Nucleus: Eight GPU Platform for Visual Simulation

David Morgan
Principal Engineer
Aechelon Technology
Session Trajectory

- Visual Simulation Background
- Monsters, Clusters, and Moore’s Law
- Nucleus Architecture
- Challenges
- Demo
Image Generation
2000: RealityMonster

- 5 Racks
- 8 Graphics “Pipelines” (GPUs)
- 24 CPUs
- 9GB RAM (NUMA)
- 140GB Storage
- Single IRIX OS
- 15kW
- $2.8M
Scalability Matters

- 128 CPUs
- 256GB RAM
- 16 GPUs
2001: GeForce 3
2002-Today: PC Clusters

- 1-3 Racks per IG
- 1 GPU per node
- 1U Diskless Renderers
- 3U Pager w/88TB Storage
- Windows OS Per Node
- Ethernet Interconnect
- 7000W (8ch)
- Unlimited Scalability
Moore’s Law is Dead
2016: 8-GPU Support
Multi-GPU is Hard
Nucleus

- 4U
- 8 Quadro GPUs
- One display per GPU
- 36 CPU Cores
- 192GB RAM
- 36TB Storage
- One Windows OS
- $100-200K
- Operates up to 35C
- 2000W
- Limited Scalability
Dual Root Complex
GPU Affinity

- Exposed in OpenGL through WGL_NV_gpu_affinity extension
- Quadro feature necessary to address individual GPUs on Windows
- pC-Nova Maps GPU device handles to screens in the Windows virtual desktop
- Beware driver crashes enumerating more than 4 screens per GPU!
EDID Management

http://johnsciacca.webs.com/apps/blog/show/16852621-installation-nightmares-9-professional-horror-stories
DWM Is...

- Independent GPUs’ video timings phase shift.
- Normally correctable by tracking the phase
- “Full-Screen Exclusive Mode” is gone.
- DWM intermediates all drawing on multi-display systems.
- One display is Primary.
DWM Is Evil

https://www.pandza.xyz/article/16/dwm,-dxgi,-swap-chains,-latency,-throughput-and-you
Workaround: Framelock

- Quadro Sync II supports 8 GPUs per system
- Shared oscillator ensures displays remain in phase with Primary
- DWM placated!
- Downside: Video timings must all match
- Downside: Wiring is delicate
Future Work

• GPU Multicast
• Or Dual Root Complex?
• VR Direct?
• Clusters of Nuclei
Thanks

• Doug Traill
• John Chaney
• Tim Woodard
• Steve Nash
• Ian Williams
Demo