3D-Chip Leaders Push the Envelope

3Dfx, ATI, Matrox, and Nvidia Tempt Both Gamers and Business Buyers

by Peter N. Glaskowsky

If you use your PC for 3D games or computer-aided design, it's time to consider buying a new graphics card. The latest graphics chips from 3Dfx, ATI, Matrox, and Nvidia are substantially better than those that were available last fall. Some are as much as 3× faster, and most offer new hardware or software features that users will welcome.

Though all these new chips share the same basic architectures as their predecessors, some have added a second pipeline for 3D rendering. Because they all run at higher internal clock speeds, they also achieve significant performance enhancements on 2D graphics and can support higher-resolution monitors than the older parts.

These new chips will also be welcomed by PC OEMs. With drivers optimized for Pentium III and support for the AGP 4× interface, these chips will help attract customers to high-end systems. Many of the chips from the previous generation of graphics accelerators, however, are likely to remain on the market. Though they offer less performance and fewer features, price reductions will make these older devices an excellent deal for entry-level PCs.

Will New Voodoo3 Outdo Voodoo2?

Last year's 3Dfx Voodoo2 (V2) was a popular choice among avid gamers. In its dual-board scan-line-interleaved (SLI) configuration, the V2 was faster than any other 3D-gaming accelerator. It was also more expensive, especially in early 1998, when an SLI board pair cost $500 or more. Because Voodoo2 provided no support for 2D graphics or video, it was also much less convenient to use than integrated 2D/3D/video chips. V2 users also needed a separate 2D card to handle the Windows desktop. But 3Dfx solved these problems with its Voodoo Banshee (see MPR 7/13/98, p. 16), and by the end of 1998, Banshee was the company's highest-volume product. Banshee's performance trailed that of the V2, as well as of Nvidia's competing TNT, which also had a better feature set. Gamers wanted more.

At Comdex, 3Dfx announced plans to give the gaming community what it wants—a single chip with the speed of the Voodoo2 SLI plus the 2D and video features of Banshee. The new Voodoo3, which began shipping in March, could also be thought of as Banshee 2—the device is essentially a Banshee with a second texture-mapping engine and a higher clock rate: 183 MHz versus 100 MHz for Banshee.

The V3 will be available on three different boards from STB Systems, the board maker 3Dfx recently acquired. The boards run at 143, 166, and 183 MHz; the slowest board lacks the TV output on the other two, while the fastest adds a flat-panel-display controller. In the past, buyers could choose among V2 boards from Diamond, Creative Labs, Canopus, and others. For the retail graphics-board market, however, the V3 will be available only through STB; 3Dfx partner Quantum3D will handle customers in the professional visual-simulation and 3D content-creation markets. It is 3Dfx's hope that the STB board assortment will help make up for the loss of other vendors' brands on retail shelves.

The company is also in the midst of a $20 million advertising effort orchestrated by the ad agency that created the "Got Milk?" campaign. The effort is meant to reinforce 3Dfx's image as a supplier of 3D accelerators for gaming, but we suspect that most of the 18–34-year-old males who make up the target audience are already aware of 3Dfx.

Even if it succeeds in making more potential customers aware of its products, 3Dfx will find it increasingly difficult to overcome the perception that it is falling behind in the 3D race. As Table 1 shows, the V3 is not dramatically faster than competing offerings, and it cannot compete on the basis of features. The V3's peak dual-textured pixel-drawing rate is slightly faster than that of Nvidia's Riva TNT2, but it is much

	3Dfx Voodoo3	ATI Rage 128 Pro	Matrox G400 MAX	Nvidia RIVA TNT2	S3 Savage4
Bus Interface	AGP 2×	AGP 2×/4×	AGP 2×/4×	AGP 2×/4×	AGP 2×/4×
Memory Width	128 bits	128 bits	128 bits	128 bits	64 bits
Max Memory Clock	183 MHz	166 MHz	200 MHz	200 MHz	143 MHz
Maximum Memory	16M	32M	32M	32M	32M
On-Chip Cache	n/a	16K	n/a	12K	8K
Core Clock Rate	183 MHz × 2	150 MHz × 2	166 MHz × 2†	175 MHz × 2	143 MHz × 2
Peak Bilinear Rate	183 MB/s	300 MB/s	333 MB/s	350 MB/s	143 MB/s
Bilinear/Dual Texture	183 MB/s	150 MB/s	166 MB/s	175 MB/s	143 MB/s
Peak Trilinear/16 bit	183 MB/s	300 MB/s	245 MB/s	175 MB/s	143 MB/s
Peak Trilinear/32 bit	not supported	300 MB/s	200 MB/s	175 MB/s	143 MB/s
RAMDAC Speed	350 MHz	300 MHz	360 MHz	300 MHz	300 MHz
Max Res at 85 Hz	2,048 × 1,536	1,920×1,200	2,048 × 1,536	1,920×1,200	1,920 × 1,440
Video Support	VMI in/out	VIP 1.1 w/ext	VIP 1.1 w/ext	VIP 1.0	VIP 1.0
LCD Support?	Yes	Yes	Yes	Yes	Yes
Availability	Now	July	Now	Now	2Q99
Volume Price	n/a	$42	n/a	$45	$25

Table 1. Nvidia's RIVA TNT2 achieves the highest pixel-fill rate for bilinear and anisotropic texture filtering. S3 claims faster trilinear rates for the Savage4, but Savage4's 64-bit memory will significantly reduce sustained throughput. n/a = not available. (Source: vendors, except †MDR testing)

slower in some modes. The 3Dfx chip does not support AGP execute-mode texturing or 32-bit rendering, nor can it perform true anisotropic texture filtering (though anisotropic filtering will be implemented in a future driver using a software-assisted technique).

These omissions are not crippling. The V3's performance on many real-world games is still unmatched—and titles written only for 3Dfx's Glide programming interface won't run on TNT2 at all. AGP texturing is relatively unimportant for cards equipped with 16M of local memory. Few 3D games look better in 32-bit color, and no current game uses anisotropic texturing. Glitzy ad campaigns can't explain these issues to potential customers, however.

Whether or not 3Dfx is ready, competitors will force it to compete on the basis of feature lists and benchmark performance. The company's current products—even the otherwise excellent Voodoo3—don't look good when evaluated on these criteria. We hope the Voodoo4, due later this year, will be more competitive in the ways that count in today's 3D market.

TNT2 Yields Nearly Twice the Power of TNT

Nvidia found a straightforward way to improve the Riva TNT—it shrank the chip from a 0.35- to a 0.25-micron process. Combined with some minor architectural enhancements, the shrink yielded excellent results. The new TNT2 runs 85% faster (up to 175 MHz vs. TNT's 95-MHz clock rate) and sets a new record for pixel-fill rate at 350 Mpixels/s, when applying just one bilinear-filtered texture to each pixel. In dual-texturing mode, the chip completes just one pixel per clock, achieving a fill rate of 175 Mpixels/s, slightly slower than the best effort of the Voodoo3.

The TNT2 supports more advanced 3D features than the Voodoo3, with support for 32-bit true-color rendering, 24 bits of depth (Z-buffer) precision, a stencil buffer, and 32M of local graphics memory. The new design is compatible with AGP 4× and adds a digital flat-panel display port. In contrast to Matrox's G400, the TNT2 has only one CRT controller, limiting its flexibility in managing a second display.

With comparable performance and a better feature set than Voodoo3, the TNT2 is sure to be popular among avid gamers as well as mainstream buyers. The Nvidia chip lacks only one important feature compared with 3Dfx's—the ability to run games written to the Glide API. New Glide titles continue to be released, and some are not ported to Direct3D or OpenGL until months after the Glide release. This consideration is enough to ensure 3Dfx a large share of the gaming market through the Christmas season, though that share is likely to shrink as Nvidia attracts more converts.

Nvidia is well aware that gaming is just a small niche market. To better service potential customers in the broader PC market, the company offers the Vanta chip—a derivative of the TNT2 design with a 64-bit memory interface (instead of the TNT2's 128-bit controller), a slower 100-MHz clock rate, and a much lower price. Vanta is priced at just $19, less than half the $45 asking price of the TNT2.

With its gentler name and greater affordability, Vanta should sell into markets for which the TNT2 isn't suited. Nongaming home computers, business desktop systems, and low-cost PCs for all applications offer good opportunities for Vanta. Only the least expensive systems—the sub-$800 machines that are likely to use integrated-graphics chip sets such as Intel's 810 (née Whitney) or sub-$10 discrete graphics chips—are outside Vanta's reach.

Nvidia's previous chips have been successful in the performance-desktop PC market, where the company faces only a few significant competitors. In such PCs, Nvidia's gaming fame is a valuable selling point.

Graphics chips for low-end and midrange PCs are available from many more suppliers, including 3Dlabs, ATI, Intel, Matrox, S3, SiS, and Trident. All have graphics chips priced below $20, some of which have feature sets superior to Vanta's. Though Vanta may offer the best 3D performance in this price range, rendering speed is not the primary determinant of success among mainstream graphics chips. Factors such as end-user brand awareness, 2D performance, video features, driver stability, and pin- or software-compatible alternatives are more important to most PC OEMs in this segment.

ATI Aims To Maintain Desktop 3D Dominance

The mainstream graphics market continues to be dominated by ATI Technologies. Though 3Dfx and Nvidia offer the best 3D-gaming accelerators, ATI's Rage Pro and Rage 128 offer better price/performance and a better feature set for most home and business users.

ATI has recently released a midlife kicker for the Rage 128, called the Rage 128 Pro. The Pro version features a speed boost from 100 to 150 MHz and incorporates a few new features, including anisotropic texturing support and on-chip digital flat-panel line drivers compatible with all the major flat-panel interface standards. The new chip is pin compatible and software compatible with the Rage 128 (see MPR 9/14/98, p. 16), making it an easy upgrade for the many PC OEMs that are already using that chip.

Like the Rage 128, the Rage 128 Pro is available with a 64-bit memory interface (the Rage 128 Pro VR). For more performance, a version with a 128-bit interface is also offered as the Rage 128 Pro GL. ATI's announced volume prices slightly undercut those of competing chips. Public pricing is not a good indication of actual selling price, however, as some vendors offer greater discounts than others.

The GL has the highest peak trilinear-filtered pixel drawing speed of any chip announced to date, but it remains slower than the Voodoo3 on the dual-textured, bilinear-filtered pixels that are commonly used in today's 3D games. ATI says the Rage 128 Pro will achieve about 50% better overall performance than the Rage 128, as measured by 3D WinBench 99, which would put it into the same performance category as the Voodoo3, G400, and TNT2.

Though it is the least improved of the chips covered here, the Rage 128 Pro is still in the top tier for 2D and 3D

performance and continues to have the best video support of any PC-graphics chip. As such, it should continue to sell well to ATI's customer base, by far the broadest and most lucrative in the graphics market.

G400 Doubles Down on G200 Design

Matrox Graphics, ATI's Canadian compatriot and competitor, has also seen its greatest success in the mainstream market. Unlike 3Dfx, which is trying to escape the gaming niche, Matrox is trying hard to make a name for itself as a provider of 3D-gaming accelerators. Matrox's new G400 Max should go a long way toward that goal and provide strong competition for both 3Dfx and Nvidia.

On the basis of Matrox's claimed 333-Mpixel/s peak rendering rate, we estimate the G400 Max runs at 166 MHz internally, with the ability to render two pixels per clock with bilinear filtering, three pixels per two clocks with 16-bit trilinear filtering, and one pixel every two clocks with 32-bit trilinear filtering. Matrox does not explain the internal architecture of its graphics chips, but these ratios suggest the G400 has some unusual design features.

For less demanding buyers (especially Matrox's loyal PC OEM customers), Matrox also offers slower speed grades of the G400 without the "Max" suffix. As it did with the G200, Matrox is likely to offer a variety of add-in boards equipped with the G400. Though the company never discusses the clock speeds of chips on its boards, the G400 Max label may help end users identify the fastest boards.

The width of the G400's memory bus has been doubled to 128 bits, eliminating one of the weaknesses of the G200 design. Matrox also doubled the G200's display support: the G400 can drive two independent displays. One monitor can be connected to the G400 Max's 360-MHz RAMDAC (the fastest of this group) for resolutions up to 2,048 × 1,536 with an 85-Hz refresh rate. A second CRT controller drives a digital output port that can be connected to an external RAMDAC, an analog video encoder, or a flat-panel display driver. The secondary port supports resolutions up to 1,280 × 1,024 pixels at 60 Hz.

The two displays can be used to create a larger Windows desktop, or the secondary output can display full-screen video on an attached television while the primary monitor is used for normal PC applications.

A Rising Tide Lifts All Boats

Every 3D-chip company claims to have some unique advantage in the market, and some do—3Dfx's Glide, ATI's video support, the unmatched performance of Nvidia's flexible dual-pipeline rendering engine, and so on—but all share the same basic process technology. All the chips described here are optimized for similar 0.25-micron ASIC processes, run at similar clock speeds, and have similar RAMDACs.

The differences among these chips are minor compared with their similarities. All are fast enough for virtually any software application short of high-end professional 3D-design tools. Indeed, some of these vendors are beginning to tout their chips' performance on professional 3D benchmarks, hoping to get some share of the small market for entry-level workstations. Though such sales are unlikely to be financially significant, they can enhance brand image—encouraging users to conclude that if a chip is fast enough for CAD, it's fast enough for them.

Avid gamers don't select chips on the basis of "fast enough," however, but look for the *fastest* 3D acceleration available. On this basis, 3Dfx's Voodoo3 and Nvidia's TNT2 are clearly the best choices, but ATI's Rage 128 Pro and Matrox's G400 Max will come within a few percentage points of the 3Dfx and Nvidia products on 3D games. Buyers willing to sacrifice this slim margin will get more features at a lower price, a tradeoff many will be willing to make.

For most home and business users, fast enough is good enough, and all of these chips have a chance for success. Sales to these customers will hinge on points other than raw speed. Among the chips described here, the ATI and Matrox products have a definite edge for most buyers, as both brands are well known and generally respected.

We believe Nvidia's Vanta brand will be more effective at gaining a share of the mainstream market than 3Dfx's gamer-targeted advertising. The products that 3Dfx is promoting aren't as well suited to average consumers as Vanta. If 3Dfx were advertising mainstream products—even if they were just defeatured Banshee or Voodoo3 chips—it would likely receive a much better return on its advertising dollar.

Despite the best efforts of both 3Dfx and Nvidia, however, neither will seriously threaten ATI's position atop the overall PC graphics-chip market this year. If ATI continues to turn out good products like the Rage 128 Pro every six months or so, there's little that its competition can do except fight for niche markets like gaming, CAD, and the relatively unprofitable deep-low-end segment. M

Pricing and Availability

3Dfx's Voodoo3 is available now on STB graphics cards. The company (*www.3Dfx.com*) has not released chip pricing, though the Voodoo3 is available in chip form for motherboard designs and other uses apart from add-in boards.

The Rage 128 Pro will be available from ATI (*www.atitech.com*) in July. The VR version is $32; pricing for the GL version is set at $42 in quantities of 10,000.

Matrox's G400 Max will ship in 2Q99 on Matrox's own graphics cards. The G400 is available for other uses, but Matrox (*www.matrox.com*) has not disclosed pricing.

Nvidia's TNT2 and Vanta chips will be in full production this quarter, priced at $45 and $19, respectively, in 10,000-unit quantities. Nvidia's Web site is at *www.nvidia.com*.