What we will cover

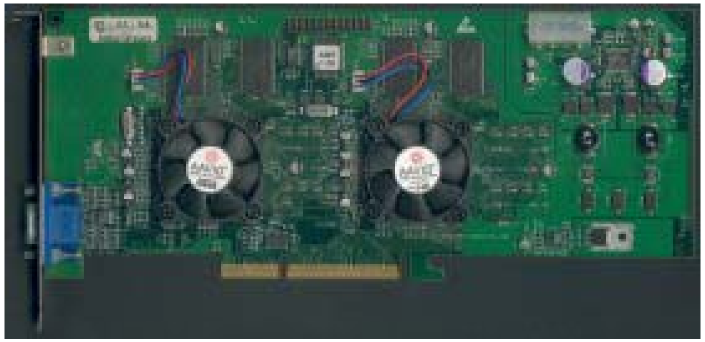
- Contour Tracking
- Surface Rendering
- Direct Volume Rendering
- Isosurface Rendering
- Optimizing DVR
- Pre-Integrated DVR
- Splatting
- Unstructured Volume Rendering
- GPU-based Volume Rendering
- Rendering using Multi-GPU

History of Multi-Processor Rendering

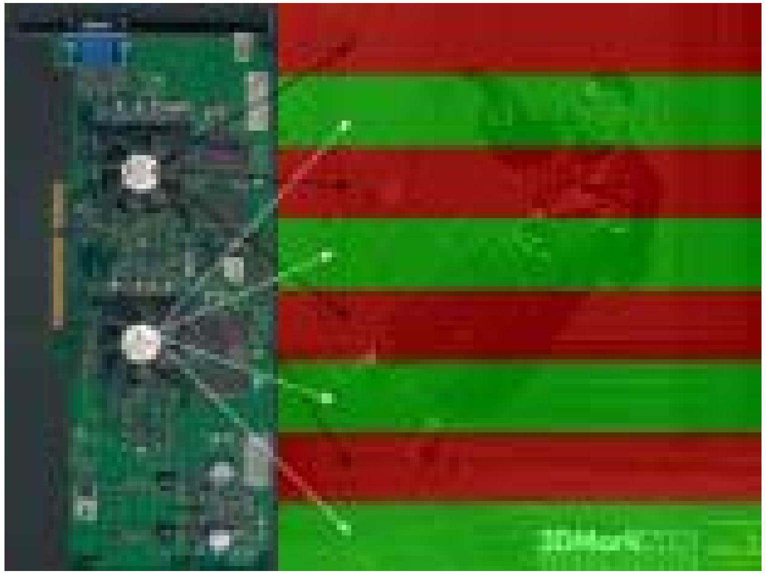
- 1993 SGI announced the Onyx series of graphics servers
  - the first commercial system with a multi-processor graphics subsystem
  - cost up to a million dollars
- 3DFx Interactive (1995)
  - introduced the first gaming accelerator called 3DFx Voodoo Graphics
  - The Voodoo Graphics chipset consisted of two chips: the Frame Buffer Interface (FBI) was responsible for the frame buffer and the Texture Mapping Unit (TMU) processed textures. The chipset could scale the performance up by adding more TMUs – up to three TMUs per one FBI.
- early 1998 3dfx introduced its next chipset, Voodoo2
  - the first mass product to officially support the option of increasing the performance by uniting two Voodoo2-based graphics cards with SLI (Scan-Line Interleave) technology.
  - the cost of the system was too high
  - 1999 NVIDIA released its new graphics chip TNT2 whose Ultra version could challenge the speed of the Voodoo2 SLI but less expensive.
  - 3DFx Interactive was devoured by NVIDIA in 2000

Voodoo

Voodoo 2 SLI



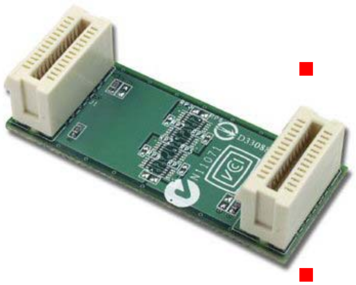
Voodoo 5 5500 - Two Single Board SLI



Voodoo 5 SLI Rendering

History of Multi-Processor Rendering

- ATI Technologies(1999)
  - MAXX –put two chips on one PCB and make them render different frames simultaneously and then output the frames on the screen alternately.
  - Failed by technical issues (sync and performance problems)
- NVIDIA SLI(200?)
  - Scalable Link Interface
  - Voodoo2 style dual board implementation
  - 3Dfx SLI was PCI based and has nowhere near the bandwidth of PCIE used by NVIDIA SLI.
  - 3Dfx SLI uses scan line interleaving (evenly splits rendering duties between the two cards) while NVIDIA has on the fly dynamic load balancing using 2 different render modes: Split Frame Rendering (SFR) and Alternative Frame Rendering (AFR).
  - instead of the old external cable, the connection is internal in form of a small circuit board, known as a bridge

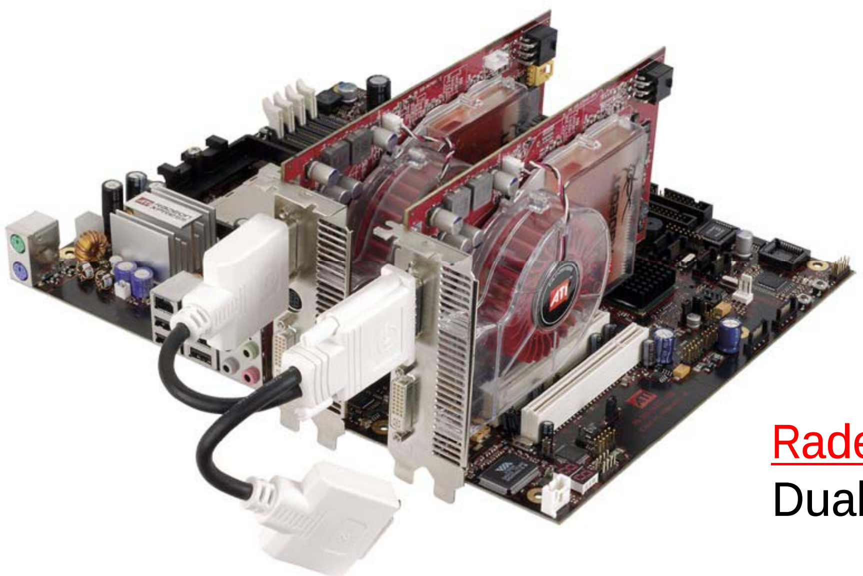


NVIDIA SLI

- An NVIDIA SLI system includes:
  - PCIE motherboard that supports two x16 physical connectors
  - Graphic boards with the exact same model number and from the same vendor. (Currently, SLI support is only for the PCIE versions of GeForce 6800 Ultra, GeForce 6800 GT, GeForce 6600 GT, or GeForce 6800)
  - Adequate Power Supply (500 Watts to be safe)

ATi's multi-GPU solution: CROSS FIRE

- CrossFire is ATI's response to nVidia's SLI technology. Basically it allows two (or more, in theory) PCI-Express graphics cards to be installed in one machine, doubling performance under ideal conditions.
- ATI will make its SLI work without a small printed circuit board (PCB) to interconnect the cards. SLI uses two PCIe graphic ports and in the case of Nforce 4, works at PCIe 8X for each card and renders your picture with both cards.



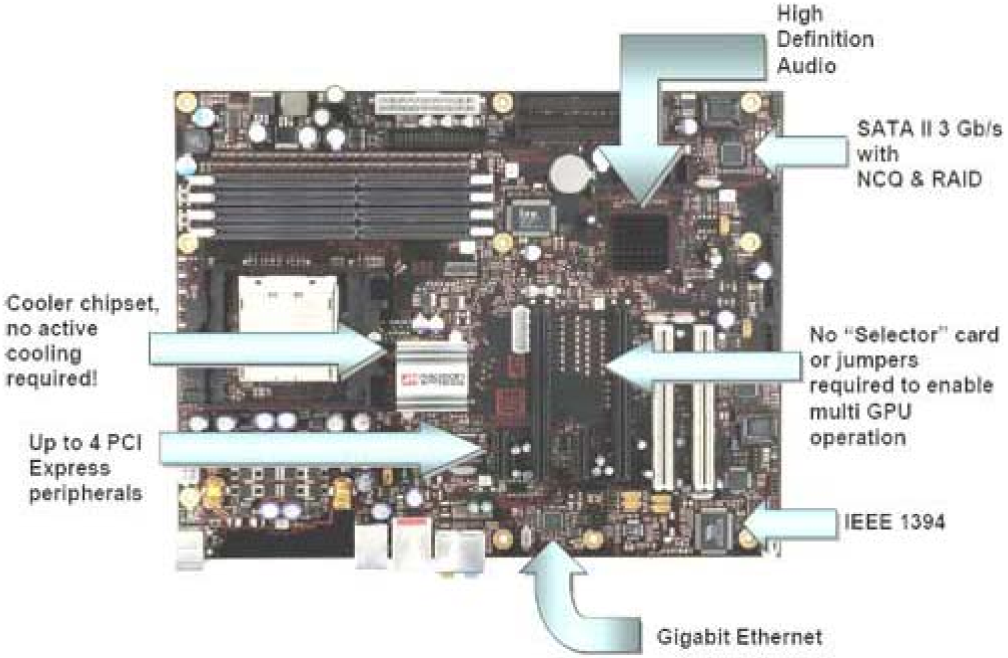
Radeon X850 XT x 2:
Dual X850s Installed

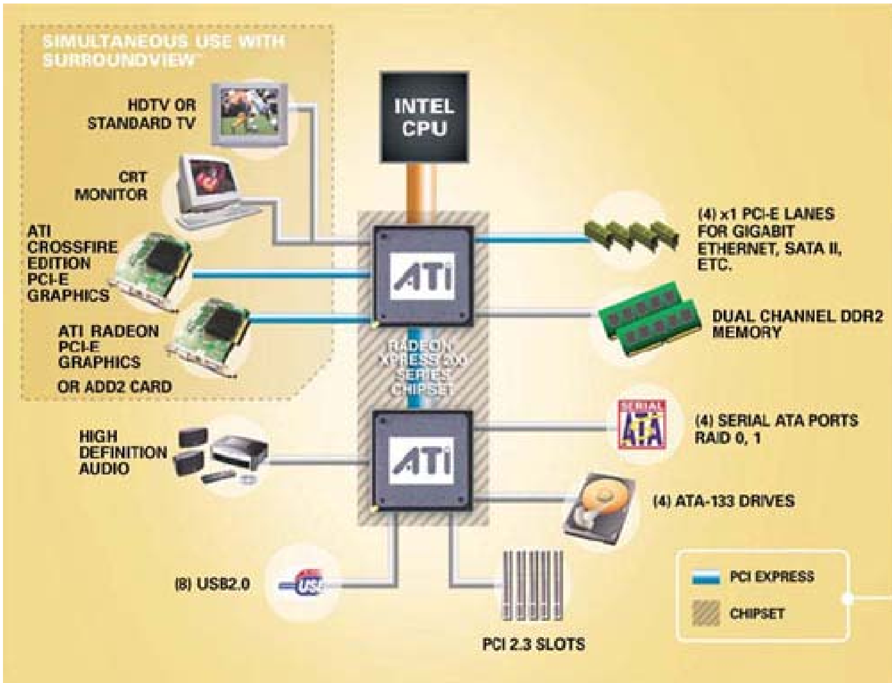
ATi's CrossFire Solution

- A CrossFire-ready ATi Radeon Xpress 200 Motherboard
  - PCI Express 16x slot
  - No manual switch on the boards to choose between single and dual video card modes
- CrossFire Edition video cards
  - Master and Slave card system
  - a Radeon X800 or X850 series Graphics card.
  - CrossFire Edition master cards house an extra chip, containing a programmable compositing engine which controls the output from the GPUs and sequences how they are shown on screen, blends them, merges them etc as is required of it. The master card can then display the final image on screen after the required compositing has been performed.
  - When a master and slave card are matched up using differing numbers of pipelines or differing quantities of RAM, the lowest common denominator will be used (except clock speed)

The Radeon X850 XT

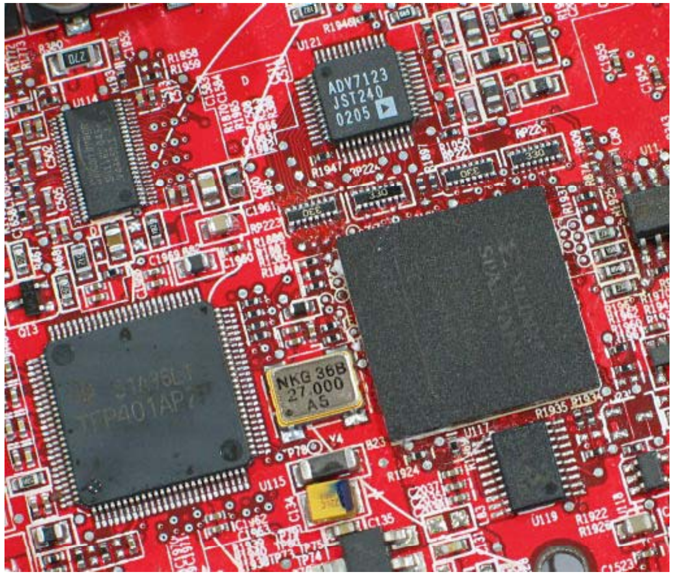
Platform of ATi CrossFire Edition Motherboard





Platform of ATi CrossFire Edition Motherboard

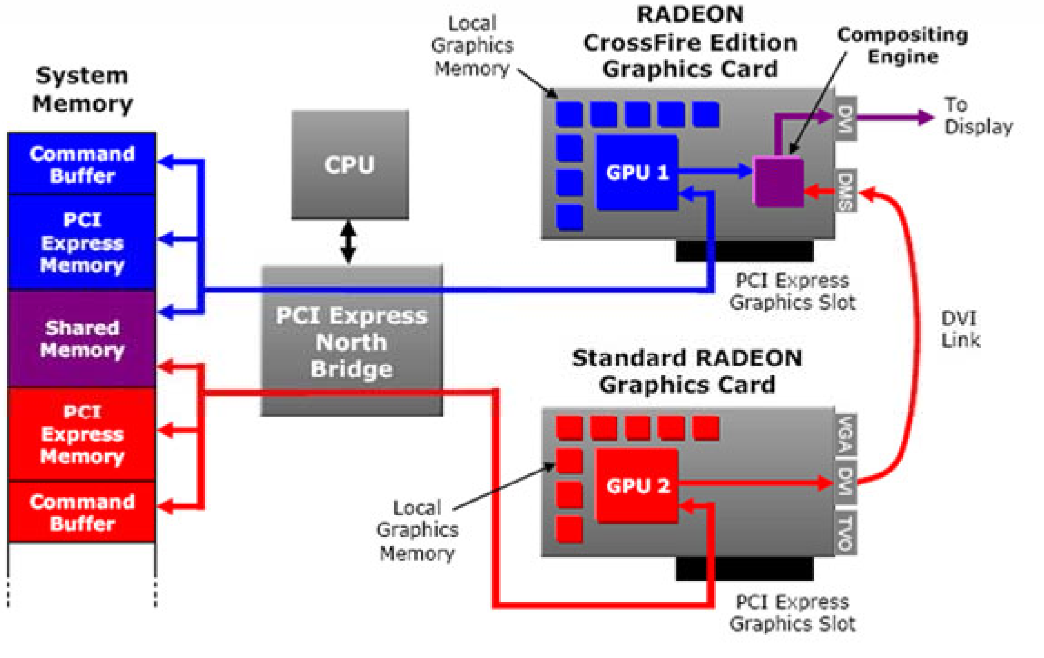
- CrossFire is a pretty slick scheme, really, given the limitations imposed by the original X800 hardware. ATI equipped its master cards with five new chips, pictured above. The second largest of the five chips there is a TMDS receiver made by Texas Instruments. To make up for the lack of a dedicated digital interconnect between GPUs, the master card can intercept and decode the DVI output of the slave card using this receiver. Next to it, the largest of the chips is a Spartan-series FPGA chip from Xilinx.
- FPGA stands for Field Programmable Gate Array, which is a fancy way of saying that this is a programmable logic chip. In this case, ATI has programmed the Xilinx FPGA to act as CrossFire's compositing engine, tasked with combining the images generated by the two Radeon GPUs into a single stream of video frames.



- The smaller chip just below the FPGA in the picture is a flash ROM; presumably, it holds the programming for the FPGA. Once the images from the slave card have been decoded by the TMDS receiver and composited with the images from the master card's GPU by the FPGA, they have to be output to a display. That's where the last two chips come into the picture.
- The little, square chip above the FPGA is a RAMDAC chip made by Analog Devices. The RAMDAC converts digital video information for output to an analog display, such as a VGA monitor. Just above the TDMS receiver is a smaller, rectangular chip from Silicon Image. That's a TMDS transmitter capable of encoding images for output to a digital display via a DVI output.
- All together, these five chips add the necessary functionality to ATI's master cards to allow a pair of graphics cards to run together in an SLI-like configuration with very little performance penalty for inter-chip communication or image compositing.

Rendering using CrossFire

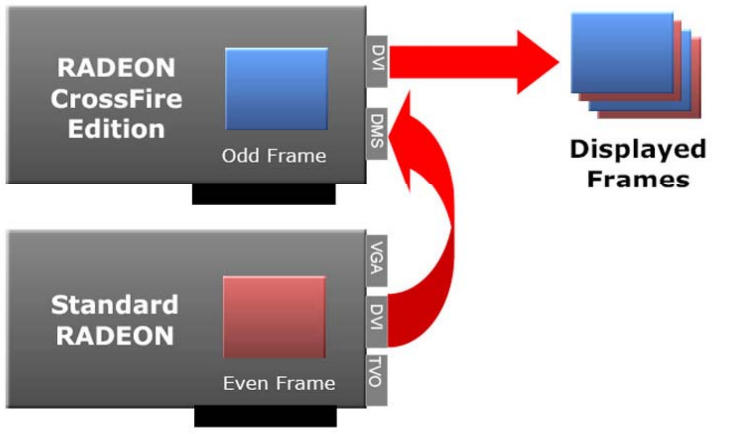
- How a CrossFire system communicates and allocates memory
  - The driver assigns a shared pool of system memory to assist with synchronization and allow access to shared textures and the like.
  - The RAM on each video card is then left to that local board to use and handle.



CrossFire Block Diagram

Rendering using CrossFire

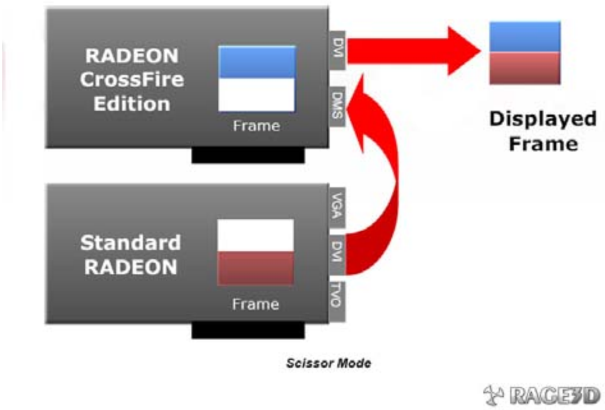
- AFR (Alternate Frame Rendering)
  - Each GPU renders every other frame
  - The advantage of this method is that all facets of the rendering process are boosted through the use of multiple GPUs
  - The downside of AFR is that if data is rendered and then held in memory to be used for a render-to-texture effect such as motion blur, then rendering suffers a performance penalty as this stored data needs to be passed between the two GPUs.



Alternate Frame Rendering (AFR) Mode

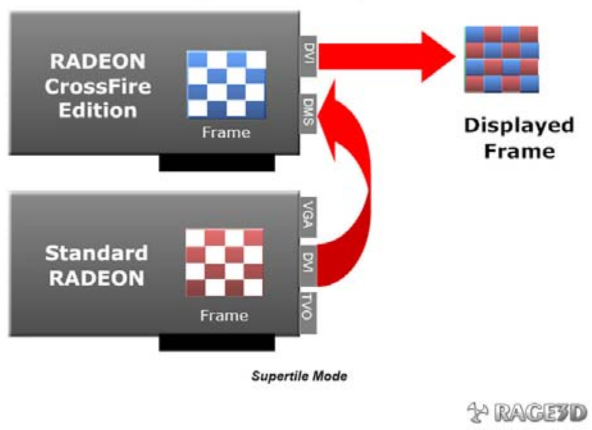
Rendering using CrossFire

- Scissor mode
  - NVIDIA calls this mode "split-frame rendering,".
  - Each frame is taken and cut in two either horizontally or vertically, and each GPU is given one portion of the frame to work on.
  - In OpenGL applications, the split between the frames is static at 50% per card. Direct3D applications get dynamic load balancing, with the split between the cards varying on a frame-by-frame basis.
  - All of the geometry set up has been performed before the scene is split between both GPUs, and thus geometry performance is not enhanced as it is when using AFR.



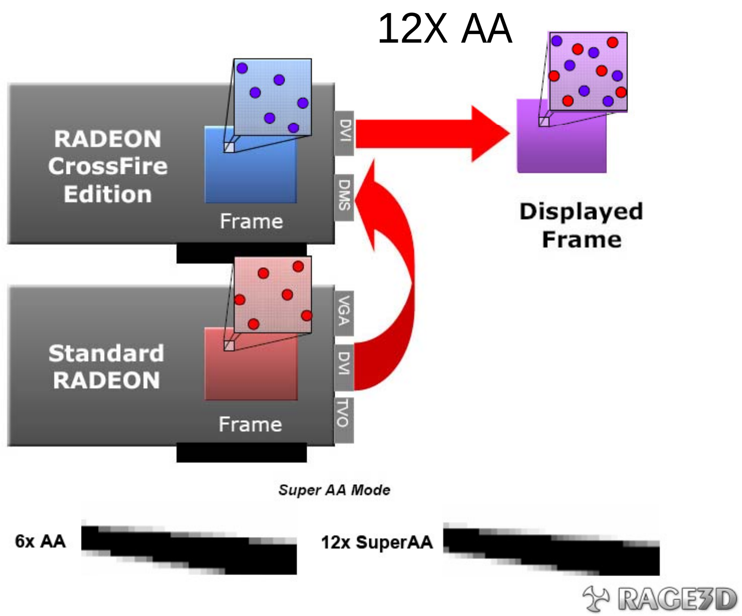
Rendering using CrossFire

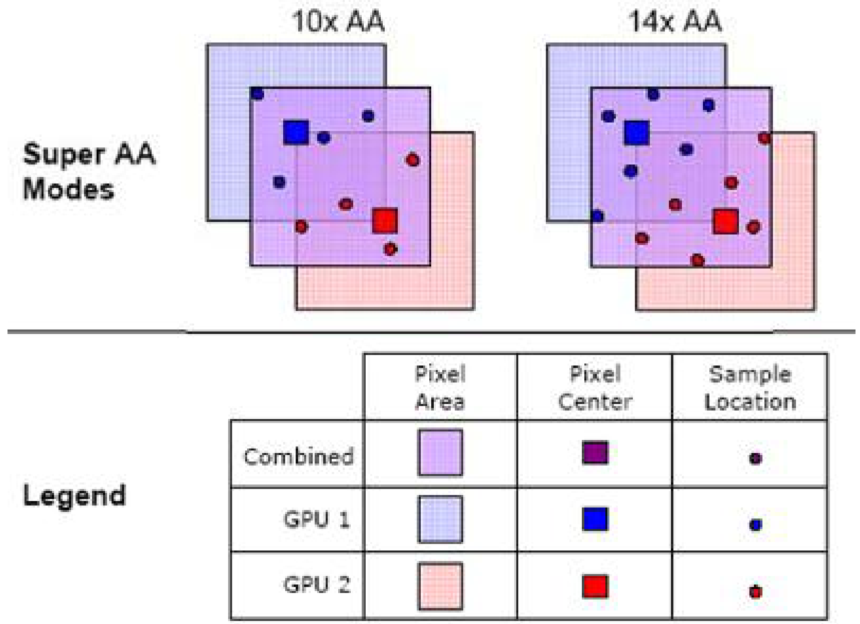
- Super-tiling
  - Supertiling breaks the frame up into a number of very small tiles (32 x 32 pixels large) and splits the rendering of these tiles between the two video cards.
  - As complicated objects or effects will span multiple tiles in such a system, the load on each GPU is thus effectively balanced without the need for any special algorithms.
  - As with Scissor mode, Supertiling requires all of the geometry calculations to be performed before the scene is split up into tiles.
  - ATI patented technology
  - This method is the default for Direct3D applications, SuperTiling is not supported in OpenGL.



ATi Super AA Modes

- ATI's Radeon Xx00 GPUs can perform up to 6x multi-sampling, compared to the maximum of four samples supported by NV4x.
- Effective antialiasing.
- CrossFire allows for the exact same scene to be rendered on both cards, but using different anti-aliasing sample patterns on each board. The two final outputs and then merged by the compositing engine, effectively giving larger quantities of AA with no real performance impact at all.

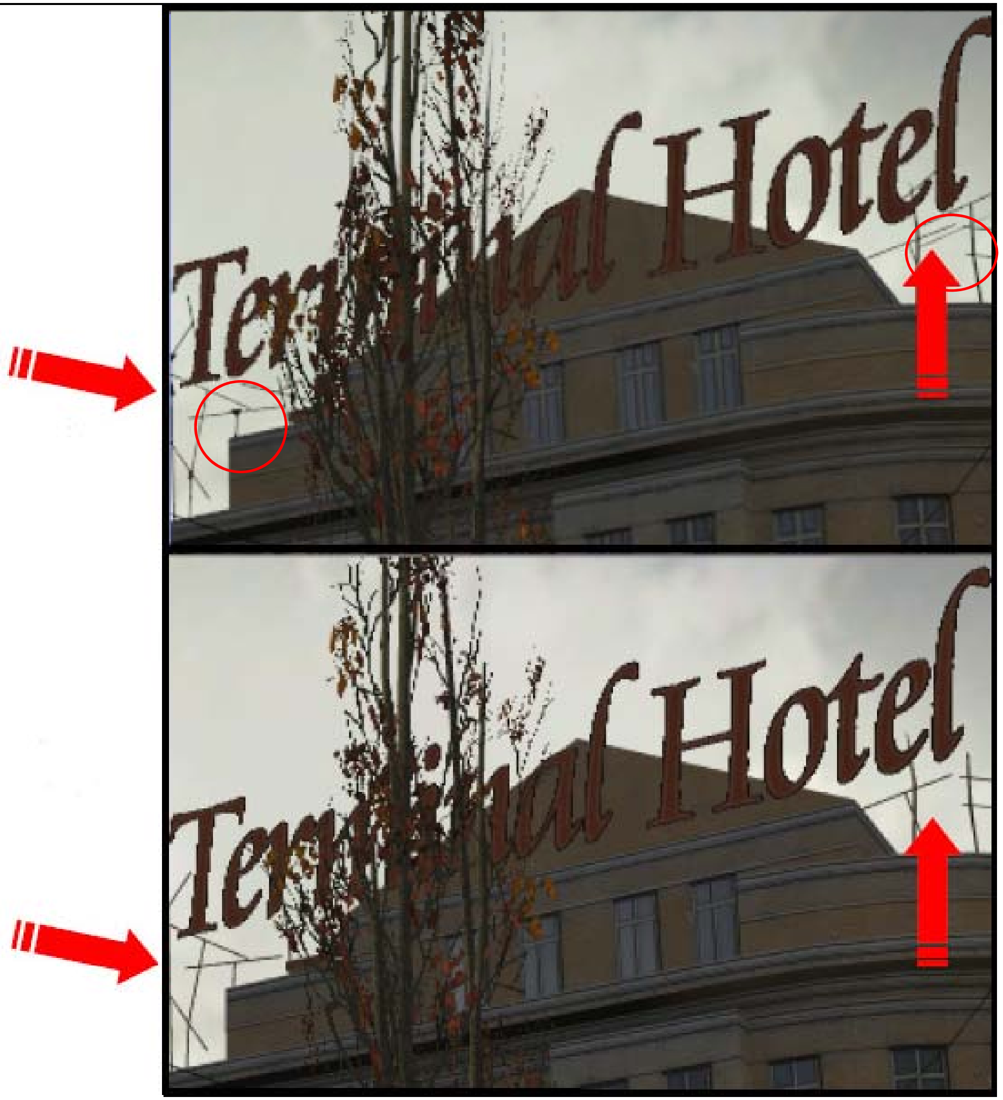




ATi Super AA Modes

ATI
14X AA

nVidia
8X AA



CrossFire vs. SLI

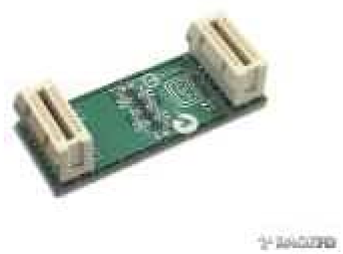
- CrossFire differs from SLI in a few areas. For SLI you need to use two identical cards. You can't, for example, link a 6800GT and a 6600GT. While CrossFire is considerably more flexable than SLI, it does have some limitations as well.
- The first is that you can't just slap any two cards in your machine and expect CrossFire to work. As the rumour mill has been suggesting for the last few weeks, there's a "Master" card that you need to install in the first PEG slot. To make things simple, ATI is calling these "CrossFire Edition" cards.
- The next limitation is that for CrossFire you need to match up cards by their "family" for it to work. For example, you can couple an X850XT CrossFire Edition card with an X850XT-PE, X850XT, or X850Pro or you can couple an X800 CrossFire with an X800XT-PE, X800XT, X800XL, X800 Pro, or X800. You can't, however, couple an X850 series card with an X800 series card. I expect the reason for this has more to do with marketing than it has to do with any technical limitations.

CrossFire vs. SLI

- It looks like that right now ATI is planning three different CrossFire Edition models, the X850XT CrossFire ($549), the X800 CrossFire 256MB ($299), and the X800 CrossFire 128MB ($249). That makes for quite a few combinations to choose from and, with regard to the X800 CrossFire choices, makes the entry price somewhat reasonable for most enthusiasts. Hopefully there will be even more CrossFire Edition cards to choose from in the future.
- Right now ATI is only focusing on ATI and Intel motherboards for CrossFire. They say that other motherboards, presumably nVidia nForce4 SLI boards, will be evaluated and qualified down the road a bit.

CrossFire vs. SLI

- Another area where CrossFire and SLI differ is in the way they are setup. nVidia's SLI uses an optional internal SLI connector that you need to install to get the most performance out of the setup (the bridge comes bundled with SLI motherboards). In most cases the SLI connector is a small PCB which "bridges" the gap between two SLI enabled cards, providing a high speed connection which the cards use to transfer data.
- ATI's CrossFire, on the other hand, uses and external dongle that goes from the DVI port on the Slave card to the DMS-59 port found on the CrossFire Master card.
- For comparison purposes, here are photos of nVidia's SLI connector, an SLI setup, and ATI's CrossFire setup.



nVidia's SLI Link

nVidia SLI Installation

ATI CrossFire Installation

Multi Graphics Processor Problems

- Render to Texture
  - a process that rather than rendering to a displayable buffer, pixels (colour values) are rendered to a texture surface ready for use later.
  - Render to texture operations can be problematic to multiple graphics rendering solutions due to both spatial and temporal issues. When a render to texture operation is occurring it is not necessarily known where in the final rendered image that texture will ultimately be used, so in the case of a Split Screen multi-graphics rendering implementation either one board performs the render to texture operation and then the results are shared over the SLI or PCI Express connection as required by the other board, or both boards can perform the render to texture operation. Some render to texture operations, such as the aforementioned Motion Blur effect, will perform the render to texture operation on one frame and then store the results for one or more frames after - this type of render to texture operation is not optimal for Alternate Frame Rendering since each board is rendering alternate frames and the rendered texture data will need to be passed from one board to the other on every frame.

Multi Graphics Processor Problems

- The eventual upshot of Render to Texture operations will ultimately be a reduction in performance. In most cases this is likely to manifest itself in the boards never really attaining their theoretical doubling of performance when Render to Texture is used, but potentially it could nullify the effects of SLI entirely depending on how it is used - in pathological theoretical cases it has the potential to make the create a performance *reduction* in relation to rendering with a single board if the texture is used extensively and is needed to be addressed from one board to the other. Render to Texture, and their uses in games, is likely to be the one of the primary causes of having two different rendering mode for NVIDIA's SLI operation such that if a title uses it in a fashion that is not desirable for one mode then at least there is another that may work. However, it is difficult to automatically know how or where a title is going to use a Render to Texture operation.

Multi Graphics Processor Problems

- Another point that has remained the same for all the multi-chip or board implementations seen in the retail PC market so far is that of memory scaling. Again, because the communication capabilities between multiple boards, or even multiple chips on a single board, usually has many times lower bandwidth than each graphics chip's own local memory, to address any static texture information from one board/chip to another during a frame would be prohibitive in terms of performance, hence static texture information is duplicated within the local memory subsystems to each graphics chip. The net result is that where the theoretical peak fill-rate, and vertex rate dependant on the mode, can double when two boards are added the memory space available to the graphics subsystem doesn’t scale similarly.

Multi Graphics Processor Problems

- 3dfx’s Voodoo 2 implementation did manage to save some memory space for the frame buffer as each board was only storing half the lines and the joining of the images was literally done at the screen; even the digital SLI of Voodoo 5 merged the images at scan-out level so that half the frame buffer space was saved on each of the two chips with the two-chip Voodoo 5 5500. With the SLI system here it's unlikely that there are any frame buffer space savings for the colour buffers as the AFR mode will need a full frame buffer as each board is rendering a full frame; the SFR mode is one where they are most likely to be able to save on some frame buffer space, but as they are load balancing on a per frame basis it is likely that a full frame of memory space has to be reserved one each board as the memory allocation is usually done the video is reset (i.e. when the 3D scene initially starts or when there is a resolution change causing a reset) and as each frame can require a variable amount of space on each board the full frame of frame buffer may have to be reserved on both boards.