

History

First, some history

GPU - Speeding up graphics computation since 1983

- Dedicated cards for handling graphical rendering
- Generally 3d acceleration today, but started off with 2d acceleration
- Arguably first card: Intel iSBX 275 Video Graphics Controller Multimodule Board

In 1983 Intel made the iSBX 275 Video Graphics Controller Multimodule Board for industrial systems based on the Multibus standard.[2] The card was based on the 82720 Graphics Display Controller and accelerated the drawing of lines, arcs, rectangles, and character bitmaps. The framebuffer was also accelerated through loading via DMA. The board was intended for use with Intel's line of Multibus industrial single board computer plugin cards.

The arrival of 3d

Dedicated 3d graphics hardware speedup

- 3dfx interactive released the Voodoo 1 card
- Dedicated 3d graphics card
- needed to be piggybacked with a 2d card
- possibly not first dedi 3d graphics card, but best at the time
- handled all the rendering of 3d primitives

Programmable pipeline

The advent programming on the graphics card

- Programmers could now define “shader programs”
- First graphics card with support: nVidia GeForce 3
- Only 12 years ago - in march 2001

The GeForce 3 (NV20) is the third-generation of NVIDIA's GeForce graphics processing units. Introduced in March 2001, it advanced the GeForce architecture by adding programmable pixel and vertex shaders, multisample anti-aliasing and improved the overall efficiency of the rendering process. The GeForce 3 family comprises 3 consumer models: the GeForce 3, the GeForce 3 Ti200, and the

GeForce 3 Ti500. A separate professional version, with a feature-set tailored for computer aided design, was sold as the Quadro DCC. A derivative of the GeForce 3, known as the NV2A, is used in the Microsoft Xbox game console.

Introduced three months after NVIDIA acquired 3dfx and marketed as the nFinite FX Engine, the GeForce 3 was the first Microsoft Direct3D 8.0 compliant 3D-card. Its programmable shader architecture enabled applications to execute custom visual effects programs in Microsoft Shader language 1.1. With respect to pure pixel and texel throughput, the GeForce 3 has four pixel pipelines which each can sample two textures per clock. This is the same configuration as GeForce 2 (not MX).

Pipeline:

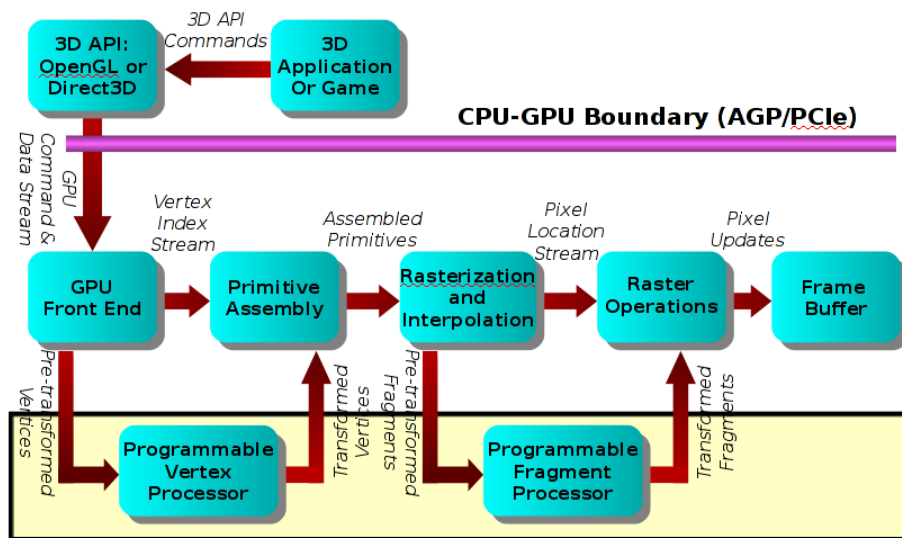
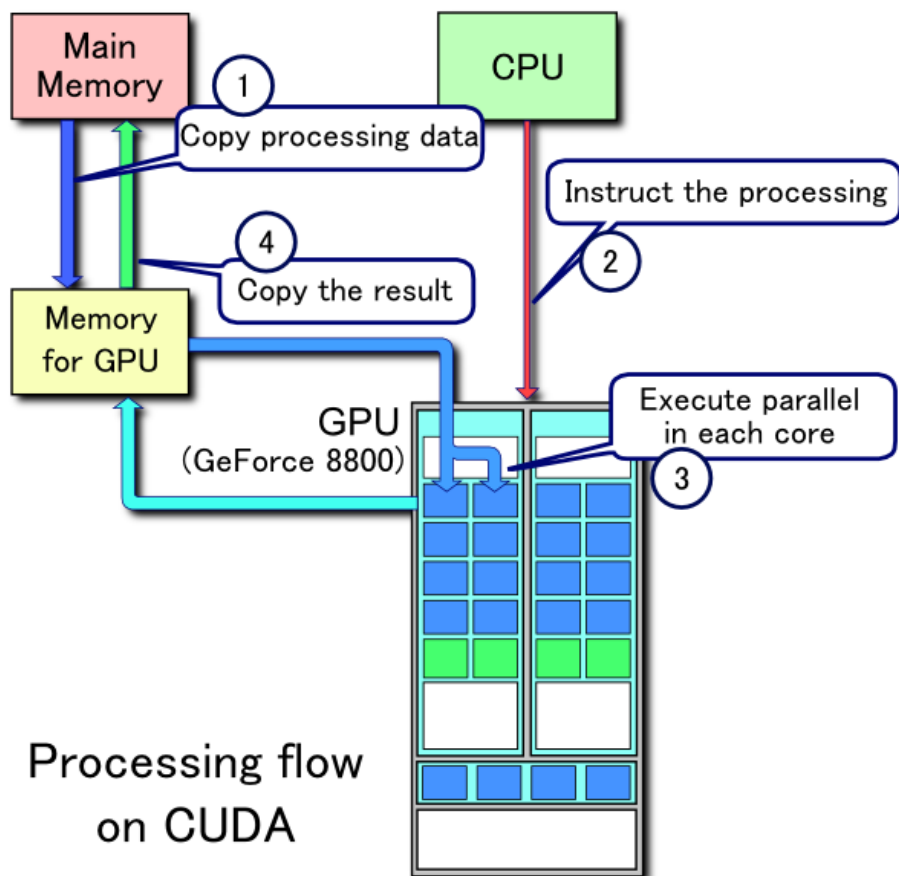


Figure 1: The programmable pipeline

CUDA arrives!

- Only available on nVidia cards (obviously)
- Released to public in
- First supporting card: 8800GTX



Processing flow
on CUDA

Figure 2: Cuda flow