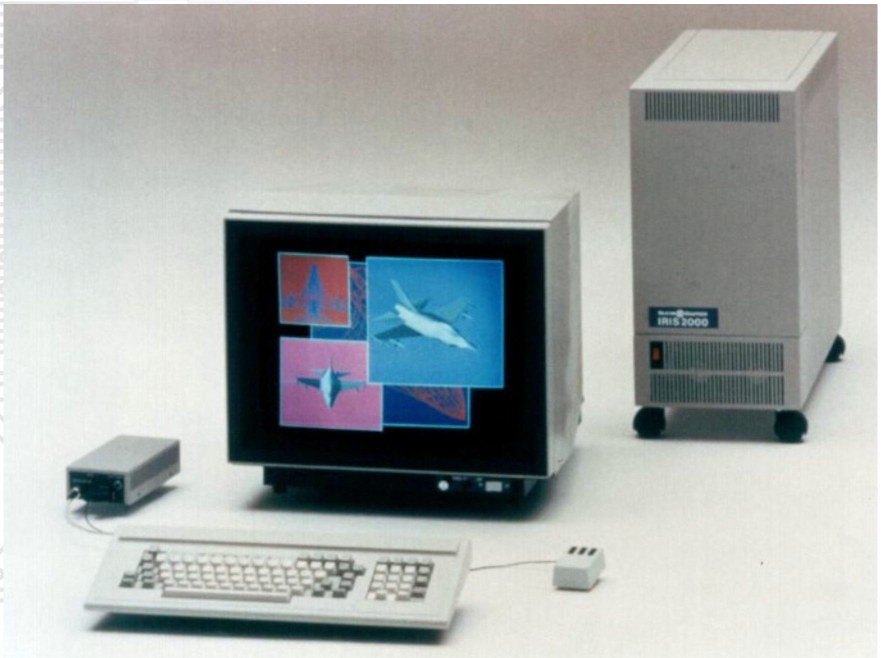


Die shot of Geometry Engine prototype.

bined the following components:

- A processor/memory board with the Motorola 68000 and 256 kB of RAM, which could be expanded to 2 MB. The 68000 microprocessor executed instructions in the on-board memory at 8 MHz. The memory was fully mapped and segmented for 16 processes. Additional memory was accessed over the Multibus at normal Multibus rates.
- A Geometry Subsystem, with Multibus interface FIFOs at the input and output of the Geometry System and from 10 to 12 copies of the Geometry Engine.
- A 1024 x 1024 Color Raster Subsystem, with high-performance hardware for polygon fill, vector drawing and arbitrary, variable-pitch characters. The hardware and firmware provide color and textured lines and polygons, character clipping, color mapping of up to 256 colors and selectable double- or single-buffered image planes.
- A 10-Mbps Ethernet interface board.

The Geometry Engine was also a breakthrough in size and performance. Matrox, one of the pioneering graphics add-in board (AIB) makers, having introduced their first PC graphics board in 1978, was the first company



SGI's IRIS 2000 graphics workstation.

to offer a single 3D AIB, the SM 640 in 1987. Matrox had been building 2D graphics boards for the PC and adopted SGI's Geometry Engine chip on a second layer mezzanine board to handle the 3D work.

But with the limited number of 3D applications for PCs, the product was a commercial flop. The migration of minicomputer and workstation applications to the PC took far longer than expected, but it lit up the imagination of all the AIB suppliers at the time.

In 1987, seeing the PC market slipping away from them due to clones, IBM introduced a proprietary operating system, OS2, and a new high-bandwidth AIB bus, the Micro Channel architecture (MCA). In 1991, SGI introduced the Micro Channel-based IrisVision, one of the first 3D accelerator cards available for the high-end PC market.

A few years later, IBM licensed both the graphics subsystem and the then-new IRIS Graphics Library (IRIS GL) API for their RS/6000 POWERstation line of POWER1-based workstations.

The Geometry System was a powerful computing system for graphics applications. It combined several useful geometric computing primitives in a custom VLSI system that was also scalable in nature. Within half a decade, the system was implemented on a single, half-million transistor integrated circuit, with a correspondingly reduced cost and increased speed.

SGI was conquered by the PC. But the legacy of the Geometry Engine has played a central role in improving graphics for games, scientific visualization, design, and VFX. Many of its basic features can be found in hundreds of millions of modern GPUs in use today.

to view this article online, ☞ *click here*

☞ **BACK TO TABLE OF CONTENTS**

CHECK OUT THESE RESOURCES FROM ELECTRONIC DESIGN AND OUR SISTER BRANDS

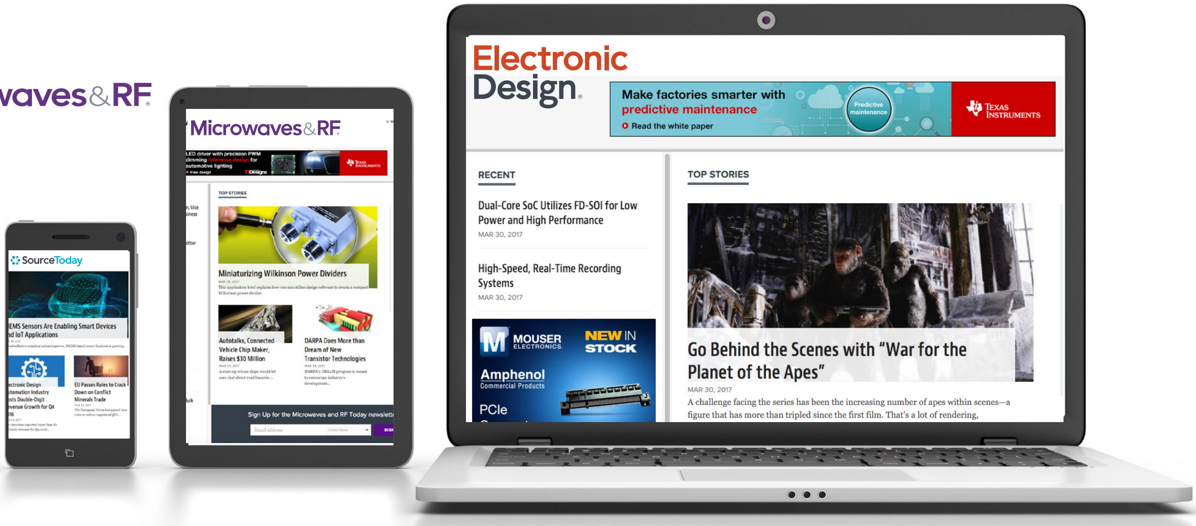
WEBSITES

Electronic Design

Microwaves&RF

sourceesb

Power Electronics


Electronic Design

Make factories smarter with predictive maintenance

Read the white paper

RECENT

Dual-Core SoC Utilizes FD-SOI for Low Power and High Performance

MAR 30, 2017

High-Speed, Real-Time Recording Systems

MAR 30, 2017

TOP STORIES

Go Behind the Scenes with "War for the Planet of the Apes"

MAR 30, 2017

A challenge facing the series has been the increasing number of apes within scenes—a figure that has more than tripled since the first film. That's a lot of rendering,

Microwaves&RF

TOP STORIES

Miniaturizing Wilkinson Power Dividers

Autotalks, Connected Vehicle Chip Maker, Raises $30 Million

DARPA Does More than Dream of New Transistor Technologies

Sign Up for the Microwaves and RF Today newsletter

SourceToday

EU Passes Rules to Crack Down on Conflict Minerals Trade

MAGAZINES

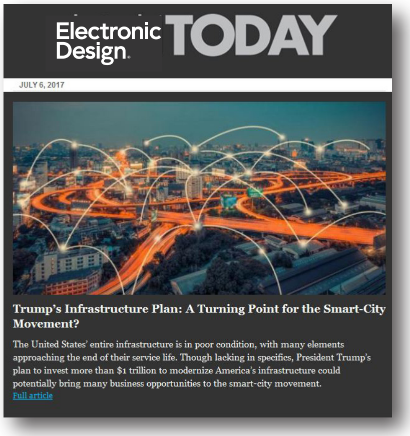
You can also apply for a subscription to one of our traditional magazines available in both print and digital formats.

ELECTRONIC DESIGN – complimentary in US and Canada - ☞ Subscribe Now
Non-qualified or Outside the US and Canada: you will be prompted to pay based on your location.

MICROWAVES & RF – complimentary internationally- ☞ Subscribe Now
Non-qualified: you will be prompted to pay based on your location.

NEWSLETTERS

Stay current with the industry with newsletters created by engineers and editors with inside connections! You'll find coverage on a wide variety of electronics technologies from Analog & Mixed Signal to Embedded and more.
☞ **Click Here** to check out what more than 200,000 engineers are reading now.


Electronic Design TODAY

JULY 6, 2017

Trump's Infrastructure Plan: A Turning Point for the Smart-City Movement?

The United States' entire infrastructure is in poor condition, with many elements approaching the end of their service life. Though lacking in specifics, President Trump's plan to invest more than $1 trillion to modernize America's infrastructure could potentially bring many business opportunities to the smart-city movement.

Full article

ABOUT US

A trusted industry resource for more than 50 years, the Penton Electronics Group is the electronic design engineer's source for design ideas and solutions, new technology information and engineering essentials. Individual brands in the group include *Electronic Design*, *Microwaves & RF* and *Power Electronics*. Also included in the group is a data product for engineers, *SourceToday.com*.

ENDEAVOR BUSINESS MEDIA™