CLOUD GAMING AND NVIDIA GRID™
Eric Young, DevTech Engineering Manager for GRID
CLOUD’IFICATION

Music, Movies, Books
< 2013

Games
2014+

amazon
Google play
HuluPlus
iTunes
kindle
NETFLIX
PANDORA
Spotify
YouTube
G-cluster Global
NETZYN
Playcast
SCALABLE GRAPHICS
ubibus
CLOUD GPUs FOR 6 MARKETS

- Cloud Gaming (GRID)
- Application Streaming (GRID SDK)
- Realtime Encode (NVENC SDK)
- Accelerated Virtual Desktops (vGPU)
- Remote Workstations (Quadro)
- High Performance Computing (TESLA)
CLOUD GAMING ADVANTAGES

- Mobility
- Uniform High Quality
- Play Instantly
- Precise Accounting
- No Piracy
- Ease of Updating Games
NVIDIA GRID GAME-STREAMING SERVICE
1080p 60 fps • Play in a Minute • AAA titles
GAME STREAMING ARCHITECTURE

GAME ENGINE

GTX SUPERCOMPUTER

60 ms

ENCODE

10 ms

15 ms

TEGRA X1

SHIELD

5 ms

DECODE

5 ms

RENDER

30 ms

CONTROLLER

10 ms

GRID

“HALF THE BLINK OF AN EYE”
Game Input Latency

<table>
<thead>
<tr>
<th>Component</th>
<th>Console</th>
<th>GRID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Game Engine</td>
<td>90 ms</td>
<td>60 ms</td>
</tr>
<tr>
<td>Encode</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Network</td>
<td>30 ms</td>
<td>30 ms</td>
</tr>
<tr>
<td>Decode</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Controller</td>
<td>30 ms</td>
<td>10</td>
</tr>
<tr>
<td>HDMI TV</td>
<td>10</td>
<td>30 ms</td>
</tr>
</tbody>
</table>

Timelines:
- 0 ms
- 25 ms
- 50 ms
- 75 ms
- 100 ms
- 125 ms
- 150 ms
- 175 ms
Supported APIs
- DirectX
- OpenGL
- PhysX (CUDA)
CAPTURE AND ENCODE

NVIDIA GRID™ SDK

H.264 Streams

Graphic Commands

Host Interface

DRAM Interface

Frame Buffer

Render Target

Front Buffer

NVENC

NVIFR

NVFBC
One Game Stream Per VM Per GPU
One VM Instance Connects to One GPU
## KEPLER VS MAXWELL

### Key Cloud Metrics

<table>
<thead>
<tr>
<th></th>
<th>Kepler GK104</th>
<th>Maxwell GM204</th>
<th>Speed Up</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GPU Specs</strong></td>
<td>1536 Cores 800 Mhz</td>
<td>2048 Cores 1100 Mhz</td>
<td>2X</td>
</tr>
<tr>
<td><strong>TFLOPS</strong></td>
<td>2.4</td>
<td>4.8</td>
<td>2X</td>
</tr>
<tr>
<td><strong>Texture Fillrate (GT/s)</strong></td>
<td>102</td>
<td>151</td>
<td>1.5X</td>
</tr>
<tr>
<td><strong>Video Memory</strong></td>
<td>4GB</td>
<td>8GB</td>
<td>2X</td>
</tr>
<tr>
<td><strong>1080p30 H.264 streams</strong></td>
<td>8</td>
<td>32</td>
<td>4x</td>
</tr>
</tbody>
</table>
WHERE TO FIND GRID?

Dedicated Capacity
- Data centers located around the world
- Choose the data center closer to user to cut down on latency

GRID Hardware Available On Amazon Web Services
- Scale on Demand and Flexible to configure
- No initial investment in hardware - Rent what you need
- Lots of Bandwidth from Servers to the Internet Backbone
BUILDING GRID SERVERS

Dual Socket Xeon E5
- 2x10 core 2.5 GHz CPUs
- SuperMicro & ASUS
- Up to 5 boards = 20 GPUs

Xeon E3 Systems
- 4x 4-Core 3.5 GHz CPUs
- 2 Boards = 8 GPUs
- Cirrascale, CARRI/GIGABYTE
# NVIDIA GRID™ PRESENTATIONS

<table>
<thead>
<tr>
<th>Time</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:00 am – 10:10 am</td>
<td>NVIDIA GRID™ Platform</td>
</tr>
<tr>
<td>10:10 am – 11:00 am</td>
<td>GRID™ Architecture Overview</td>
</tr>
<tr>
<td></td>
<td>GameStream - Low Latency Streaming</td>
</tr>
<tr>
<td>11:30 am – 12:00 pm</td>
<td>Data Analytics from the NVIDIA GRID™</td>
</tr>
<tr>
<td>12:15 pm – 1:15 pm</td>
<td>Best Practices for Games on NVIDIA GRID™</td>
</tr>
<tr>
<td></td>
<td>NVIDIA GRID™ Link SDK</td>
</tr>
</tbody>
</table>
RESOURCES

• Signup for GameWorks!
  • https://developer.nvidia.com/gameworks-registered-developer-program

• Android TV Developer Guide
  • http://developer.android.com/guide/practices/screens_support.html

• NVGamePad Library
  • http://developer.nvidia.com/cross-platform-gamepad-api

• Contact:
  • GRID-developer-support@nvidia.com