CITRIX

Getting Started with HDX 3D Pro

Reviewer's Guide for Remote 3D Graphics Apps

Part 2: vSphere GPU Pass-through

with XenDesktop 7 Apps,
Nvidia GRID K1/K2 cards,
Dell R720 Server

CITRIX

Getting Started with HDX 3D Pro

Contents

Audience

In the first part of this guide, we saw how to physically install Nvidia GRID cards with graphics processing units (GPU) in compatible server hardware. Part 1 also discusses how to enable GPU pass-through in XenServer and test that it is working in the VM. In this part, we list the steps necessary to enable shared GPU acceleration for 3D applications using Citrix XenDesktop 7 Apps on VMware vSphere. VMware calls this virtual Dedicated Graphics Acceleration (vDGA).

This guide walks through the following topics:

- Configuration of GPU acceleration on the hypervisors – vSphere 5.1
- Install, configure and assign GPU to a XenDesktop 7 WIndows Server VM
- Verify 3D applications are using the GPU
- Install and publish 3D applications and hosted shared desktop(s) using Desktop Studio
- Access 3D applications from Citrix Receiver on any device

It is assumed that the reader has good knowledge of networking, virtualization, server hardware, and Windows administration. Familiarity with Citrix and Nvidia products is recommended but not essential to complete these steps. Please see the resources section for more information.

Related Documents in this Series

Part 1: XenServer GPU pass-through for Citrix XenDesktop 7 (includes, physical installation of GPU cards)
Part 2: vSphere GPU pass-through (a.k.a vDGA) for Citrix XenDesktop 7
Part 3: XenServer GPU virtualization (a.k.a vGPU) for Citrix XenDesktop 7
Part 4: vSphere shared GPU (a.k.a vSGA) for Citrix XenDesktop 7

About the Authors

Pushpal Ray and **Mayunk Jain** in the Technical Marketing team of Citrix XenDesktop produced this guide.

Pushpal (@pushpalray) is a Technical Marketing Engineer with over 10 years experience in 3D graphics, infrastructure management, and virtualization. Mayunk (@mayunkj) is responsible for competitive marketing, technical demos, and sales enablement for the desktop and cloud solutions at Citrix.

Lab Environment

Hardware	
Graphical Processing Unit (GPU)	NVIDIA GRID K1 (K1 and K2 Specs)
Server hardware	Dell R720 (PowerEdge R720 Technical Guide)
GPU Installation Kit	• Power Cables (2 – Internal for GPU) • Heat Sink
Storage	Local/ NFS

Software	
Hypervisor(s)	VMware ESXi 5.1.0 build 838463 XenServer 6.2.0-rc4 build 69934c
NVIDIA GPU driver	320.00 (GRID K1)
Guest OS	Windows Server 2008 R2 Standard Service Pack 1 Windows 7 Service Pack 1

Go to *Control Panel* → *Add/Remove Programs* and ensure the following components are updated on your target virtual machine before you begin the 3D optimization process.

Tools and Applications	
Hypervisor Tools (latest)	VMware Tools XenServer Tools
Windows Applications	Adobe Flash Player Adobe Reader Java Plugin Microsoft .NET Framework 4 (latest)
GPU statistics (free third-party utilities)	TechPowerUp GPU-Z OpenGL Viewer

The process for enabling 3D acceleration for shared user-sessions is quite similar with certain customizations for each hypervisor. Figure 1 and Figure 2 shows the difference between Citrix and VMware implementation of GPU pass-through.

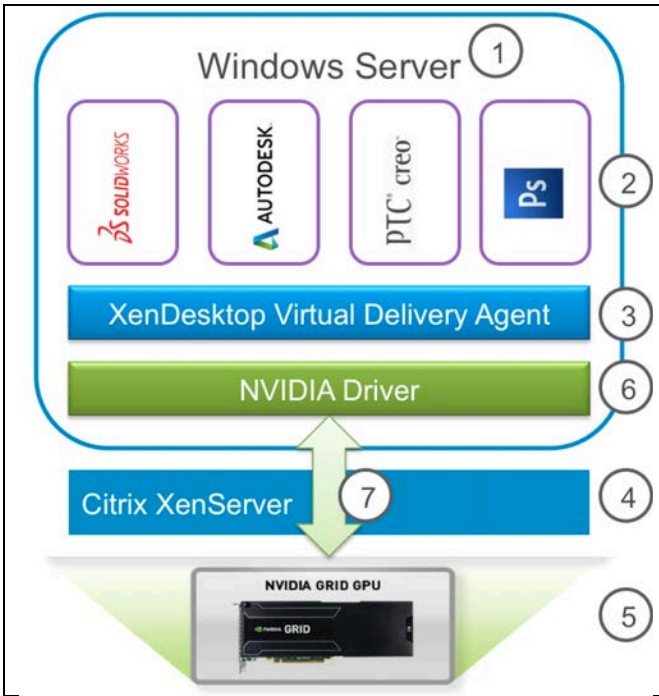


Figure 1 Apps sharing GPU on XenServer

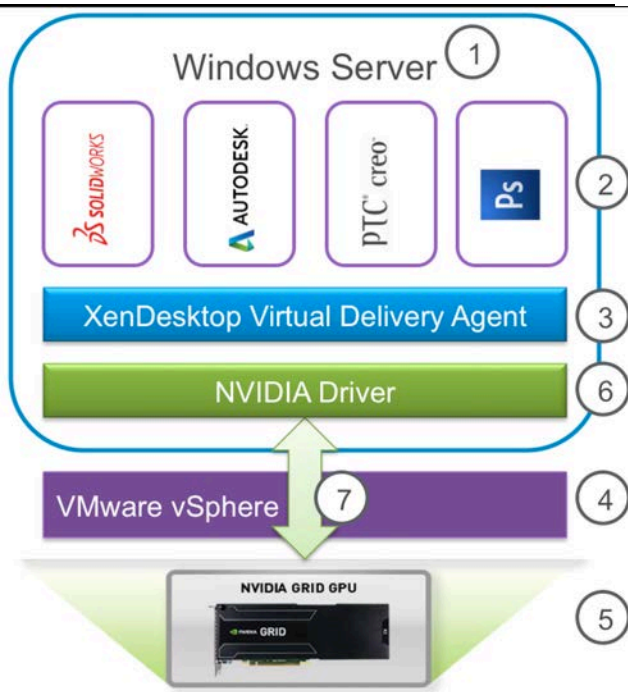


Figure 2 Apps sharing GPU on vSphere (vDGA)

Legend		
1	Operating system (OS) type	Windows Server is a multi-user OS, while Windows Desktop is a single-user OS
2	User-sessions	Each application is running in its own user-session, within the same OS instance
3	Delivery Agent	The Citrix software that enables connections between end-users and the published applications
4	Hypervisor	This is the platform to virtualize the applications, and enable sharing of resources such as GPU
5	GPU card	GRID architecture is the next-generation hardware from Nvidia that supports multiple GPU cards on a single board, and has been designed to work with virtualized workloads.
6	Graphics driver	Renders the graphics commands from the 3D applications to the display.
7	Graphics Virtualization type	In the case of GPU-pass-through or vDGA, the virtual machine has direct and full access to the underlying GPU hardware.

VMware: Virtual Dedicated Graphics Acceleration (vDGA)

Source: VMware Horizon View Graphics Acceleration Deployment Guide [PDF]

As seen in Figure 2 above, enabling vDGA on VMware vSphere allows VM full and direct access to the underlying GPU hardware. We will enable vDGA (also called GPU pass-through) for a Windows Server virtual machine that will host the 3D applications to be delivered using XenDesktop 7.

To configure an ESXi host with only a single GPU, first find the PCI ID of the graphics device by running the following command:

```
~ # lspci | grep -i display
00:07:00.0 Display controller: nVidia Corporation GK107 [VGX K1]
00:08:00.0 Display controller: nVidia Corporation GK107 [VGX K1]
00:09:00.0 Display controller: nVidia Corporation GK107 [VGX K1]
00:0a:00.0 Display controller: nVidia Corporation GK107 [VGX K1]
00:10:00.0 Display controller: Matrox Electronics Systems Ltd. G200eR2
~ #
```

`00:07:00.0` is the PCI ID of the graphics card.

Confirm Successful Installation

To check if the Graphics Adapter has been installed correctly, run the following command on the ESXi host. In case of GRID K1, it shows the 4 GPU cards available on the single board

```
~ # esxcli hardware pci list -c 0x0300 -m 0xff
```

See the Appendix for detailed command output.

VMware vSphere vDGA Configuration

This section takes you through enabling GPU pass-through at the host level and preparing the virtual machines for 3D rendering.

Enable the Host for GPU Pass-through

To enable an ESXi host for GPU pass-through, follow the documented checks and steps in the following section.

(Optional Step) Check VT-d or AMD IOMMU Is Enabled

[**Note**: This step is only required when the server hardware is new and hypervisor is not yet installed.]

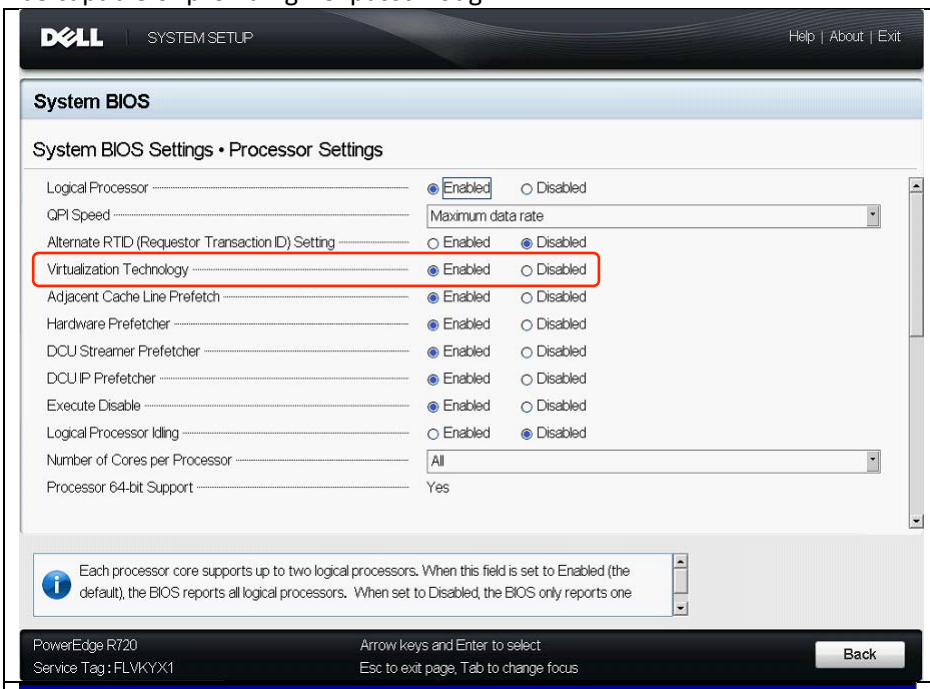
Before pass-through can be enabled, check if VT-d or AMD IOMMU is enabled on the host by running the following command, either via SSH or on the console. (**Note**: replace `<module_name>` with the name of the module: **vtddmar** for Intel, **AMDiommu** for AMD).

```
# esxcfg-module –l | grep <module_name>
```

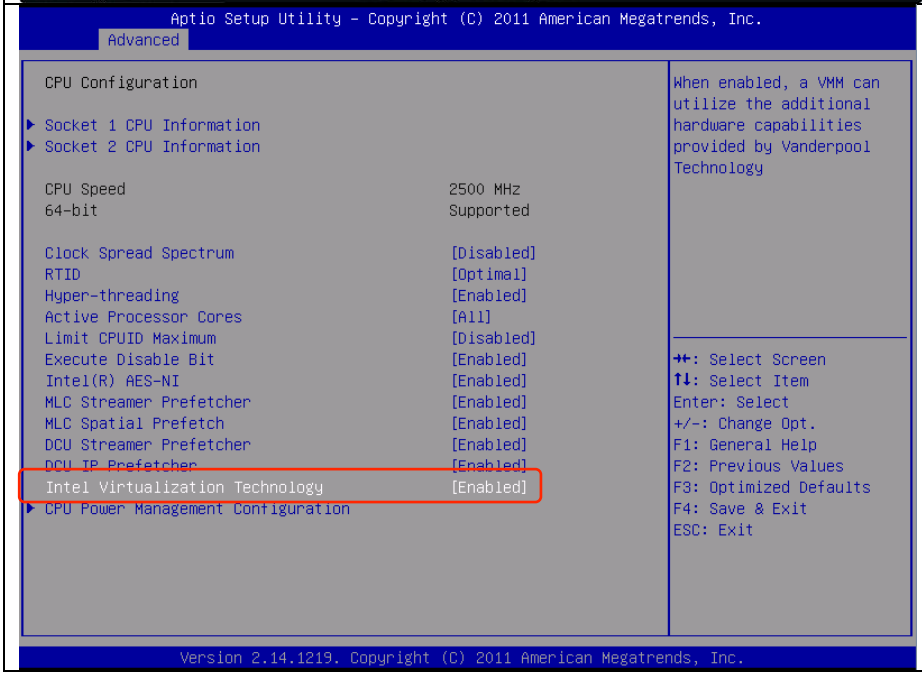
If above does not give any output, then browse to the below location to verify either vtddmar or AMDiommu is listed depending on your server hardware.

```
/usr/lib/vmware/vmkmod # ls
AMDIommu                filedriver              megaraid_mbox
aacraid                 fnic                    megaraid_sas
adp94xx                 forcedeth               migrate
ahci                    hbr_filter              mpt2sas
```

If the appropriate module is not present, you might have to enable it in the BIOS, or your hardware might not be capable of providing PCI passthrough.



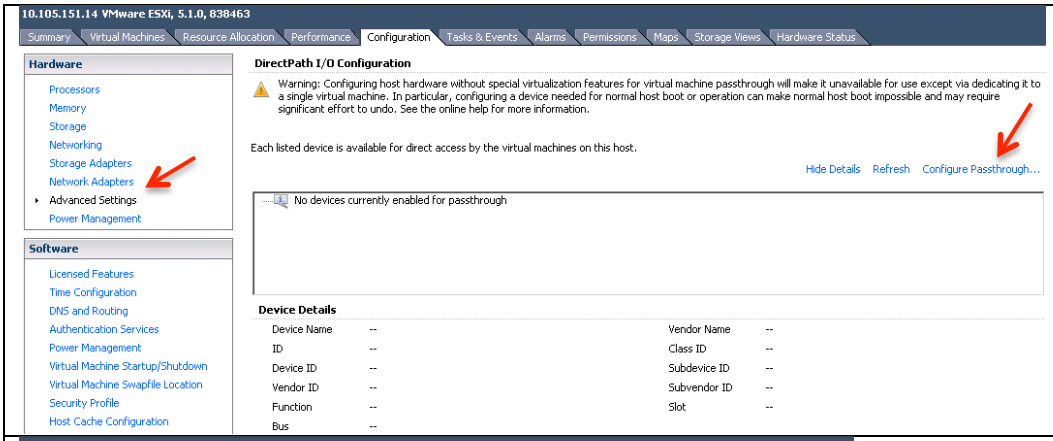
BIOS check for AMD-V on a Dell R720 server



BIOS check for Intel-VT on a Supermicro server

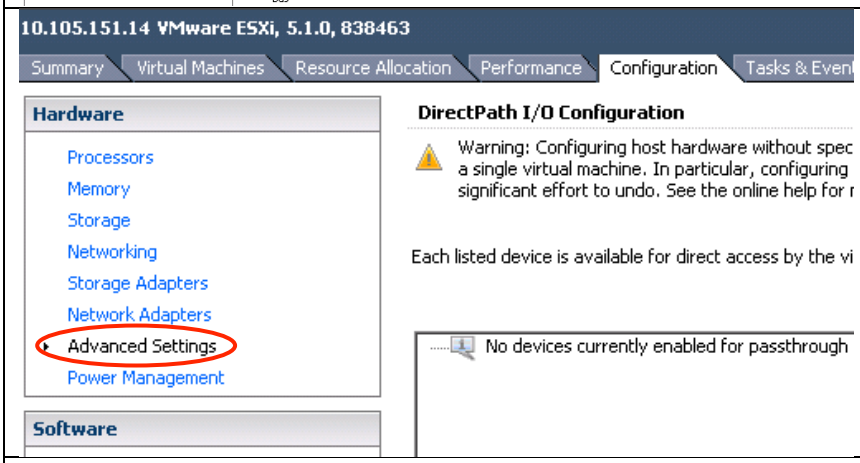
Enable Device Pass-through

Using the vSphere Client, connect to VMware vCenter and select the host with the GPU card installed.

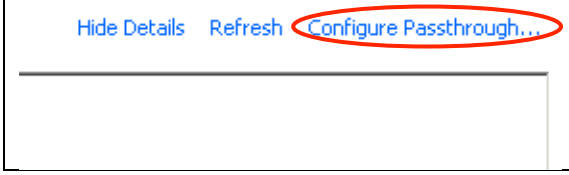


Select the **Configuration** tab for the host.

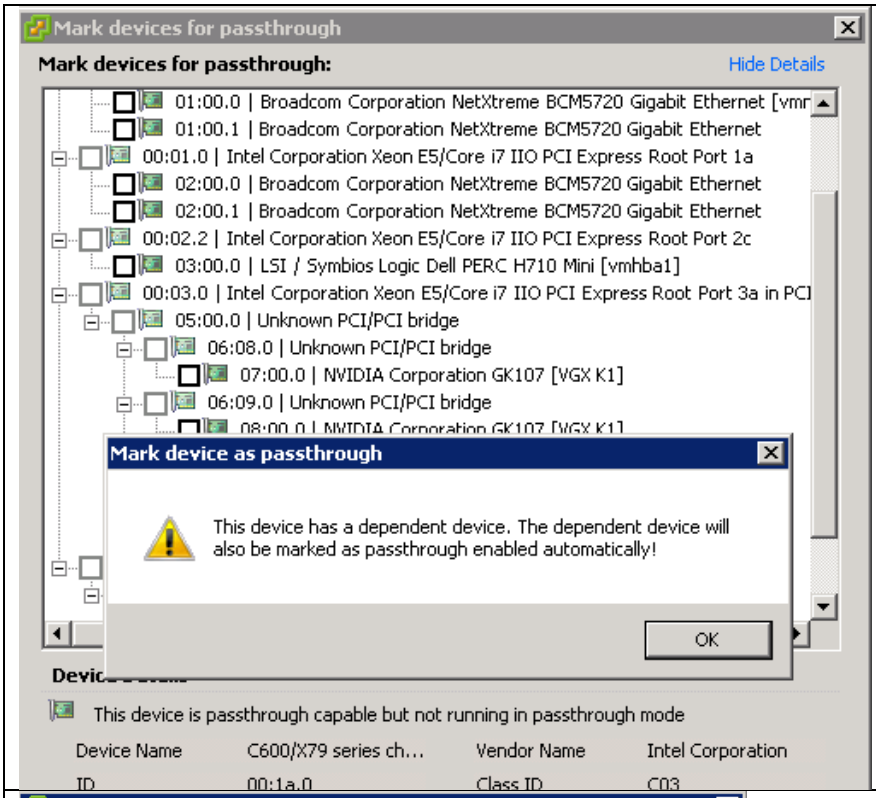
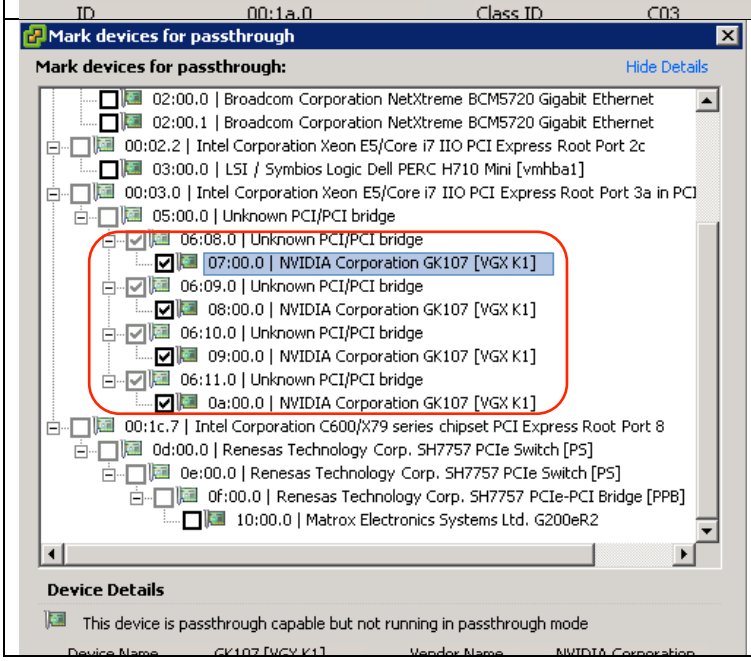
Note: If the host already has devices enabled for passthrough, these devices will be listed here.

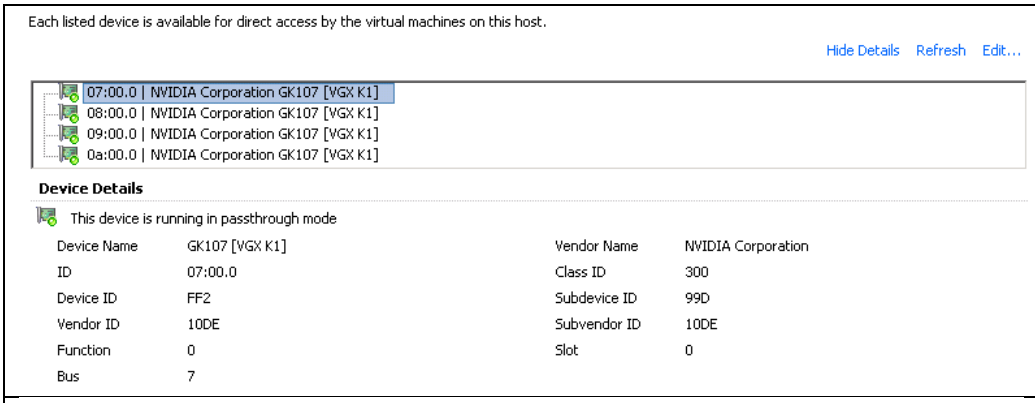


Click **Advanced Settings** (in the top left Hardware section).

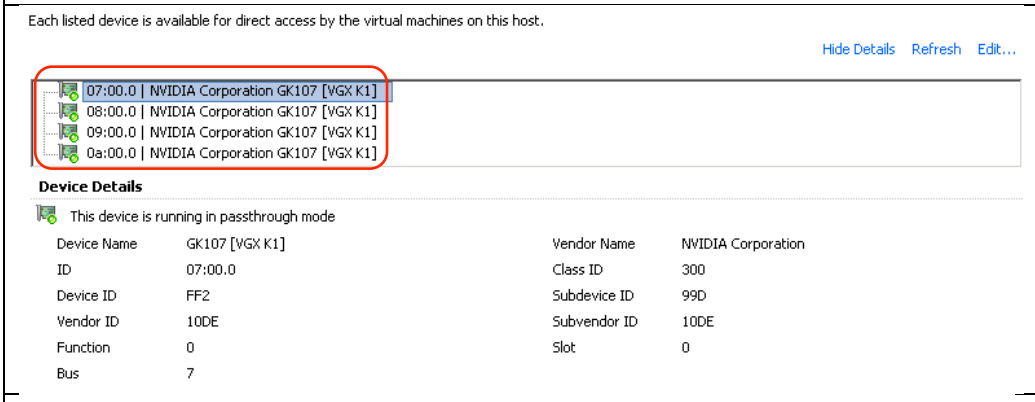


To configure passthrough for the GPU, click **Configure Passthrough**.

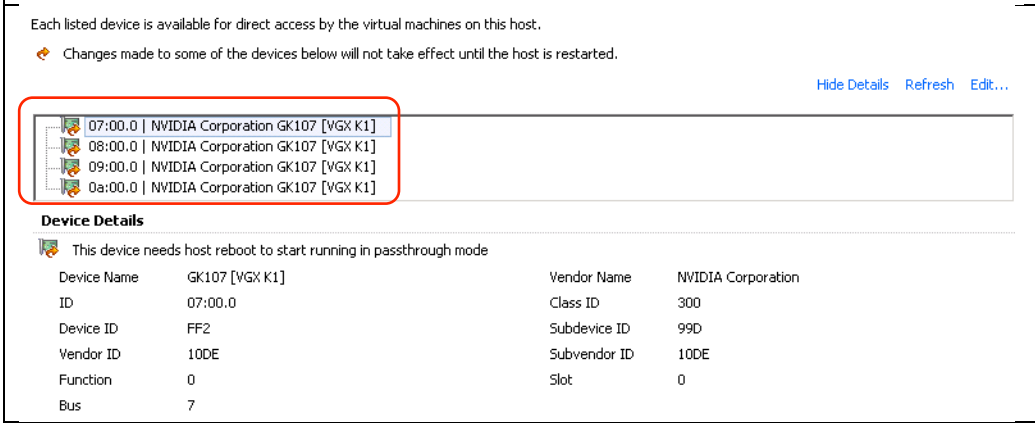
		In the Mark Devices for Passthrough window, check the box that corresponds to the GPU adapter installed in the host. You'll receive a Warning: *The device has a dependent device. The dependent device will also be marked as passthrough enabled automatically!* You may ignore the warning message and click **OK**
		In the Mark Devices for Passthrough window, check the box that corresponds to the GPU adapter installed in the host. Click **OK**



The GPU should now be listed in the Window on the **Advanced Settings** page.



If the device icon is **green**, passthrough is enabled.



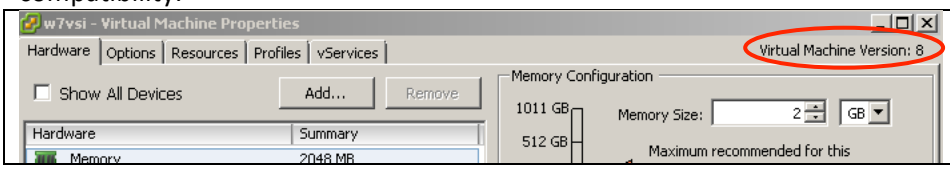
If the device has an **orange** arrow displayed on the icon, the host needs to be rebooted before passthrough will function.

Enable the Virtual Machine for GPU Pass-through

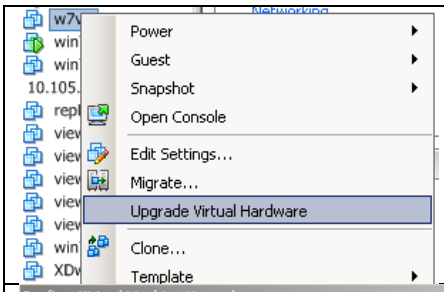
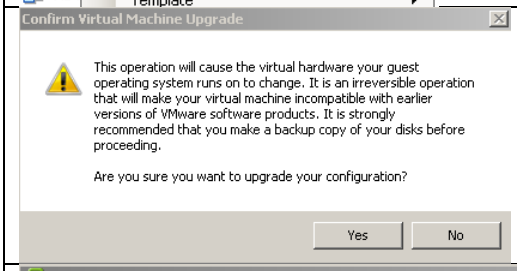
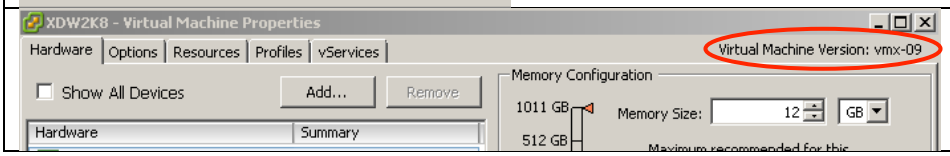
To enable a virtual machine for GPU pass-through, follow the documented checks and steps in the following section.

Update to Hardware Version 9

You must upgrade all 3D virtual machines to Hardware version 9 (HWv9 shows as *vmx-09*) to ensure maximum compatibility.

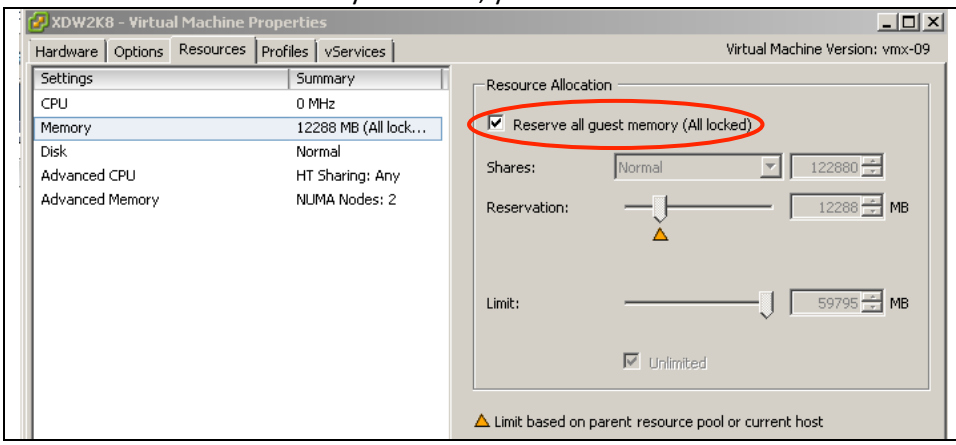


Pre Virtual Hardware upgrade: Virtual Machine Version is *8*

<table>
<tr><td> 
</td><td>From vCenter:
→Right-click the virtual machine to be upgraded
→Select Upgrade Virtual Hardware</td></tr>
<tr><td> 
</td><td>Upgrade Warning
The virtual hardware version upgrade is an irreversible process. You may ignore this message.</td></tr>
<tr><td> 
</td><td>Post Virtual Hardware upgrade: Virtual Machine Version is vmx-09</td></tr>
</table>

Reserve All Configured Memory

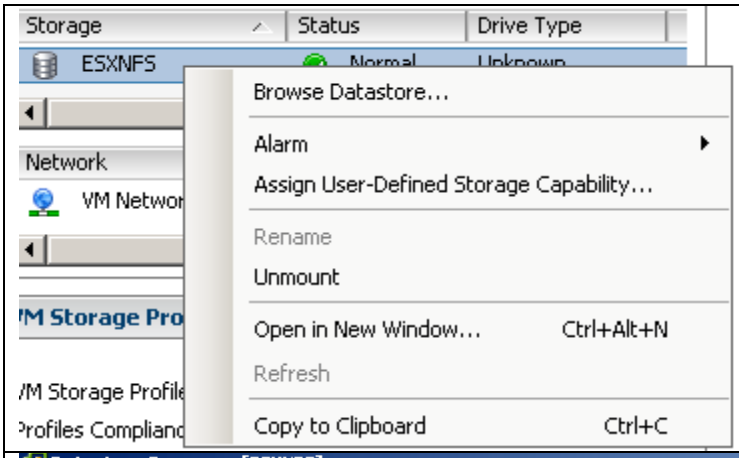
For vDGA to function, all the virtual machine configured memory must be reserved. If each virtual machine has 2GB of memory allocated, you should reserve all 2GB. To do this:

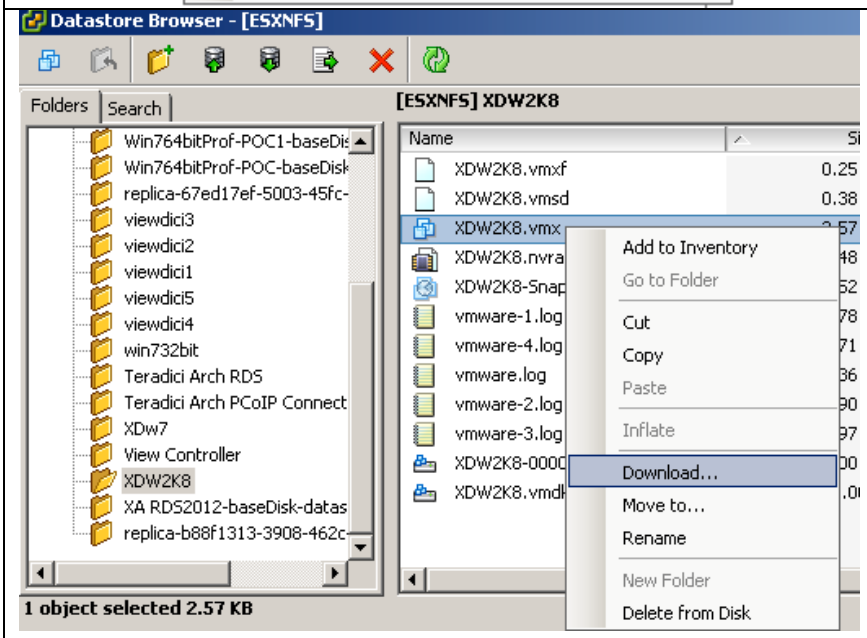
<table>
<tr><td> 
</td><td>Select the Reserve all guest memory option when you view the Memory option under the Resources tab in a virtual machine's settings window.

By default, memory of VMs are unreserved i.e. Reserver all guest memory check-box is unchecked.</td></tr>
</table>

Adjust pciHole.start

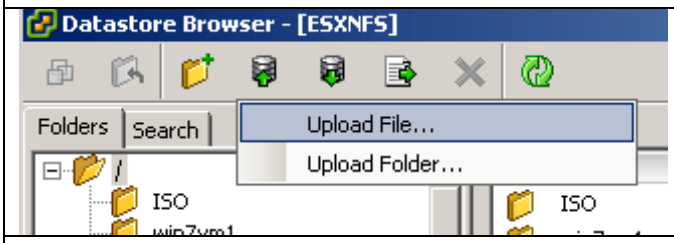
Note: This is required only if the virtual machine has more than 2GB of configured memory. Before you start, ensure that the virtual machine is shut down completely.

<table>
<tr>
<td> 
</td>
<td>From vCenter:
Browse to the datastore where the VM resides.

Right-click the datastore → Select Browse Datastore…</td>
</tr>
<tr>
<td> 
</td>
<td>On the left panel:
Highlight the VM folder in the datastore
On the right panel:
Right-click the .vmx file of the VM
Select Download… and Save it to your local machine.
Open the file using WordPad
Add the following parameter to the .vmx file of the virtual machine (you can add this at the end of the file):

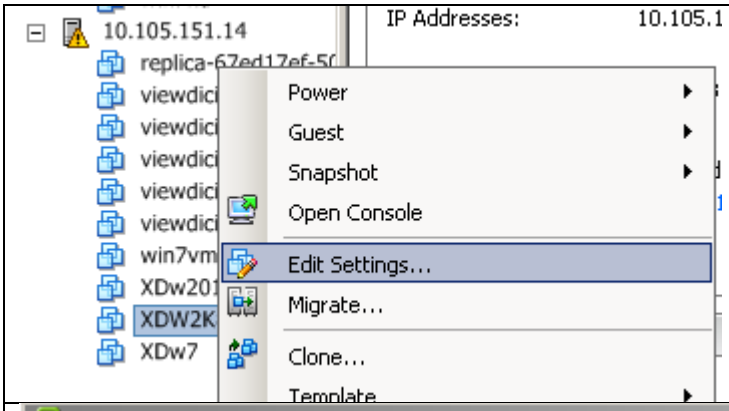
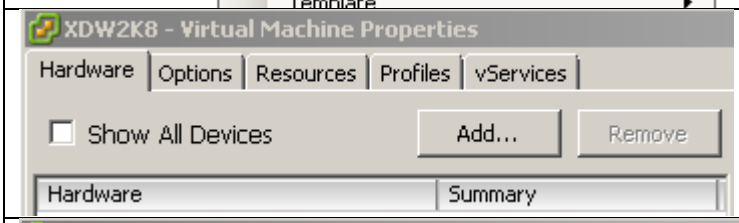
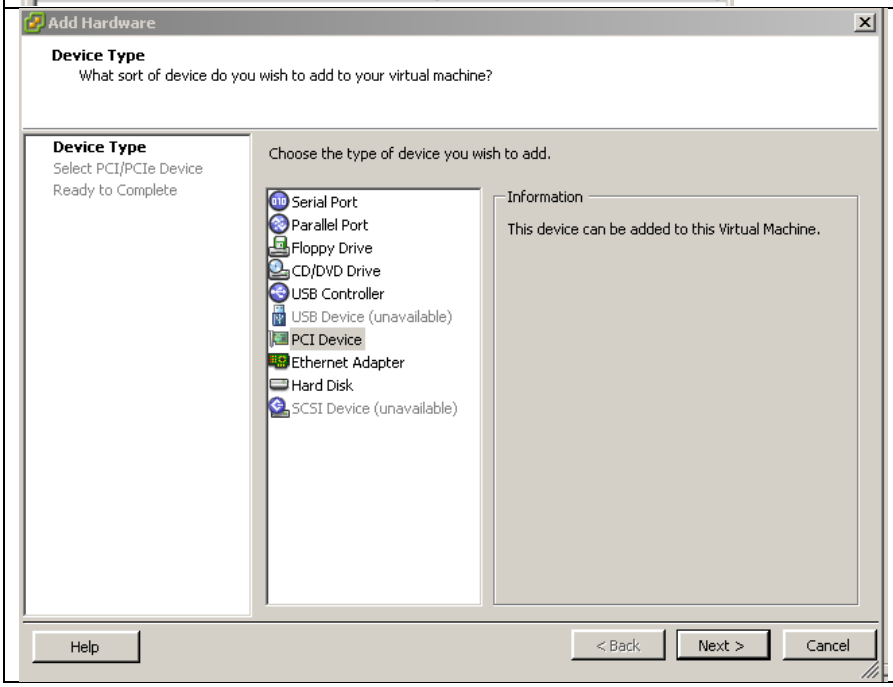
<code>pciHole.start = “2048”</code>

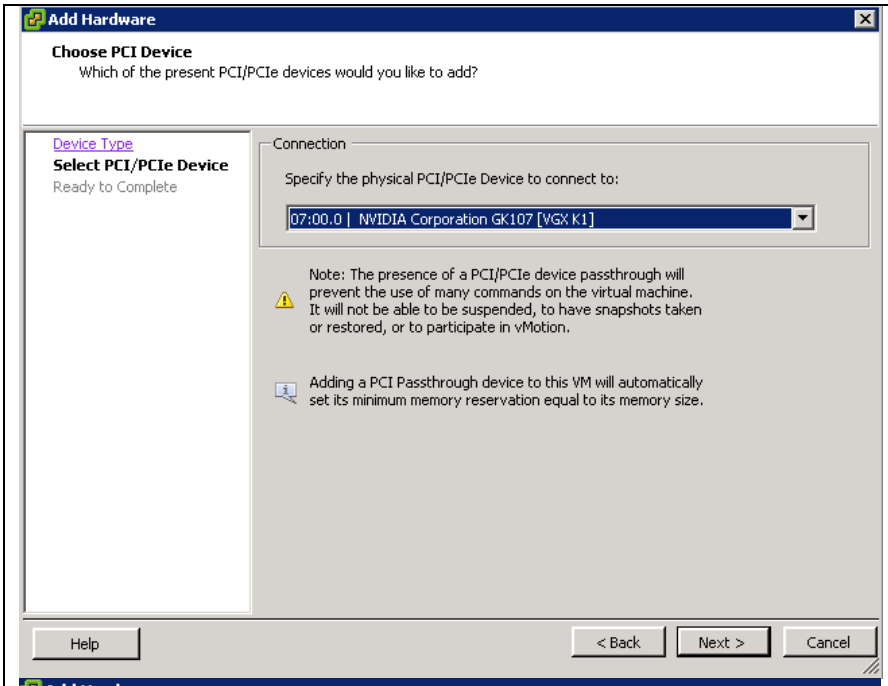
This can also be done using ESX CLI or vSphereCLI tool.

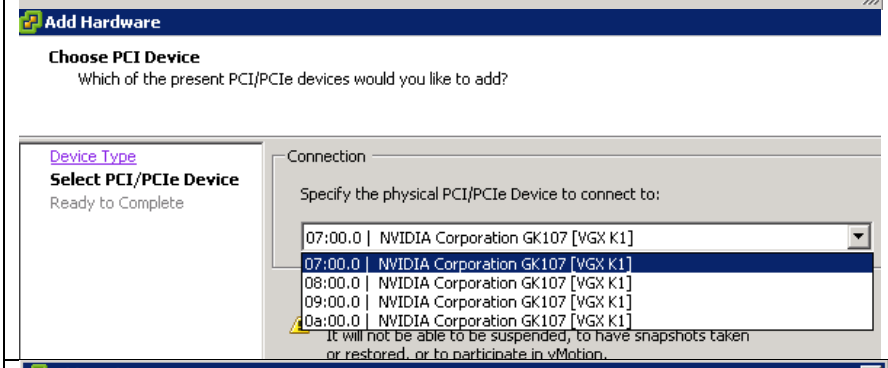
CLI stands for Command Line Interface</td>
</tr>
<tr>
<td> 
</td>
<td>Using Upload File… option from Datastore Browser window, upload and replace the modified .vmx file with the existing one in the VM folder.</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</table>

Add the PCI Device

To enable vDGA for a virtual machine, the PCI device needs to be added to the virtual machine’s hardware. Using the vSphere Client, connect directly to the ESXi host with the GPU card installed, or select the host in vCenter.

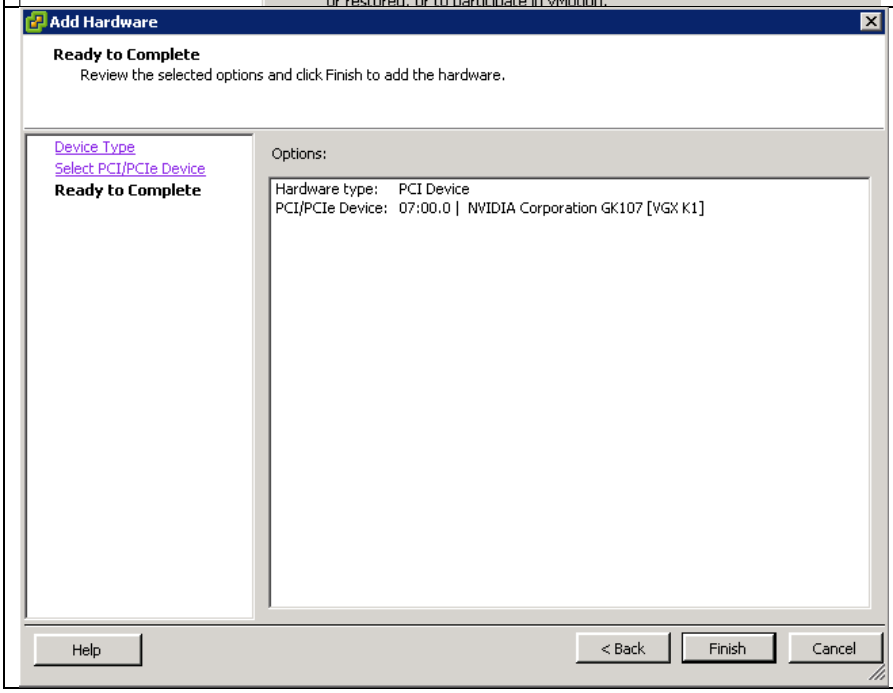
	Right-click the virtual machine and t **Edit Settings**.
	Under Hardware tab, click **Add…**
	**Add** a new device by selecting **PCI Device** from the drop-down list, and click **Next**. Select **PCI Device**, click **Next**



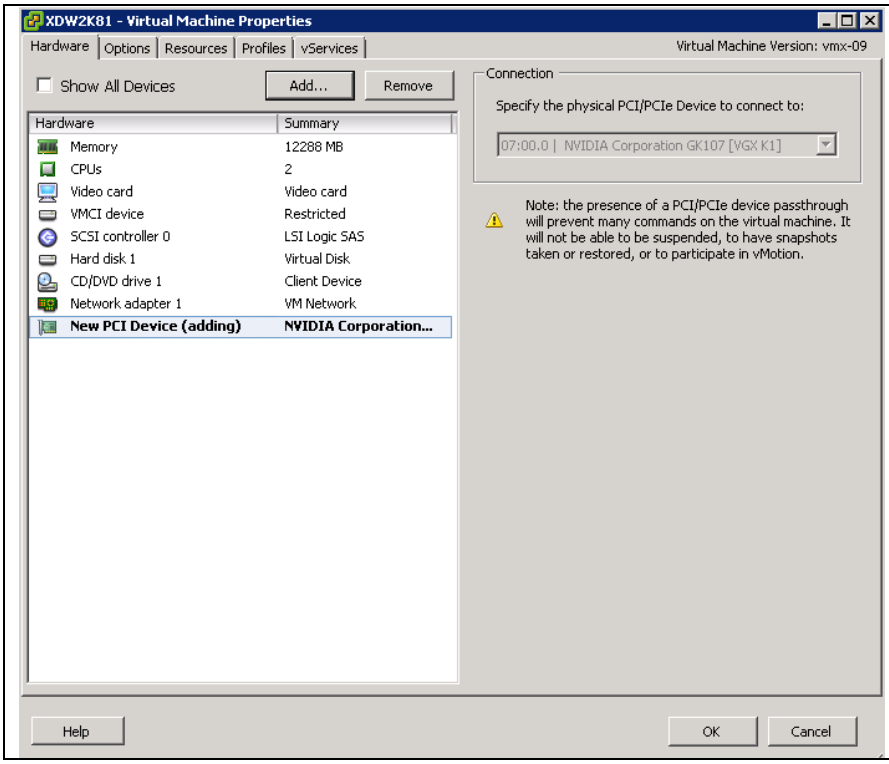


Select one of the GPUs from the GRID as the passthrough device to connect to the virtual machine from the drop-down list, and click **Next**.

All the four GPUs in the GRID are listed under PCIe device connection list



Click **Finish**

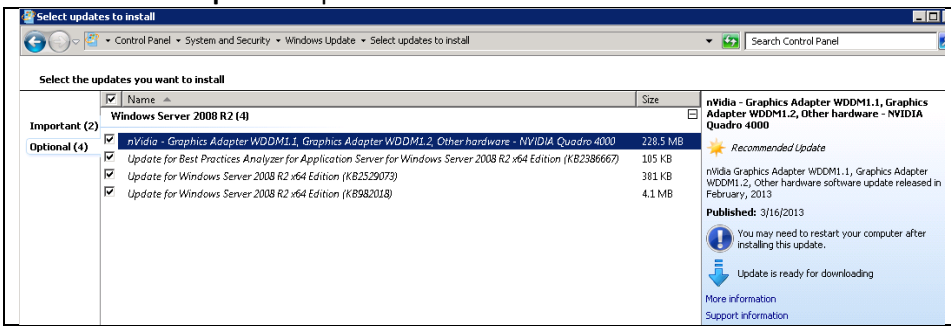


Click **OK**

Install the NVIDIA Driver

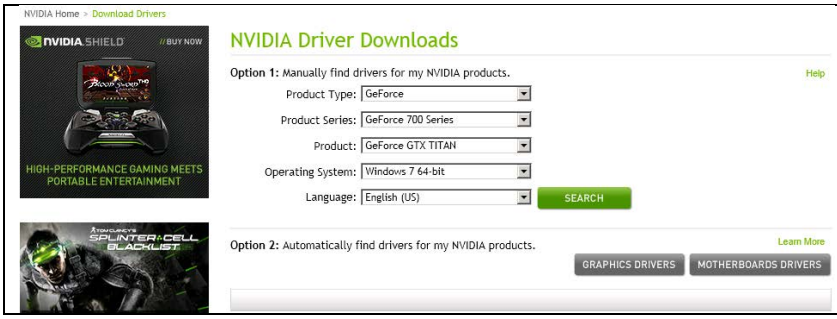
Two ways to install NVIDIA driver on the guest OS (Desktop VDA and/or XenApp server):

- **Microsoft Windows Update**: Run windows update and NVIDIA driver will be available for download under **Optional** updates.



Microsoft Windows Update

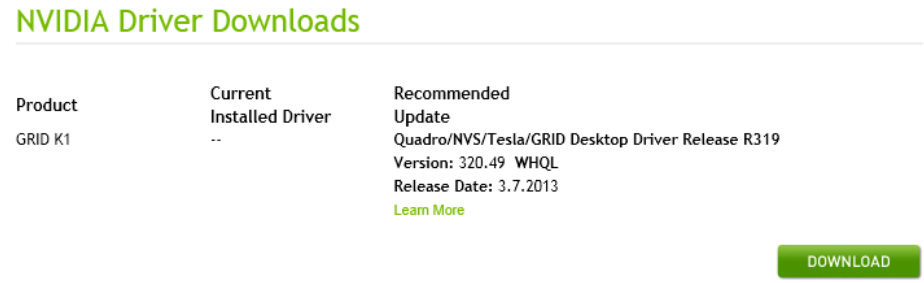
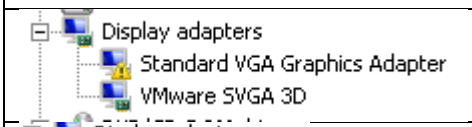
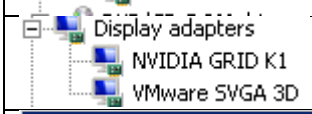
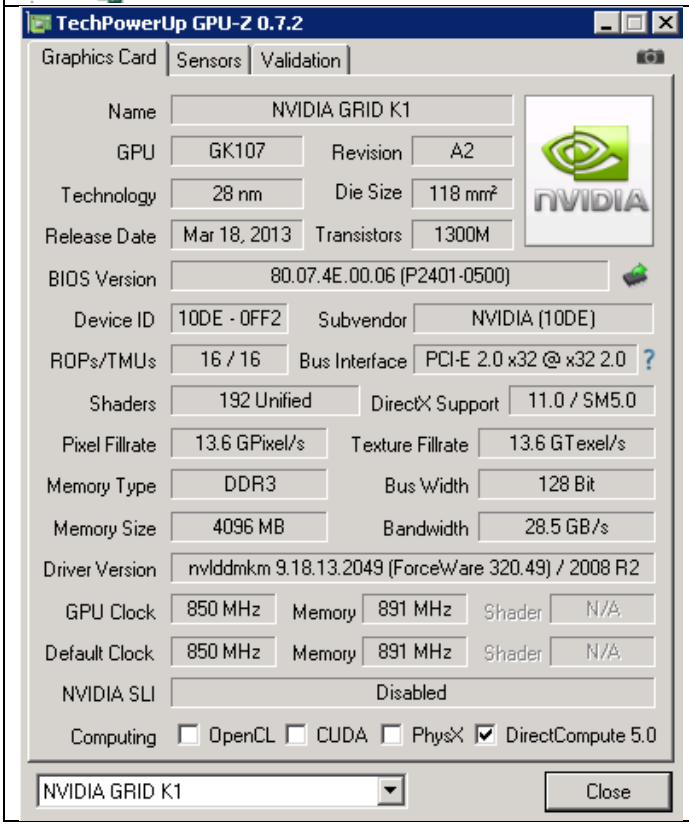
- **NVIDIA website**: Download and install the latest NVIDIA Windows driver on the virtual machine. All NVID IA drivers can be downloaded from the NVIDIA Download Drivers page.



NVIDIA website

Select **Option 2**

Click **Graphics Drivers** button

	**Pre-requisite to scan : Latest Java update** **The nvidia.com website scans automatically and shows the appropriate latest driver to download** **Direct download for GRID K1 URL:** **http://www.nvidia.com/object/quadr tesla-grid-win8-win7-winvista-64bit-320.49-whql-driver.html [189 MB]**
	Before: Display adapter has warning before NVIDIA guest OS driver is installed
	After: Display adapter with NO warning after NVIDIA guest OS driver is installed
	GPU-Z shows NVIDIA GRID K1 running o the VM

After the driver is installed, reboot the virtual machine.

XenServer GPU-Passthrough

Please see the **Part 1 of this guide** for step-by-step instructions on enabling GPU pass-through on Citrix XenServer.

Installation of XenDesktop 7 and Delivering 3D Apps from Windows Server OS

- Install Virtual Desktop Agent (VDA) on the guest OS. For publishing hosted applications and shared desktops, install VDA on Windows Server 2008 R2 or Server 2012.
- **Please see the Reviewer's Guide for step-by-step instructions on installing the virtual desktop agent and other Citrix XenDesktop components such as the Studio.**

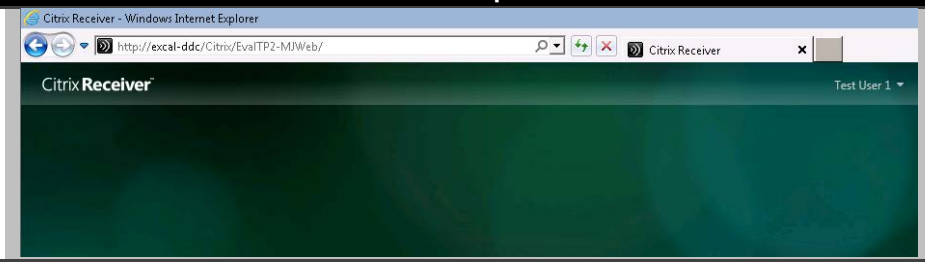
End-user Experience from Citrix Receiver

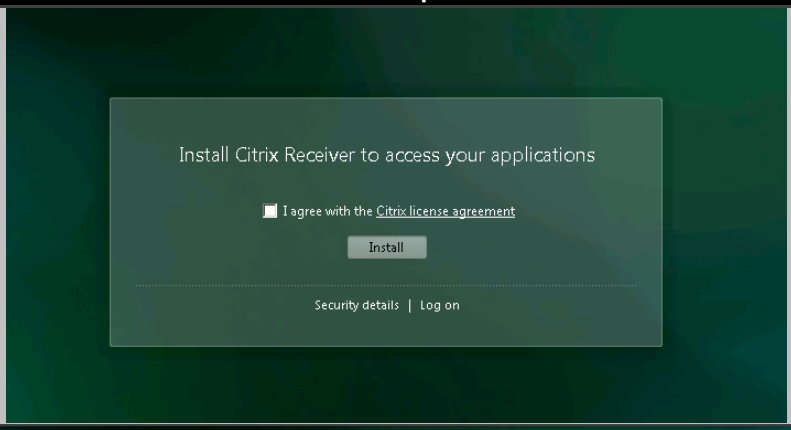
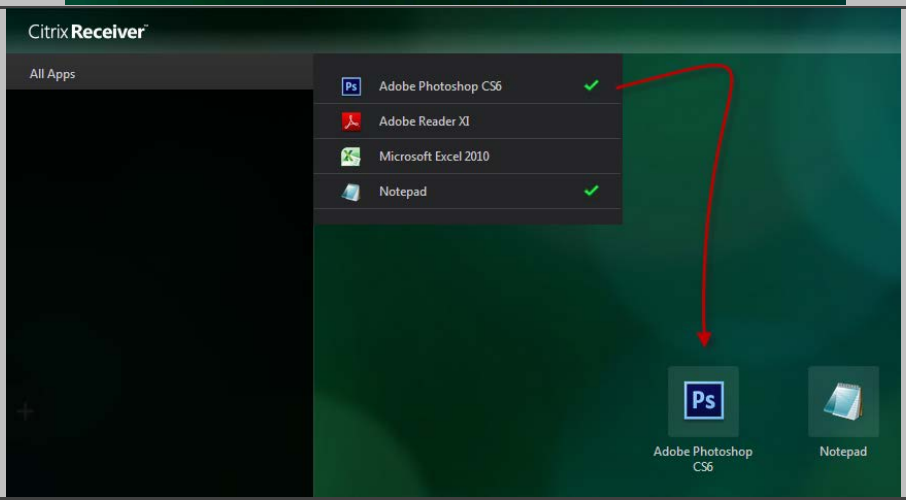
This section shows the users launching 3D applications published with XenDesktop 7 Apps (formerly, XenApp) using Citrix Receiver on the end-point devices. In this example, we launch multiple sessions of Unigine Heaven 3D and Google Earth, freely available demo apps, from XenDesktop server hosted on both VMware vSphere and Citrix XenServer (with GPU enabled, as seen previously).

3D Application	Unigine Heaven, Google Earth, eDrawings
Monitoring Tools used	o Process Explorer with GPU monitoring enabled o GPU-Z o Furmark o GPU Shark
No. of XenApp sessions (users) tested	2 and 4
GPU Card	GRID K1

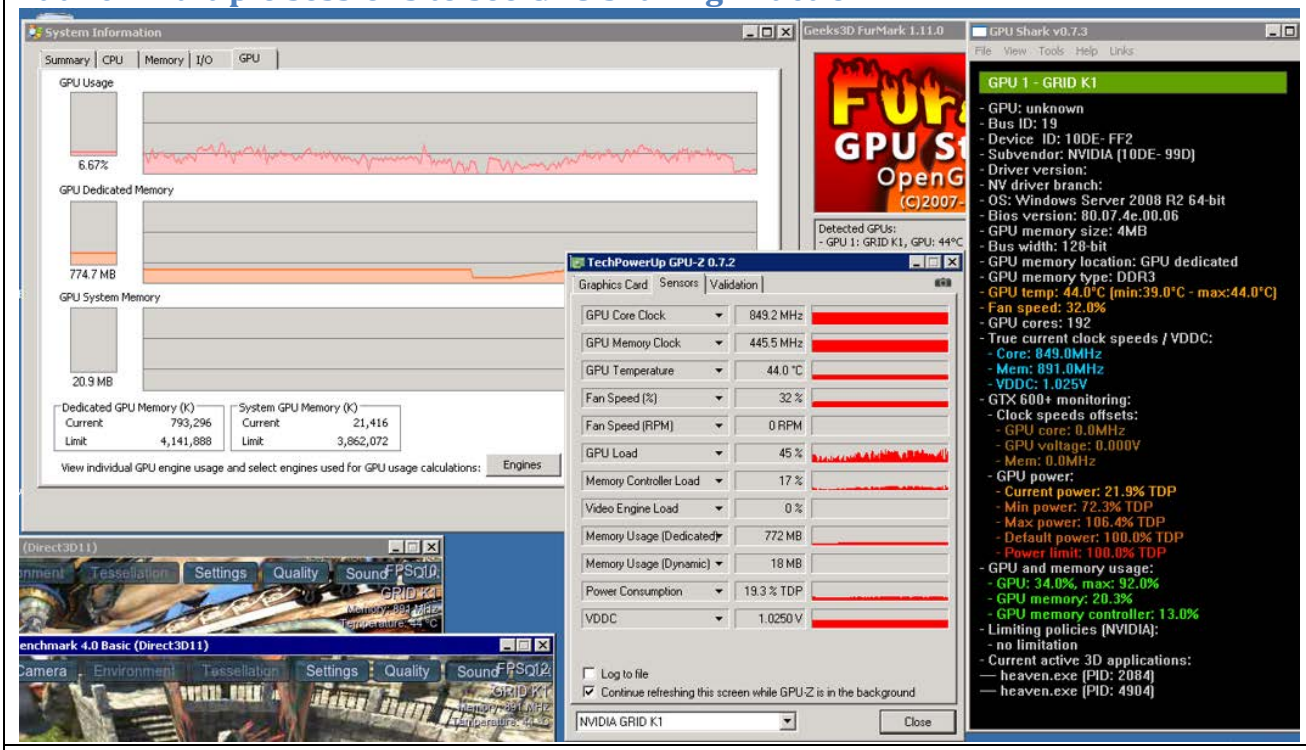
Launch desktops and applications on Windows client

Citrix Receiver is the unified access client to access applications and desktops from StoreFront. With a user account, you will access those applications and desktops.

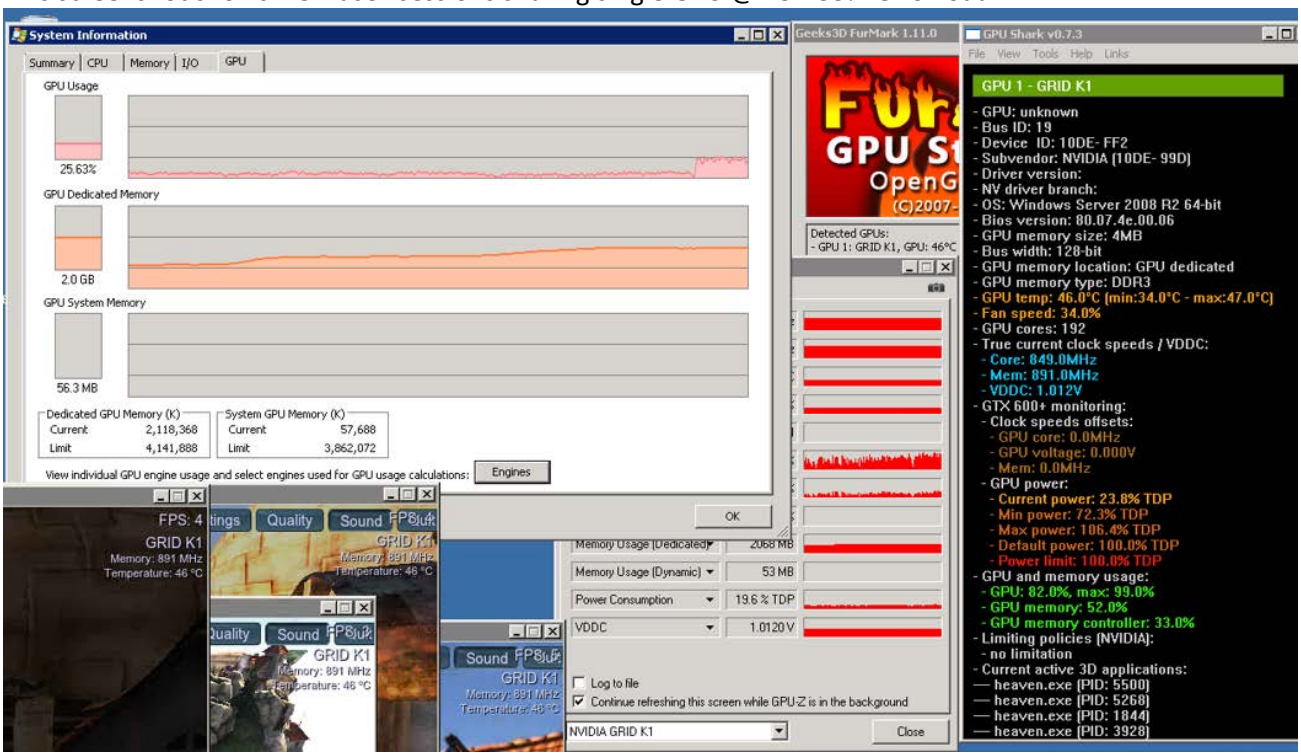
#	Screen capture	Instructions
		On a client machine, Windows 7 in this case, open a browser and go to the default **Storefront URL** http://<yourservername>/Citrix/StoreWeb

#	Screen capture	Instructions
		If **Citrix Receiver** is not already installed on the client, you are prompted to install it. **Accept** the EULA, Click **Install** and follow the installation process. Return to the login page once it is installed.
		Login as a domain user. Click the **+ sign** at the left edge of the screen, and click **All Apps** to see list of available apps. Click a few apps and **add** them to your self-service portal.

Launch multiple sessions to see GPU sharing in action



This screenshot shows 4 3D user-sessions sharing single GPU @ ~52-55% GPU Load



Summary

In this first part of the HDX 3D Pro Reviewer's Guide, we learnt how to identify the different hardware components of HDX 3D Pro solution and complete the physical installation. We also saw how to enable GPU pass-through on XenServer. In this document, we configured GPU pass-through on VMware's vSphere hypervisor, and tested the GPU being ready for use inside the virtual machine (VM). Using a Windows Server VM, this GPU can be shared by multiple users through XenDesktop 7 Apps. Please refer to the XenDesktop 7 Reviewer's Guide to learn how these VMs act as the base image for HDX 3D delivery using Citrix XenDesktop. It explains the steps for setting up the XenDesktop infrastructure and accessing applications from thin-clients and standard PCs using Citrix Receiver.

In the next two parts, we learn the steps to enable shared GPU access for desktops using the hardware virtualization technology in XenServer (vGPU) and software implementation in vSphere (vSGA).

Appendix

Third-party 3D applications and GPU benchmark tools and blogs

[**Note**: These are utilities found on the Internet and not provided by Citrix. Citrix does not guarantee or support use of these tools.]

Third-party tools	URLs
3DMark	http://www.3dmark.com/ Download: location1 or location2
Geeks3D	http://www.geeks3d.com/ http://www.geeks3d.com/20130719/furmark-1-11-0-gpu-vga-videocard-burn-in-stress-test-opengl-benchmark-utility-nvidia-geforce-amd-radeon/ http://www.geeks3d.com/20110408/download-tessmark-0-3-0-released/ http://www.geeks3d.com/20130308/fluidmark-1-5-1-physx-benchmark-fluid-sph-simulation-opengl-download/ http://www.geeks3d.com/20120511/geexlab-0-4-0-ultim8-edition-available-gtx-600-opengl-bindless-textures-support-added/ http://www.geeks3d.com/20110719/quick-test-process-explorer-15-0-with-gpu-support/
Aquamark	http://downloads.guru3d.com/download.php?det=673
3dmark	http://www.futuremark.com/benchmarks/
Lightsmark	http://dee.cz/lightsmark/
Furmark	http://www.ozone3d.net/benchmarks/fur/ GPU Shark: http://www.ozone3d.net/gpushark/ GPU –Z: http://www.techpowerup.com/gpuz

Demo Apps	
Unigine	http://unigine.com/products/heaven/download/
Google Earth	http://www.google.com/earth
eDrawings	http://www.edrawingsviewer.com/ed/edrawings-samples.htm
Adobe Photoshop (trial)	http://www.adobe.com/photoshop
Autodesk Inventor	http://www.autodesk.com/inventor

Command to check if GPU is installed properly

To check if the Graphics Adapter has been installed correctly, run the following command on the ESXi host. In case of GRID K1, it shows the 4 GPU cards available on the single board:

```
~ # esxcli hardware pci list -c 0x0300 -m 0xff
000:007:00.0
   Address: 000:007:00.0
   Segment: 0x0000
   Bus: 0x07
   Slot: 0x00
   Function: 0x00
   VMkernel Name:
   Vendor Name: NVIDIA Corporation
   Device Name: GK107 [VGX K1]
   Configured Owner: Unknown
   Current Owner: VMkernel
   Vendor ID: 0x10de
   Device ID: 0x0ff2
   SubVendor ID: 0x10de
   SubDevice ID: 0x099d
   Device Class: 0x0300
   Device Class Name: VGA compatible controller
   Programming Interface: 0x00
   Revision ID: 0xa1
   Interrupt Line: 0x0f
   IRQ: 15
   Interrupt Vector: 0xc0
   PCI Pin: 0xc0
   Spawned Bus: 0x00
   Flags: 0x0201
   Module ID: -1
   Module Name: None
   Chassis: 0
   Physical Slot: 8
   Slot Description:
   Passthru Capable: true
   Parent Device: PCI 0:6:8:0
   Dependent Device: PCI 0:6:8:0
   Reset Method: Bridge reset
   FPT Sharable: true

000:008:00.0
   Address: 000:008:00.0
   Segment: 0x0000
   Bus: 0x08
   Slot: 0x00
   Function: 0x00
   VMkernel Name:
   Vendor Name: NVIDIA Corporation
   Device Name: GK107 [VGX K1]
   Configured Owner: Unknown
   Current Owner: VMkernel
```

```
   Vendor ID: 0x10de
   Device ID: 0x0ff2
   SubVendor ID: 0x10de
   SubDevice ID: 0x099d
   Device Class: 0x0300
   Device Class Name: VGA compatible controller
   Programming Interface: 0x00
   Revision ID: 0xa1
   Interrupt Line: 0x0e
   IRQ: 14
   Interrupt Vector: 0xc8
   PCI Pin: 0xc8
   Spawned Bus: 0x00
   Flags: 0x0201
   Module ID: -1
   Module Name: None
   Chassis: 0
   Physical Slot: 9
   Slot Description:
   Passthru Capable: true
   Parent Device: PCI 0:6:9:0
   Dependent Device: PCI 0:6:9:0
   Reset Method: Bridge reset
   FPT Sharable: true

000:009:00.0
   Address: 000:009:00.0
   Segment: 0x0000
   Bus: 0x09
   Slot: 0x00
   Function: 0x00
   VMkernel Name:
   Vendor Name: NVIDIA Corporation
   Device Name: GK107 [VGX K1]
   Configured Owner: Unknown
   Current Owner: VMkernel
   Vendor ID: 0x10de
   Device ID: 0x0ff2
   SubVendor ID: 0x10de
   SubDevice ID: 0x099d
   Device Class: 0x0300
   Device Class Name: VGA compatible controller
   Programming Interface: 0x00
   Revision ID: 0xa1
   Interrupt Line: 0x0f
   IRQ: 15
   Interrupt Vector: 0xc0
   PCI Pin: 0x63
   Spawned Bus: 0x00
   Flags: 0x0201
   Module ID: -1
   Module Name: None
```

```
   Chassis: 0
   Physical Slot: 16
   Slot Description:
   Passthru Capable: true
   Parent Device: PCI 0:6:16:0
   Dependent Device: PCI 0:6:16:0
   Reset Method: Bridge reset
   FPT Sharable: true

000:00a:00.0
   Address: 000:00a:00.0
   Segment: 0x0000
   Bus: 0x0a
   Slot: 0x00
   Function: 0x00
   VMkernel Name:
   Vendor Name: NVIDIA Corporation
   Device Name: GK107 [VGX K1]
   Configured Owner: Unknown
   Current Owner: VMkernel
   Vendor ID: 0x10de
   Device ID: 0x0ff2
   SubVendor ID: 0x10de
   SubDevice ID: 0x099d
   Device Class: 0x0300
   Device Class Name: VGA compatible controller
   Programming Interface: 0x00
   Revision ID: 0xa1
   Interrupt Line: 0x0e
   IRQ: 14
   Interrupt Vector: 0xc8
   PCI Pin: 0x00
   Spawned Bus: 0x00
   Flags: 0x0201
   Module ID: -1
   Module Name: None
   Chassis: 0
   Physical Slot: 17
   Slot Description:
   Passthru Capable: true
   Parent Device: PCI 0:6:17:0
   Dependent Device: PCI 0:6:17:0
   Reset Method: Bridge reset
   FPT Sharable: true

~ #
```

If the NVIDIA GPU is not listed in the above output, then GPU card is either not installed correctly and/or is malfunctioning. Also, ensure the Xorg service is up and running.

Related Documents in this Series

Part 1: XenServer GPU pass-through for Citrix XenDesktop 7 (includes, physical installation of GPU cards)
Part 2: vSphere GPU pass-through (a.k.a vDGA) for Citrix XenDesktop 7
Part 3: XenServer GPU virtualization (a.k.a vGPU) for Citrix XenDesktop 7
Part 4: vSphere shared GPU (a.k.a vSGA) for Citrix XenDesktop 7