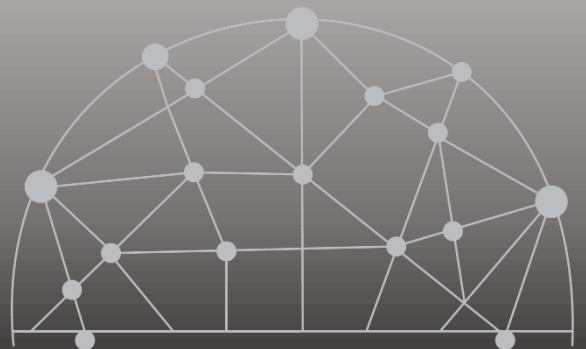
Implementing NVIDIA GRID with XenDesktop

Technical Deep Dive

THE KANAVEL GROUP

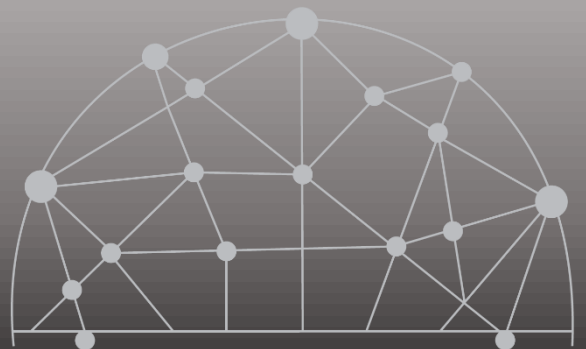
Who are we?

- Garrett Taylor
  - CIO of The Kanavel Group
  - Citrix CCIA/CCE for Virtualization
- Kanavel Group
  - Citrix Partner
  - NVIDIA Partner
  - VMWare Partner
  - Microsoft Partner



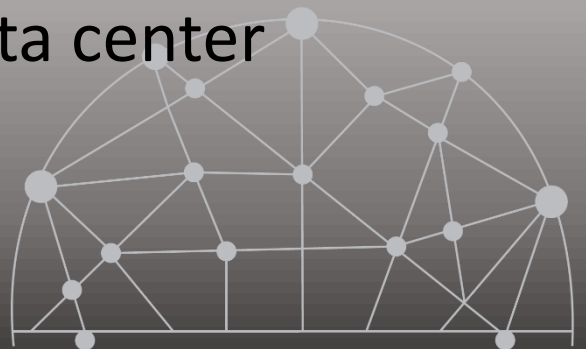
What is GRID?

- 3D Hardware, Software and Delivery services from a cloud
  - Cloud? Yours or Mine?
    - GRID-Enabled Public Clouds
      - Amazon EC2
      - IBM SoftLayer
    - OR… you can build your own!



Why GRID?

- GRID is the proverbial “missing link” between performance of a real desktop and the much-touted benefits of desktop virtualization
- Task/Knowledge workers can share “small” GPUs to get acceleration on video decoding, Windows Aero, Google Earth, etc…
- Time-zone sharing of expensive GPU resources for designers
- Expensive GPUs for part-time GPU users
- Deliver high-end applications to low-end devices
- Keep your data in the data center

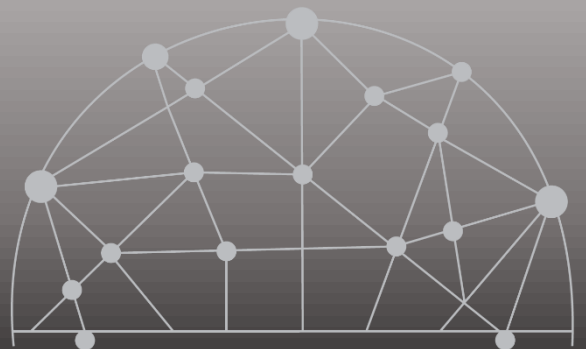


THE KANAVEL GROUP

NVIDIA

Who's using GRID?

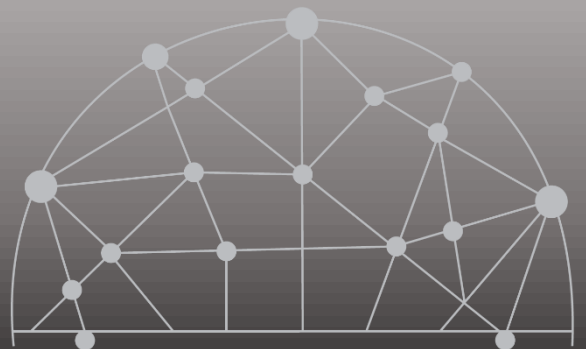
- Boeing
- Peugot/Citron
- Jellyfish Pictures
- Little Rock School District



NVIDIA GRID Hardware

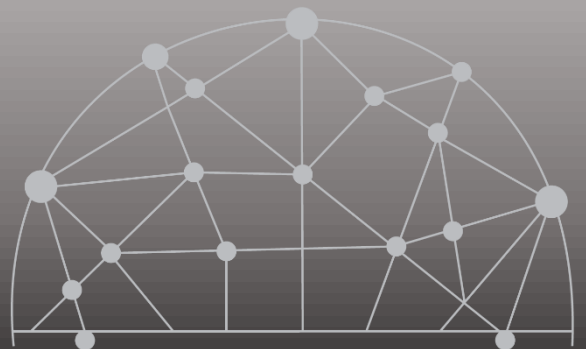
NVIDIA GRID Hardware

- K1
  - 4 x Quadro K1100 (GK107) at 850 Mhz with 192 CUDA Cores
  - 16GB RAM (4 per GPU)
- K2
  - 2 x Quadro K5000 (GK104) at 745 Mhz with 1536 CUDA Cores
  - 8GB RAM (4 per GPU)



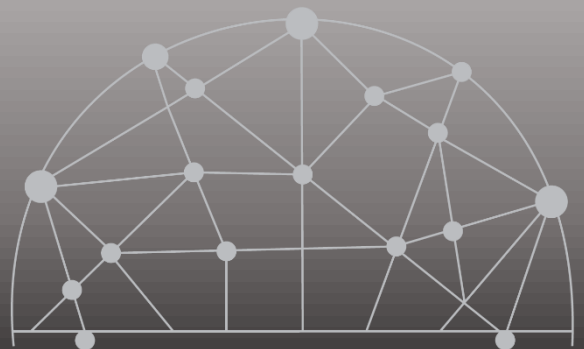
NVIDIA GRID Hardware

- Passive Cooling
  - No Fans – Server **MUST** be designed for Kepler GRID cards
- Power Requirements
  - K1 = 130W
  - K2 = 225W



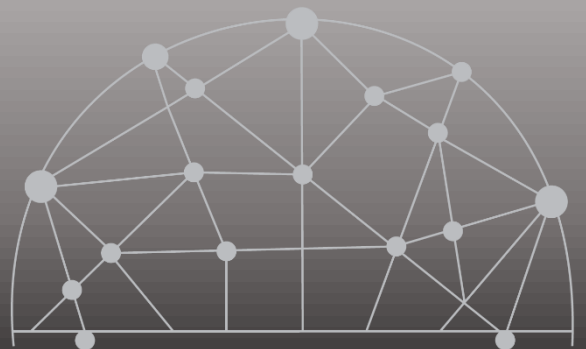
NVIDIA GRID Hardware

- Approved list of severs at (google://NVIDIA grid certified servers)
- Cisco, Dell, SuperMicro were the first
- Most major vendors have supported platforms now
- Do not mix K1 and K2 cards in the same system (per-OEM)

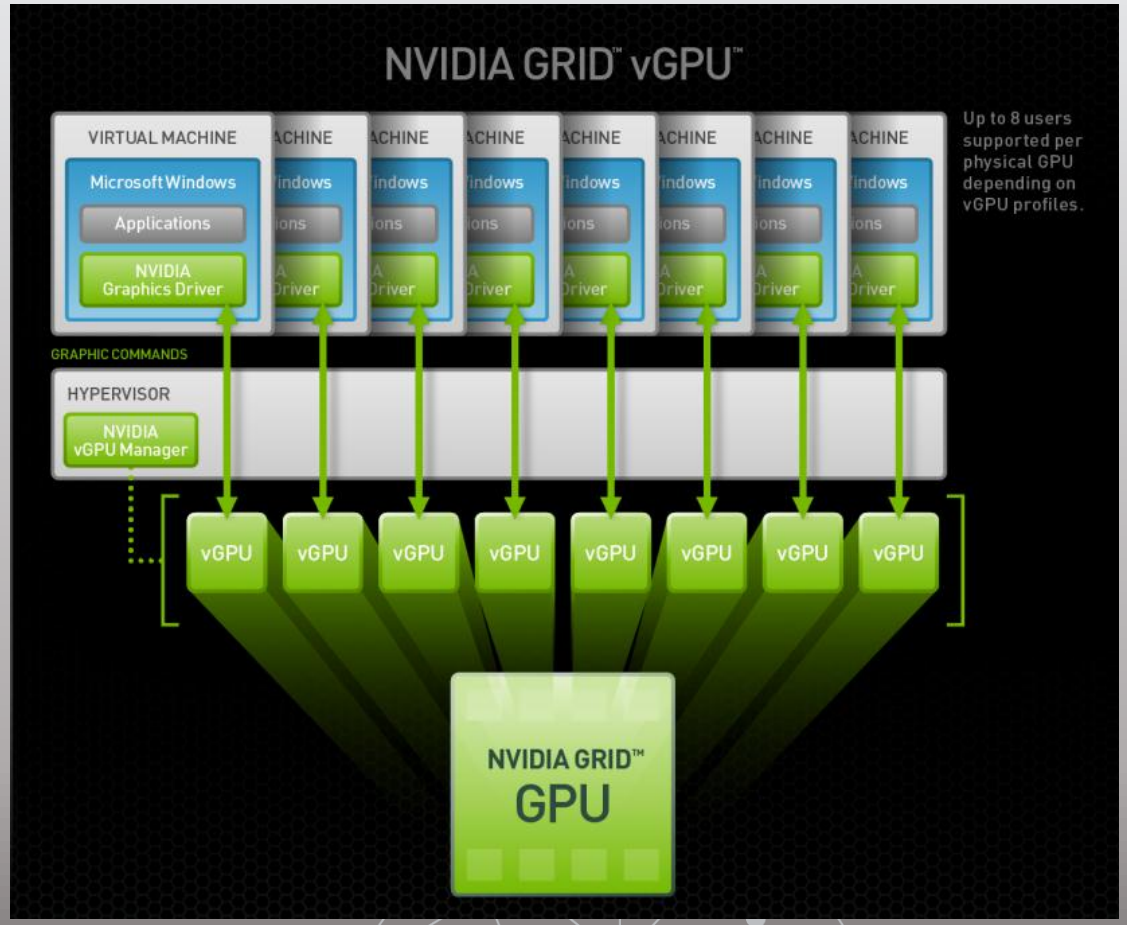


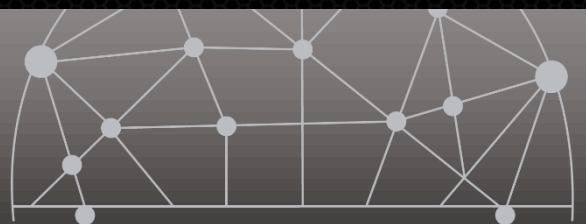
What is vGPU

- GPU Virtualization – Shared GPUs in a virtual environment
- Remote GPU – Delivering GPU-enabled applications to users



How does vGPU Work?



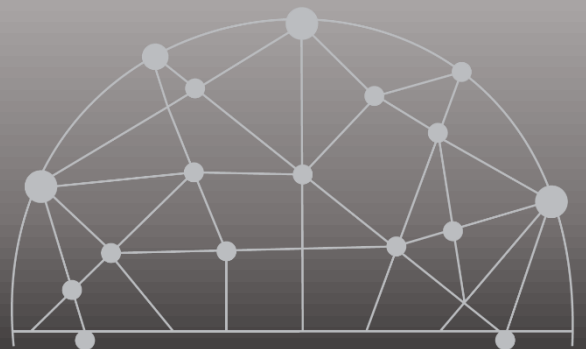


THE KANAVEL GROUP

NVIDIA

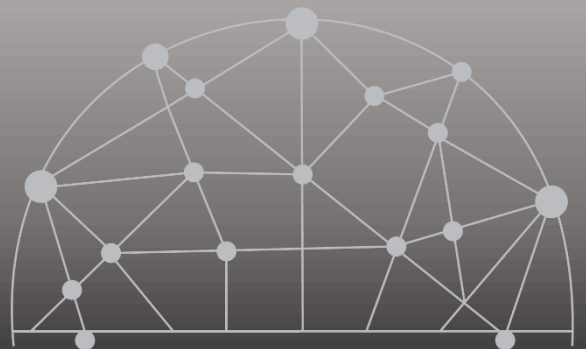
How does vGPU Work?

- Divide a GPU in to between 1 and 8 **EQUAL** pieces
- Divide the RAM by *n*
- Divide pipelines by *n*
- Custom GPU scheduler runs in hardware
- Proprietary driver on Guest VM is virtualization-aware
- Guesses about the future?



How does vGPU Work?

- Virtualization profiles
  - 1, 2, 4, 8 or pass-through
  - 1 vs pass-through?
    - pass-through enables guest to take full control of GPU (incl. CUDA and OpenCL)
    - “1” vGPU profile allows the hypervisor to monitor and control the GPU (DirectX, and OpenGL only)



THE KANAVEL GROUP

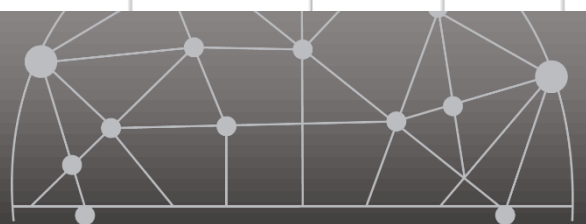
NVIDIA

How does vGPU Work?

vGPU Profiles

KxxxQ = Quadro Certified

Card	Physical GPUs	Virtual GPU	Intended Use Case	Frame Buffer (Mbytes)	Virtual Display Heads	Max Resolution per Display Head	Maximum vGPUs	
							Per GPU	Per Board
GRID K1	4	GRID K180Q	Power User	4096	4	2560x1600	1	4
		GRID K160Q	Power User	2048	4	2560x1600	2	8
		GRID K140Q	Power User	1024	2	2560x1600	4	16
		GRID K120Q	Power User	512	2	2560x1600	8	32
		GRID K100	Knowledge Worker	256	2	1920x1200	8	32
GRID K2	2	GRID K280Q	Designer	4096	4	2560x1600	1	2
		GRID K260Q	Power User, Designer	2048	4	2560x1600	2	4
		GRID K240Q	Power User, Designer	1024	2	2560x1600	4	8
		GRID K220Q	Power User, Designer	512	2	2560x1600	8	16
		GRID K200	Knowledge Worker	256	2	1920x1200	8	16

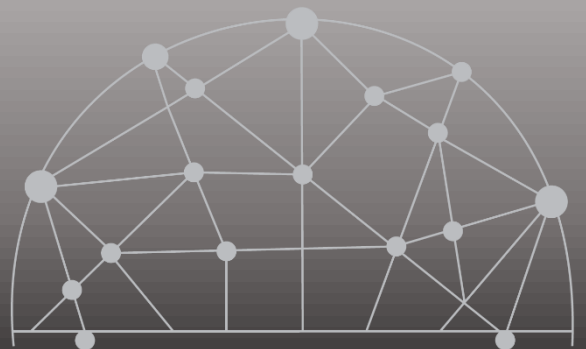


THE KANAVEL GROUP

NVIDIA

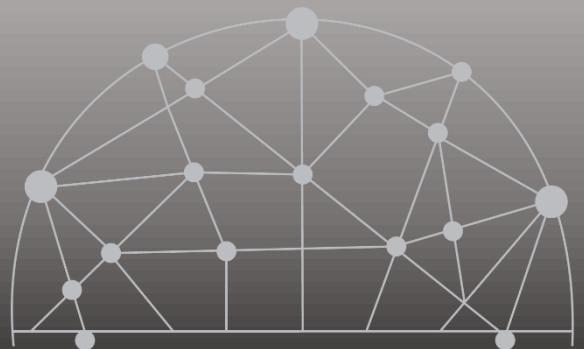
GRID Software Components

- Hypervisor
  - XenServer
  - VMWare
  - Hyper-V
- NVIDIA VGX
  - Driver and Tools
- VM Drivers



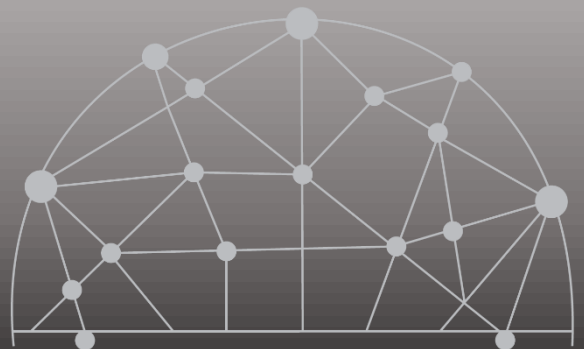
XenServer

- First hypervisor to support GRID (since 2013)
- Version 6.2
  - Service Pack 1, Hotfixes 9 and 11
- Version 6.5
  - Day-0 support planned
  - Enterprise or Desktop+ License Required



XenServer – Installation Overview

- XenServer: Installed and patched
- Install NVIDIA GRID RPM (nvidia-vgx-xenserver-6.2-340.57.i386.rpm)
- lsmod | grep -i NVIDIA (determine driver is loaded)
- nvidia-smi (determine the vGPU Manager is running)

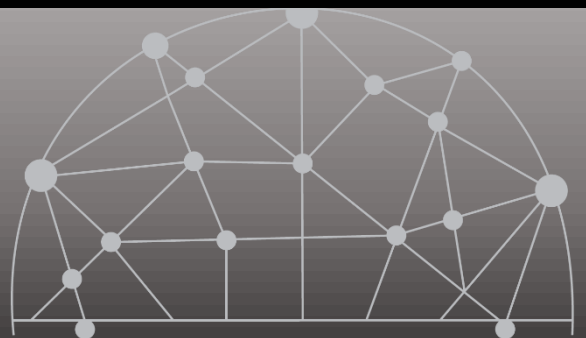


XenServer – Installation Overview

NVIDIA-smi sample output

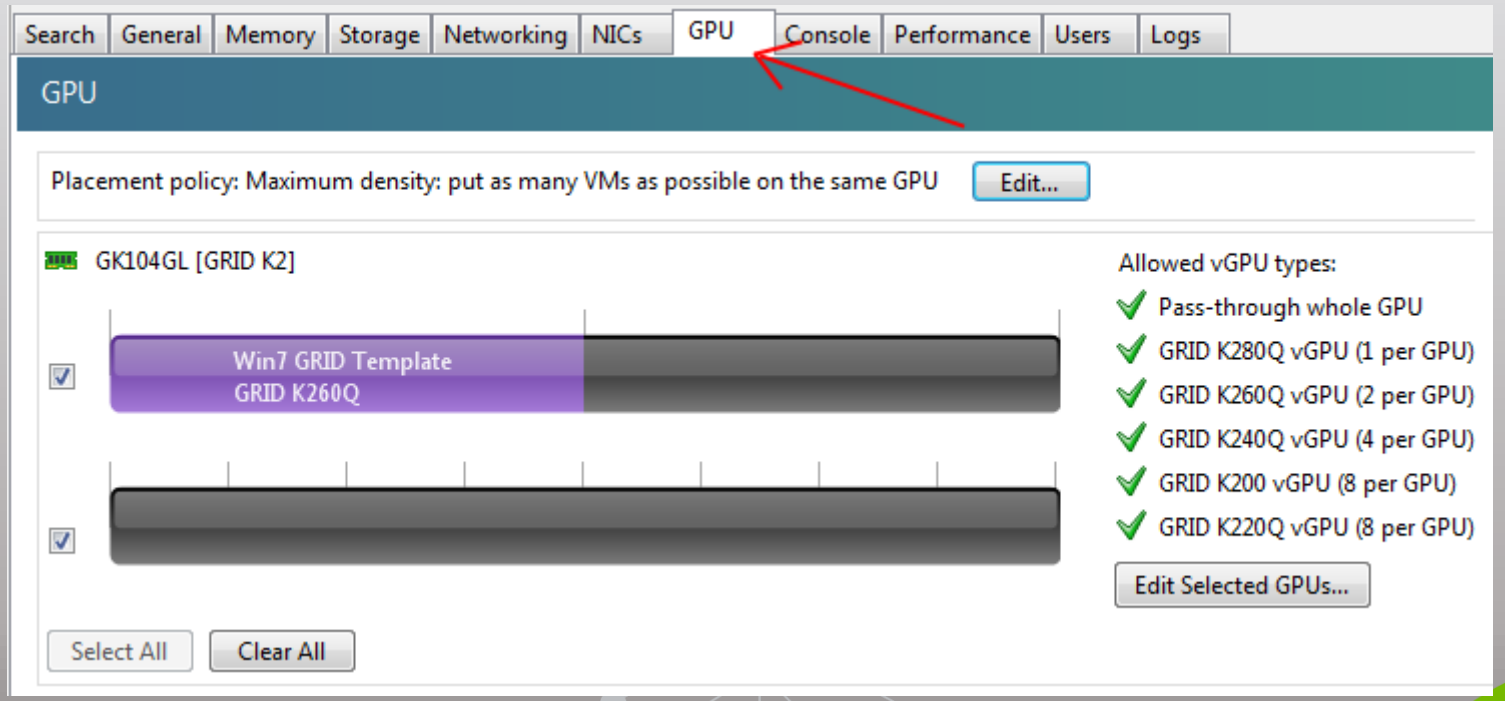
```
+-----------------------------------------------------+
| NVIDIA-SMI 340.57     Driver Version: 340.57        |
|-------------------------------+----------------------+----------------------+
| GPU  Name        Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf  Pwr:Usage/Cap|         Memory-Usage | GPU-Util  Compute M. |
|===============================+======================+======================|
|   0  GRID K2             On   | 0000:05:00.0     Off |                  Off |
| N/A   41C    P8    26W / 117W |   1904MiB /  4095MiB |      0%      Default |
+-------------------------------+----------------------+----------------------+
|   1  GRID K2             On   | 0000:06:00.0     Off |                  Off |
| N/A   34C    P8    26W / 117W |     11MiB /  4095MiB |      0%      Default |
+-------------------------------+----------------------+----------------------+

+-----------------------------------------------------------------------------+
| Compute processes:                                               GPU Memory |
|  GPU       PID  Process name                                     Usage      |
|=============================================================================|
|    0     16319  /usr/lib/xen/bin/vgpu                             1856MiB |
+-----------------------------------------------------------------------------+
```



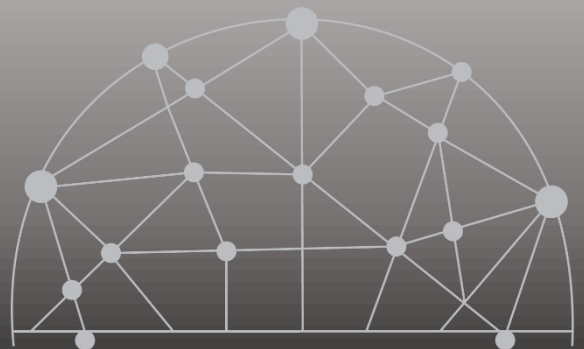
XenServer – Installation Overview

"GPU" tab on the host



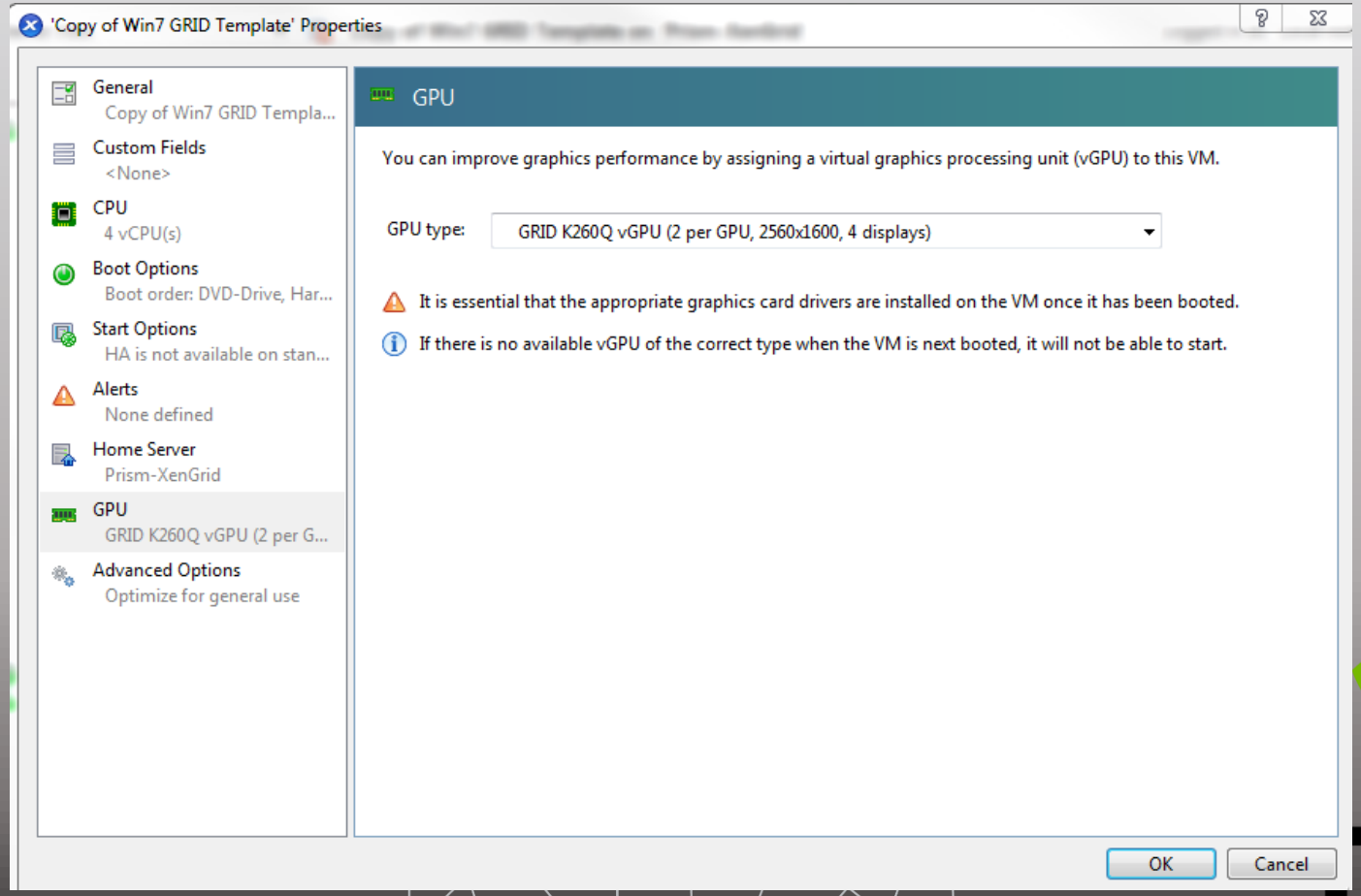
XenServer – Installation Overview

- Install Guest OS (Windows 7/8) normally
- Install XenServer Guest Tools
- Assign GPU
- Install NVIDIA GRID driver for guest (next, next, finish)



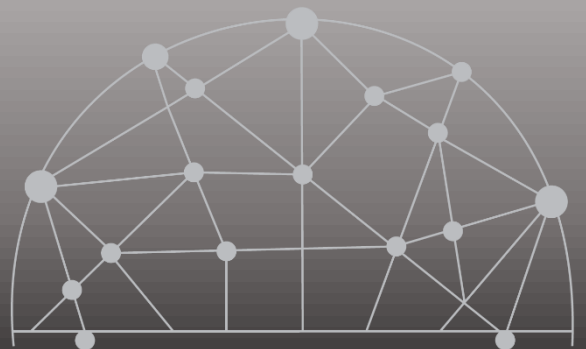
XenServer – Installation Overview

Assigning a GPU to a VM



XenServer – Installation Overview

- Disable VGA Console on VM
  - **First, make sure Remote Desktop is enabled**
  - `xe vm-list name-label=VM\ Name` (use backslash-space for a space)
  - `xe vm-param-set uuid=[`from above`] platform:vgpu_extra_args="disable_vnc=1"`

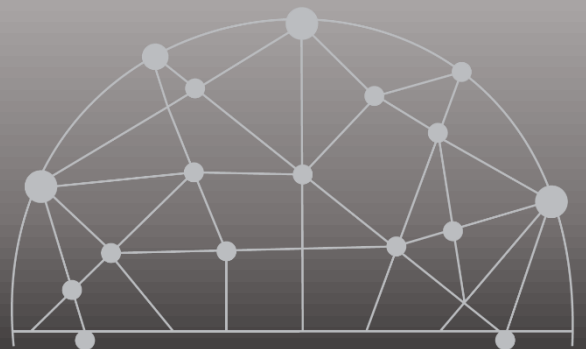


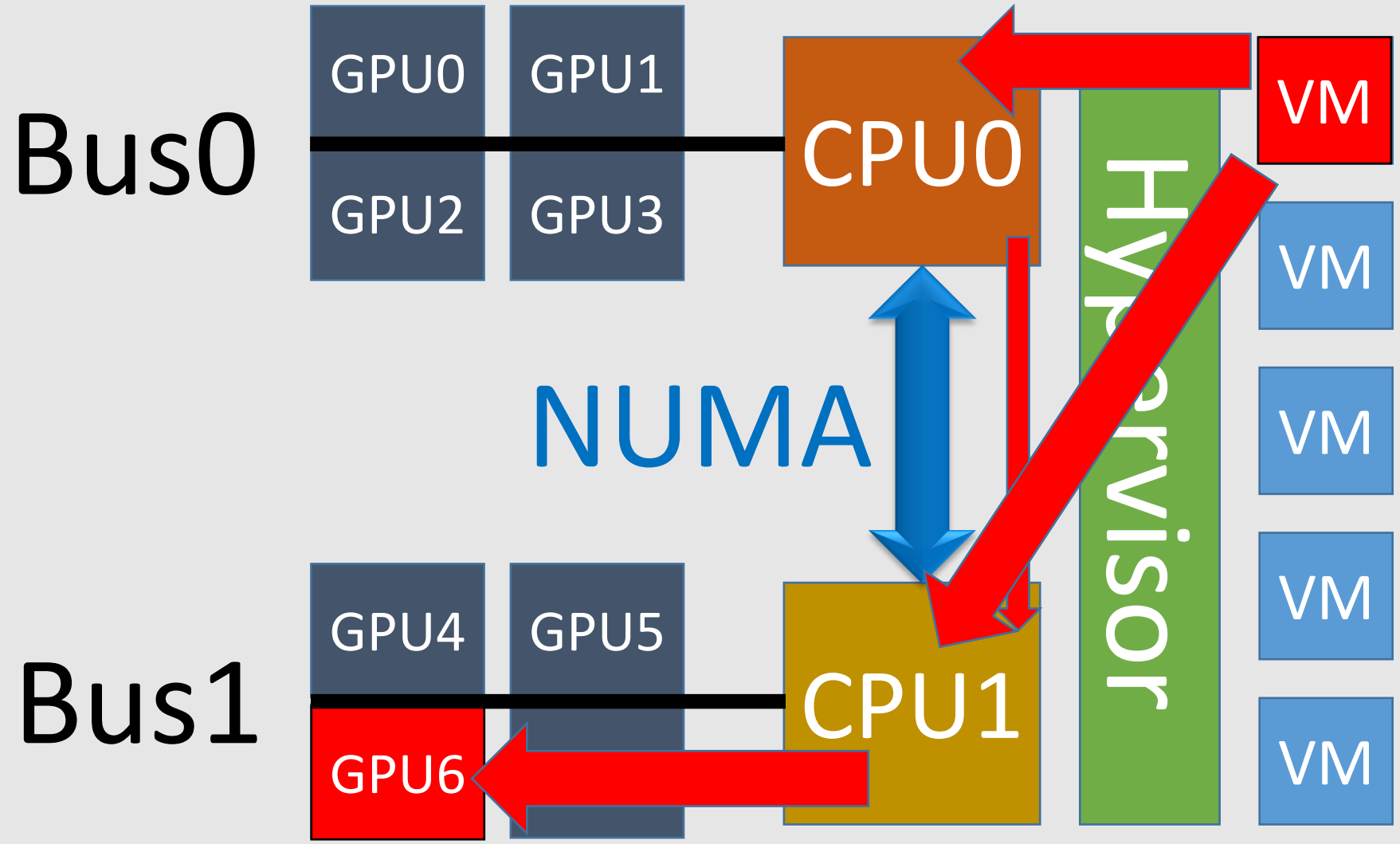
XenServer – Tuning

- GPU <> CPU pinning
  - Each CPU in a system controls a PCI bus
  - Make sure all GPU-enabled VMs are using the appropriate CPU to prevent sending requests across QPI/HT bus (see: NUMA)

```
[redacted]# lspci |grep -i nvidia
05:00.0 VGA compatible controller: NVIDIA Corporation GK104GL [GRID K2] (rev a1)
06:00.0 VGA compatible controller: NVIDIA Corporation GK104GL [GRID K2] (rev a1)
```

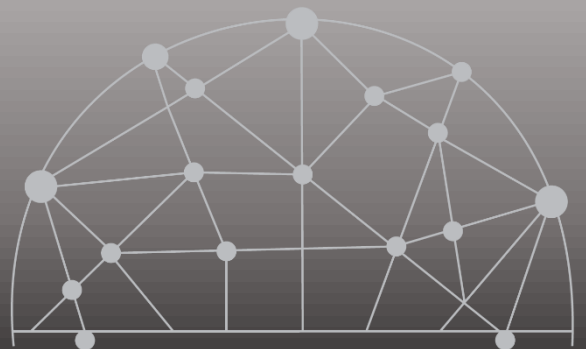
  - If the bus address starts with 0x:, GPU is on CPU0




Bus0
GPU0
GPU1
GPU2
GPU3
CPU0
VM
Hypervisor
VM
VM
VM
VM
NUMA
Bus1
GPU4
GPU5
GPU6
CPU1

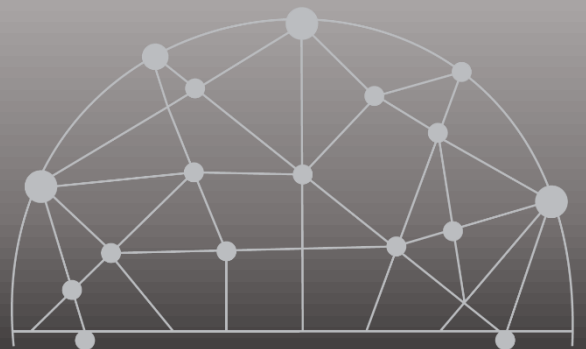
XenServer – Tuning

- Determining physical GPU
  - `xe vm-list name-label=VM\ Name` (returns UUID)
  - `xe vgpu-list vm-uuid=[from previous]` (returns UUID)
  - `xe vgpu-param-get uuid=[from previous]param-name=resident-on` (returns UUID)
  - `xe pgpu-param-list uuid=[from previous]`
  - Look for pci-id parameter



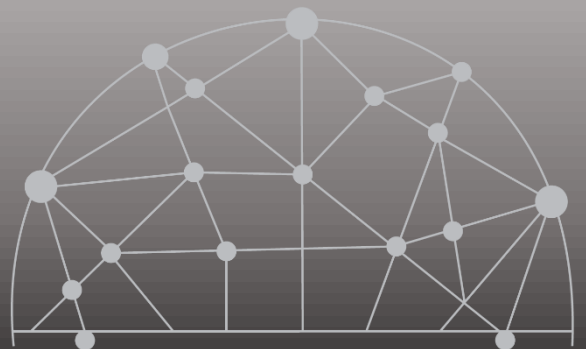
XenServer – Tuning

- To set CPU preference
  - `xe vm-param-set uuid=[UUID] VCPUs-params:mask=`n_0,n_1,n_2,n_3…
  - where n_0 is the starting core number
  - `xe vm-param-set uuid=[UUID] VCPUs-params:mask=0,1,2,3,4,5`
    - For CPU0 of a 6-core system (don't forget about hyperthreading)
- Cores versus sockets
  - XenServer presents sockets by default



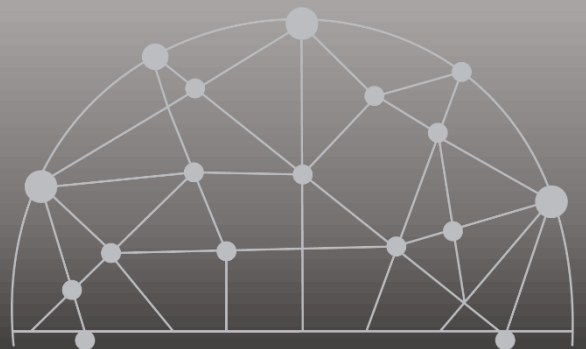
XenServer – Tuning

- Cores versus sockets
  - Windows licenses by socket but not by core
  - `xe param-set uuid=[uuid] platform:cores-per-socket=[2,4,…] VCPUs-max=[cores] VCPUs-at-startup=[cores]`
  - `xe param-set uuid=[uuid] platform:cores-per-socket=4 VCPUs-max=4 VCPUs-at-startup=4`
  - This is finally in the GUI for XenServer 6.5 (XenCenter)



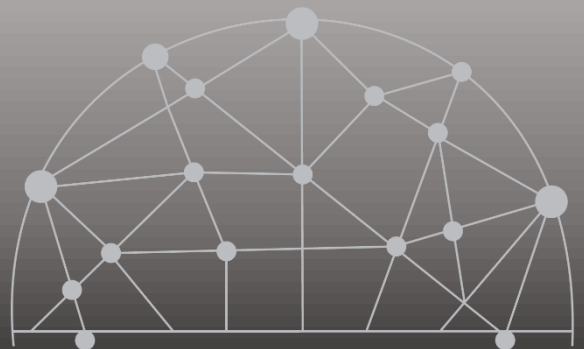
VMWare vSphere

- Pass-through is fully supported
- Virtualization *presently* available through vSGA
  - Does not support NVIDIA extensions (APIs)
  - DirectX 10,11 not supported, OpenGL 3.0+ not supported
  - Software virtualization severely hampers performance
  - Does allow virtualization of non-GRID hardware (GTX, Quadro, AMD)



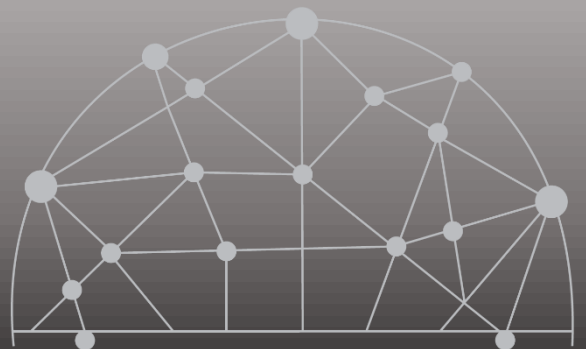
VMWare vSphere

- Full GRID support announced for vSphere 6
  - Tech preview supposedly available if you ask nicely
  - GPU Profiles supported by Horizon View 6 with PCoIP



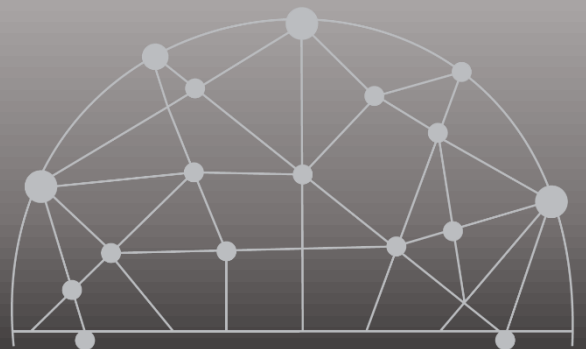
Hyper-V

- Pass-though support only using RemoteFX
- Useful only for Remote Desktop over LAN
- Some RemoteFX encapsulation in Citrix HDX



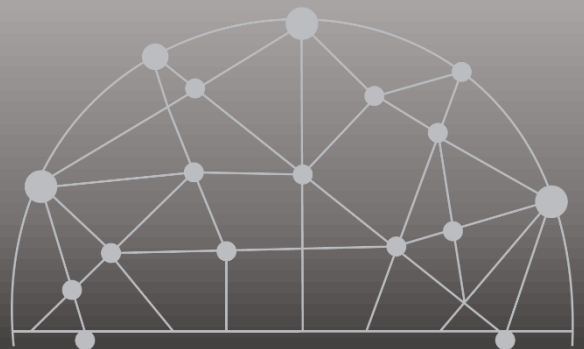
Delivering vGPU Enabled Workloads

- Citrix XenDesktop
- Citrix XenApp
- Horizon View



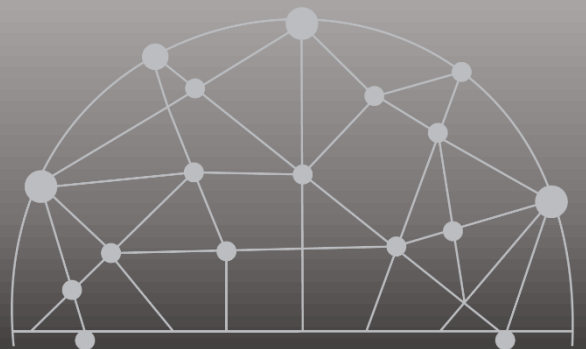
Citrix XenDesktop

- Only XenDesktop includes HDX 3D Pro
  - Adaptive H.264 compression and encapsulation
  - Some rendering is performed client-side when appropriate
  - Most scalable protocol for WAN deployments (lowest bandwidth)
  - FrameHawk acquisition will further increase scalability



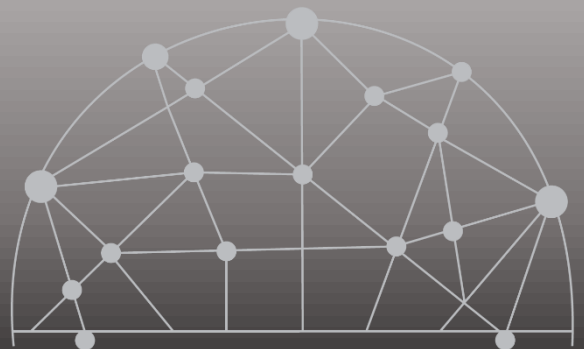
Citrix XenDesktop

- Enable HDX 3D Pro by installing the HDX 3D Pro component when installing the Virtual Desktop Agent (VDA)
- 3D hardware is automatically detected and utilized
- Client-side GPU hardware is automatically detected and utilized as long as the client is running the latest Citrix Receiver



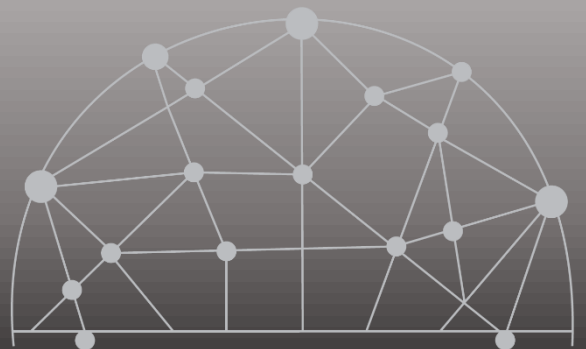
Citrix XenDesktop Lossless HDX

- Workloads can be set to “lossless” mode
- When combined with a Quadro-certified workload it is suitable for all 3D purposes including medical
- Every frame will be rendered on the client no matter the latency
- This will kill performance on WAN. Never use unless needed.



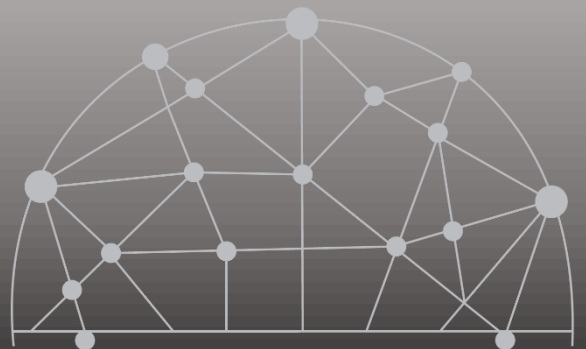
XenDesktop on vSphere

- XenDesktop has always been supported on vSphere
- There is no reason to think that Citrix will not support GRID-enabled desktops on vSphere 6



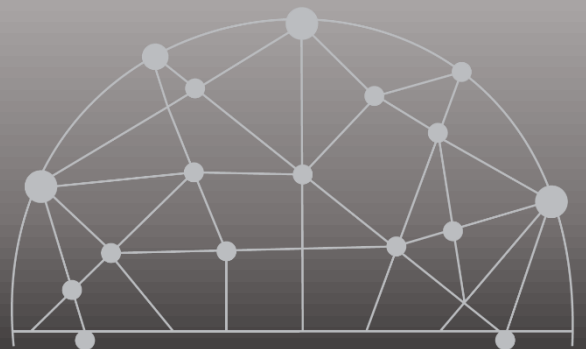
XenApp

- Built on Windows Server with Terminal Services/Remote Desktop
- Many users on one VM
- No VDA license required, TS/RDS CAL only
- No USB peripherals
- Users share resources
- Your mileage will vary



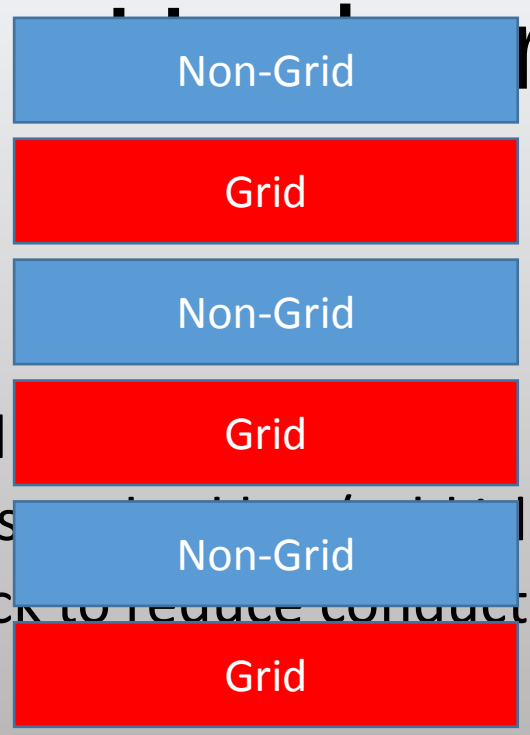
Horizon View

- The Teradici PCoIP protocol has been delivering pass-through GPU workloads as well as vSGA for some time. Little is expected to change with the addition of vGPU.
- When deciding between Citrix and VMWare, bear in mind the cost of bandwidth. PCoIP requires significantly more bandwidth and can increase the cost of deployment to branch offices.



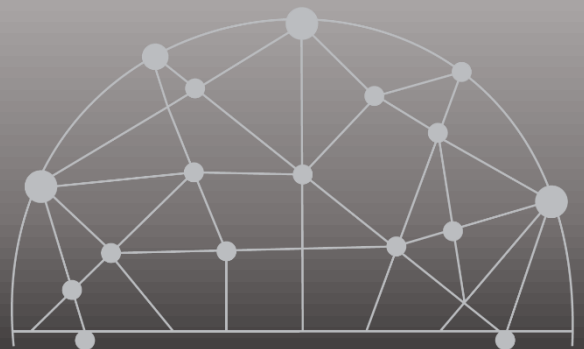
Design Considerations [illegible]re

- Server cooling
  - A K2 card can consume up to 225W
  - 3 K2 cards means up to 675W additional
  - A large GRID farm should be setup with s[illegible]les
  - Consider staggering workloads in the rack to reduce conductive heat transfer



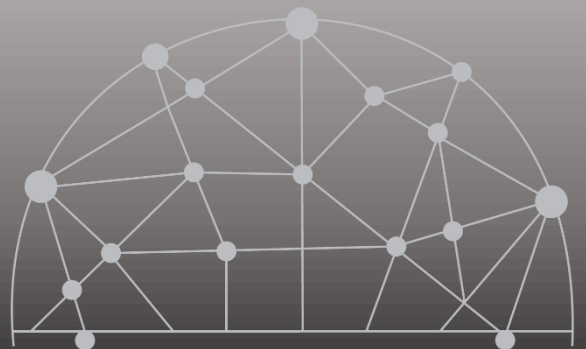
Design Consideration - General

- 3D workload-enabled VMs require more storage (80+ GB)
- Swap files will be larger and require more IOPS
- GPU-enabled VMs *cannot* be xenMotioned/vMotioned to other hosts
  - Therefore: storing static images on a SAN may not be needed
  - Local SSD can give more IOPS than a SAN for less cost
  - If the server fails, VM may be lost or unavailable



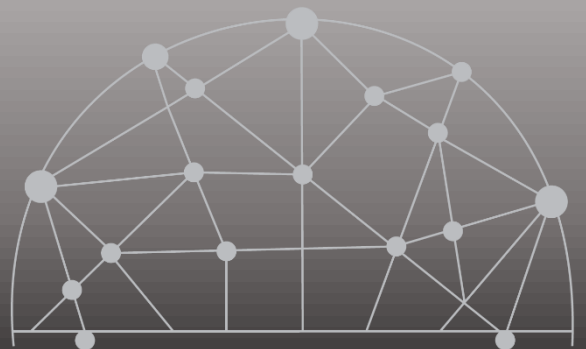
Design Consideration

- If you have “pinned” a VM to a GPU<>CPU, you will have to re-do it when a VM is moved to another host
- VM density per physical server will probably be GPU limited
  - 3 x K2 Cards with K260Q Profile = 12 VMs
  - 12 VMs x 16GB RAM = 192GB (up to 768GB in a server)
  - CPU constraining is rare
  - Stagger GPU and non-GPU workloads for balance



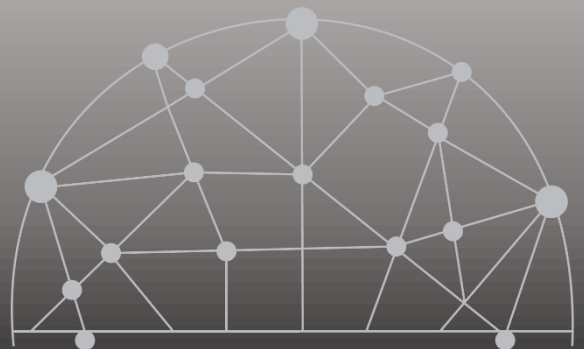
Design Considerations

- User Profile Management
  - If users will be sharing VMs or moving between them
  - Uncoupling user settings from the app workload
  - Microsoft Roaming Profiles and Folder Redirection
  - Citrix UPM and VMWare Persona
  - AppSense and RES
- Use your newest servers with the fastest CPUs



Design Considerations

- Memory
  - Application Memory + GPU Memory
  - i.e. 4GB of RAM for Apps/OS + K280Q GPU with 4GB = 8GB per VM



Thank You

Thank you for attending my session.

Questions?

Ask me now

or email me: gtaylor@kanavelgroup.com

THE KANAVEL GROUP