

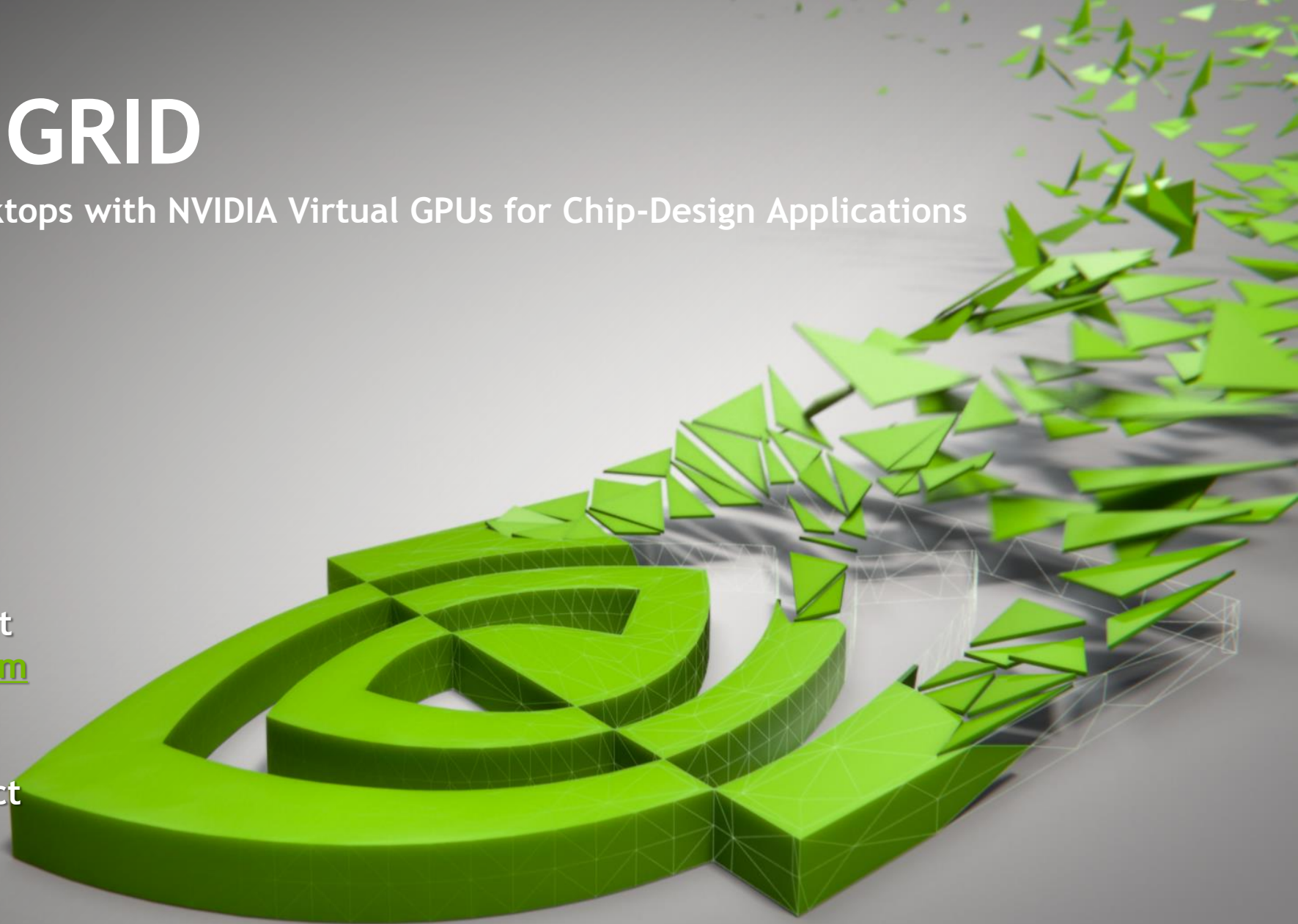
# NVIDIA GRID

Linux Virtual Desktops with NVIDIA Virtual GPUs for Chip-Design Applications



Shailesh Deshmukh  
Sr Solutions Architect  
[ShaileshD@Nvidia.com](mailto:ShaileshD@Nvidia.com)

Eric Kana  
Sr. Solutions Architect  
[Ekana@Nvidia.com](mailto:Ekana@Nvidia.com)



# AGENDA

- *Introducing VDI Linux Desktops*
- *Use Cases for VDI*
- *Remote Access Solutions*
- *Supporting Hypervisor Stacks*
- *Virtualized GPU Options*
- *Demonstration*
- *Troubleshooting*
- *Summary*

# LINUX VDI DESKTOPS

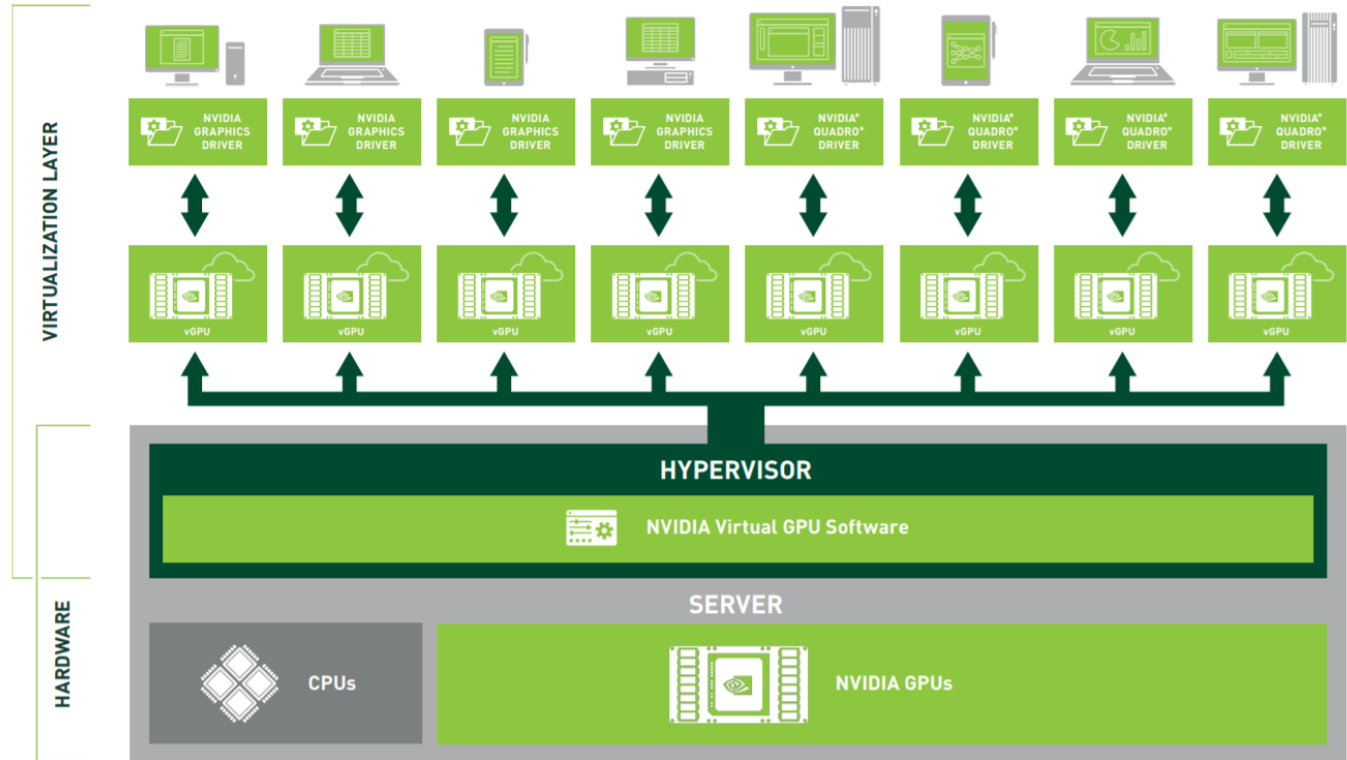
## Virtual Desktop Infrastructure

End Points/Users

Virtualized Components

Hypervisor

Chassis Hardware



# LINUX VDI DESKTOPS

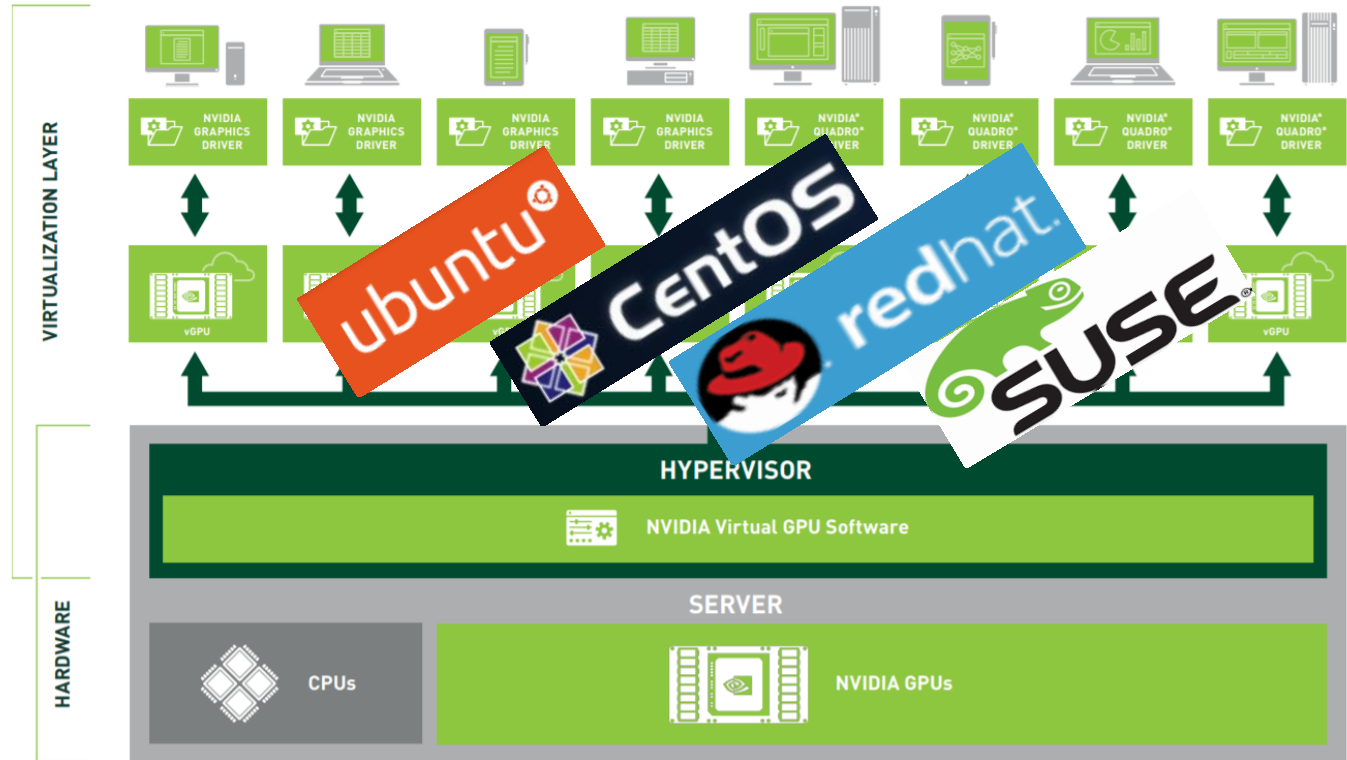
## Virtual Desktop Infrastructure

End Points/Users

Virtualized Components

Hypervisor

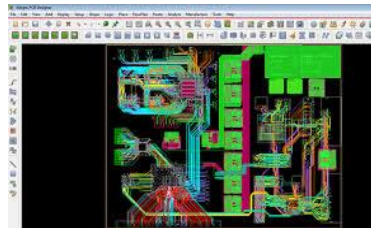
Chassis Hardware



# WHY LINUX VDI ??

## Common Linux Use Cases

- Batch processing / job scheduling workloads ( HPC / DL / AI )
- **2D application rendering and compute** ( ANSYS / Cadence / etc .. )
- Rendering / Transcoding servers ( Wowza / VLC Live / ESPN Live .. )
- Streaming Servers ( Netflix / Hulu Note: GPU in HPC mode )
- Design shops / visual effects ( VFX / Pixar / Dreamworks .. )



# AVAILABLE REMOTE ACCESS SOLUTIONS

## Virtual Desktop Agents

- **VNC**
  - Pros: Enterprise and Free versions are available
  - Cons:
    - No management console
    - JPEG based transcoding may introduce jitter, poor video quality, performance, etc.
    - Not ideal for low bandwidth connections
- **Exceed On Demand**
  - Pros: Proven technology and widely used at many locations
  - Cons:
    - Slightly older technology
    - License cost may be higher ?
    - Sometime has issues with high resolution monitors ( 4K )
- **HP RGS**
  - Pros: HP Proprietary technology
  - Cons:
    - Individual management of instances
    - May have compatibility issues with new graphics cards and drivers

# VIRTUALIZATION STACK OPTIONS

Pass-through / VMware vSphere / CITRIX XenServer / KVM / Nutanix AHV

## Pass-through

- Traditional method of exposing hardware to a Virtual Machine
- High performance
- Poor Scaling
- Supported by most NVIDIA Tesla and (high-end Quadro) Solutions

## Commercial Hypervisor Solutions for virtualized GPU

- VMware Hypervisor
- CITRIX XenServer
- Nutanix AHV
- KVM

# VMWARE VSPHERE

## vSphere 6.0+

Hardware: Vmware Compatibility Guide

<https://www.vmware.com/resources/compatibility/search.php?deviceCategory=sptg>

Software : Horizon View 7.X Family

Horizon for Linux: <https://www.vmware.com/products/horizon/horizon-linux.html>

Horizon View Agent supported OS for Linux:

Ubuntu: 12.04 , 14.04 , 16.04 ( Note for 14\* and 16\* you must do <http://kb.vmware.com/kb/2151294>.)

Cent / RHEL: 6.6 , 6.7 , 6.8 , 7.2 , 7.3 , 7.4

SLES: 12 SP1 / SP2 ( version GRID 6.0 onwards )

SLED: 11 SP3/SP4, SLED 12 SP1/SP2

NeoKylin 6 Update 1



# CITRIX XENSERVER

Hardware: XenServer Compatibility Guide

<http://hcl.xenserver.org/gpus/>

Software : XenDesktop 7.11 and above

Citrix Linux VDA: <https://www.citrix.com/blogs/tag/linux-vda/>

XenDesktop VDA agent:

Ubuntu: 16.04 ( with Kernel 4.4.X)

RHEL: 6.6 , 6.7 , 6.8 , 7.2 , 7.3 , 7.4

CentOS: 6.7 , 6.8 , 7.2

SLES: 11 SP4 / 12 SP1

# NVIDIA TESLA GPUS

All Pascal GPUs Supported, P40, P100, P4 and V100 are Recommended

	M10	M60	P40	M6	P6
<b>GPU</b>	4 NVIDIA Maxwell GPUs	2 NVIDIA Maxwell GPUs	1 NVIDIA Pascal GPU	1 NVIDIA Maxwell GPU	1 NVIDIA Pascal GPU
<b>CUDA Cores</b>	2,560 (640 per GPU)	4,096 (2,048 per GPU)	3,840	1,536	2,048
<b>Memory Size</b>	32 GB GDDR5 (8 GB per GPU)	16 GB GDDR5 (8 GB per GPU)	24 GB GDDR5	8 GB GDDR5	16 GB GDDR5
<b>H.264 1080p30 streams</b>	28	36	24	16	24
<b>Max vGPU instances</b>	64 (512 MB Profile)	32 (512 MB Profile)	24 (1 GB Profile)	16 (512 MB Profile)	16 (1 GB Profile)
<b>vGPU Profiles</b>	0.5 GB, 1 GB, 2 GB, 4 GB, 8 GB	0.5 GB, 1 GB, 2 GB, 4 GB, 8 GB	1 GB, 2 GB, 3 GB, 4 GB, 6 GB, 8 GB, 12 GB, 24 GB	0.5 GB, 1 GB, 2 GB, 4 GB, 8 GB	1 GB, 2 GB, 4 GB, 8 GB, 16 GB
<b>Form Factor</b>	PCIe 3.0 Dual Slot (rack servers)	PCIe 3.0 Dual Slot (rack servers)	PCIe 3.0 Dual Slot (rack servers)	MXM (blade servers)	MXM (blade servers)
<b>Power</b>	225W	240W / 300W (225W opt)	250 W	100W (75W opt)	90 W (70W opt)
<b>Thermal</b>	passive	active / passive	passive	bare board	bare board
	<b>USER DENSITY</b> Optimized	<b>PERFORMANCE</b> Optimized		<b>BLADE</b> Optimized	

# NVIDIA TESLA GPUS CONTINUED ..

All Pascal GPUs Supported, P40, P100, P4 and V100 are Recommended

	P4	P100	New V100
GPU	1 Pascal GPU	1 Pascal GPU	1 Volta GPU
CUDA Cores	2560	3584	5120
Memory Size	8 GB GDDR5	16 GB GDDR5	16 GB GDDR5
H.264 1080p30 streams	24	36	36
Max vGPU instances	8	16	16
vGPU Profiles	1 GB, 2 GB, 4 GB, 8 GB	1 GB, 2 GB, 4 GB, 8 GB, 16GB	1 GB, 2 GB, 4 GB, 8 GB, 16GB
Form Factor	Low Profile PCI Express 3.0 x16	PCIe 3.0 Dual Slot	PCIe 3.0 Dual Slot
Power	75W	250W	250W
Thermal	passive	passive	passive
PERFORMANCE Optimized			

# SUGGESTED GPU SOLUTIONS

## VDI for Linux

**Pass-through:** M60 , M6 , P40 , P4 , P6 , P100 , soon .. V100

(why not M10 ? :- this card is best for high density, light graphical workloads)

**vGPU:** M60 , P4, P6, P40 , P100 soon .. V100

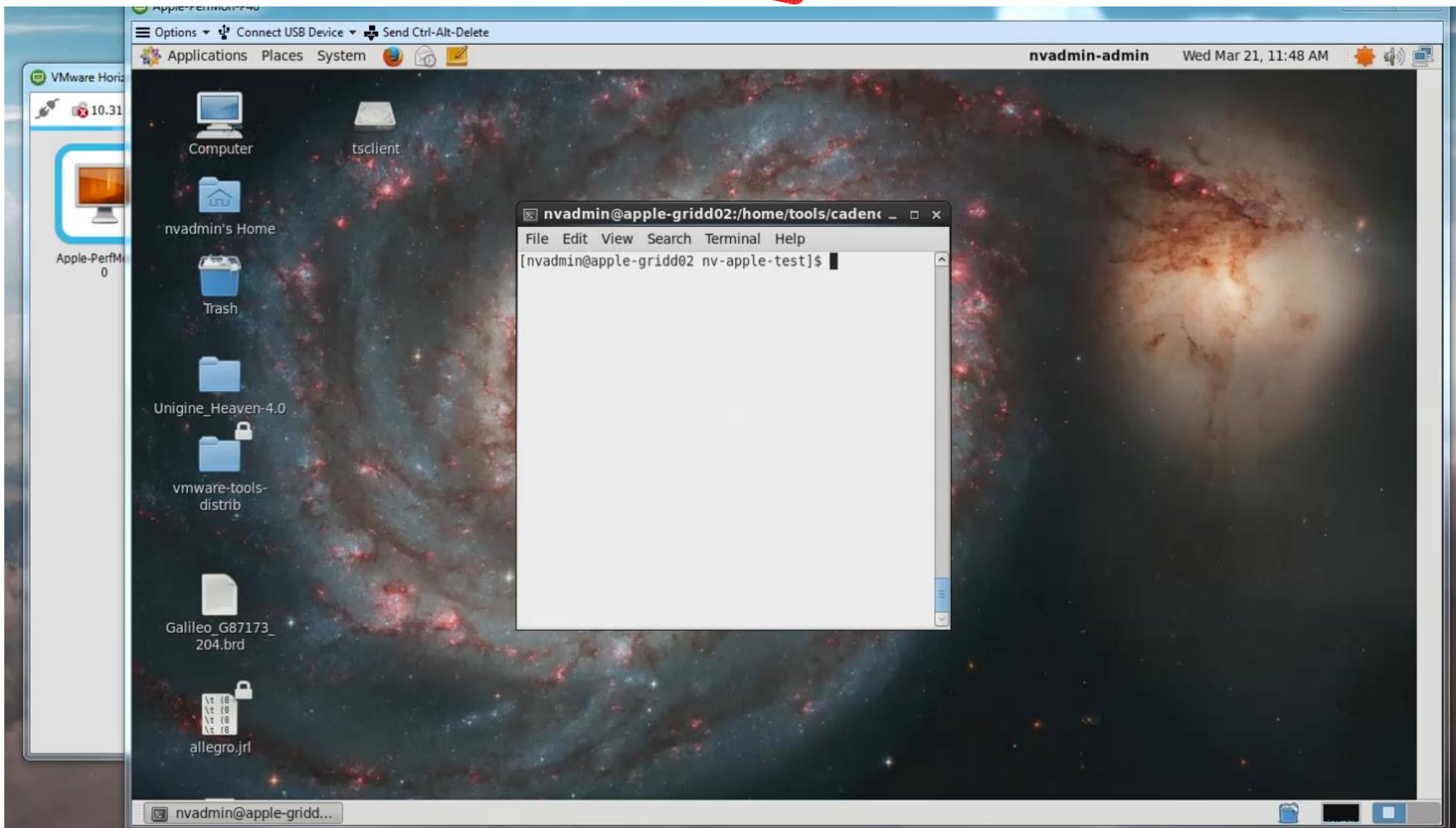
### Linux Variants:

- RHEL / Cent : 6.6 , 6.7 , 6.8, 6.9 , 7.X
- Ubuntu: 12.04, 14.04 , 16.04 , 17.04 ( make sure you are running latest kernel )

### GRID Profiles:

- Supported GRID Profiles: All Q profiles, 1B and 2B Profiles,
- No support for vAPP profiles

DEMO



# TROUBLESHOOTING

## Ubuntu Login Loop with vGPU Driver installation !!!

Symptoms:

After installing the Nvidia vGPU driver, user can no longer login, the login screen refreshes after the username and password, prompting it again ...

Yes this is a the login loop....

Affected distros: 14.04, 16.04, may be 17.04 ?

Possible solutions: Run the following commands-

```
# ldconfig
```

```
# ubuntu-drivers list
```

```
# ubuntu-drivers autoinstall
```

```
# reboot
```

# TROUBLESHOOTING

Ubuntu only showing in one screen , Citrix XenDesktop!!!

Symptoms:

Everything is installed correctly but the XenDesktop session is only showing in one screen

Affected distros: All Linux VDIs

Possible solutions: Unlink the Windows VDA, Citrix's linux VDA is stupid ... that's right, you need to bring the session in "windowed mode" , then stretch it across all the monitors and hit "FullScreen" it will automatically resize 😊 , this may not work if you have 4K monitors .. The solution is on the next screen -

# TROUBLESHOOTING

Citrix VDA with Nvidia vGPU multiple 4K screens ( customers may be getting GRAY bars in the VDA session):

Run this command on the VDI :

```
# ./ctxreg dump | grep MaxFbHeight
```

```
# ./ctxreg dump | grep MaxFbWidth
```

This will tell you the session max resolution, now to adjust it for 4K displays, run the following -

```
/opt/Citrix/VDA/bin/./ctxreg create -k " HKEY_LOCAL_MACHINE  
\System\CurrentControlSet\Control\Citrix\Thinwire" -t "REG_DWORD" -v " MaxFbWidth" -d " 16384" -force
```

```
/opt/Citrix/VDA/bin/./ctxreg create -k " HKEY_LOCAL_MACHINE  
\System\CurrentControlSet\Control\Citrix\Thinwire" -t "REG_DWORD" -v " MaxFbHeight" -d " 16384" -force
```

This will stretch the session to 4X4 K display 😊



# TROUBLESHOOTING

## Citrix VDA with Nvidia vGPU multi-screen support :

Run this command on the VDI :

```
# ./ctxreg dump | grep MaxScreenNum
```

This will tell you the number of screens supported , not to adjust run the following -

```
#/opt/Citrix/VDA/bin/./ctxreg create -k "HKLM\System\CurrentControlSet\Control\Citrix\Thinwire" -t "REG_DWORD" -v "MaxScreenNum" -d "0x00000004" -force
```

This will force the number of screen to 4

# TROUBLESHOOTING

## Citrix how to check if Hardware encoding or H.264 is used ?

Run this command on the VDI :

```
./ctxreg dump | grep AdvertiseH264
```

```
./ctxreg dump | grep HardwareEncoding
```

This will tell you the number of screens supported , not to adjust run the following -

```
/opt/Citrix/VDA/bin/./ctxreg create -k "HKLM\System\CurrentControlSet\Control\Citrix\Thinwire" -t "REG_DWORD" -v "AdvertiseH264" -d "0x00000001" --force
```

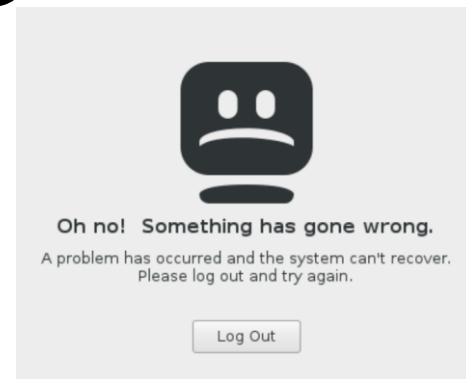
```
/opt/Citrix/VDA/bin/./ctxreg create -k "HKLM\System\CurrentControlSet\Control\Citrix\Thinwire" -t "REG_DWORD" -v "HardwareEncoding" -d "0x00000001" -force
```

This will force the hardware encoding and H.264 codec, be careful as this might result in high B/W for the users and may not be a good option for long distance users

More good stuff: <https://docs.citrix.com/en-us/linux-virtual-delivery-agent/7-13/configuration/configuring-graphics.html>

# TROUBLESHOOTING

CentOS 7.X, Oh snap something went wrong !!:



Possible solution :

There are some known issues with SELinux and Graphics driver, so we recommend to make sure the SELINUX is either in “permissive” or “disabled” mode.

You can check the current mode using

```
# cat /etc/sysconfig/selinux
```

Using you editor of choice (vi or nano or vim), edit the above file and configure the selinux mode. In this case, we have disabled the option

A screenshot of a terminal window titled "CentOS73 - VMware Remote Console". The terminal shows the command `cat /etc/sysconfig/selinux` and its output. The output includes comments about SELinux modes and the current configuration: `SELINUX=disabled`. The word "disabled" is highlighted in yellow in the original image. Other lines include `SELINUXTYPE=targeted`.

# TROUBLESHOOTING

Ubuntu, CentOS or RHEL , You do not appear to

Be using the Nvidia X drivers .....

Possible solutions:

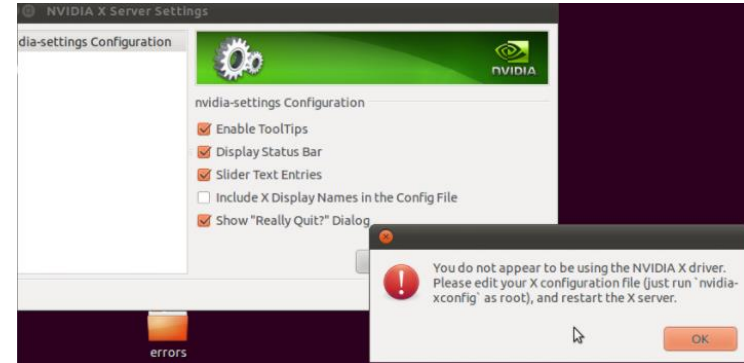
Step 1: run `# nvidia-smi` and make sure you get the correct response back

Step 2: run `# lspci | grep -i nvidia` ( the purpose is to find the bus ID )

Step 3: edit the `xorg.conf` , adjust the bus-

```
Section "Device"
Identifier "Device0"
Driver "nvidia"
VendorName "NVIDIA Corporation"
BusID "PCI:2:1:0"
Option "IgnoreDisplayDevices" "CRT-0"
Option "NoFlip" "TRUE"
Option "ConnectedMonitor" "DFP-0, DFP-1, DFP-2, DFP-3"
Option "CustomEDID" "DFP-0:/etc/X11/edid.bin; DFP-1:/etc/X11/edid.bin; DFP-2:/etc/X11/edid.bin; DFP-3:/etc/X11/edid.bin"
Option "ModeValidation"
"noVesaModes,noPredefinedModes,noXserverModes,noMaxSizeCheck,noEdidMaxPclkCheck,noHorizSyncCheck,noVertRefreshCheck,allowNonE
didModes"
EndSection
```

( Note: if your `xorg.conf` does not have this section, you can copy paste the above to adjust the `xorg.conf` , reboot the VM afterword's )



# FINAL THOUGHTS

## VDI for Linux

- Linux in a VDI environment is supported by both commercial hypervisors and NVIDIA GPUs
- Offers an alternative to commonly deployed Windows VDI
- Choose the right GPU and virtualized GPU profile for the job
- Consider Pascal or Volta GPUs for best VDI-friendly operation
- Consider Profiling Workloads for appropriate resource utilization
- Linux in VDI enables 'lite' HPC and compute for the masses

# NVIDIA GRID RESOURCES



GRID Test Drive

[www.nvidia.com/trygrid](http://www.nvidia.com/trygrid)

---



GRID Website

[www.nvidia.com/grid](http://www.nvidia.com/grid)

---



GRID News

<http://tinyurl.com/gridinfo>

---



GRID YouTube Channel

<http://tinyurl.com/gridvideos>

---



Questions? Ask on our Forums

<https://gridforums.nvidia.com>

---



NVIDIA GRID on LinkedIn

<http://linkd.in/QG4A6u>

---



Follow us on Twitter

[@NVIDIAGRID](https://twitter.com/NVIDIAGRID)

# Thank-You

