DELIVERING HIGH-PERFORMANCE REMOTE GRAPHICS WITH NVIDIA GRID VIRTUAL GPU

Andy Currid
NVIDIA
WHAT YOU’LL LEARN IN THIS SESSION

- NVIDIA's GRID Virtual GPU Architecture
  - What it is and how it works

- Using GRID Virtual GPU on Citrix XenServer

- How to deliver great remote graphics from GRID Virtual GPU
WHY VIRTUALIZE?

- ENGINEER / DESIGNER
  - Workstation
- POWER USER
  - High-end PC
- KNOWLEDGE WORKER
  - Entry-level PC

WHY VIRTUALIZE?
WHY VIRTUALIZE?

- Awesome performance!
- High cost
- Hard to fully utilize, limited mobility
- Challenging to manage
- Data security can be a problem
... CENTRALIZE THE WORKSTATION

- Awesome performance!
- Easier to fully utilize, manage and secure
- Even more expensive!

Datacenter
  - Desktop workstation
  - Quadro GPU

Remote Graphics

Notebook or thin client
... VIRTUALIZE THE WORKSTATION

GPU-enabled server

Virtual Machine
Guest OS
Apps
NVIDIA Driver
Virtual Machine
Guest OS
Apps
NVIDIA Driver

Datacenter

Direct GPU access from guest VM

NVIDIA GRID GPU

Notebook or thin client

Remote Graphics

Citrix XenServer
VMware ESX
Red Hat Enterprise Linux
Open source Xen, KVM

Dedicated GPU per user
... SHARE THE GPU
NVIDIA GRID VIRTUAL GPU

- Standard NVIDIA driver stack in each guest VM
  - API compatibility
- Direct hardware access from the guest
  - Highest performance
- GRID Virtual GPU Manager
  - Increased manageability
VIRTUAL GPU RESOURCE SHARING

- Frame buffer
  - Allocated at VM startup

- Channels
  - Used to post work to the GPU
  - VM accesses its channels via GPU Base Address Register (BAR), isolated by CPU’s Memory Management Unit (MMU)

- GPU Engines
  - Timeshared among VMs, like multiple contexts on single OS
VIRTUAL GPU ISOLATION

- GPU MMU controls access from engines to framebuffer and system memory
- vGPU Manager maintains per-VM pagetables in GPU’s framebuffer
- Valid accesses are routed to framebuffer or system memory
- Invalid accesses are blocked
NVIDIA GRID vGPU on Citrix XenServer

- First hypervisor to support GRID vGPU
  - Also supports GPU passthrough
- Open source
- Full tools integration for GPU
- GRID certified server platforms
XENSERVER SETUP

- Install XenServer
XENSERVER SETUP

- Install XenServer
- Install XenCenter management GUI on PC
- Install GRID Virtual GPU Manager

```
rpm -i NVIDIA-vgx-xenserver-6.2-331.30.i386.rpm
```
ASSIGNING A VGPU TO A VIRTUAL MACHINE

- Citrix XenCenter management GUI
- Assignment of virtual GPU, or passthrough of dedicated GPU
VM’s console accessed through XenCenter

Install NVIDIA guest vGPU driver
NVIDIA driver now loaded, vGPU is fully operational

Verify with NVIDIA control panel
DELIVERING GREAT REMOTE GRAPHICS

- Use a high performance remote graphics stack
- Tune the platform for best graphics performance
NVIDIA GRID SDK

- Available on vGPU and passthrough GPU
- Fast readback of desktop or individual render targets
- Hardware H.264 encoder
- Citrix XenDesktop
- VMware View
- NICE DCV
- HP RGS
TUNING THE PLATFORM

- Platform basics
- GPU selection
- NUMA considerations
PLATFORM BASICS

- Use sufficient CPU!
  - Graphically intensive apps typically need multiple cores

- Ensure CPUs can reach their highest clock speeds
  - Enable extended P-states / TurboBoost in the system BIOS
  - Set XenServer’s frequency governor to performance mode
    ```
    xenpm set-scaling-governor performance
    
    /opt/xensource/libexec/xen-cmdline --set-xen
cpufreq=xen:performance
    ```

- Use sufficient RAM! - don’t overcommit memory

- Fast storage subsystem - local SSD or fast NAS / SAN
**MEASURING UTILIZATION**

- **nvidia-smi** command line utility
- Reports GPU utilization, memory usage, temperature, and much more

```
[root@xenserver-vgx-test2 ~]# nvidia-smi
```

```
Mon Mar 24 09:56:42 2014
+------------------------------------------------------------------------+-------------------------+-------------------------+-------------------------+-------------------------+-------------------------+-------------------------+-------------------------+-------------------------+-------------------------+-------------------------+-------------------------+-------------------------+
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0 GRID K1 On P0 20W / 31W 530MiB / 4095MiB 61% Default</td>
<td>N/A</td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1 GRID K1 On P0 19W / 31W 270MiB / 4095MiB 46% Default</td>
<td>N/A</td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2 GRID K1 On P0 15W / 31W 270MiB / 4095MiB 7% Default</td>
<td>N/A</td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3 GRID K1 On P0 19W / 31W 270MiB / 4095MiB 46% Default</td>
<td>N/A</td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4 GRID K1 On P0 19W / 31W 270MiB / 4095MiB 45% Default</td>
<td>N/A</td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5 GRID K1 On P0 15W / 31W 10MiB / 4095MiB 0% Default</td>
<td>N/A</td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>6 GRID K1 On P0 19W / 31W 270MiB / 4095MiB 53% Default</td>
<td>N/A</td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>7 GRID K1 On P0 19W / 31W 270MiB / 4095MiB 46% Default</td>
<td>N/A</td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
```
MEASURING UTILIZATION

- GPU utilization graph in XenCenter
PICK THE RIGHT GRID GPU

GRID K2
2 high-end Kepler GPUs
3072 CUDA cores (1536 / GPU)
8GB GDDR5 (4GB / GPU)

GRID K1
4 entry Kepler GPUs
768 CUDA cores (192 / GPU)
16GB DDR3 (4GB / GPU)

ENGINEER / DESIGNER
POWER USER
KNOWLEDGE WORKER
SELECT THE RIGHT VGPU

GRID K260Q
2GB framebuffer
4 heads, 2560x1600

ENGINEER DESIGNER

GRID K240Q
1GB framebuffer
2 heads, 2560x1600

POWER USER

GRID K200
256MB framebuffer
2 heads, 1920x1200

KNOWLEDGE WORKER

GRID K2
2 high-end Kepler GPUs
3072 CUDA cores (1536 / GPU)
8GB GDDR5 (4GB / GPU)
SELECT THE RIGHT VGPU

GRID K260Q
2GB framebuffer
4 heads, 2560x1600

GRID K240Q
1GB framebuffer
2 heads, 2560x1600

GRID K200
256MB framebuffer
2 heads, 1920x1200

GRID K2
2 high-end Kepler GPUs
3072 CUDA cores (1536 / GPU)
8GB GDDR5 (4GB / GPU)
SELECT THE RIGHT VGPU

GRID K140Q
1GB framebuffer
2 heads, 2560x1600

GRID K100
256MB framebuffer
2 heads, 1920x1200

POWER USER

GRID K1
4 entry Kepler GPUs
768 CUDA cores (192 / GPU)
16GB DDR3 (4GB / GPU)

KNOWLEDGE WORKER
Non-Uniform Memory Architecture

Memory and GPUs connected to each CPU

CPUs connected via proprietary interconnect

CPU/GPU access to memory on same socket is fastest

Access to memory on remote socket is slower

TAKE ACCOUNT OF NUMA

CPU Socket 0

Memory 0

CPU Socket 1

Memory 1

CPU Interconnect

PCI Express

GPU

GPU

CPU

CPU

GPU

GPU
VM pinned to CPU socket by restricting its vCPUs to run only on that socket

- `xe vm-param-set uuid=<vm-uuid> VCPUs-params:mask=0,1,2,3,4,5`
SELECTING A VGPU ON A SPECIFIC SOCKET

CPU Socket 0
- Memory 0
- Grid K2
- GPU 1
- GPU 2
- GPU 3
- GPU 4

CPU Socket 1
- Memory 1
- Grid K2
- GPU 5
- GPU 6
- GPU 7
- GPU 8
GPU GROUPS

- XenServer manages physical GPUs by means of GPU groups

- Default behavior: all physical GPUs of same type are placed in one GPU group

- GPU group allocation policy:
  - Depth first: allocate vGPU on most loaded GPU
  - Breadth first: allocate vGPU on least loaded GPU
GPU GROUPS

XenServer manages physical GPUs by means of GPU groups:

- Default behavior: all physical GPUs of the same type are placed in one GPU group.

- GPU group allocation policy:
  - Depth first: allocate vGPU on the most loaded GPU.
  - Breadth first: allocate vGPU on the least loaded GPU.
GPU GROUPS

- Default GPU group takes no account of where a VM is running.
- Your VM may end up using a vGPU that’s allocated on a GPU on a remote CPU socket.
GPU GROUPS

- Create custom GPU groups
  - Per socket, or per GPU for ultimate control

- `xe gpu-group-create`
  - `name-label= "GRID K2 Socket 0"

- `xe pgpu-param-set`
  - `uuid=<pgpu-uuid>
    gpu-group-uuid=
    <group-uuid>

- `xe gpu-group-param-set`
  - `uuid=<group-uuid>
    allocation-algorithm= breadth-first"
GPU GROUPS

- Create custom GPU groups
  - Per socket, or per GPU for ultimate control

  - `xe gpu-group-create name-label= "GRID K2 Socket 0"

  - `xe pgpu-param-set uuid=<pgpu-uuid> gpu-group-uuid= <group-uuid>

  - `xe gpu-group-param-set uuid=<group-uuid> allocation-algorithm= breadth-first
WRAP UP

- NVIDIA's GRID Virtual GPU Architecture
- GRID Virtual GPU on Citrix XenServer
- Remote graphics performance
RESOURCES

- NVIDIA GRID vGPU User Guide
  - Included with GRID vGPU drivers
  - Visit http://www.nvidia.com/vgpu, look for driver download link

- Citrix XenServer with 3D Graphics Pack
  - Visit http://www.citrix.com/go/vgpu

- Qualified server platforms
  - Visit http://www.nvidia.com/buygrid
RESOURCES

- Remote Graphics
  - Citrix XenDesktop
    http://www.citrix.com/xendesktop
  - HP Remote Graphics Software (RGS)
  - NICE Desktop Cloud Visualization (DCV)
    https://www.nice-software.com/products/dcv

- XenServer CPU performance tuning
THANK YOU!

- NVIDIA GRID Forum
  https://gridforums.nvidia.com/
- Twitter
  @NVIDIAGRID