Shootout: GeForce FX 5800 Ultra / Radeon 9800 Pro

By Brian Neal – April 2003

Next Generation Graphics Cores Face Off

Back in November of last year, NVIDIA officially announced the GeForce FX, the company's next-generation 3D accelerator. Codenamed NV30, the GeForce FX incorporates the most comprehensive programmable DirectX 9 vertex and pixel shader implementations currently on the market. But what about the rest of the GPU? How does the GeForce FX stack up to the competition?

Though announced last November, the GeForce FX didn't appear on shelves until earlier this year. So while NVIDIA has maintained performance leadership over its competitors for several product cycles, there was another GPU leading pack: the ATI Radeon 9700. Shipping last August with a 256-bit memory interface capable of nearly 20 GB/s and 8 pixel pipelines with a peak fillrate of 2.6 GPixels/s, the Radeon 9700 easily surpassed the competition of its time.

Despite the delays, the GeForce FX Ultra has arrived. With a core clockrate of 500 MHz, it is the highest-clocked GPU in its class, and its 128-bit DDR-II memory interface comes very close (16 GB/s) to ATI's 256-bit behemoth.

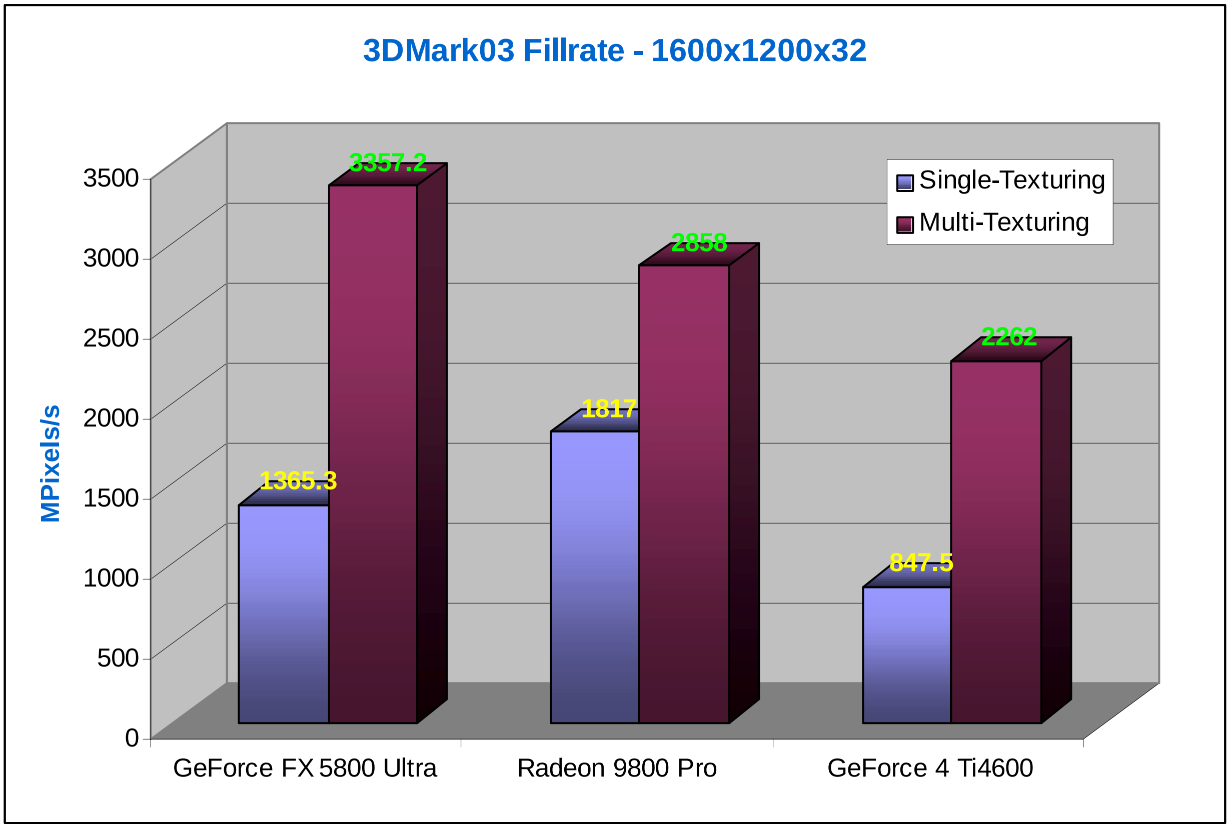
But ATI isn't sitting still, either. The tweaked R350 core at the heart of the Radeon 9800 Pro has been boosted to 380 MHz, increasing peak fillrate just past 3 GPixels/s. The memory clockrate has also increased to 340 MHz DDR (680 MT/s), thereby increasing memory bandwidth to 21.76 GB/s -- roughly a 9% boost over the previous Radeon 9700. These improvements have been made at the same 0.15µ process node as the previous R300 core. ATI's first 0.13µ chip is the RV350: the mid-range Radeon 9600.

In this review, we'll take a look at some of the architectural elements of both the GeForce FX and Radeon 9800 Pro and we'll also run a variety of benchmarks from first person shooters like Ghost Recon to role playing games like NeverWinter Nights. We'll also talk a bit about image quality and anti-aliasing. So, let's start by taking a look at the technical specifications of the GeForce FX 5800 Ultra and the Radeon 9800 Pro:

	NVIDIA GeForce FX 5800 Ultra	**ATI Radeon 9800 Pro**	**NVIDIA GeForce 4 Ti4600**
Core	NV30	R350	NV25
Transistors	125M	110M	63M
Process	0.13µ	0.15µ	0.15µ
Core Clockrate	500 MHz	380 MHz	300 MHz
Pixel Pipelines	8 (4)*	8	4
Pixel Fillrate (MPixels/s)	4000 (2000)*	3040	1200
Texture Units / Pipeline	1 (2)*	1	2
Texel Fillrate (MTexels/s)	4000 (4000)*	3040	2400
Vertex Pipelines	N/A**	4	2
Memory Clockrate	500 MHz (1 GT/s)	340 MHz (680 MT/s)	300 MHz (600 MT/s)
Memory Bus Width	128-bit	256-bit	128-bit
Peak Bandwidth	16 GB/s	21.76 GB/s	10.4 GB/s
DX Vertex Shader Rev	2.0+	2.0	1.1
DX Pixel Shader Rev	2.0+	2.0	1.3

** Though the official GeForce FX specifications indicate an 8x1 pipeline arrangement, this is only true in some circumstances, and in fact, it more closely resembles a 4x2 arrangement in most cases.*
*** NVIDIA documents the GeForce FX vertex shader as a large array of FP units. This organization differs from the more typical vertex pipelines of the R350 and NV25.*

From the specs above, the GeForce FX appears to have the edge in fillrate, however, this is a more complex issue than one might initially assume. Let's start with some fillrate benchmarks:



As you can clearly see from these results, the measured pixel fillrate of the GeForce FX (1365.3 MPixels/s) is well below the stated 4000 MPixels/s. It's also well below what the Radeon 9800 Pro achieves, despite the fact that the R350 core is rated with a lesser peak fillrate. In the case of multitexturing, however, the GeForce FX's fillrate is much closer to its theoretical peak.

The performance characteristics exhibited by the GeForce FX in fillrate benchmarks and other tests sparked quite a bit of discussion across the internet regarding the validity of NVIDIA's claims and the true nature of the NV30's architecture. One such discussion began in **a thread** at Beyond3D. Several days later, The Inquirer **reported** that the GeForce FX was in fact a 4x2 renderer in most cases, as opposed to the 8x1 configuration that had earlier been disclosed. Tech-Report **followed-up** on the story, and published the following statement from NVIDIA:

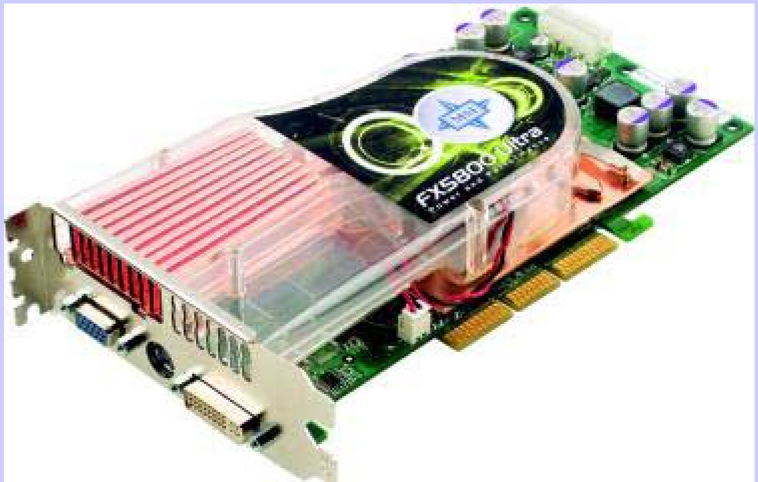
> GeForce FX 5800 and 5800 Ultra run at 8 pixels per clock for all of the following:
>
> a) z-rendering
> b) stencil operations
> c) texture operations
> d) shader operations
>
> For advanced applications (such as Doom3) *most* of the time is spent in these modes because of the advanced shadowing techniques that use shadow buffers, stencil testing and next-generation shaders that are longer and therefore make the apps "shading-bound" rather than "color fill-rate" bound.
>
> **Only color+Z rendering is done at 4 pixels per clock, all other modes (z, stencil, texture, shading) run at 8 pixels per clock.**
>
> The more advanced the application, the less percentage of total rendering is color, because more time is spent texturing, shading and doing advanced shadowing/lighting.

The quote above makes things quite clear. When rendering color+Z, the GeForce FX operates as though it has four pixel pipelines with two texture units each. Since this is exactly what we are benchmarking in the above fillrate test, the GeForce FX writes four color pixels to the framebuffer per clock. As a result, the true fillrate of 2 GPixels/s for single-texturing is indicated in parentheses in the specifications table above.

Conversely, the Radeon 9800 Pro is a true 8x1 renderer, and therefore has twice the pixel fillrate per clock as the GeForce FX, enabling it to outperform its competitor in single-textured fillrate despite running at a lower clockrate. In multi-texturing, the GeForce FX pulls ahead because it has two TMUs per pipeline, along with a higher clockrate.

With that issue out of the way, let's take a look at the video cards benchmarked in this review:

MSI FX5800 Ultra



The GeForce FX 5800 Ultra and its large two-slot cooler

The GeForce FX tested in this review is one of the first commercially available boards, the MSI FX5800 Ultra. It is one of the 100,000 Ultras that will be produced and, as such, features a 500 MHz core clockspeed. For comparison, the normal GeForce FX 5800 runs at 400 MHz, with a memory clockrate of 800 MHz (12.8 GB/s).

Another picture, note the molex connector at the edge of the board

Much has been made of the GeForce FX's cooling system in terms of size and noise. The GeForce FX employs a variable speed cooler that runs at full-speed when the GPU is in 3D mode and at a reduced-speed when in 2D mode. This is to reduce the noise output from the cooler when it's not needed. While gaming, it is assumed that the extra noise will not be noticed over the soundtrack of the game itself. Johan humorously noted while running the benchmarks for this review that the variable speed fan was actually a useful feature when benchmarking, as in a way, it audibly informs the reviewer that a test has completed when the cooler becomes silent. The MSI FX5800 Ultra retails for US$399.00. The non-Ultra version is $100 cheaper at US$299.00.

ATI Radeon 9800 Pro

The Radeon 9800 Pro features a smaller, single-slot cooler

As you can see in the picture above, the Radeon 9800 Pro has a noticably smaller heatsink and fan that fits within a single slot. Subjectively speaking, it is quieter than the GeForce FX and some Radeon 9700 Pro boards. The 128 MB Radeon 9800 Pro lists at US$399.00 from ATI.

Benchmarked Configurations

For all benchmarks we used NVIDIA's 43.00 drivers and ATI's official 7.84 drivers (not the beta 7.85). Vsync was disabled at all times. Driver configuration options were set to their defaults, with the Radeon drivers using "Performance" mode and the NVIDIA drivers using "Balanced" mode. The benchmark platform was a 2.8 GHz Pentium 4 on a Gigabyte Granite Bay motherboard, as detailed below:

Hardware

- **2.8 GHz Pentium 4**
- **Gigabyte GA-8INXP**
- **MSI FX5800 Ultra (GeForce FX)**
- **ATI Radeon 9800 Pro**
- **512 MB Corsair PC3200 XMS (DDR-SDRAM) running at 333 MHz CAS 2 (2-3-3-6)**
- **Maxtor 80 GB DiamondMax 740X (7200 rpm, ATA-100/133)**
- **Sound Blaster Live!**

Software

- **NVIDIA Detonator 43.00 Drivers**
- **ATI Catalyst Radeon 3.2 (7.84)**
- **Intel chipset INF update 4.09.1012**
- **Windows XP Service Pack 1**
- **DirectX 9**

Firstly, I want to thank **Johan De Gelas** for running all the benchmarks seen in this review. This article would not be possible without his efforts.

Additionally, we'd like to thank the following helpful people for their support and important contributions to this review:

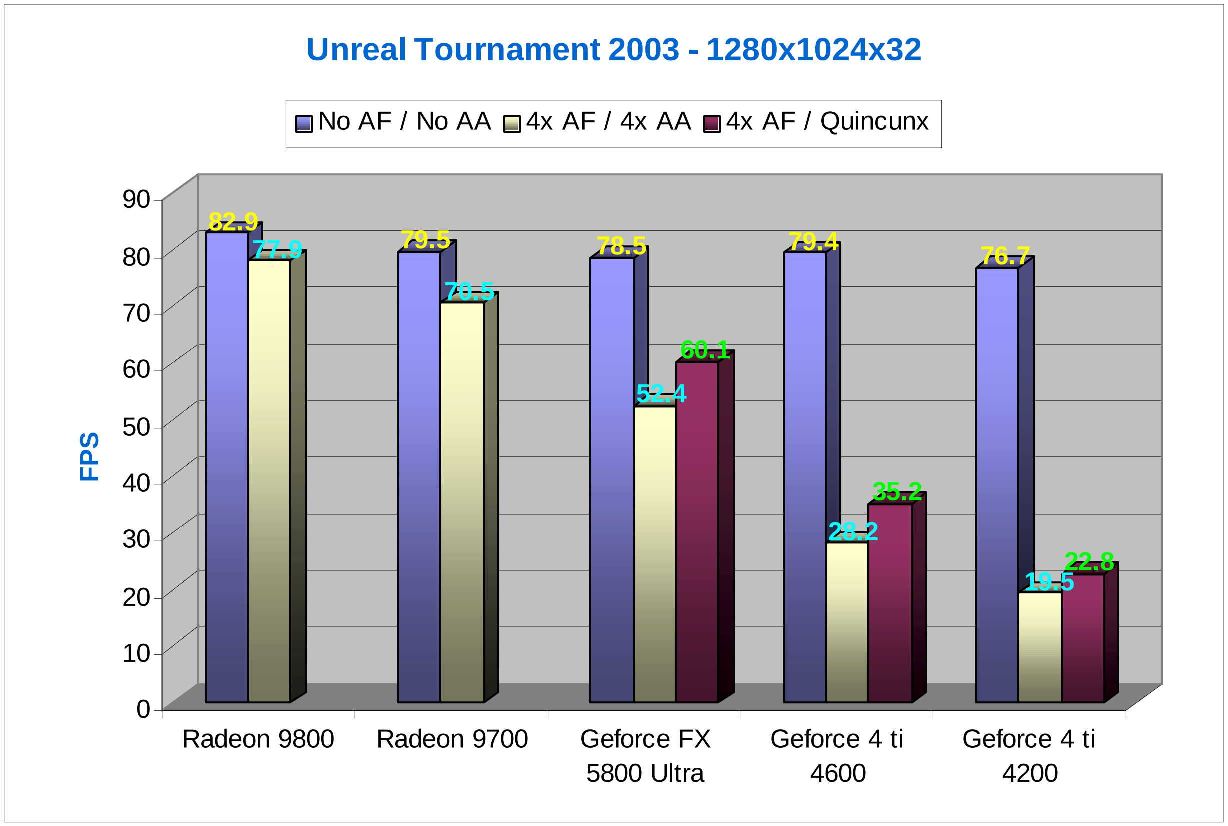
- Saskia Verhappen and Angelique Berden (**MSI**)
- Brenda Chen and Leo Chu (**Gigabyte**)
- Robert Pearce (**Corsair**)
- Kristof Semhke and George Alfs (**Intel**)

On to the benchmarks!

Gaming Benchmarks: Unreal Tournament 2003

We'll start with Unreal Tournament 2003, the popular online first person shooter. We'll begin by testing at 1280x1024x32bpp with a variety of filtering and anti-aliasing configurations.

Image courtesy ***UnrealTournament2003.com***

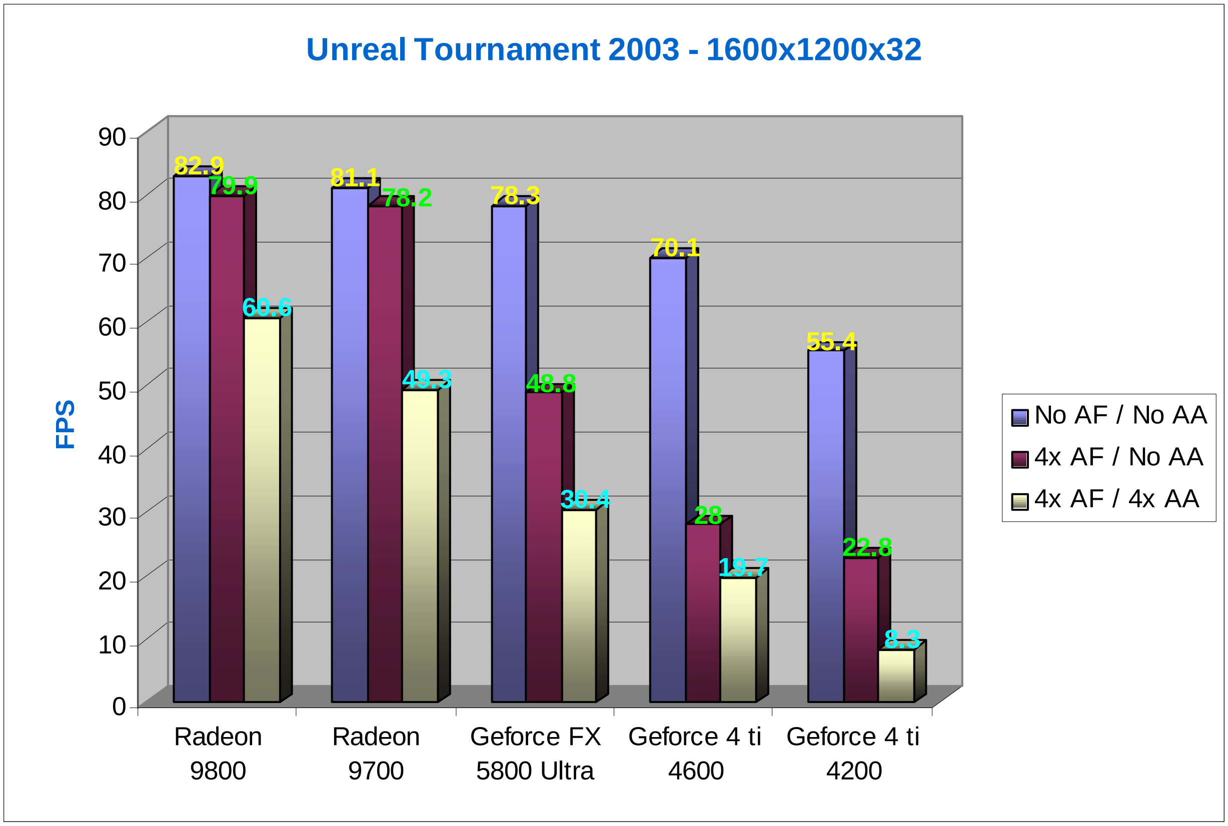


In this graph, anisotropic filtering and anti-aliasing are denoted by "AF" and "AA," respectively.

As you can see above, the Radeon 9800 Pro is the leader with both 4x anisotropic filtering and 4x anti-aliasing enabled (77.9 FPS), as well as with both AA and AF disabled (82.9 FPS). The GeForce FX 5800 Ultra is within a few frames per second of both Radeons with AA and AF disabled, though it actually falls short of the GeForce 4 Ti4600 in this test by around 1 FPS.

Obviously, one doesn't purchase a top of the line $400 video card to play games with all quality enhancing features disabled, so the results with no AA and AF are presented here for comparison purposes and to give you an idea of how each different architecture is affected by different workloads. In the case of the 4xAF/4xAA and 4xAF/Quincunx tests, the GeForce FX 5800 Ultra manages to almost double the performance of its predecessor, though the Radeon 9800 Pro still outperforms it by a solid 33% in 4x/4x mode.

Also shown here are results using NVIDIA's Quincunx anti-aliasing. The GeForce FX manages to roughly double the performance of the GeForce 4 Ti4600 at 1280x1024. The GeForce 4 cards were not tested at 1600x1200, though the GeForce FX turns in 35.6 FPS at this resolution. Compared to the GeForce FX's 48.8 FPS with 4x anisotropic filtering and no anti-aliasing, there is a 28% performance hit associated with enabling Quincunx at this resolution.

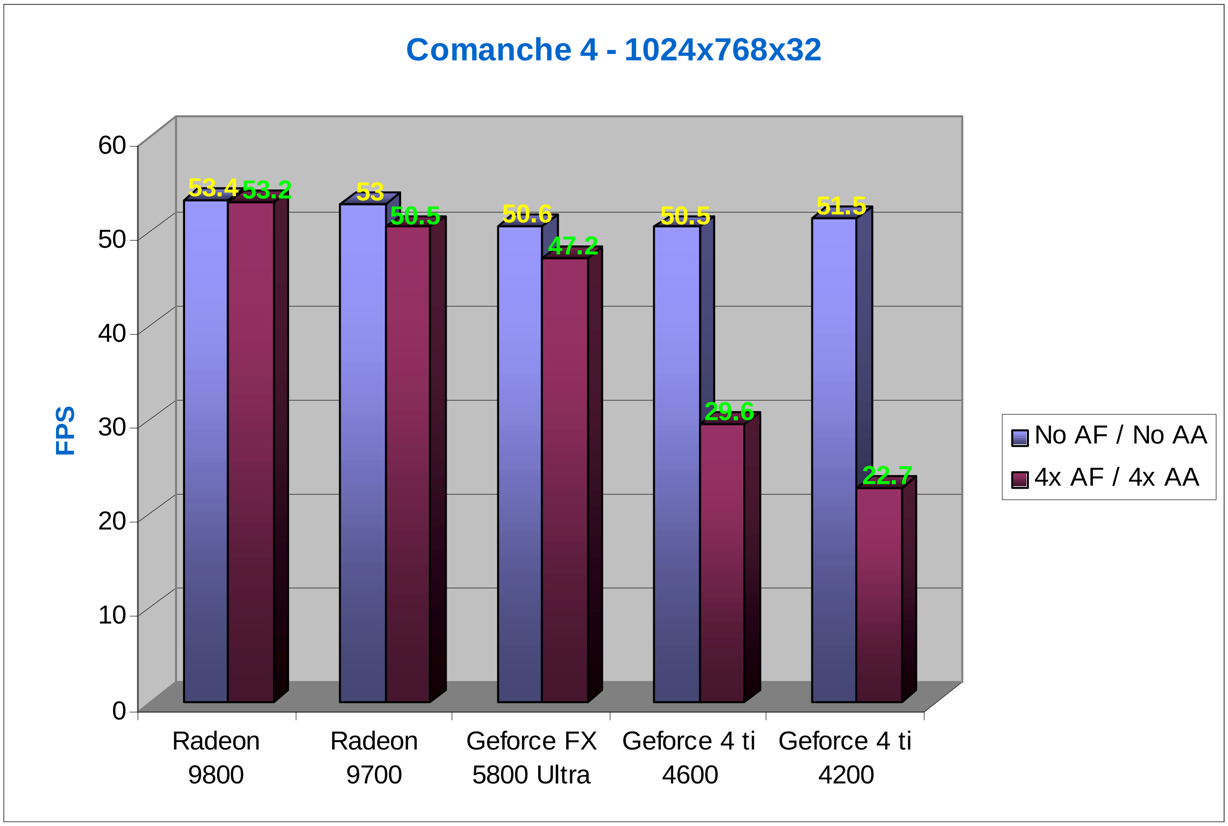


Here we're looking at results from 1600x1200, and an additional mode has been added: 4x anisotropic filtering with no anti-aliasing. Once again, all three of the top-end cards (R9800, R9700, and GFFX) deliver performance that is within a few frames per second of each other when AF and AA are disabled. When 4x anisotropic filtering is enabled, the GeForce FX 5800 Ultra takes a 39% drop in framerate. As part of NVIDIA's Intellisample technology, the GeForce FX implements adaptive texture filtering, but here we see the card's performance is being significantly impacted by 4x AF, while the impact to the Radeon 9800 Pro is negligible at best. Later on in the review, we'll discuss the image quality of these various anisotropic filtering and anti-aliasing implementations. For now, however, we'll focus on the performance.

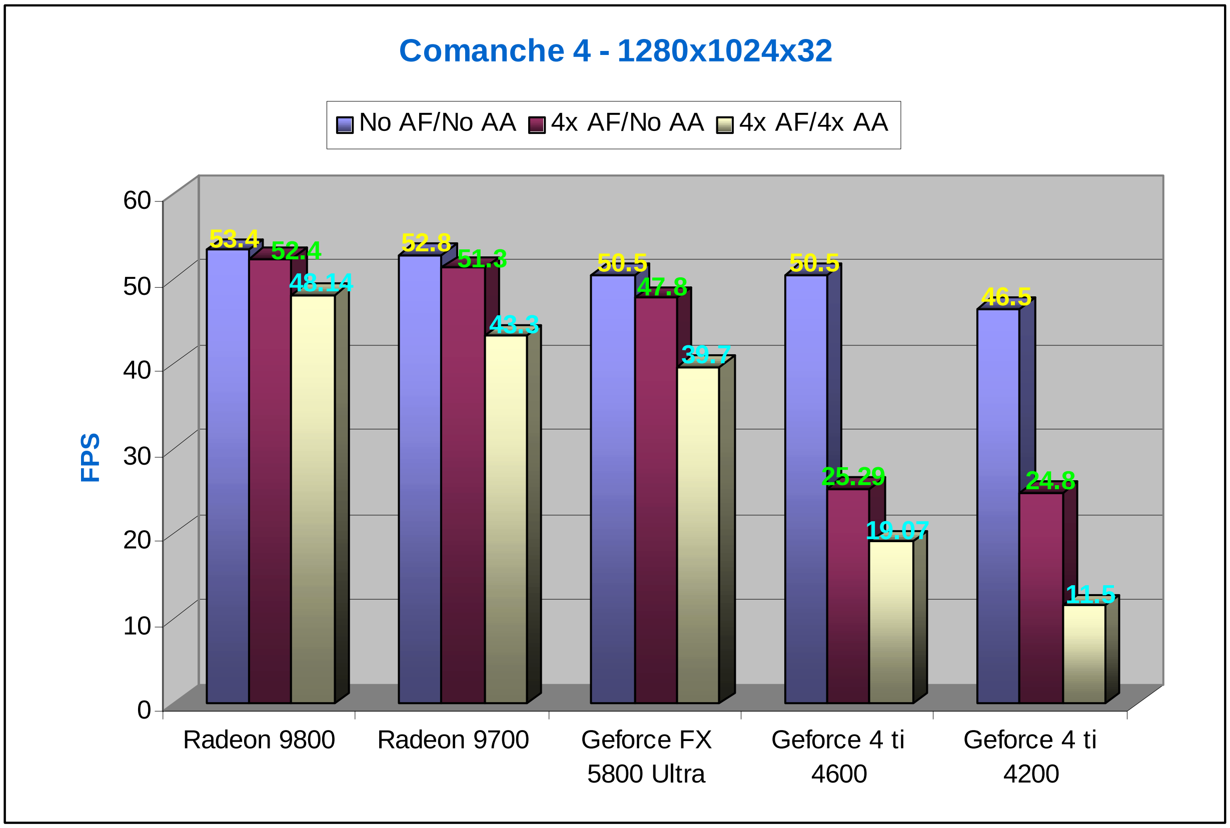
It's worth noting that the Radeon 9800 Pro sees a noticeable 19% improvement in framerate with both 4x AF and 4x AA compared to the previous R9700 Pro. The other modes are nearly identical in performance on both cards.

Comanche 4

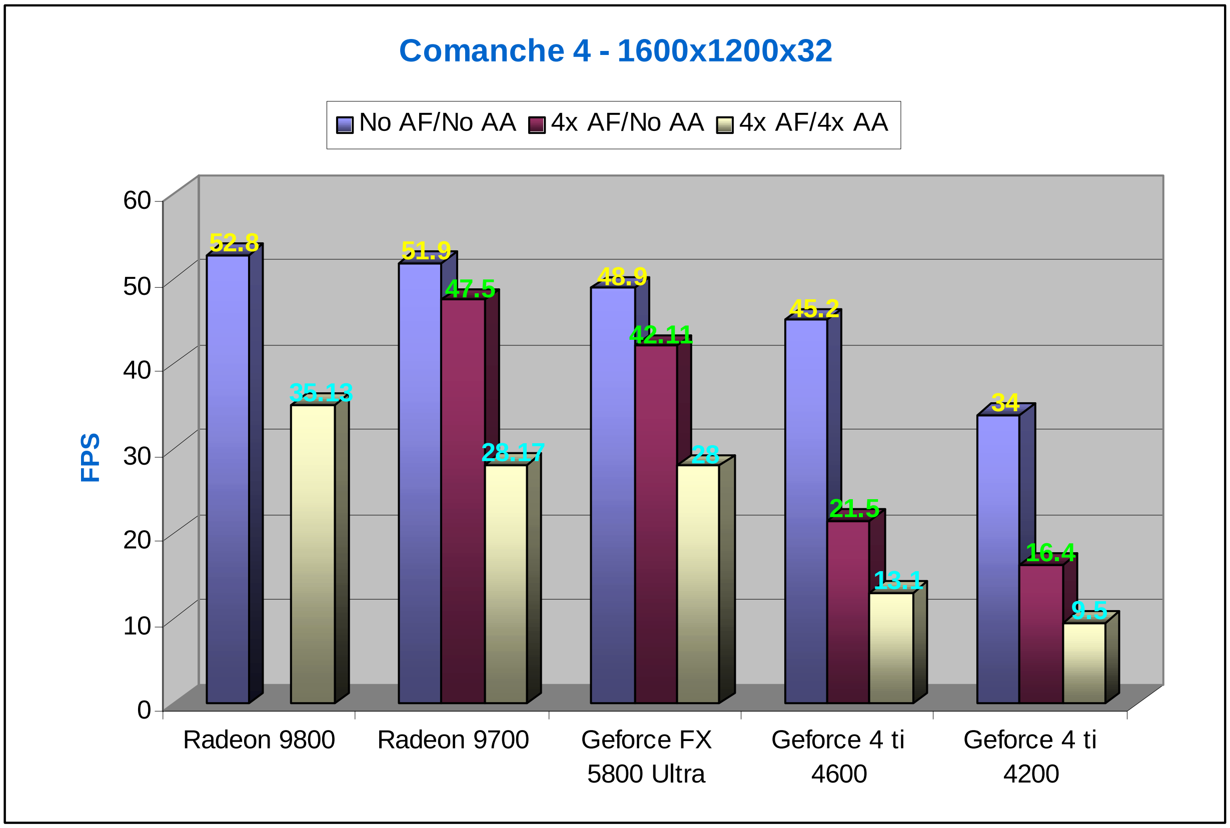
Shifting gears for a moment, let's take a look at the performance of the GeForce FX 5800 and Radeon 9800 Pro in a simulation game: Comanche 4. We start out at 1024x768x32, testing with the same AF/AA modes as before:



As a simulation, Comanche 4 is far more demanding on the CPU than what we saw previously with Unreal Tournament 2003. As a result, the framerates are much closer to each other, with only the GeForce 4 boards showing a significant decline in performance when 4x AF and 4x AA are enabled. 1024x768 is a bit low, however, so let's increase the resolution to 1280x1024:



Like before, benchmark results with 4x anisotropic filtering enabled with anti-aliasing disabled are now included to give us an idea of the performance drop associated with anisotropic filtering. Enabling 4x AF incurs only a minor performance hit for all modern cards in this benchmark. Enabling 4x AA as well causes performance to drop a bit more, with the Radeon 9800 Pro falling 9% from 52.4 to 48.1 FPS and the GeForce FX 5800 Ultra dropping 17% from 47.8 to 39.7 FPS. The Radeon 9800 Pro shows an improvement over its predecessor in this benchmark, as the Radeon 9700 Pro suffers an 18% drop when 4x AA is enabled in addition to 4x AF.

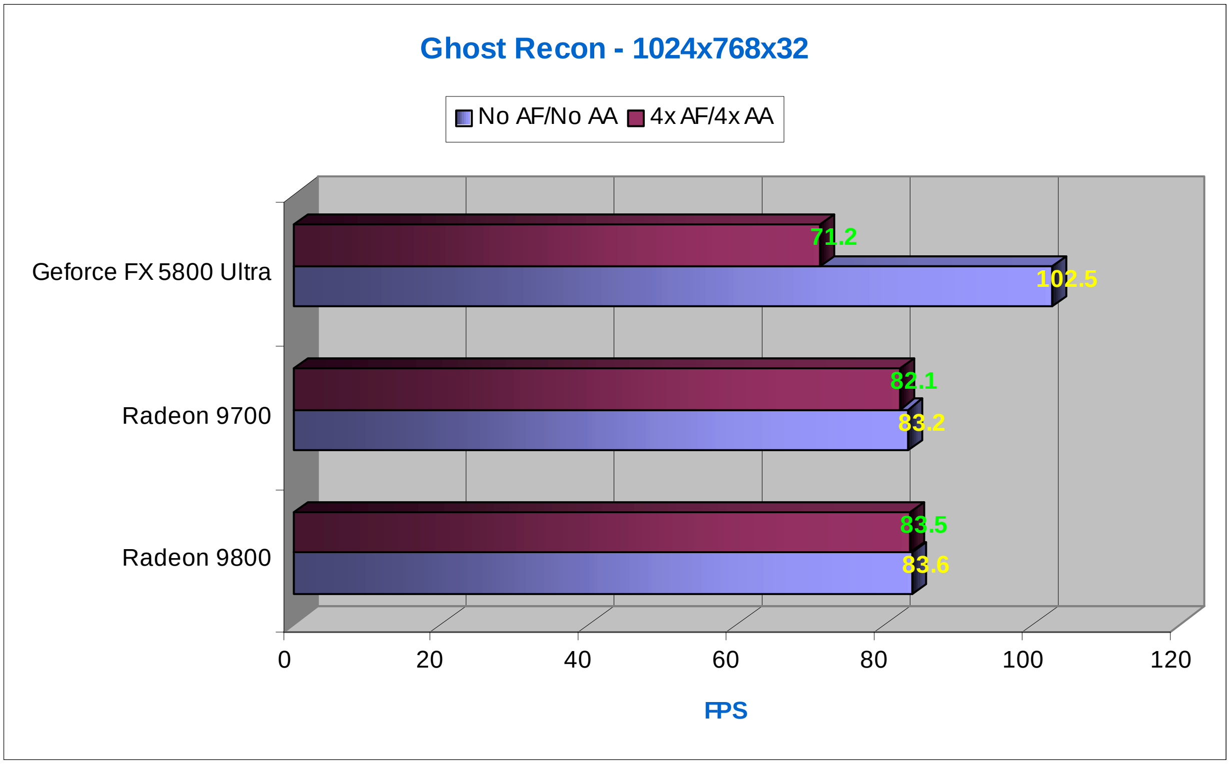


The picture looks very similar at 1600x1200, with the only notable drops in performance occuring with both anisotropic filtering and anti-aliasing enabled. Unfortunately, we don't have a result for the Radeon 9800 Pro with 4x AF enabled and AA disabled. However, we see that the GeForce FX 5800 Ultra performs very similarly to the Radeon 9700 Pro, both of which see roughly 40-45% drops when enabling 4x AF and 4x AA from running with both options disabled. This is an area where the Radeon 9800 Pro makes a slight improvement, as it is running about 7 FPS faster than its predecessor with 4x AF and 4x AA.

We've been testing in "Performance" mode for the Radeon 9x00, as noted earlier in the Benchmarked Configurations section. However, when testing with "Quality" mode on the Radeon 9700 Pro, the card delivers 44.5 FPS with 4x AF and no AA. This is roughly a 7% decline from the 47.5 FPS the R9700 Pro achieves under "Performance" mode. With 4x AF and 4x AA, the Radeon 9700 Pro's performance falls from 28.17 FPS to 23.4 FPS (17%), when enabling "Quality" mode.

Ghost Recon

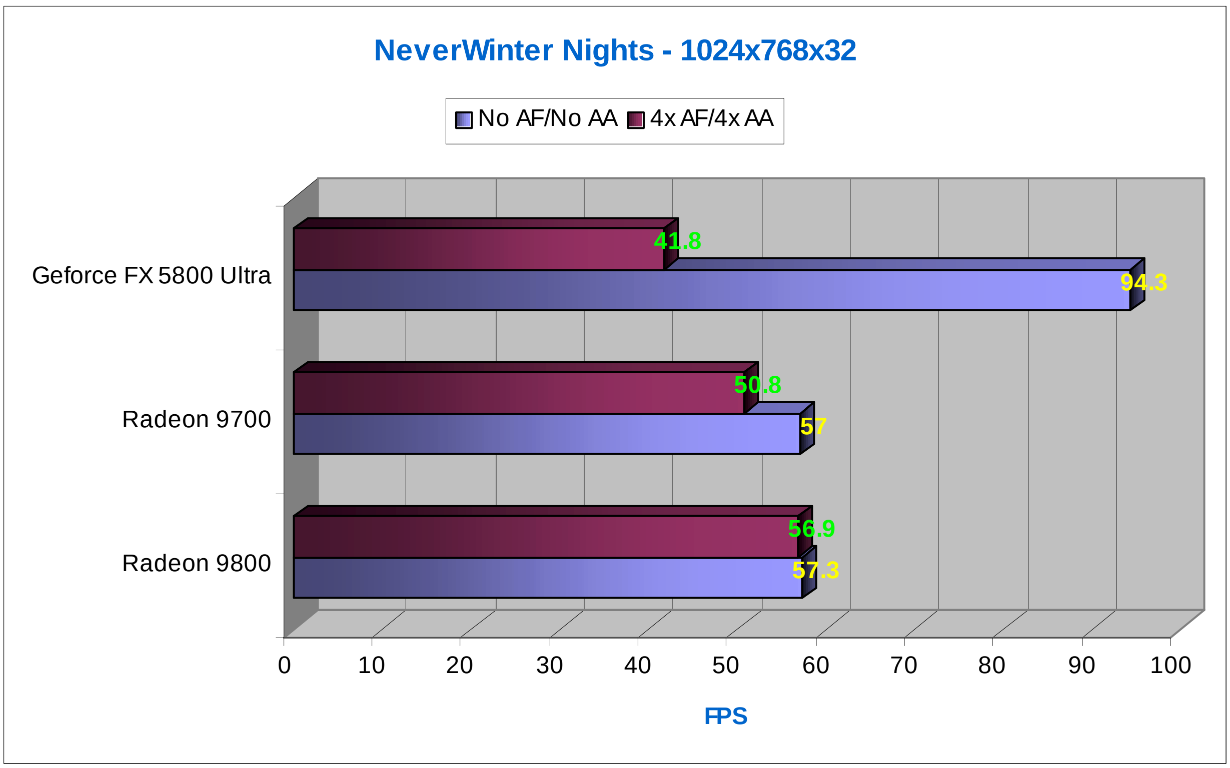
Let's switch back to first person shooters for a moment. Specifically, the tactical shooter Ghost Recon.



At last, the GeForce FX 5800 Ultra takes the lead, turning in a result that's 19% faster than the Radeon 9800 Pro with both anisotropic filtering and anti-aliasing disabled. Of course, the second mode -- 4x AF and 4x AA -- is far more likely to be the most common of the two (especially at 1024x768), and there the GeForce FX falls behind by about 15%.

NeverWinter Nights

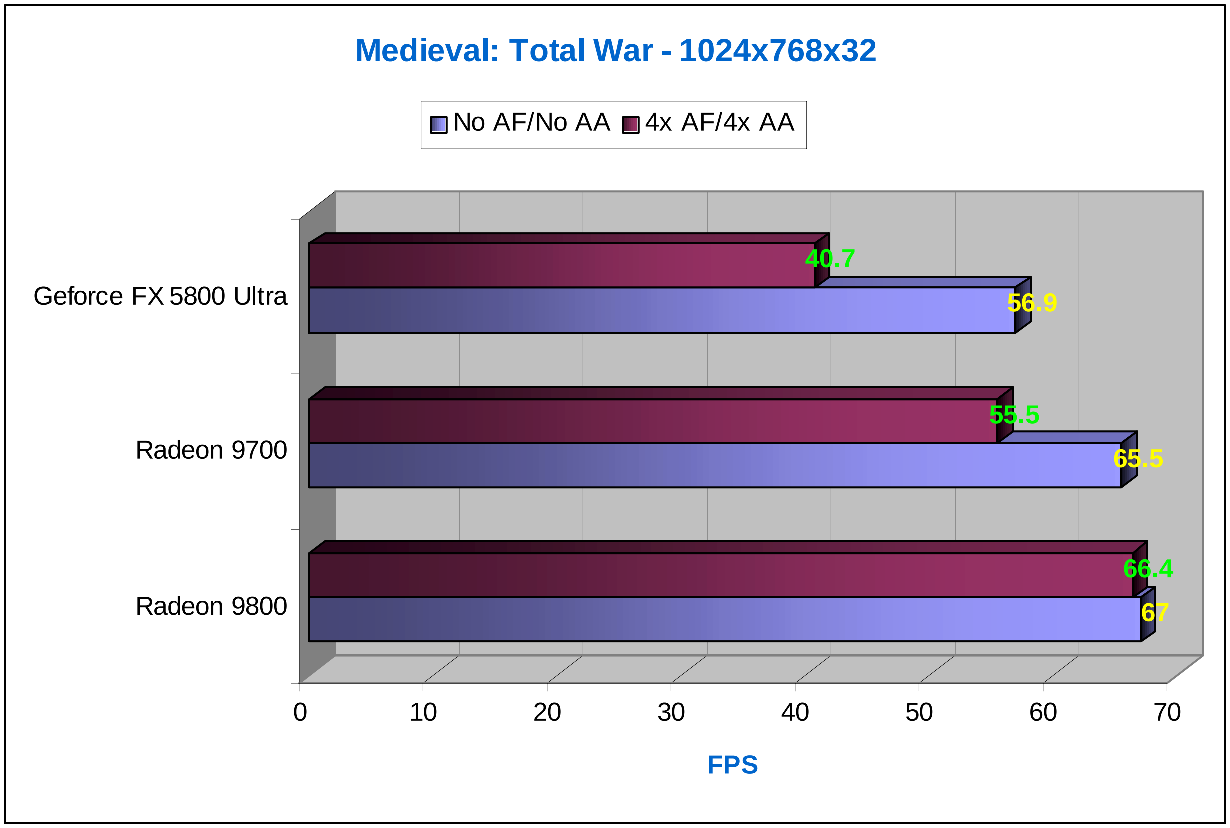
Let's take a look at RPG performance with NeverWinter Nights:



The GeForce FX surprises again with a significant 40% lead over the Radeon 9800 Pro in NeverWinter Nights with AF and AA disabled. Again, with 4x AF and 4x AA, the performance of the GeForce FX falls 56% to 41.8 FPS. Some explanation for the performance differential may lie in the game code, as there have been **reports** that performance is particularly bad on ATI GPUs with shadows set to maximum, and our **2003 Gamers' Hardware Upgrade Guide** showed similar results. A game patch may be able to resolve this issue, and we are looking into it. For the time being, however, these results indicate the GeForce FX is definitely the best solution for NeverWinter Nights.

Medieval: Total War

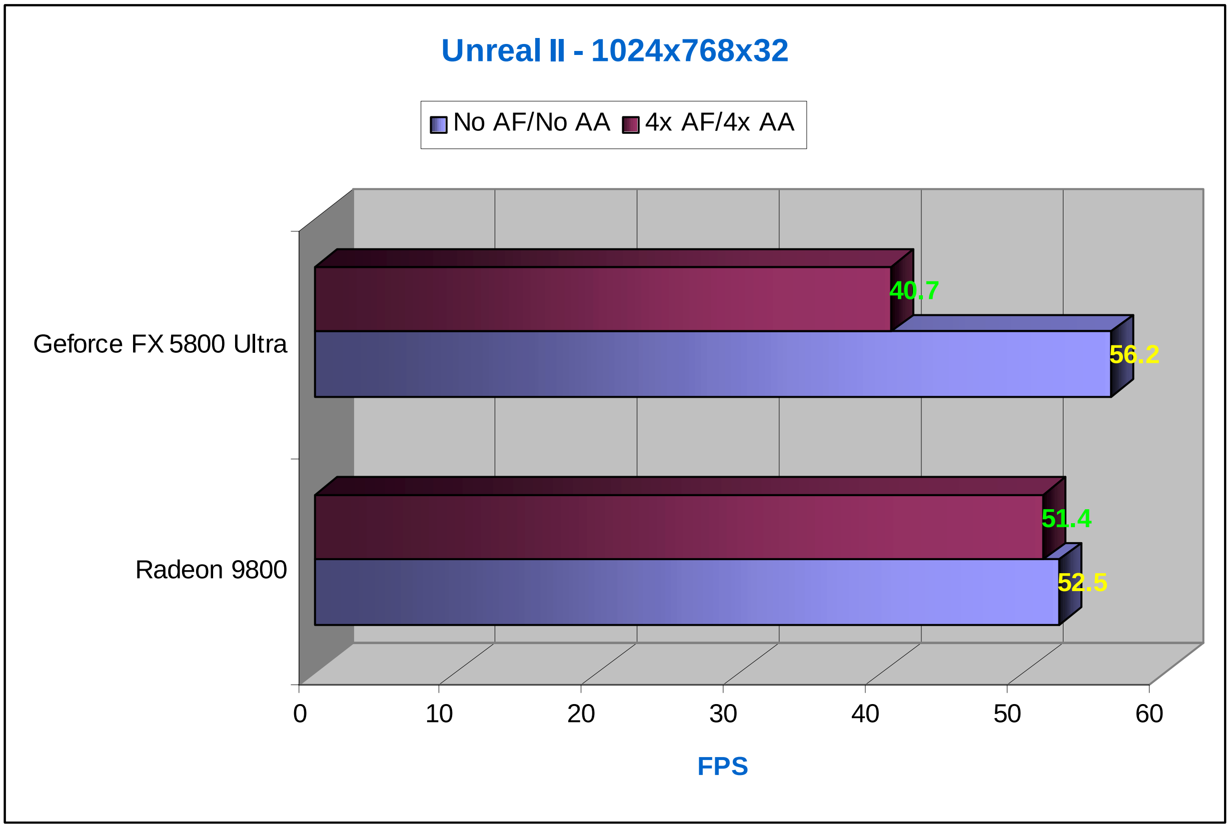
Like the rest of the Total War series (Shogun and the upcoming Rome: Total War), this game is a strategy game that often involves managing thousands of units at any given time. Besides exercising one's video card, the game also relies heavily on your CPU, as the AI opponent has to manage equally large forces.



As mentioned above, this is a very CPU intensive game, and as such even the highest frame rates fall under the default refresh rate of most monitors at this resolution. Despite leading in NeverWinter Nights and Ghost Recon, we find that the GeForce FX 5800 Ultra is falling behind in Medieval: Total War. The difference is about 10 FPS with AF and AA disabled, but with 4x AF and 4x AA, the GeForce FX trails the leader by 39%.

Unreal II

Next we return to first person shooters with Unreal II. The Swamp demo was benchmarked and World Texture Detail was set to "High," as was Skin Detail. Shadows were set to "Medium" and High Detail Texturing was enabled. FRAPS was used for the actual benchmarking, since we have had stability issues with the internal frame counter.



Unreal II presents another opportunity for the GeForce FX to take a slight lead, but again, only with both anisotropic filtering and anti-aliasing disabled. The performance impact associated with 4x AF and 4x AA on the GeForce FX is 28% in this test, while the drop on the Radeon 9800 Pro is practically within the margin of error.

Image Quality

One interesting new feature of 3DMark03 is the ability to capture a specific frame from any of the tests in the benchmark. Therefore, to compare the image quality between different modes and different GPUs, we can simply capture a particular frame and compare the differences. For the following image quality comparison, we have captured a frame from the Mother Nature test. A small area of this frame has been selected for comparison between several anti-aliasing and anisotropic filtering modes for both the GeForce FX 5800 Ultra and the Radeon 9800 Pro. Additionally, the lower left quadrant of this area has been magnified to provide a closer view (see the images under "Zoom"). Remember, as before, we're using "Performance" and "Balanced" driver options for the Radeon 9800 and GeForce FX, respectively.

3DMark03 - Mother Nature: Frame 1669 (0:55.66)
The area indicated in red is used for the image quality comparison below

	No AF, No AA	4x AF, No AA	4x AF, 4x AA
GeForce FX			
Radeon 9800			
GeForce FX Zoom			
Radeon 9800 Zoom			

FSAA Sample Patterns

Last year we had a discussion with David Naegle, senior engineer at Sun Microsystems working on the **SAGE architecture** (XVR-4000), regarding anti-aliasing and high-end visualization. SAGE is a topic for another article, but nevertheless, this discussion prompted me to investigate into comparing the sample patterns of different anti-aliasing implementations. A simple program, a **Direct3D FSAA Viewer** created by "**Colourless**" provides the means to do so.

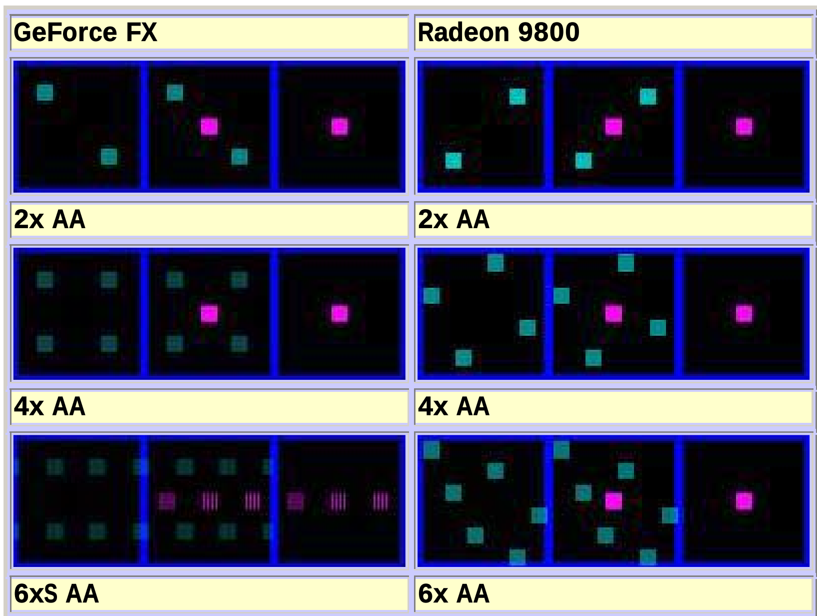
Below, you'll see a set of three blocks for each different AA mode. Within the blocks are the geometry and texture sampling patterns. Geometry sampling points are indicated in light blue, while texture sampling points are indicated in purple. As you can see, the first block shows the geometry samples, the third shows the texture samples, and the second shows the combined sample pattern.

To provide a quick overview: there are two common anti-aliasing methods implemented by today's video cards, supersampling and multisampling. Supersampling is an obvious form of anti-aliasing that works by rendering to a higher resolution and then resampling down to the target resolution. Supersampling requires a number of texture samples that is equivalent to the degree of antialiasing, so 4x supersampling AA would require four texture samples per pixel. This can make supersampling expensive in terms of performance.

Conversely, multisampling works by using a single texture sample repeatedly for all the samples in a pixel. This makes it significantly faster and more efficient than supersampling at the expense of sacrificing texture quality. So, in the case of 4x multisampling, you only need one sample for final color value of the pixel. While this method is fast and is commonly used, as you will see below, it does rely more on texture filtering for texture quality (as opposed to just the edges) than supersampling.

With the exception of the GeForce FX's 6xS mode, multisampling is used exclusively for all the anti-aliasing modes shown below. So, you will see one texture sample and a number of geometry (edge) samples depending upon the degree of anti-aliasing in use.

Considering that higher quality texture filtering implementations are becoming more and more common and efficient, with 4x anisotropic filtering used throughout this review, multisampling in tandem with anisotropic filtering can result in a very nice looking image.



Looking at this comparison, the differences between implementations emerge quickly. Both the GeForce FX and Radeon 9800 use a rotated grid for 2x FSAA. In the case of 4x AA, however, the GeForce FX uses an ordered grid, as opposed to the rotated one used by the Radeon 9800.

The sample pattern of the 6xS mode of the GeForce FX may be difficult to see, so you may want to click on it to view the full image. Nevertheless, it shows six geometry samples and three texture samples, as the 6xS mode is a combination of both mutisampling and supersampling. The Radeon 9800's 6x mode, on the other hand, is multisampling with an irregular sampling pattern, resulting in increased image quality.

Final Thoughts

After reviewing the benchmark results, it becomes clear that while the GeForce FX 5800 Ultra can equal the Radeon 9800 Pro in some benchmarks (and even beat it in a few with both anisotropic filtering and anti-aliasing disabled), it falls behind its competition in at least as many others and never manages to outperform the Radeon 9800 Pro in any of the benchmarked games with both 4x AA and 4x AF enabled.

Architecturally speaking, the GeForce FX is very interesting. Despite all the specifications and performance figures, we still don't know precisely how it is organized internally, hence the 8x1/4x2 pipeline issue. The R300/R350 core, on the other hand, is a known quantity with performance that is very much expected from looking at its specs. Perhaps the GeForce FX will improve as drivers and applications are optimized to better suite the architecture, or visa versa.

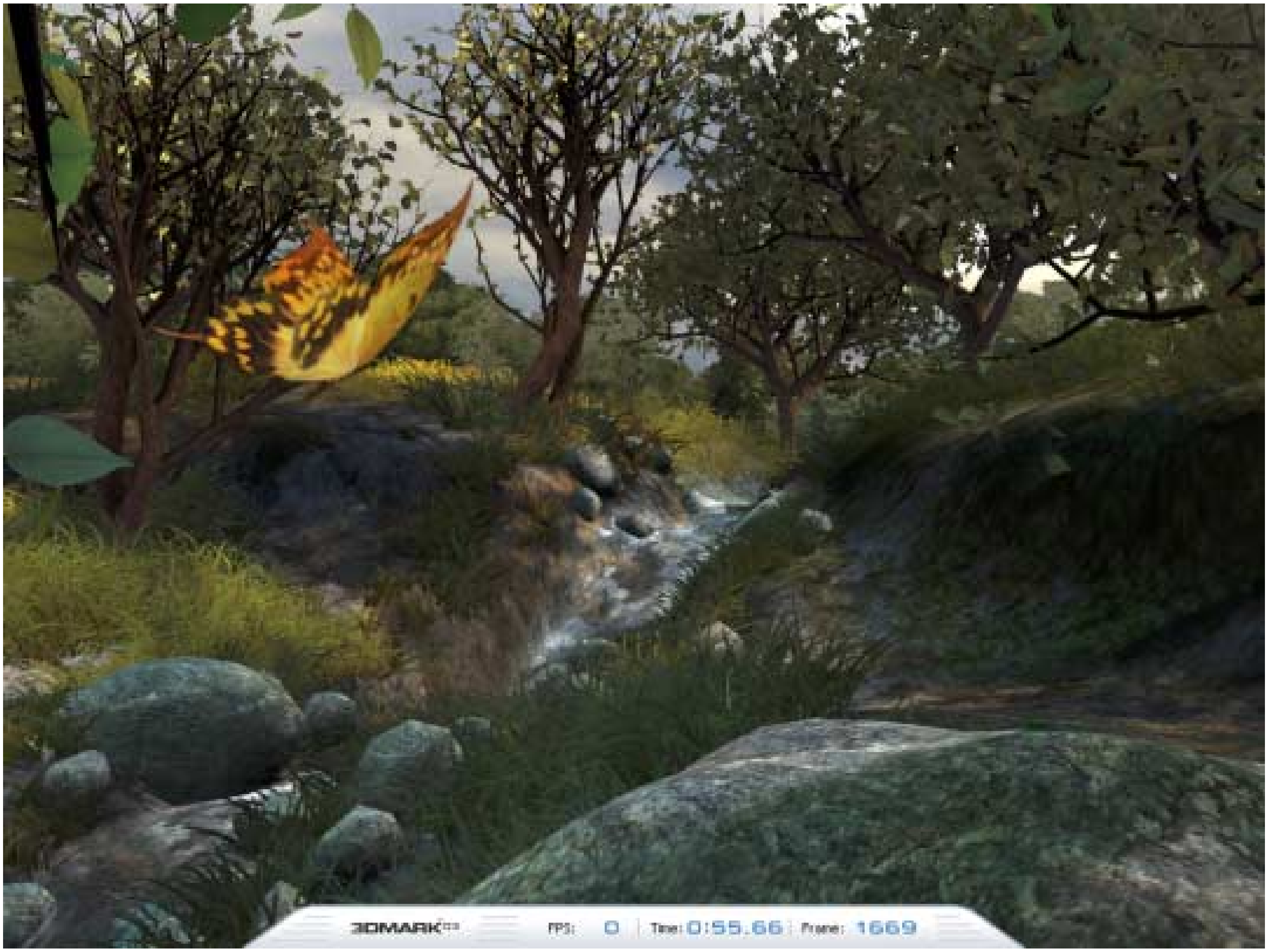
The performance of the GeForce FX is quite respectable in relation to the GeForce 4, so perhaps the real surprise is not the NV30, but rather ATI's R300/350 core, and the impressive performance it has demonstrated thus far. The Radeon 9800 Pro makes incremental performance improvements over the Radeon 9700 Pro thanks to its increase core and memory clockspeeds. With its $399 price tag, the Radeon 9800 Pro is unlikely to entice current Radeon 9700 or GeForce FX owners, though it should provide a significant upgrade for GeForce 4 and Radeon 8500 owners.

Boasting vertex and pixel shader specifications that exceed DirectX 9 VS and PS 2.0, the GeForce FX is likely the most future-proof GPU currently on the market in terms of feature set. Where it's not so future-proof is the performance. NVIDIA claims the GeForce FX is at its best when running more advanced applications like the upcoming DOOM3. The Radeon 9800 Pro, on the other hand, delivers better performance for today's games, and should deliver solid framerates well into the future.

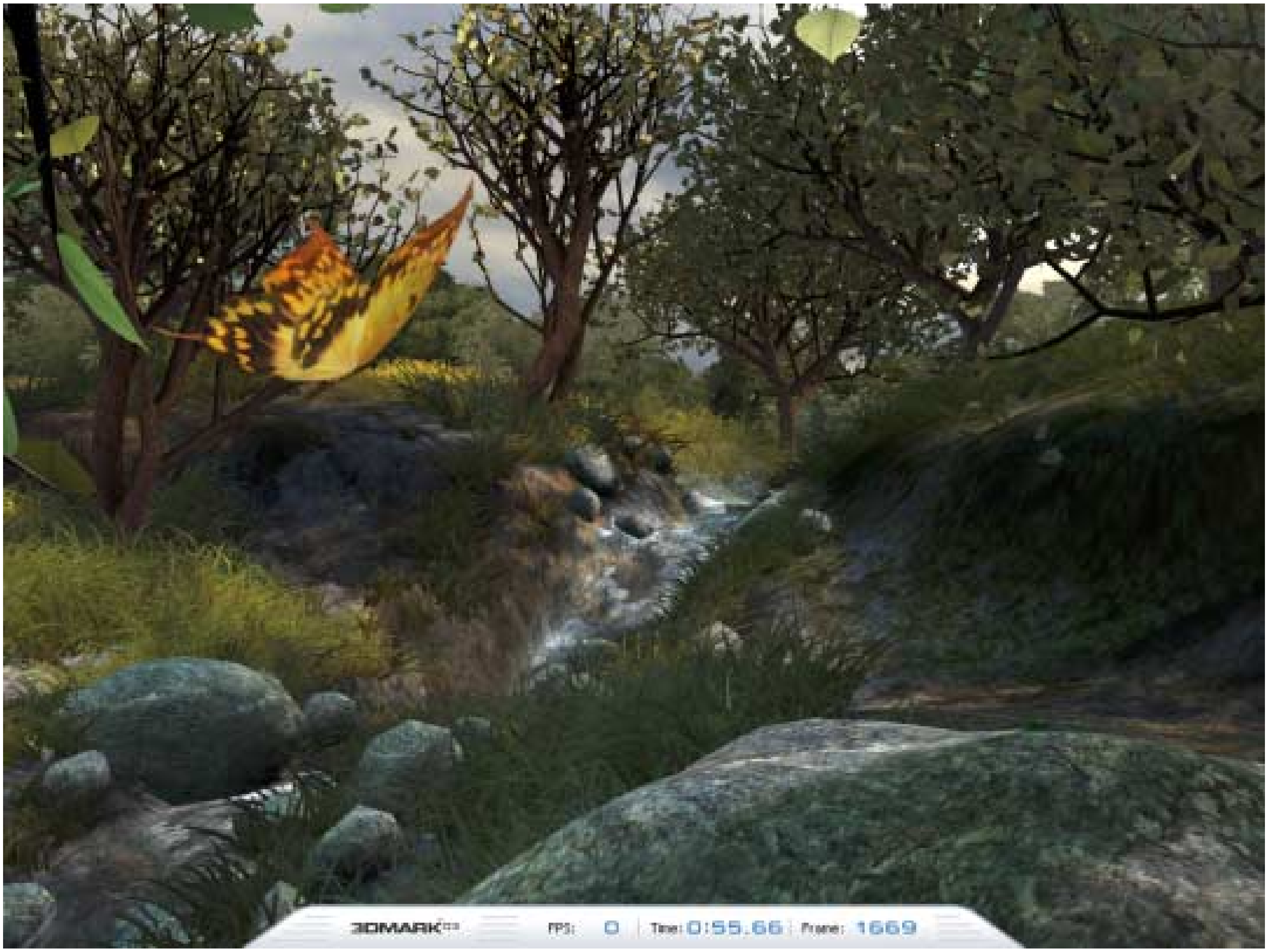
Appendix A

Image Quality Comparison Images

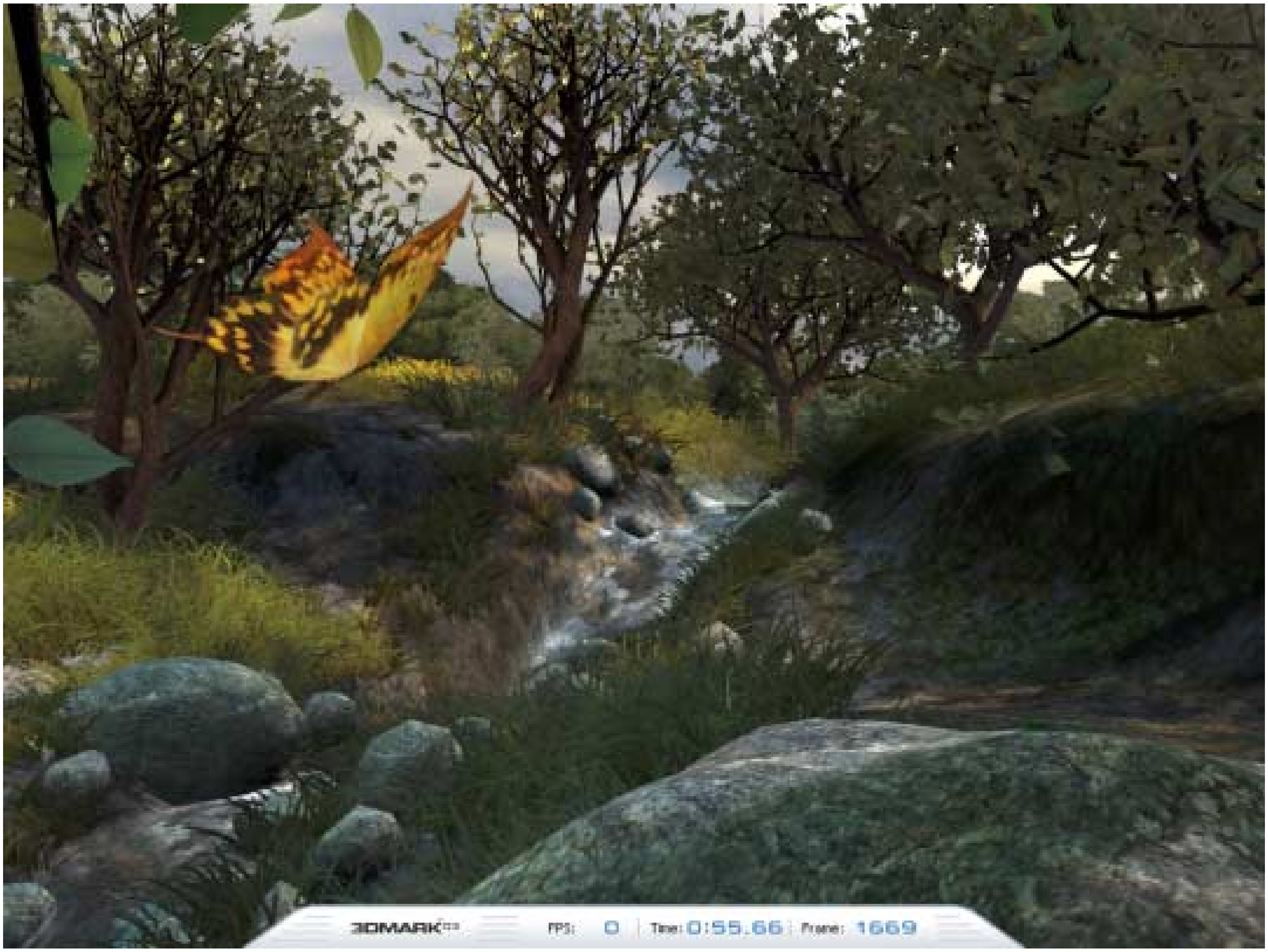
GeForce FX – No AF, No AA



Radeon 9800 Pro – No AF, No AA



GeForce FX - 4x AF, 4x AA



Radeon 9800 - 4x AF, 4x AA