Graphics Hardware

- Display (CRT, LCD, ...)
- Graphics accelerator
  - Scan controller
  - Video Memory (frame buffer)
  - Display/Graphics Processor
- CPU/Memory/Disk ...
Display Technologies

- Front projection
- Back projection
- Direct view
- Backlit
Display Technologies

- CRT
- LED
- LCD
- Plasma Panels
- DLP
- OLED
- Etc.
Trade-offs

- Cost, Weight, Size
- Power consumption
- Spatial & Color resolution
- Peak brightness, Black, contrast
- Etc.
1. Filament (generate heat)
2. Cathode (emit electrons)
3. Control grid (control intensity)
4. Focus
5. Deflector
6. Phosphor coating
Operation of a delta-delta, shadow-mask CRT. Three electron guns, aligned with the triangular color-dot patterns on the screen, are directed to each dot triangle by a shadow mask.
Color CRT

3 electron guns, 3 color phosphor dots at each pixel

Color = (red, green, blue)

Red = 0 to 100%
Green = 0 to 100%
Blue = 0 to 100%

Black = (0,0,0)
White = (1,1,1)
Red = (1,0,0)
Green = (0,1,0)
Blue = (0,0,1)

...
LED
- Direct view
- Backlight source
LCD: backlit
Plasma Panels: emit light; soon extinct?

A schematic matrix electrode configuration in an AC PDP
DLP: http://www.dlp.com/includes/video_demo.aspx

For digital projection

Digital Micromirror Device
Trade-offs

Peak brightness
Black level
Contrast
Screen brightness
Motion artifaces
Aging
Maximum resolution
Thickness
Weight
Power consumption

http://www.displaymate.com/ShootOut_Comparison.htm
Old way: No pixels - The electron gun draws straight lines from location to location on the screen (vector graphics)

- a.k.a. calligraphic display,
- Random scan device, vector drawing display

Use either display list or storage tube technology
Vector graphics

Display list
Move (100,200)
Draw(200,200)
Draw(200,100)
Draw(100,100)
Raster Display graphics

Digital Display

- Based on (analog) raster-scan TV technology
- The screen (and a picture) consists of discrete pixels
How CRT draws a picture

- We have only one electron gun but many pixels in a picture need to be lit simultaneously...
**Refresh of CRT**

- **Refresh** – the electron gun needs to come back to hit the pixel again before it fades out.

- An appropriate fresh rate depends on the property of phosphor coating.
  - **Phosphor persistence**: the time it takes for the emitted light to decay to 1/10 of the original intensity.

- Typical **refresh rate**: 60 – 80 times per second (Hz).
  (What will happen if refreshing is too slow or too fast?)
Frame Buffer

- Frame buffer: the memory to hold the pixel intensity values

- Properties of a frame buffer that affect the graphics performance:
  - Size: screen resolution
  - Depth: color level
    - 1 bit/pixel: black and white
    - 8 bits/pixel: 256 levels of gray
    - 24 bits/pixel: 16 million colors
  - Speed: refresh speed
What we do now: the electron gun will scan through the pixels from left to right, top to bottom (scanline by scanline)
Raster Scan Order

- The electron gun will scan through the pixels from left to right, top to bottom (scanline by scanline)

Horizontal retrace
Raster Scan Order

- The electron gun will scan through the pixels from left to right, top to bottom (scanline by scanline).
Progressive vs. Interlace

- **Progressive**: Scan every scan line
- **Interlace**: Scan only every other scan line (even -> odd -> even -> odd …)

- so the refresh rate becomes twice as fast
Standards

- NTSC: Interlaced, 525 line, 59.94Hz, 4:3
- DTV: Digital version of NTSC, 480i, 4:3
- HDTV: e.g. 720p, 60Hz, 16:9
- Blu-ray: 1080p
- Other standards?
Raster Scan Control

- **Scan Controller** (video adaptor) and frame buffer
Color is expensive …

- At least used to be
- The more color you want, the more bits you will need for each pixel
- Exercise: 1024 x 1280 screen with 24 bits per pixel, how many bytes in the frame buffer?

\[
1024 \times 1280 \times 24 / 8 = 4\text{M Byte}
\]

@30fps = 120MB/sec
Color Lookup Table

- Say I am a poor man … I only have 3 bits per pixel
- But I insist on having high quality pictures …
- Use Color Look Up Table (LUT)

You can still have 24 bits in each of the color table entries.
A simple graphics system

Frame buffer can be part of the main memory

Problem?
Dedicated memory

Video memory: On-board frame buffer: much faster to access

- CPU
- Main Memory

Scan Controller
Frame buffer

System bus
Graphics Accelerator

A dedicated processor for graphics processing

- Graphics Memory/Frame buffer
- Graphics Processor
- Scan Controller

- CPU
- Main Memory

System bus
Graphics Accelerator

NVIDIA Quadro FX 5600 and FX 4600
The Revolutionary Visual Computing Solution
Graphics Accelerator
## NVIDIA GPUs

<table>
<thead>
<tr>
<th>Feature</th>
<th>Quadro FX 5600</th>
<th>Quadro FX 4600</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory Size</td>
<td>1.5GB GDDR3</td>
<td>768MB GDDR3</td>
</tr>
<tr>
<td>Memory Interface</td>
<td>384-bit</td>
<td>384-bit</td>
</tr>
<tr>
<td>Memory Bandwidth</td>
<td>76.8 GB/sec.</td>
<td>67.2 GB/sec.</td>
</tr>
<tr>
<td>Max Power Consumption</td>
<td>171W</td>
<td>134W</td>
</tr>
<tr>
<td>Number of Slots</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Display Connectors</td>
<td>DVI-I DVI-I Stereo</td>
<td>DVI-I DVI-I Stereo</td>
</tr>
<tr>
<td>Dual-Link DVI</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Price</td>
<td>$2,999.00</td>
<td>$1,999.00</td>
</tr>
</tbody>
</table>
## AMD GPUs

### Desktop vs Mobility Radeon Graphics

<table>
<thead>
<tr>
<th></th>
<th>Desktop Radeon HD 6990</th>
<th>Desktop Radeon HD 6870</th>
<th>Radeon HD 6990M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transistors</td>
<td>5.28 billion</td>
<td>1.7 billion</td>
<td>1.7 billion</td>
</tr>
<tr>
<td>Engine Clock</td>
<td>830 MHz</td>
<td>900 MHz</td>
<td>715 MHz</td>
</tr>
<tr>
<td>Shader (ALUs)</td>
<td>3072</td>
<td>1120</td>
<td>1120</td>
</tr>
<tr>
<td>Texture Units</td>
<td>192</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>ROP Units</td>
<td>64</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>Compute Performance</td>
<td>5.1 TFLOPS</td>
<td>2.01 TFLOPS</td>
<td>1.60 TFLOPS</td>
</tr>
<tr>
<td>DRAM Type</td>
<td>GDDR5-5000</td>
<td>GDDR5-4200</td>
<td>GDDR5-3600</td>
</tr>
<tr>
<td>DRAM Interface</td>
<td>256-bits per GPU</td>
<td>256-bits</td>
<td>256-bits</td>
</tr>
<tr>
<td>Memory Bandwidth</td>
<td>160 GB/s per GPU</td>
<td>134 GB/s</td>
<td>115.2 GB/s</td>
</tr>
<tr>
<td>TDP</td>
<td>375 W</td>
<td>151 W</td>
<td>100 W</td>
</tr>
</tbody>
</table>
Evolution of Performance

<table>
<thead>
<tr>
<th>Year</th>
<th>PCI (133 MB/s)</th>
<th>AGP (266 MB/s)</th>
<th>AGP2x (533 MB/s)</th>
<th>AGP4x (1.06 GB/s)</th>
<th>AGP8x (2.1 GB/s)</th>
<th>PCIe (4 GB/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4 MB</td>
<td>32 MB</td>
<td>64 MB</td>
<td>128 MB</td>
<td>256 MB</td>
<td>512 MB</td>
</tr>
<tr>
<td>1995</td>
<td>DirectX 1, OpenGL 1.1</td>
<td>DirectX 2, OpenGL 1.2</td>
<td>DirectX 3, DirectX 5</td>
<td>DirectX 6, OpenGL 1.3</td>
<td>DirectX 7, OpenGL 1.4</td>
<td>DirectX 9, OpenGL 1.5</td>
</tr>
<tr>
<td>1996</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Graph shows the increase in performance metrics over time.
GPUs

The Evolution of GPUs


THE GPU COMPUTING ERA

The Graphics Pipeline

Application Stage → 3D Triangles → Geometry Stage → 2D Triangles → Rasterization Stage → Pixels

For each triangle vertex:
- Transform 3D position into screen position
- Compute attributes

For each triangle:
- Rasterize triangle
- Interpolate vertex attributes across triangle
- Shade pixels
- Resolve visibility
Graphics Bus Interface

PCI based technology

- Graphics Memory/Frame buffer
- Graphics Processor
- Scan Controller
- Other Peripherals

PCI Bus - 132 MB/s

System Bus - 800 MB/s

- CPU
- Main Memory
Graphics Bus Interface (2)

- PCI Bus becomes the bottleneck!
  - Many devices are using it
  - There is a lot of stuff needs to be transmitted from main memory to graphics memory (geometry, textures, etc)
  - Example: 2M triangle, 90 Bytes each – 180MB > 132 MB (PCI bandwidth)
Accelerated Graphics Port (AGP)

A dedicated bus that allows direct access of main memory

- Graphics Memory/Frame buffer
- Graphics Processor
- Scan Controller
- Other Peripherals

AGP 1x: 518 MB/s

Fast!!!

PCI Bus – 132 MB/s

CPU

Main Memory
AGP

- AGP 1x is four times as fast compared to PCI! (now we have AGP 8x)
- No more local bus congestion!
- More geometry can be processed!
- Direct execution of many graphics operations from main memory
PCI Express

- Bandwidth?
Reading and Lab 1

- Textbook Chapter 1, 2

Lab 1: Compile and run the sample OpenGL program posted on the class web site