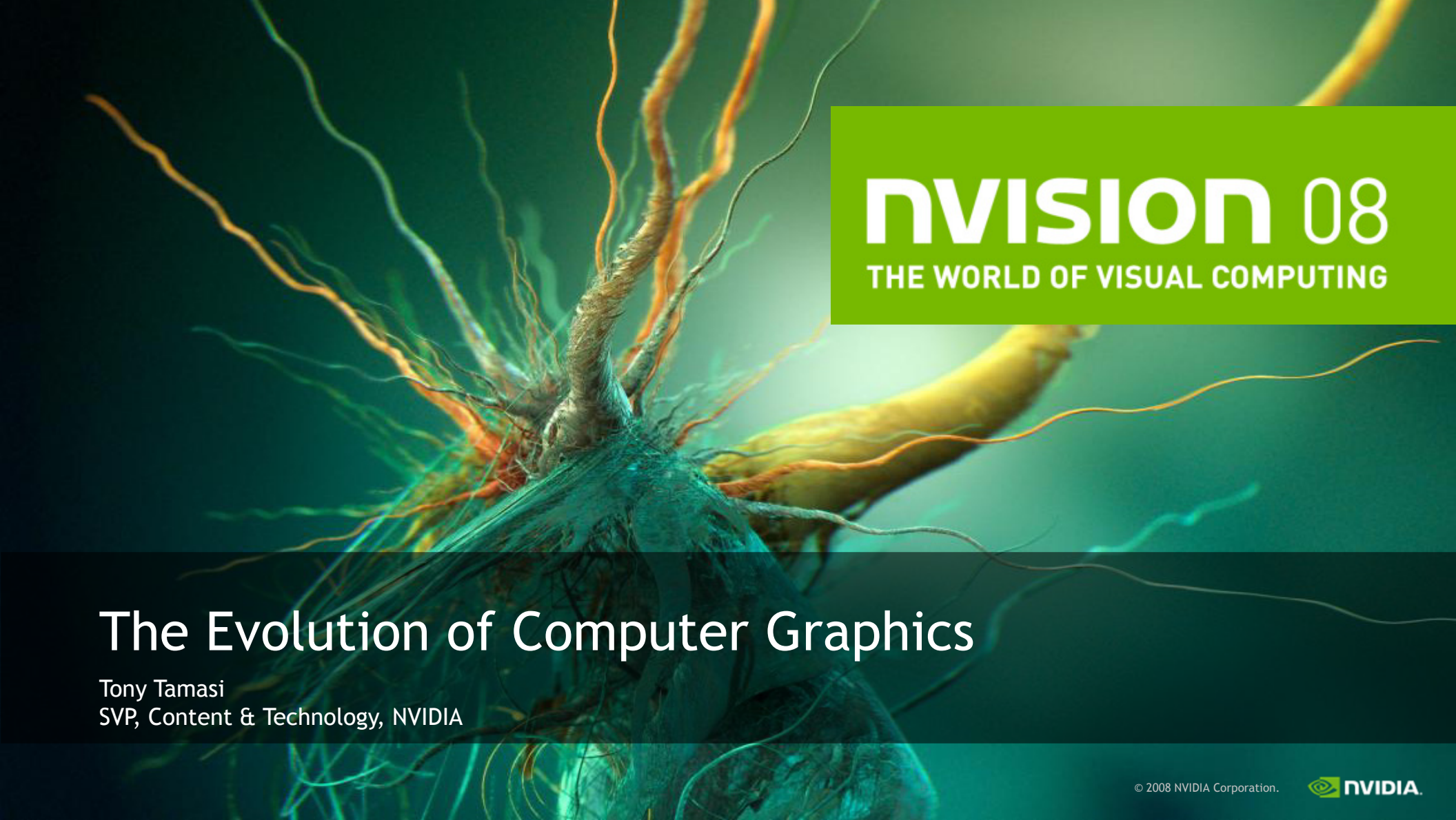
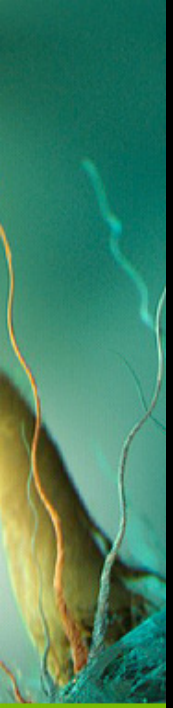



# The Evolution of Computer Graphics

Tony Tamasi
SVP, Content & Technology, NVIDIA



NVIDIA.

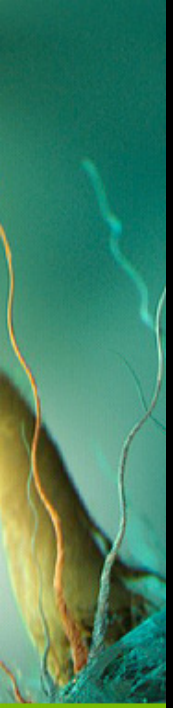




# Graphics

- Make great images
  - intricate shapes
  - complex optical effects
  - seamless motion

- Make them fast
  - invent clever techniques
  - use every trick imaginable
  - build monster hardware

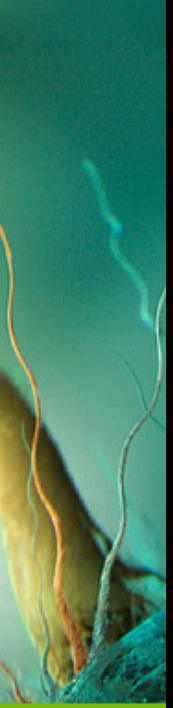


NVISION 08
THE WORLD OF VISUAL COMPUTING

# What is a Graphics Processing Unit?

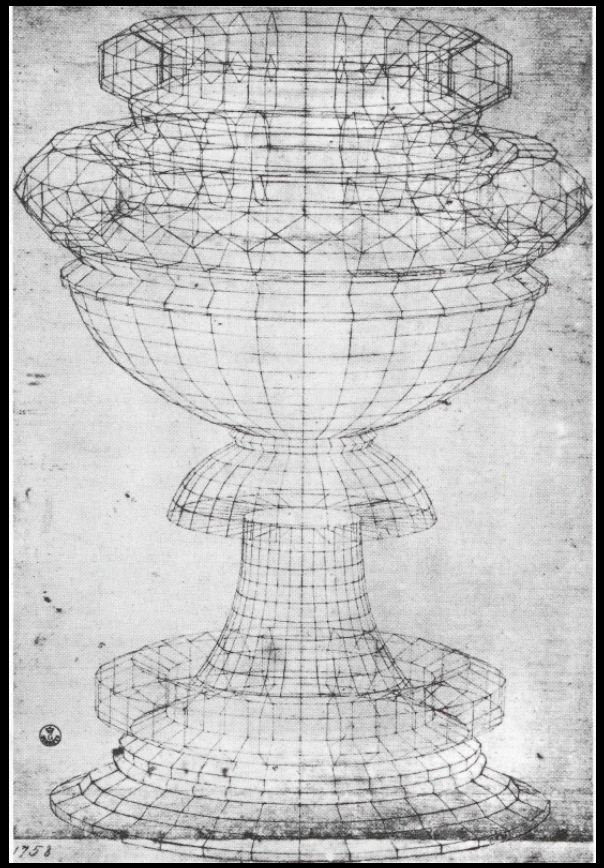
- Massively Parallel - 1000s of processors (today)
- Hardware managed parallelism – achievable performance
- Specialized processing – order of magnitude more efficient
- Latency Tolerant – throughput oriented
- Memory Bandwidth – saturate 100's GB/sec
- Not cache dependent – ALU heavy architecture


NVIDIA

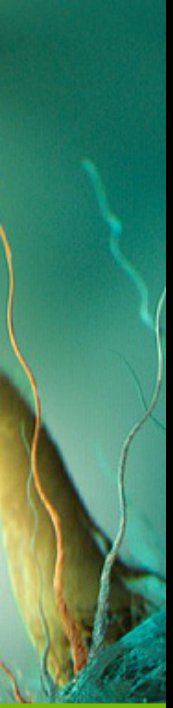




# Or, we could do it by hand…



*Perspective study of a chalice*
Paolo Uccello, circa 1450





# What is reality?

- Reality is 80 million polygons.

Alvy Ray Smith
1999

Image © and courtesy Digital Domain

Image © and courtesy Digital Domain

# Evolution of GPU Game Computing

## Constrained Fixed Function

Limited characters & animation
Simplistic AI
Indoors / basic environment
Multipass Rendering
Multi-Texture

## Dynamic Programmable Shading

Increased character count
IK skeletons / rigid bodies
Outdoors
Dynamic lighting
Programmable pixels

## Programmable Graphics

Armies of characters
Complex physical simulations
Sophisticated AI
Procedural generation
Custom renderers / lighting

# GPU Architecture Progression

Test Drive 6

Far Cry

**1998**
16-bit depth, Color, and textures

**1999**
Multi-texture, 32b rendering

**2000**
T&L, cube maps, Texture compression, Anisotropic filtering

**2001**
Programmable vertex, 3D textures, shadow maps, multisampling

**2002**
Early z-cull, Dual-monitor

**2003**
Fragment programs, Color and depth compression

**2004**
Flow control, FP textures, VTF

**2005**
Transparency antialiasing

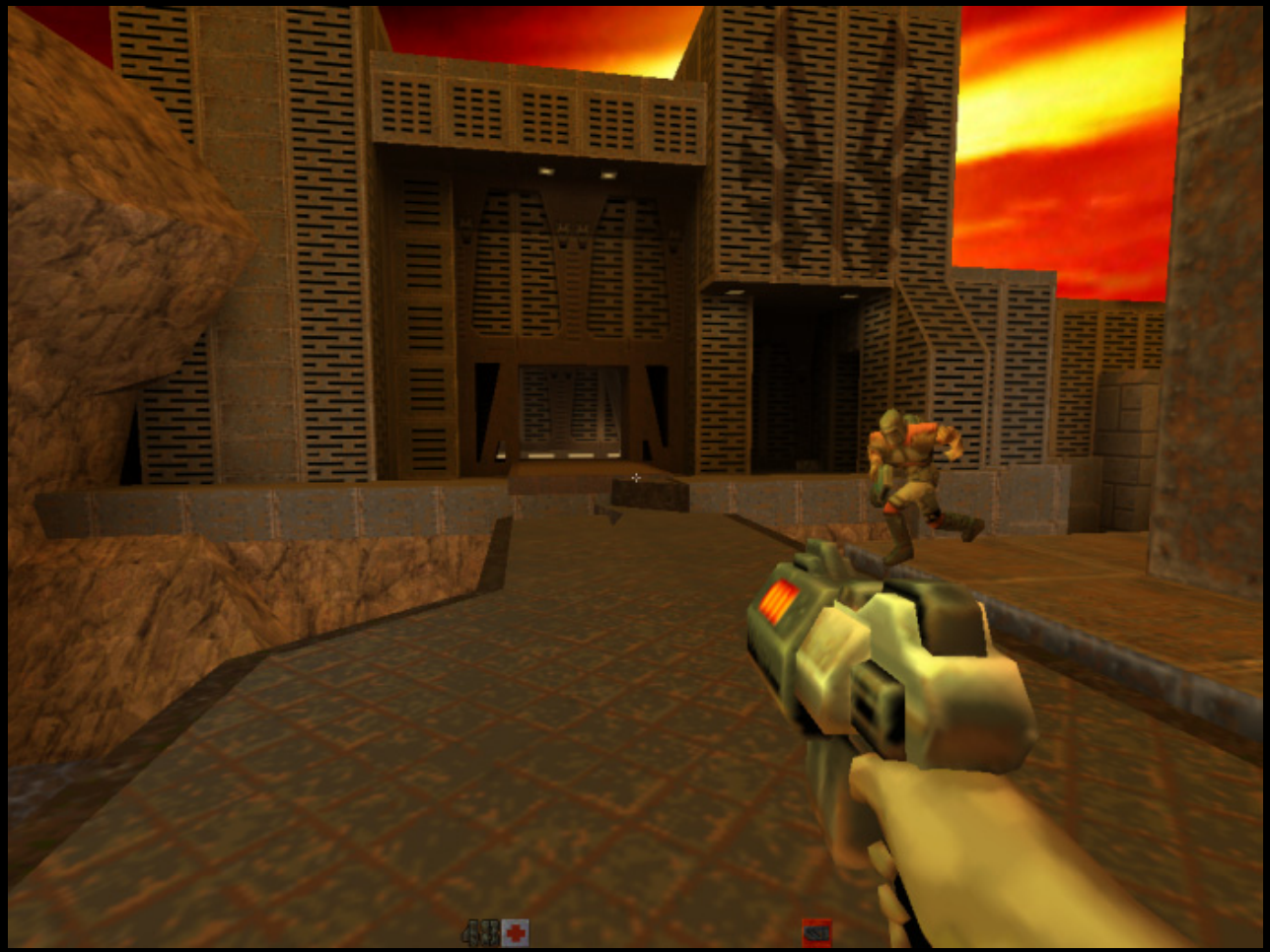
**2006**
Unified shader, geometry shader, CUDA/C

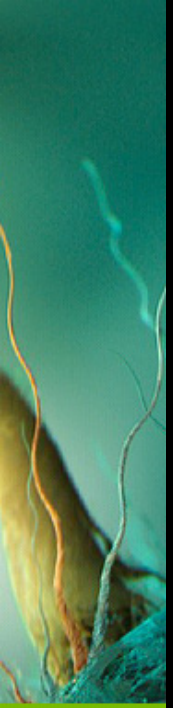
**2007**
Double Precision

Ballistics

Crysis

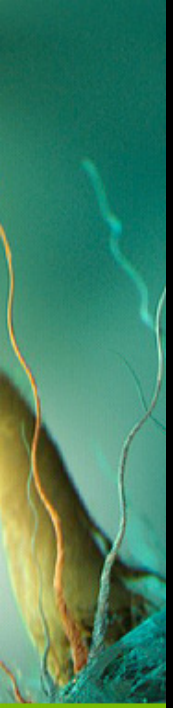
# Graphics 1998





# GPU 2003

- Dawn of programmable shading
- Vertex & Pixel Shaders
- More (8) pixels per clock
- Faster/Wider/Larger memory
- 32-bit and higher precision



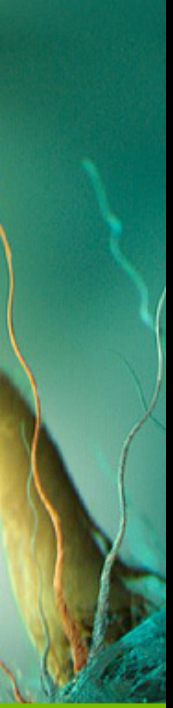
NVISION 08
THE WORLD OF VISUAL COMPUTING

# GPU 2003

- GeForce® FX 5800
  - Shader Model 2
  - 4 ppc / 8tpc @ 500 Mhz
  - Single precision
  - AGP8x
  - 128b 16GB/sec memory system
  - 4 processors @ 500 MHz
  - Single threaded – massive SIMD width

  - Programmed in OpenGL, DirectX or assembly


NVIDIA

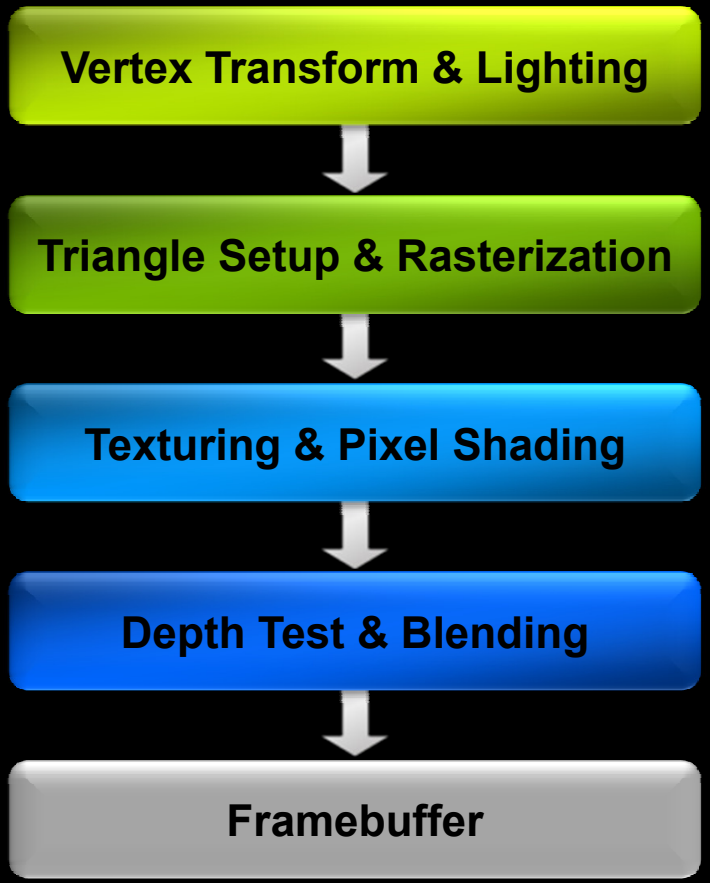


NVISION 08
THE WORLD OF VISUAL COMPUTING

# GPU 2003

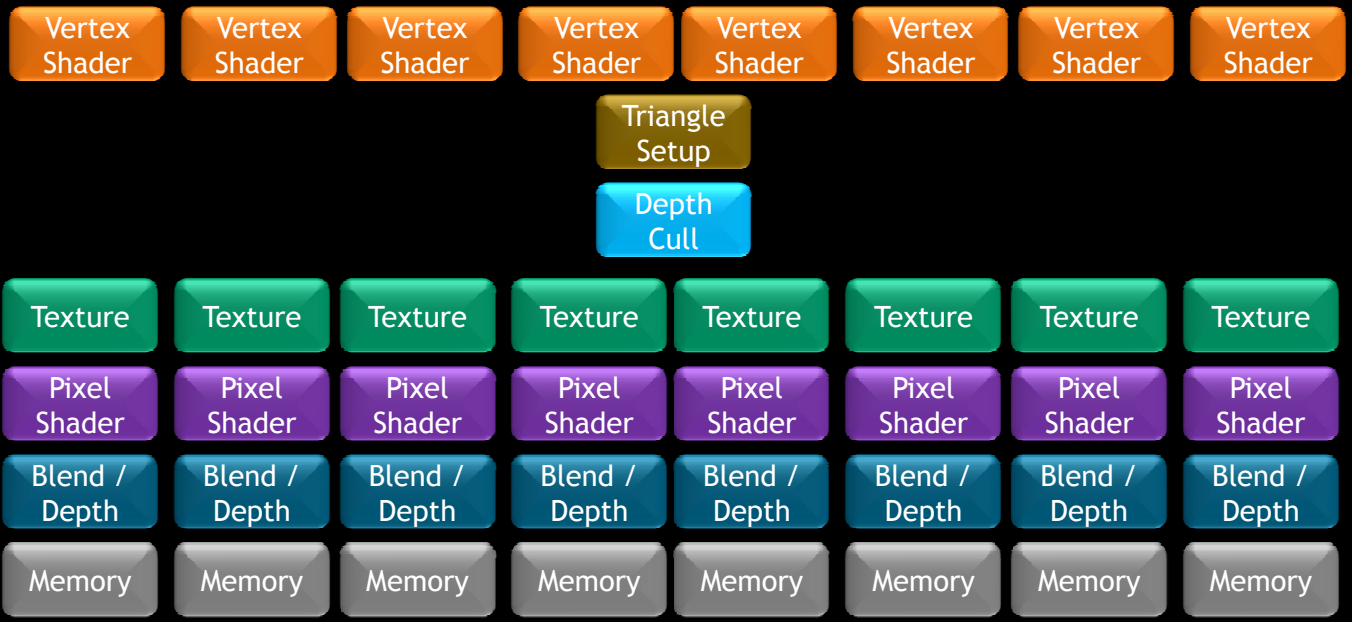
| | Riva 128ZX 1998 | GeForce FX5800 2003 |
|---|---|---|
| Texture Performance (tex/s) | 0.1G | 4 G |
| Antialiasing Performance (sam/s) | 0.1G | 2 G |
| Depth Performance (pix/s) | 0.1G | 4 G |
| Floating Point Performance (FLOPS) | n/a | 12 G |
| Memory Bandwidth (B/s) | 1.6G | 16 G |

# The Graphics Pipeline

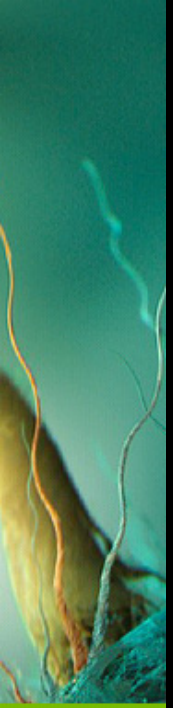



- Key abstraction of real-time graphics
- Hardware used to look like this
- Distinct chips/boards/units per stage
- Fixed data flow through pipeline

# GPU Architecture 2003

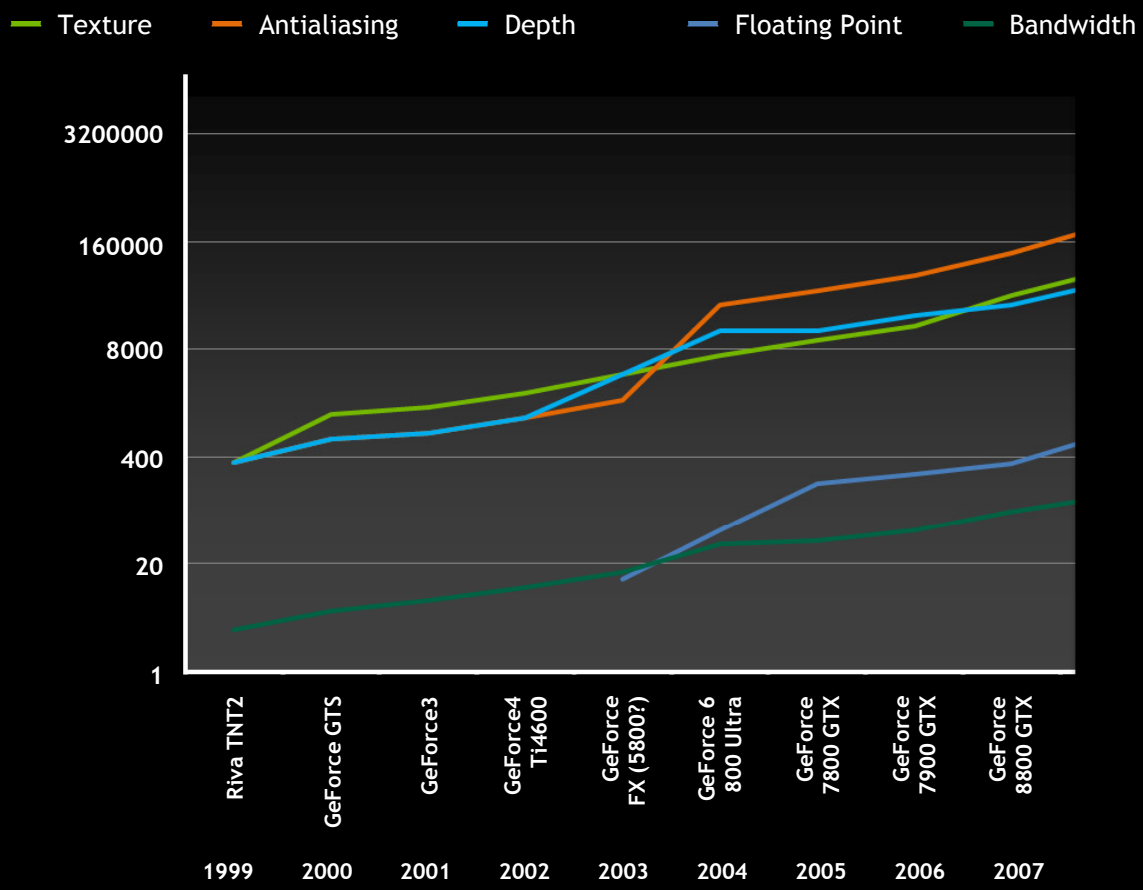


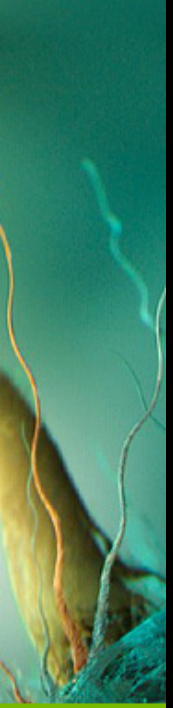
# Graphics 2003



NVISION 08
THE WORLD OF VISUAL COMPUTING

# Performance Trends





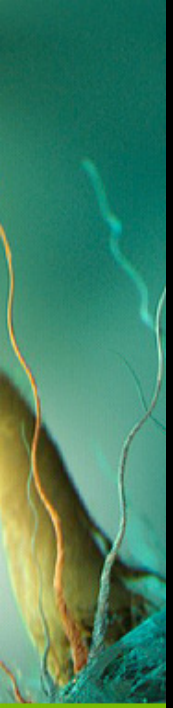
NVISION 08
THE WORLD OF VISUAL COMPUTING

# GPU Processing CAGR 1998 - 2007

| | Compound Annual Growth Rate |
|---|---|
| Texture Performance | 1.9 |
| Antialiasing Performance | 2.2 |
| Depth Performance | 1.9 |
| Floating Point Performance | 2.3 |
| Memory Bandwidth | 1.6 |



NVIDIA



NVISION 08
THE WORLD OF VISUAL COMPUTING

# GPU 2008 Projection

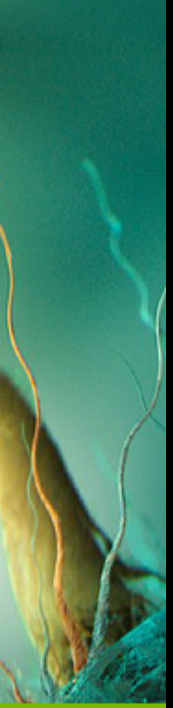
| | CAGR 1998 – 2007 | GeForce FX (2003) |
|---|---|---|
| Texture Performance (tex/s) | 1.9 | 4 G |
| Antialiasing Performance (sam/s) | 2.2 | 2 G |
| Depth Performance (pix/s) | 1.9 | 4 G |
| Floating Point Performance (FLOPS) | 2.3 | 12 G |
| Memory Bandwidth (B/s) | 1.6 | 16 G |

Within 10%

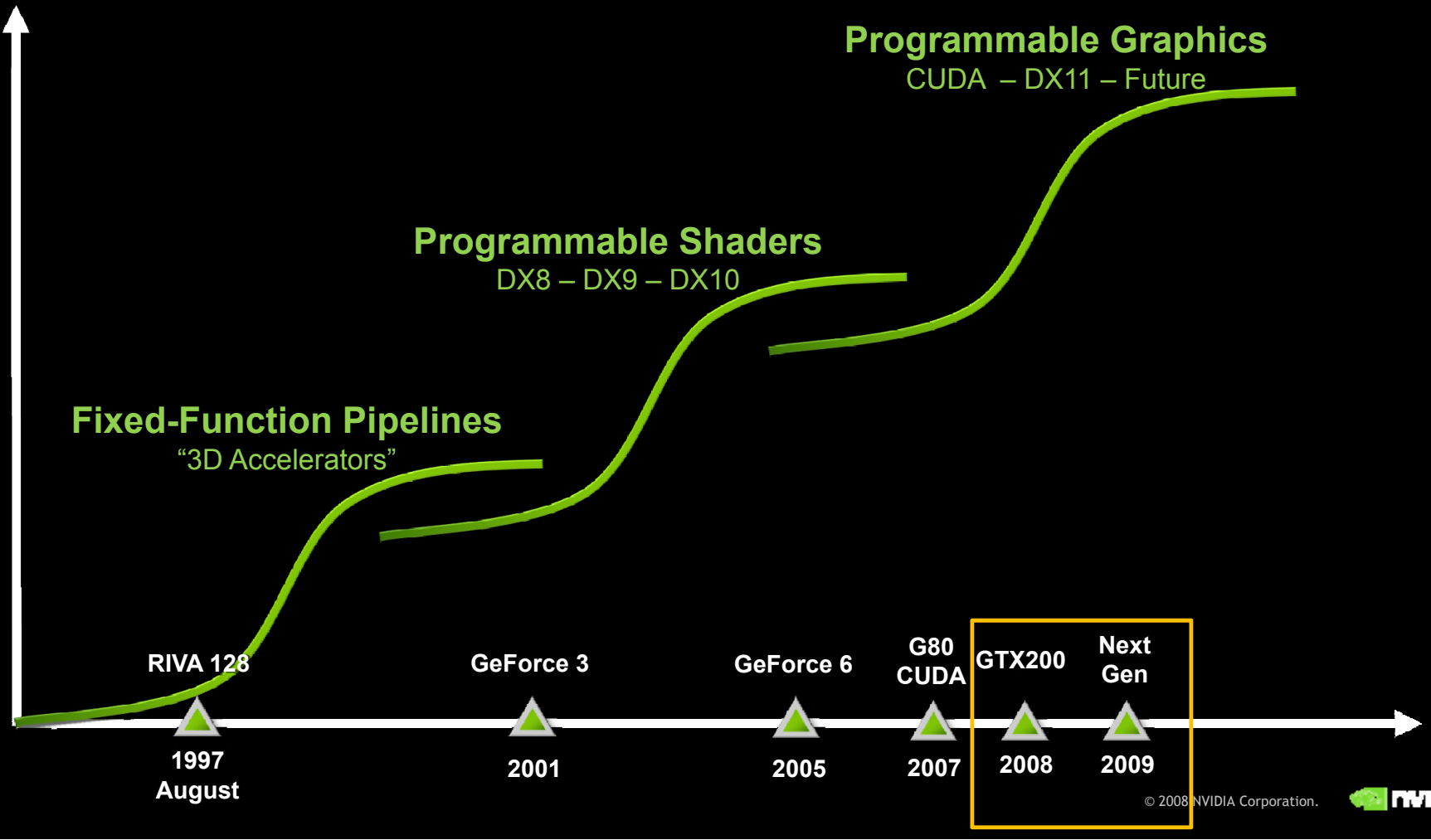


NVIDIA

# Era of Visual Computing

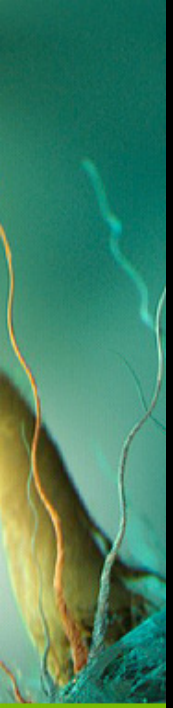


NVISION 08
THE WORLD OF VISUAL COMPUTING





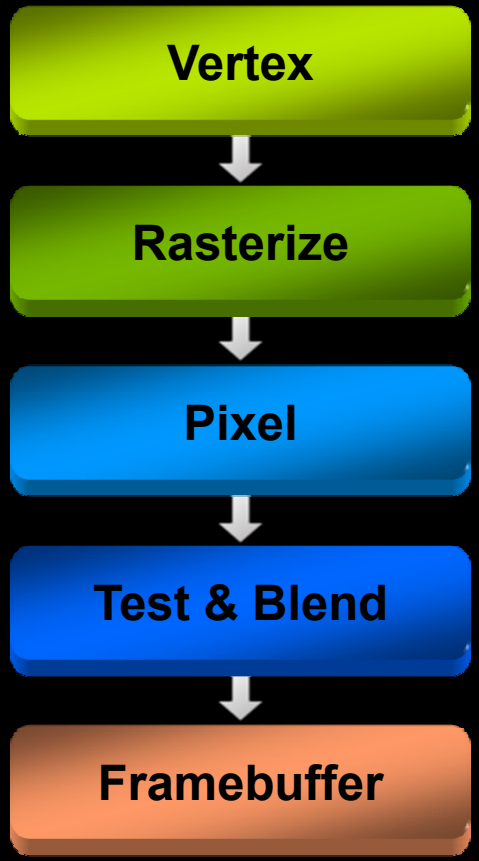

NVIDIA



# GPU 2008

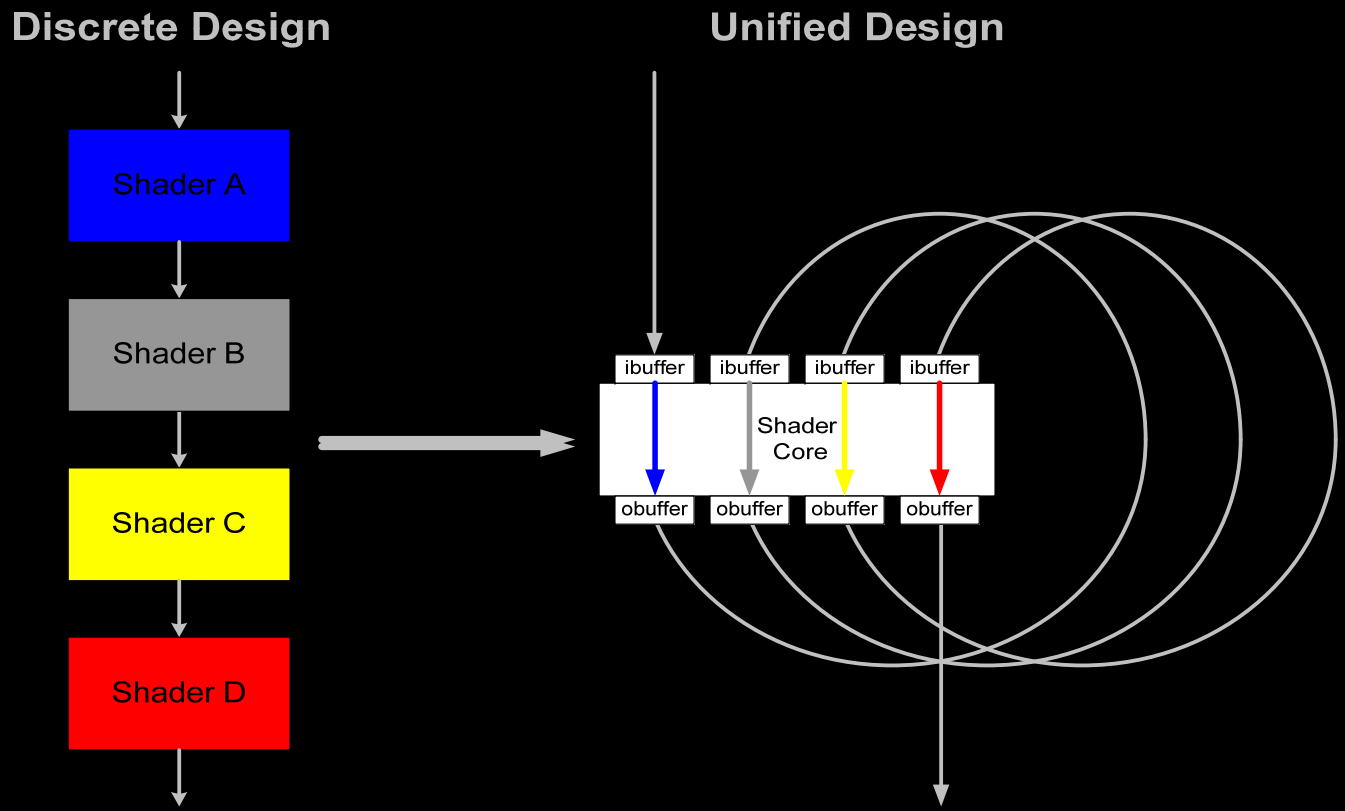
- Dawn of fully programmable generation
- Unified architectures
- New graphics functionality – geometry shading
- Programmable in C

# The Graphics Pipeline



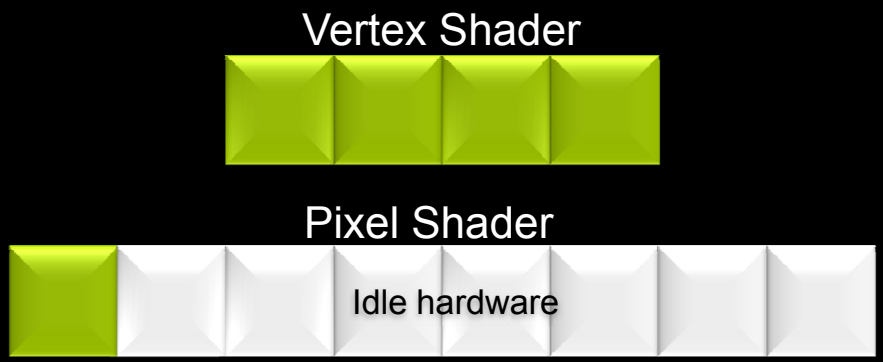

- Remains a useful abstraction
- Hardware used to look like this

# Modern GPU's: Unified Architecture

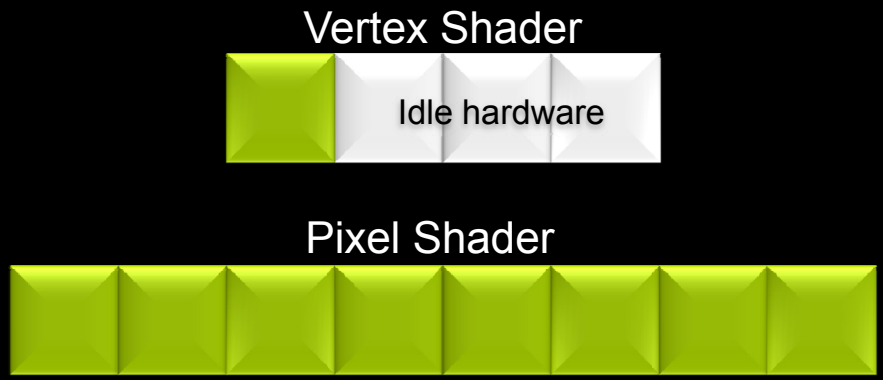



Vertex shaders, pixel shaders, etc. become *threads* running different programs on flexible cores

# Why unify?



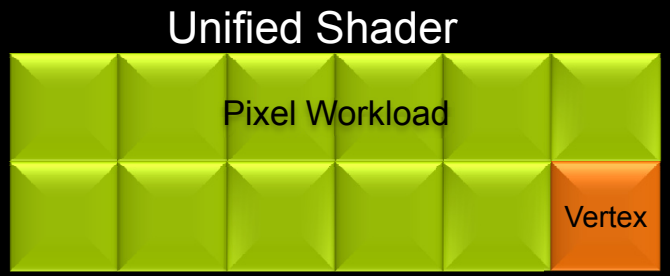

Heavy Geometry
Workload Perf = 4



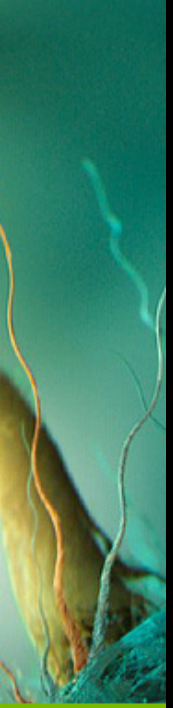

Heavy Pixel
Workload Perf = 8

# Why unify?


Heavy Geometry
Workload Perf = 11




Heavy Pixel
Workload Perf = 11



NVISION 08
THE WORLD OF VISUAL COMPUTING

# GPU 2008

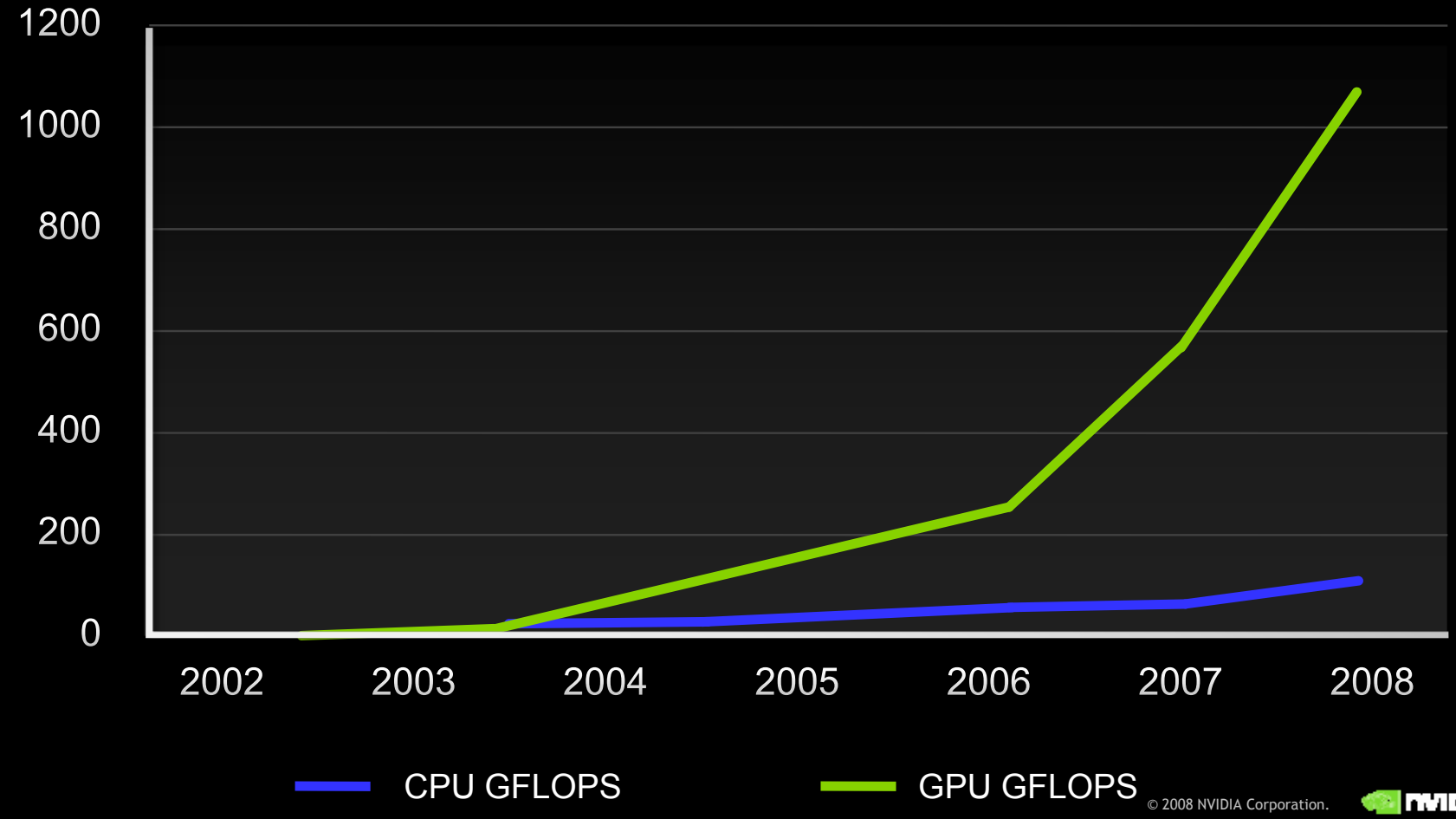
- GeForce® GTX 200
  - Shader Model 4
  - Double precision
  - PCI-Express Gen2
  - 512b 140GB/sec memory system
  - 32ppc / 80 tpc @ 600 Mhz
  - 240 processors @ 1.5 GHz
  - Many threaded – scalar

  - Programmed in OpenGL, DirectX, CUDA™ C



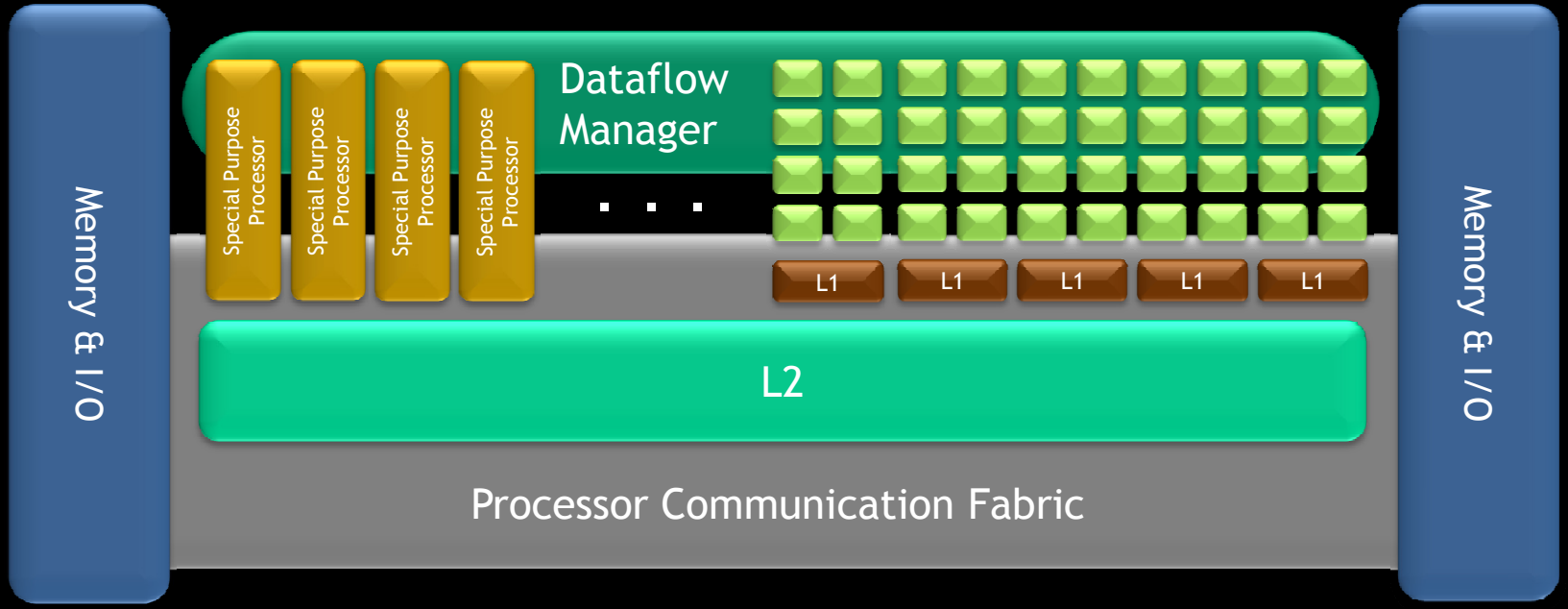
NVIDIA

# GPU Processing



# GPU Architecture 2008

Processing Cores



# State of the art 2008


Grid - Codemasters

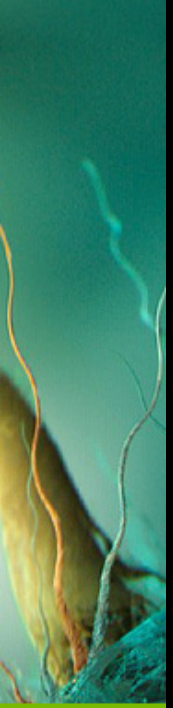

Far Cry 2 - Ubisoft

Call of Duty 5 - Activision

Bionic Commando - Capcom



NVISION 08 THE WORLD OF VISUAL COMPUTING



NVISION 08
THE WORLD OF VISUAL COMPUTING

# GPU 2013

**2003**
Fragment programs, Color and depth compression

**2004**
Flow control, FP textures, VTF

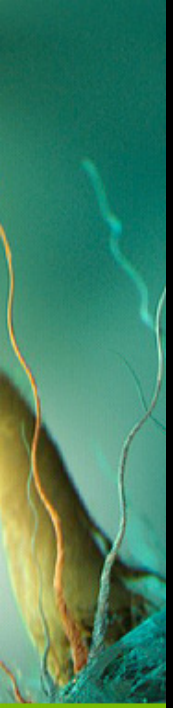
**2005**
Transparency, antialiasing

**2006**
Unified shader, geometry shader, CUDA/C

**2008**
Double precision

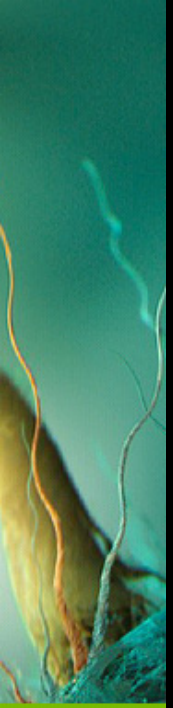


NVIDIA



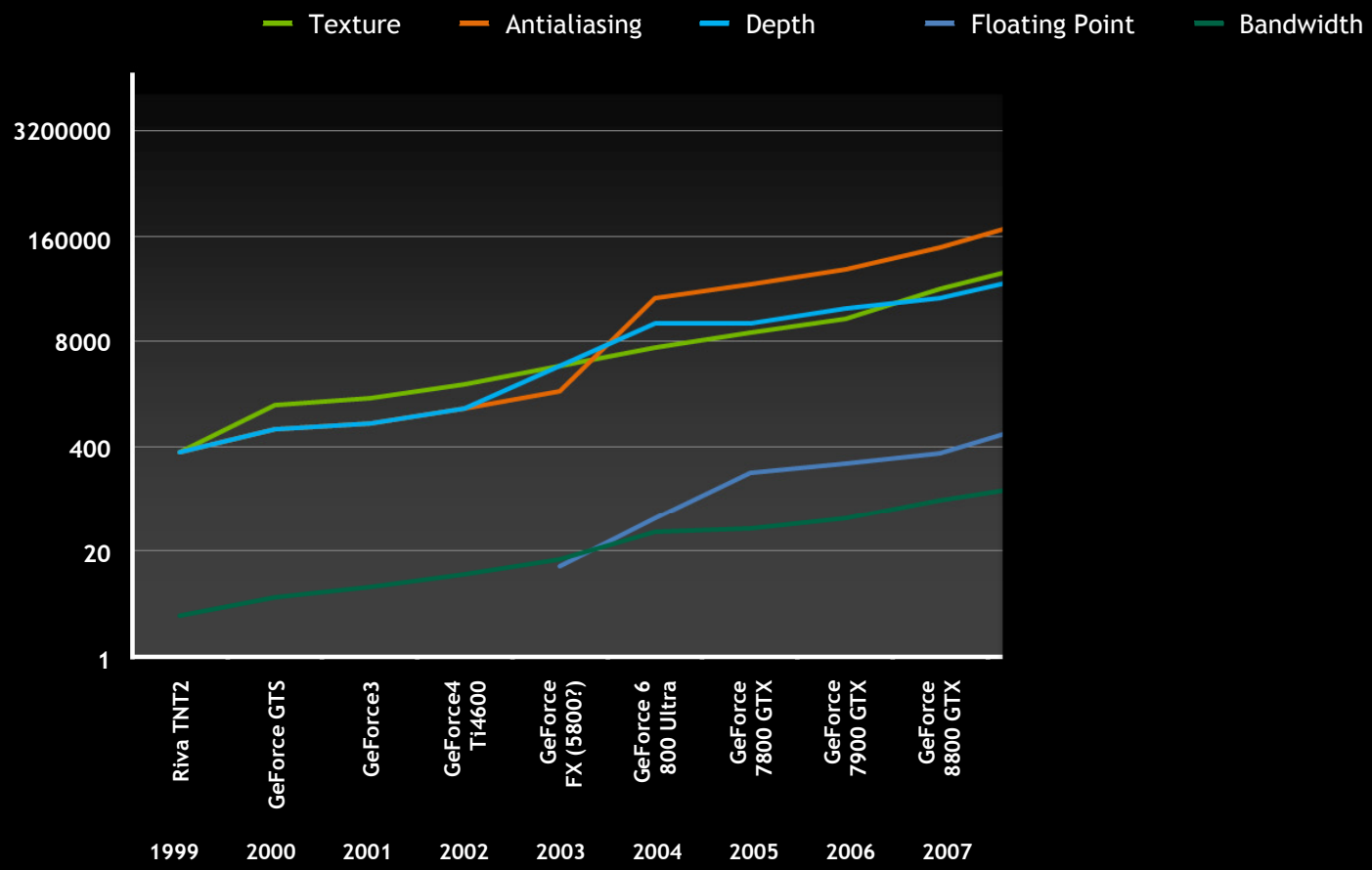
# GPU 2013

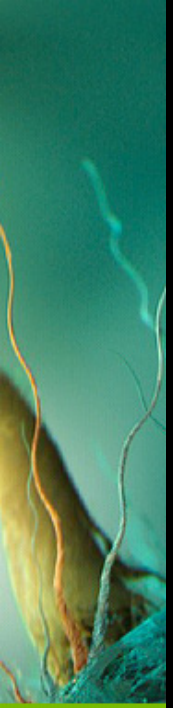
- Arbitrary dataflow
- General purpose programming model
- Special purpose hardware
- Hardware managed threading and pipelining
- Freely intermingle “graphics” & computation





# Performance Trends





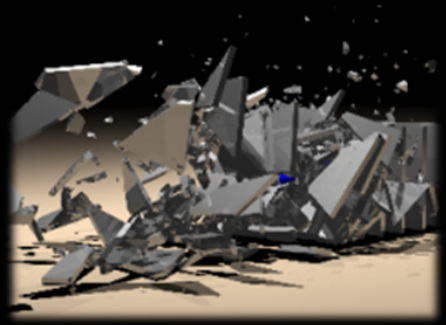
NVISION 08
THE WORLD OF VISUAL COMPUTING

# GPU 2013

| | CAGR 1998 - 2007 | GeForce GTX 200 |
|---|---|---|
| Texture Performance (tex/s) | 1.9 | 48 G |
| Antialiasing Performance (sam/s) | 2.2 | 153 G |
| Depth Performance (pix/s) | 1.9 | 38 G |
| Floating Point Performance (FLOPS) | 2.3 | 1080 G |
| Memory Bandwidth (B/s) | 1.6 | 141 G |



NVIDIA



Fracture

Soft Shadows

Detailed Characters

Rich Environments

Indirect Lighting

Subsurface Scatter

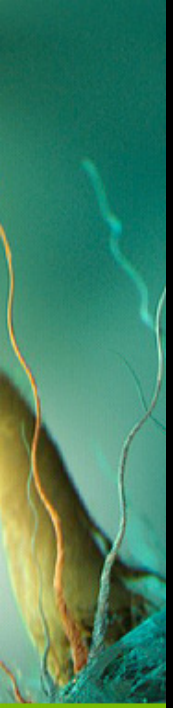
Ambient Occlusion

Turbulence

Participating Media

Simulations

Fluids



NVISION 08
THE WORLD OF VISUAL COMPUTING

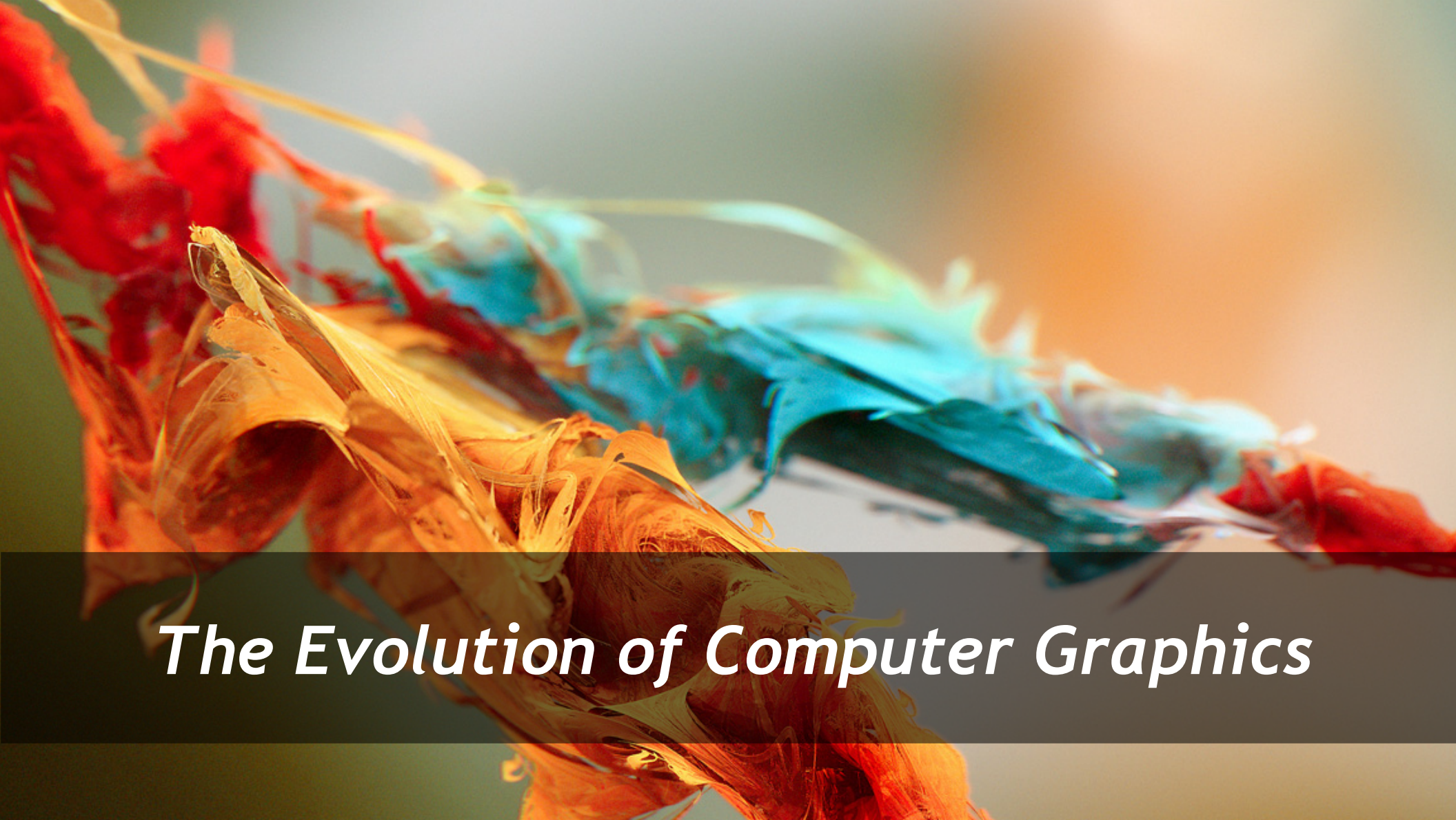
# The Future of Graphics Processing

- Programmable and Specialized Processing
- Monster performance and power efficiency
- Graphics and Arbitrary C/C++ Programming
- Global Illumination and Rasterization
- Rendering and Simulation

…Evolution



NVIDIA


The Evolution of Computer Graphics