Technical Brief

AGP 4X With Fast Writes – A Unique NVIDIA GeForce 256™ Feature

NVIDIA

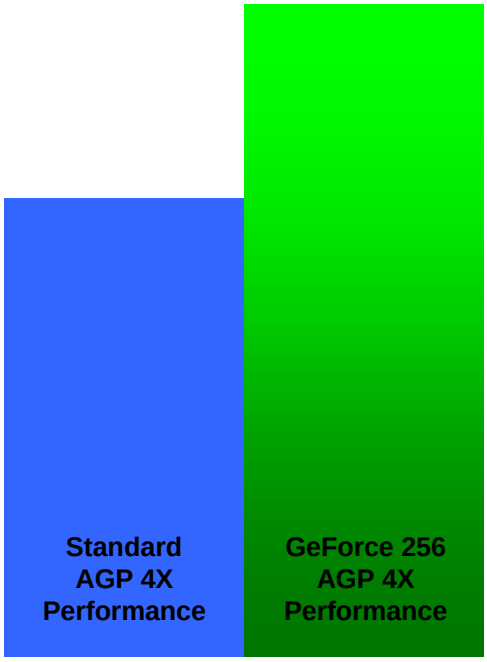
AGP 4X With Fast Writes – A Unique NVIDIA GeForce 256 Feature

Executive Summary

AGP 4X with Fast Writes is a unique feature implemented on the GeForce 256 GPU™ (Graphics Pocessor Unit). NVIDIA is the only vendor to take advantage of this feature, even though Fast Writes is part of the AGP 2.0 specification. Fast Writes improves all writes from the CPU to the graphics chip including:

- All 2D operations
- Operations involving writing to the frame buffer or sending any data to the graphics chip.
- Loading textures in Direct3D® into local memory.
- Writing push buffers to graphics local memory – this is where most of the performance boost is generated.

As such, systems with Fast Writes can have a dramatic increase over systems without Fast Writes.



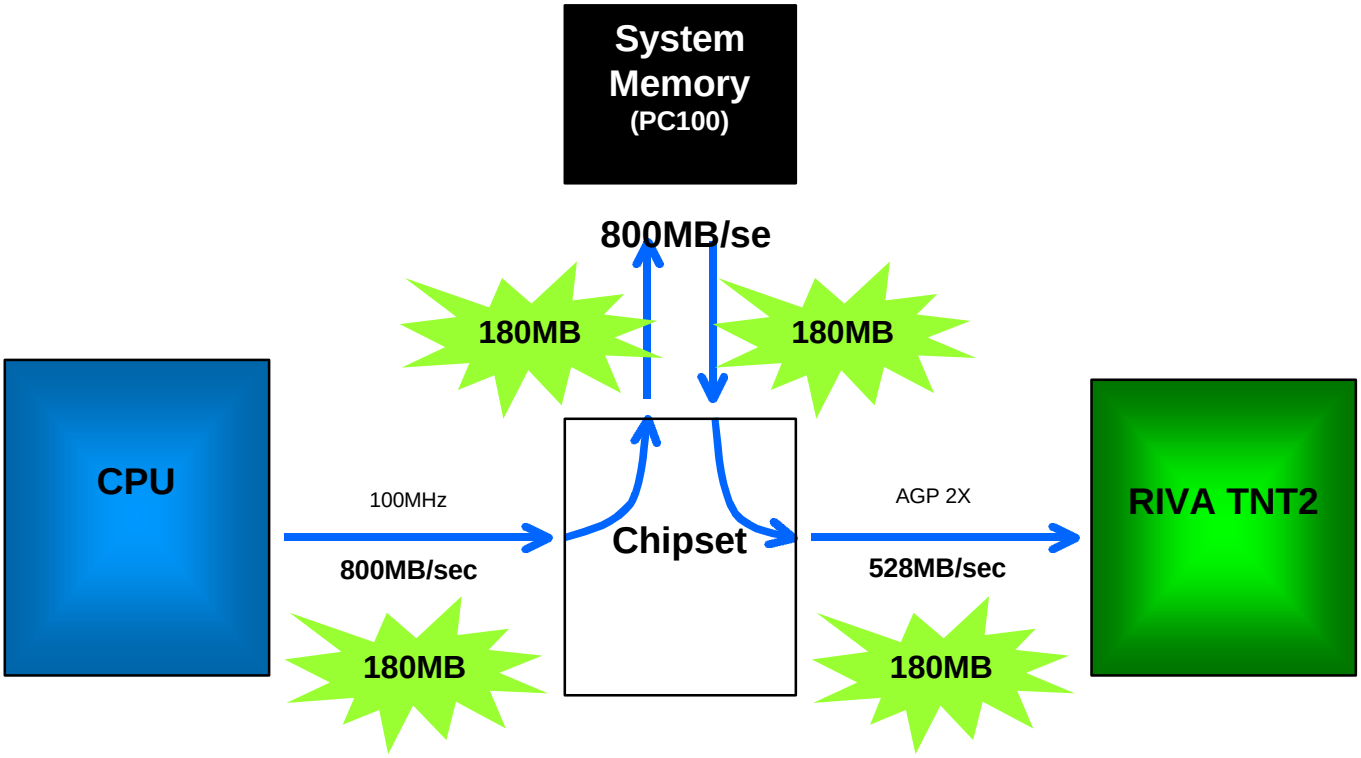
This paper explains how Fast Writes works, the benefits of this compelling feature, and why it is unique to NVIDIA's GeForce 256 GPU.

Today's Model with RIVA TNT2™

Today's system, shown below, typically uses an Intel®440® BX chipset (AGP 2X) and 100MHz system memory. This system's graphics performance is limited by the geometry performance of the CPU.

Assuming the CPU can process triangles at a rate of 2M triangles per second, and an average triangle is 90 bytes, you will see a data transfer of 180 MB/second – well under the maximum bandwidth the system can provide.

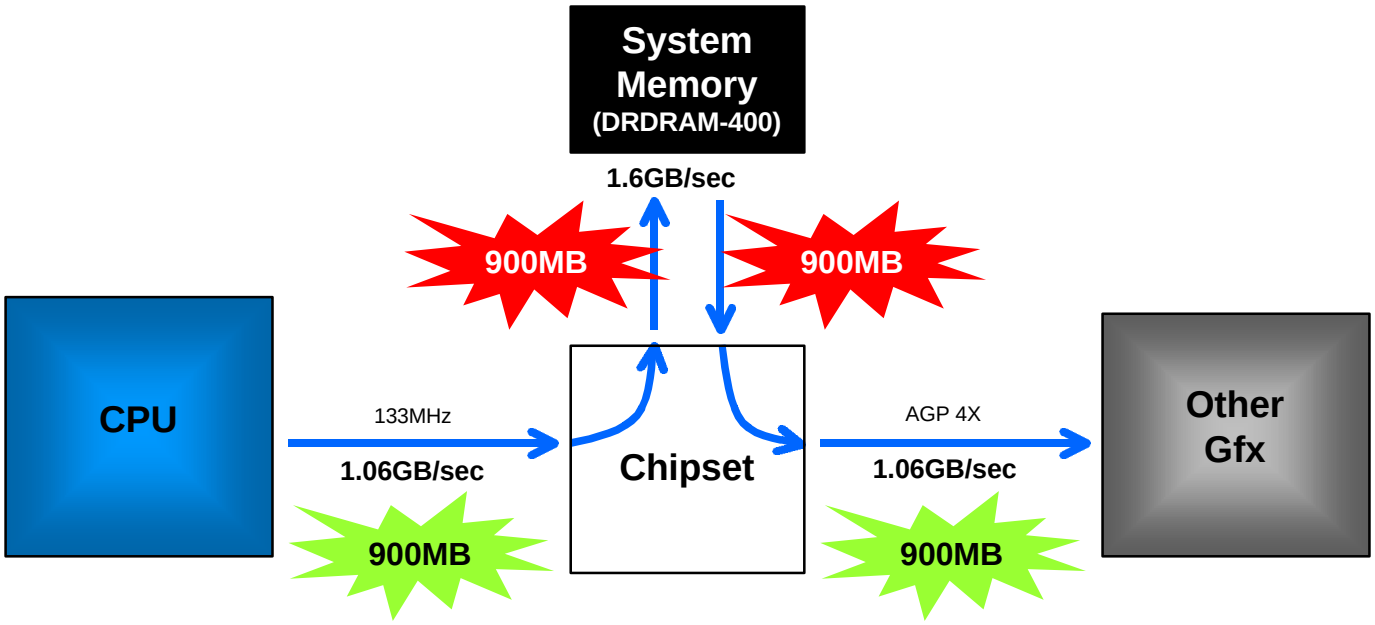
Thus, system memory bandwidth is <u>not</u> the bottleneck.



However, next-generation 3D applications will need up to five times greater triangle performance. This means that for the same triangle size (in terms of bytes transferred per triangle, e.g. 90 Bytes/triangle), the required bandwidth for data transfers will increase five times or up to 900 MB/sec! This exceeds the bandwidth of the current system.

Graphics Without Fast Writes Implementation

The picture below shows the “next-generation” system, implemented using Intel® 820 chipset (AGP 4X) and Direct RDRAM. With the system front side bus operating at 133MHz (1.06 GB/sec) and system memory at 1.6 GB/sec, a bottleneck still occurs when 900MB/sec data transfers are required.



In this case, the graphics controller does not implement Fast Writes. This causes the CPU to stall while it waits for the data to be transferred over the system. Even with 1.6 GB/sec, the system memory bus is unable to handle the constant shuffling of 900MB/sec data in and out of system memory. This thrashing on the system memory bus will also cause overall system performance to degrade.

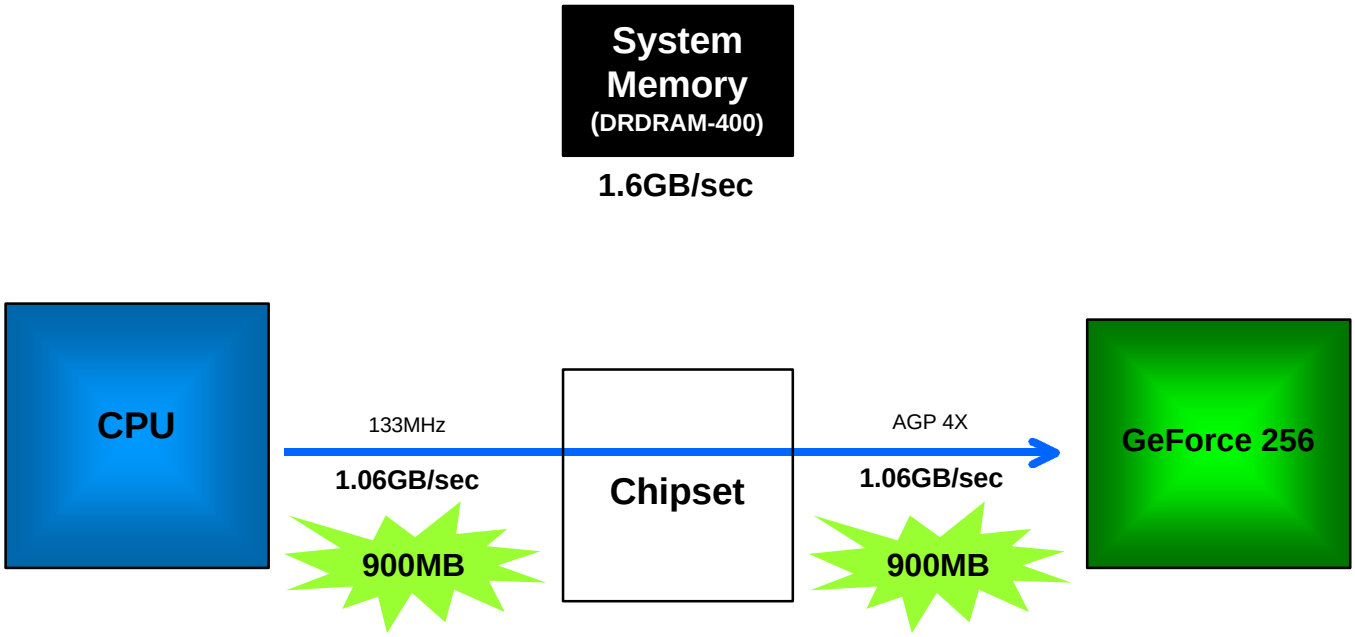
The next section explains how a graphics processor with Fast Writes implementation will eliminate this problem.

GeForce 256 With Fast Writes Implementation

Fast Writes capability enables the CPU to send data directly to the graphics bus without going through the system memory. This benefits overall system performance in two ways:

1. Removes the bottleneck previously that exists in systems without Fast Writes.
2. Frees up system memory bus to perform other functions and increase system efficiency.

The picture below shows how the GeForce 256 graphics processor work with Fast Writes implemented:



Fast Writes improves all writes from the CPU to the graphics chip including:

- All 2D operations.
- Operations involving writing to the frame buffer or sending any data to the graphics chip.
- Loading textures in D3D into local memory.
- Writing push buffers to graphics local memory – this is where most of the performance boost is generated.

As a result, systems with Fast Writes can have a dramatic increase over systems without Fast Writes.

Fast Writes – Why Doesn't Everyone Else Have It?

It is clear that AGP with Fast Writes gives a significant performance boost. So why hasn't everyone taken advantage of this feature?

The simple answer to this simple question is as follows:

1. Implementing Fast Writes is architecturally challenging. The front-end engine of the graphics processor must be architected to handle 4X the CPU write bandwidth while maintaining compatibility to non-Fast Writes capable chipsets (e.g. Intel 440BX).

2. A Fast Writes implementation must meet tight AC timing specifications. Meeting AGP 4X timings without Fast Writes is difficult already – Fast Writes makes it even harder.

So how is NVIDIA's implementation unique? NVIDIA designed our own AGP Fast Writes north bridge (per the AGP 2.0 specification) and used this as a test vehicle. We emulated this test chip in combination with theGeForce 256 GPU. We are the only graphics vendor with this unique feature and we have the performance advantage over all other AGP 4X implementations. NVIDIA is the only partner that Intel has to validate Fast Writes capability with the Intel 820 (aka Camino) chipset.

Summary

- Next-generation content will need a 5X increase in triangle performance and will saturate system memory bandwidth.
- Graphics controllers ***without*** Fast Writes will be bottlenecked by the system memory bandwidth, causing the CPU to stall and system performance to suffer.
- GeForce 256 with Fast Writes will eliminate this problem and dramatically boost overall system performance.
- Fast Writes is a feature unique to the GeForce 256 GPU because it requires a superior architectural implementation and industry-leading design skills.