



Configuring Citrix XenServer® 7.3 for Graphics

Published December 2017
1.0 Edition



Configuring Citrix XenServer ® 7.3 for Graphics

© 1999-2017 Citrix Systems, Inc. All Rights Reserved.
Version: 7.3

Citrix Systems, Inc.
851 West Cypress Creek Road
Fort Lauderdale, FL 33309
United States of America

Disclaimers

This document is furnished "AS IS." Citrix Systems, Inc. disclaims all warranties regarding the contents of this document, including, but not limited to, implied warranties of merchantability and fitness for any particular purpose. This document may contain technical or other inaccuracies or typographical errors. Citrix Systems, Inc. reserves the right to revise the information in this document at any time without notice. This document and the software described in this document constitute confidential information of Citrix Systems, Inc. and its licensors, and are furnished under a license from Citrix Systems, Inc.

Citrix Systems, Inc., the Citrix logo, Citrix XenServer and Citrix XenCenter, are trademarks of Citrix Systems, Inc. and/or one or more of its subsidiaries, and may be registered in the United States Patent and Trademark Office and in other countries. All other trademarks and registered trademarks are property of their respective owners.

Trademarks

Citrix®
XenServer ®
XenCenter ®



Contents

1. Introduction	1
1.1. GPU Pass-Through	1
1.2. Shared GPU	1
1.2.1. Licensing Note	1
2. Guest Support and Constraints	2
3. NVIDIA GRID vGPU	3
3.1. Licensing Note	3
3.2. Available NVIDIA GRID vGPU Types	3
3.3. NVIDIA GRID System Requirements	3
3.4. Preparation Overview	4
3.5. Installation on XenServer 7.3	4
3.5.1. Licensing note	4
3.6. Installing the NVIDIA GRID vGPU Manager for XenServer	4
4. Intel GVT-d and GVT-g	7
4.1. Intel GVT-g System Requirements and Configuration	7
4.2. Enabling Intel GPU Pass-Through	7
5. Create a Virtual GPU or GPU Pass-Through Enabled VM	9
5.1. Install the XenServer Tools	9
5.2. Install the In-guest Drivers	10
5.2.1. Install the NVIDIA drivers	10
5.2.2. Install the AMD Drivers	11
5.2.3. Install the Intel Drivers	12
5.3. Configure VDA for HDX 3D Pro	14

Chapter 1. Introduction

Citrix XenServer is leading the way in the virtual delivery of 3D professional graphics applications and workstations. Its offerings include GPU Pass-through (for NVIDIA, AMD and Intel GPUs) as well as hardware-based GPU sharing with NVIDIA GRID™ vGPU™ and Intel GVT-g™.

1.1. GPU Pass-Through

Unlike the rest of the physical system components, which are shared and represented as multiple virtual instances to multiple clients by the hypervisor, a pass-through GPU is not abstracted at all, but remains one physical device. Each hosted virtual machine (VM) gets its own dedicated GPU, eliminating the software abstraction and the performance penalty that goes with it.

XenServer allows you to assign a physical GPU (in the XenServer host) to a Windows or HVM Linux VM running on the same host. This GPU Pass-Through feature is intended for graphics power users, such as CAD designers, who require high performance graphics capabilities.

1.2. Shared GPU













Shared GPU, where one physical GPU which can be used by multiple VMs concurrently. As a portion of a physical GPU is used, performance will be greater than emulated graphics, and there is no need for one card per VM. This enables resource optimization, boosting the performance of the VM. The graphics commands of each virtual machine are passed directly to the GPU, without translation by the hypervisor.

1.2.1. Licensing Note

Graphics Virtualization is available for XenServer Enterprise Edition customers, or those who have access to XenServer through their XenApp/XenDesktop entitlement. To learn more about XenServer editions, and to find out how to upgrade, visit the Citrix website [here](#). For detailed information on Licensing, refer to the *XenServer Licensing FAQ* on the [Citrix Product Documentation](#) website.

Additionally, customers may need a vendor subscription or a license depending on the graphics card used.

The following table summarizes GPU and shared GPU support for guests at XenServer 7.3:

	GPU for Windows VMs	GPU for HVM Linux VMs	Shared GPU for Windows VMs	Virtual GPU for Linux VMs
AMD				
Intel				
NVIDIA				



Chapter 2. Guest Support and Constraints

XenServer 7.3 supports the following guest operating systems for virtual GPU.

Technology	Supported Guests
NVIDIA vGPU	<p>Windows Guests</p> <ul style="list-style-type: none">• Windows 7 (32-bit/64-bit)• Windows Server 2008 R2 SP1 (64-bit)• Windows 8 (32-bit/64-bit)• Windows 8.1 (32-bit/64-bit)• Windows 10 (64-bit)• Windows Server 2012 (64-bit)• Windows Server 2012 R2 (64-bit)• Windows Server 2016 (64-bit) <p>HVM Linux Guests</p> <ul style="list-style-type: none">• RHEL 7• CentOS 7• Oracle Linux 7• Scientific Linux 7• Ubuntu 14.04
Intel GVT-g	<p>Windows guests</p> <ul style="list-style-type: none">• Windows 7 (32-bit/64-bit)• Windows 8.1 (64-bit)• Windows 10 (64-bit)• Windows Server 2008 R2 SP1 (64-bit)• Windows Server 2012 R2 (64-bit)• Windows Server 2016 (64-bit)

Note:

- XenServer supports only one GPU per VM.
- Virtual GPU and GPU Pass-Through are **not** compatible with High Availability, XenMotion, Storage XenMotion or VM Suspend. This means VMs cannot be migrated to avoid downtime. However, VMs using GPU Pass-through or virtual GPU can still be started on any host that has the appropriate resources.
- XenServer automatically detects and groups together identical physical GPUs across hosts in the same pool. Once assigned to a group of GPUs, a VM may be started on any host in the pool that has an available GPU in the group.

Chapter 3. NVIDIA GRID vGPU

NVIDIA GRID vGPU enables multiple Virtual Machines (VM) to have simultaneous, direct access to a single physical GPU, using the same NVIDIA graphics drivers that are deployed on non-virtualized Operating Systems. Under the control of NVIDIA's GRID Virtual GPU Manager, which runs in the XenServer Control Domain (dom0), GRID physical GPUs are capable of supporting multiple virtual GPU devices (vGPUs) that can be assigned directly to VMs.

Guest VMs use GRID virtual GPUs in the same manner as a physical GPU that has been passed through by the hypervisor; an NVIDIA driver loaded in the guest VM provides direct access to the GPU for performance critical fast paths, and a paravirtualized interface to the GRID Virtual GPU Manager.

3.1. Licensing Note

NVIDIA vGPU is available for XenServer Enterprise edition customers, or those who have access to XenServer through their XenApp/XenDesktop entitlement. To learn more about XenServer editions, and to find out how to upgrade, visit the Citrix website [here](#). For detailed information on Licensing, refer to the *XenServer Licensing FAQ* on the [Citrix Product Documentation](#) website.

Depending on the NVIDIA graphics card used, customers may need NVIDIA subscription or a license.

3.2. Available NVIDIA GRID vGPU Types

NVIDIA GRID cards can contain multiple Graphics Processing Units (GPU). For example, TESLA M10 cards contain four GM107GL GPUs, and TESLA M60 cards contain two GM204GL GPUs. Each physical GPU (pGPU) can host several different types of virtual GPU (vGPU). vGPU types have a fixed amount of framebuffer, number of supported display heads and maximum resolutions, and are targeted at different classes of workload.

For a list of the most recently supported NVIDIA cards, refer to the [XenServer Hardware Compatibility List](#) in conjunction with [NVIDIA product information](#).

Note:

At any given time, vGPUs hosted on a physical GPU **must all be of the same type**. However, there is no corresponding restriction between physical GPUs on the same card. This restriction is automatic and may cause unexpected capacity planning issues.

For example, a TESLA M60 card has two physical GPUs, and can support eleven types of vGPU: GRID M60-1A, GRID M60-2A, GRID M60-4A, GRID M60-8A, GRID M60-0B, GRID M60-1B, GRID M60-0Q, GRID M60-1Q, GRID M60-2Q, GRID M60-4Q and GRID M60-8Q.

3.3. NVIDIA GRID System Requirements

- NVIDIA GRID card:
 - For a list of the most recently supported NVIDIA cards, refer to the [XenServer Hardware Compatibility List](#) in conjunction with [NVIDIA product information](#).
- Depending on the NVIDIA graphics card used, customers may need NVIDIA subscription or a license. Refer to the [NVIDIA product information](#) for details.
- XenServer Enterprise edition (or access to XenServer through a XenDesktop/XenApp entitlement).
- A server capable of hosting XenServer and NVIDIA GRID cards.
- NVIDIA GRID vGPU software package for XenServer, consisting of the GRID Virtual GPU Manager for XenServer, and NVIDIA drivers.
- To run XenDesktop with VMs running NVIDIA vGPU, you will also need: XenDesktop 7.6 or later, full installation.



Note:

Customers are advised to review the *NVIDIA GRID Virtual GPU User Guide* (Ref: DU-06920-001) available from the [NVIDIA website](#). Customers will need to register with NVIDIA to gain access to these components for additional performance optimization steps.

3.4. Preparation Overview

1. Install XenServer
2. Install the NVIDIA Virtual GPU Manager for XenServer
3. Reboot the XenServer host

3.5. Installation on XenServer 7.3

XenServer 7.3 is available for download from the [XenServer Downloads](#) page.

Customers should install:

- **XenServer 7.3 Base Installation ISO**
- **XenCenter 7.3 Windows Management Console**

Refer to the *XenServer Installation Guide* on the [Citrix Product Documentation](#) website for comprehensive details on installation.

3.5.1. Licensing note

vGPU is available for XenServer Enterprise edition customers, or those who have access to XenServer through their XenApp/XenDesktop entitlement. To learn more about XenServer editions, and to find out how to upgrade, visit the Citrix website [here](#). For detailed information on Licensing, refer to the *XenServer Licensing FAQ* on the [Citrix Product Documentation](#) website.

Depending on the NVIDIA graphics card used, customers may need NVIDIA subscription or a license. Refer to the [NVIDIA product information](#) for details.

For information on licensing NVIDIA cards, refer to the [NVIDIA website](#).

3.6. Installing the NVIDIA GRID vGPU Manager for XenServer

Customers should install NVIDIA GRID vGPU Software available from [NVIDIA](#). The NVIDIA GRID software consists of:

1. GRID vGPU Manager
(for example: NVIDIA-vGPU-XenServer-7.2-367.64.x86_64.rpm)
2. Windows Display Driver (The Windows display driver depends on the Windows version)
(for example: 369.71_grid_win10_server2016_64bit_international.exe)

The *GRID vGPU Manager* runs in the XenServer Control Domain (dom0). It is provided as either a supplemental pack or an RPM file. Please refer to the User Guide included in the NVIDIA GRID vGPU Software for detailed installation steps and specifics.

Note:

The Update, RPM names and versions are examples and will most likely be different in your environment.



The Update can be installed using XenCenter (Tools > Install Update > Select update or supplemental pack for disk) or using the xe CLI command `xe-install-update`. The Update will be named `NVIDIA-vGPU-XenServer-7.2-367.64.x86_64.iso` or similar.

Note:

If you are installing the GRID vGPU Manager using an RPM file, you must ensure to copy the RPM file to dom0 and then install.

1. Use the rpm command to install the package:

```
[root@xenserver ~]# rpm -iv NVIDIA-vGPU-XenServer-7.2-367.64.x86_64.rpm
Preparing packages for installation...
NVIDIA-vGPU-XenServer-7.2-367.64
[root@xenserver ~]
```

2. Reboot the XenServer host:

```
[root@xenserver ~]# shutdown -r now
Broadcast message from root (pts/1) Mon May 09 13:05:31 2016):
[root@xenserver ~]
```

3. After rebooting the XenServer host, verify that the GRID package has installed and loaded correctly by checking for the NVIDIA kernel driver in the list of kernel loaded modules:

```
[root@xenserver ~]# lsmod |grep nvidia
nvidia                8152994 0
i2c_core              20294 2 nvidia,i2c_i801
[root@xenserver ~]#
```

4. Verify that the NVIDIA kernel driver can successfully communicate with the GRID physical GPUs in your host by running the `nvidia-smi` command, which produces a listing of the GPUs in your platform similar to:

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

Thu Jan 26 13:48:50 2017
+-----+
NVIDIA-SMI 367.64 Driver Version: 367.64 |
+-----+
 GPU Name Persistence-M| Bus-Id  Disp.A | Volatile Uncorr. ECC |
 Fan Temp Perf Pwr:Usage/Cap| Memory-Usage | GPU-Util  Compute M. |
+-----+-----+-----+-----+-----+-----+
|  0 Tesla M60          On | 0000:05:00.0  Off |   Off |
| N/A  33C  P8          24W / 150W |  7249MiB /  8191MiB |    0%      Default |
+-----+-----+-----+-----+-----+-----+
|  1 Tesla M60          On | 0000:09:00.0  Off |   Off |
| N/A  36C  P8          24W / 150W |  7249MiB /  8191MiB |    0%      Default |
+-----+-----+-----+-----+-----+-----+
|  2 Tesla M60          On | 0000:85:00.0  Off |   Off |
| N/A  36C  P8          23W / 150W |   19MiB /  8191MiB |    0%      Default |
+-----+-----+-----+-----+-----+-----+
|  3 Tesla M60          On | 0000:89:00.0  Off |   Off |
| N/A  37C  P8          23W / 150W |   14MiB /  8191MiB |    0%      Default |
+-----+-----+-----+-----+-----+-----+
+-----+
| Processes:                      GPU Memory |
| GPU  PID  Type  Process name      Usage   |
+-----+-----+-----+-----+-----+
| No running compute processes found |
+-----+
[root@xenserver ~]#
```




Note:

When using NVIDIA vGPU with XenServer hosts with more than 768GB of RAM, add the parameter `iommu=dom0-passthrough` to the Xen command line. To do this:

1. Run the following command in the control domain (Dom0):

```
/opt/xensource/libexec/xen-cmdline --set-xen iommu=dom0-passthrough
```

2. Restart the host.

Chapter 4. Intel GVT-d and GVT-g

XenServer supports Intel's virtual GPU (GVT-g); a graphics acceleration solution that requires no additional hardware. It uses the Intel Iris Pro functionality embedded in certain Intel processors, and utilizes a standard Intel GPU driver installed within the VM.

Note:

As the Intel Iris Pro graphics functionality is embedded within the processors, CPU-intensive applications can cause power to be diverted from the GPU. As a result, customers may not experience the same level of graphics acceleration as purely GPU-intensive workloads.

4.1. Intel GVT-g System Requirements and Configuration

In order to use Intel GVT-g, the XenServer host must be equipped with a CPU with Iris Pro graphics listed as supported for Graphics on the [XenServer Hardware Compatibility List](#) (for example, Xeon E3 1285 v4) and a motherboard with a graphics-enabled chipset (for example, C226 for Xeon E3 v4 CPUs or C236 for Xeon E3 v5 CPUs).

Note:

Customers must ensure to reboot the hosts when switching between Intel GPU pass-through (GVT-d) and Intel Virtual GPU (GVT-g).

When configuring Intel GVT-g, the number of Intel virtual GPUs supported on a specific XenServer host depends on the GPU bar size of that XenServer host (known as the 'Aperture size' in the BIOS). Citrix recommends that you set the Aperture size to 1024MB in order to support a maximum of seven virtual GPUs per host.

If you configure the Aperture size to 256MB, only one VM can be started on the host. Setting it to 512MB can result in only three VMs being started on the XenServer host. Setting the Aperture size higher than 1024MB is not supported and it **does not** increase the number of VMs which can be started on a host.

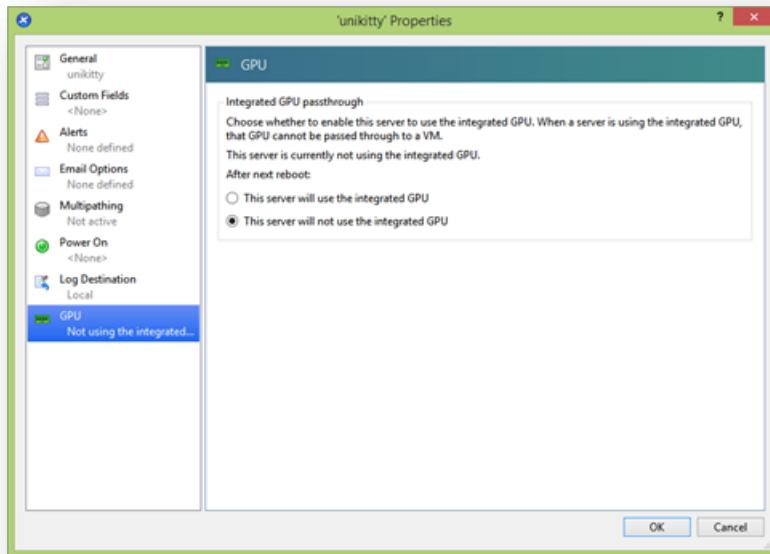
4.2. Enabling Intel GPU Pass-Through

XenServer supports the GPU Pass-Through feature for Windows 7 and Windows 8 (32-/64-bit) VMs using an Intel integrated GPU device. For more information on supported hardware, refer to the [XenServer Hardware Compatibility List](#).

When using Intel GPU on Intel servers, the XenServer host's Control Domain (dom0) will have access to the integrated GPU device. In such cases, the GPU will not be available for pass-through. Customers who wish to use the Intel GPU Pass-through feature on Intel servers must disable the connection between dom0 and the GPU before passing through the GPU to the VM.

To do this:

1. On the **Resources** pane, select the XenServer host.
2. On the **General** tab, click **Properties**, and in the left pane, click on **GPU**.
3. In the **Integrated GPU passthrough** section, select **This server will not use the integrated GPU**.



This disables the connection between dom0 and the Intel integrated GPU device.

4. Click **OK**.
5. Reboot the XenServer host for the changes to take effect.

The Intel GPU will now be visible on the GPU type drop-down list during new VM creation, and on the VM's **Properties** tab.

Note:

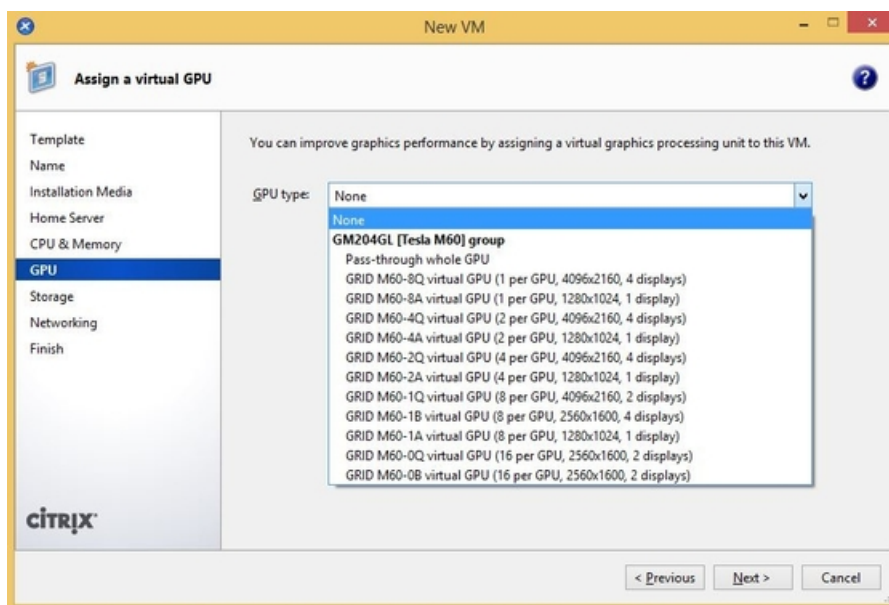
The XenServer host's external console output (for example, VGA, HDMI, DP) will not be available after disabling the connection between dom0 and the GPU.

Chapter 5. Create a Virtual GPU or GPU Pass-Through Enabled VM

Note:

Customers using Intel GPU Pass-through feature should first refer to the previous section *Enabling Intel GPU Pass-through* for additional configuration, and then follow the steps below.

1. Create a VM using XenCenter. To do this, select the host on the Resources pane and then select **New VM** on the **VM** menu.
2. Follow the instructions on the New VM wizard and select the **Installation Media, Home Server, CPU, and Memory**.
3. GPU-enabled servers will display a **GPU** configuration page:



4. From the **GPU Type** drop-down list, select either **Pass-through whole GPU**, or a **virtual GPU type**.
(Unavailable virtual GPU types will be greyed-out)
5. Click **Next** to configure **Storage** and then **Networking**.
6. When you complete your configuration, click **Create Now**.

5.1. Install the XenServer Tools

1. Install the XenServer Tools

Without the optimized networking and storage drivers provided by the XenServer Tools, remote graphics applications running on GRID vGPU will **not** deliver maximum performance.

- a. Select the VM in the **Resources** pane, right-click, and then click **Install XenServer Tools** on the shortcut menu. Alternatively, on the **VM** menu, click **Install XenServer Tools**.
- b. Click **Install XenServer Tools** on the message dialog to go to the VM's console and begin the installation.
- c. If Autoplay is enabled for the VM's CD/DVD drive, installation will be started automatically after a few moments. This process installs the I/O drivers and the Management Agent. Restart the VM when prompted to get your VM to an optimized state. If Autoplay is not enabled, the XenServer Tools installer displays the installation options. Click **Install XenServer Tools** to continue with the installation. This mounts the XenServer Tools ISO (guest-tools.iso) on the VM's CD/DVD drive.



- d. Click **Run setup.exe** to begin XenServer Tools installation and restart the VM when prompted to get your VM to an optimized state.

5.2. Install the In-guest Drivers

When viewing the VM console in XenCenter, the VM will typically boot to the desktop in VGA mode with 800 x 600 resolution. The standard Windows screen resolution controls can be used to increase the resolution to other standard resolutions. (Control Panel > Display > Screen Resolution)

Note:

When using GPU Pass-Through, we recommend that you install the in-guest drivers through RDP or VNC over the network (that is, not through XenCenter).

5.2.1. Install the NVIDIA drivers

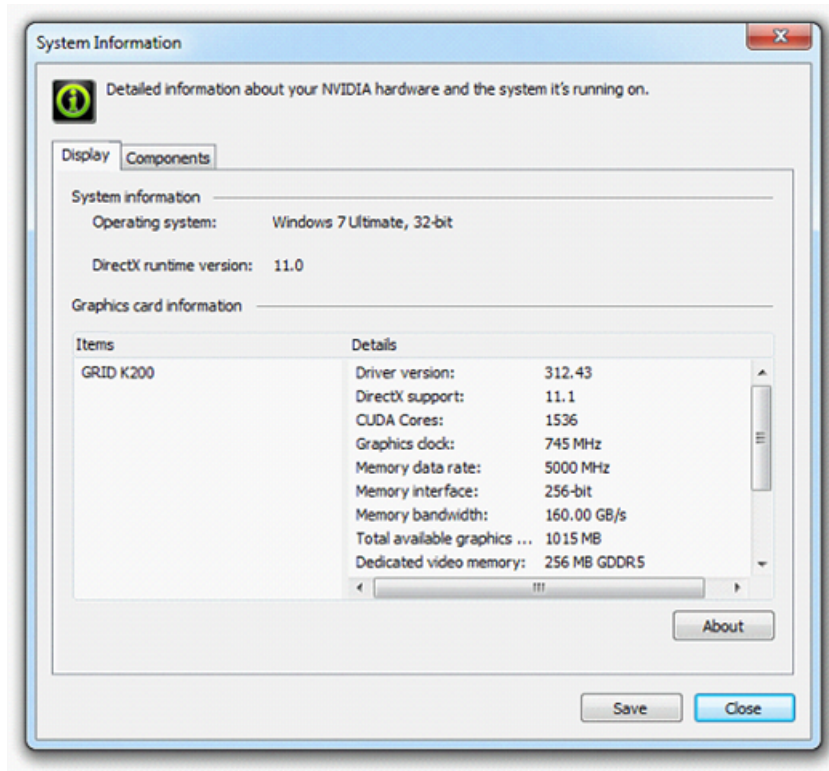
To enable vGPU operation (as for a physical NVIDIA GPU), NVIDIA drivers must be installed into the VM.

The following section provides a generic overview of the procedure. Please refer to the NVIDIA User Guide for detailed instructions.

1. Start the VM. In the **Resources** pane, right-click on the VM, and click **Start**.

During this boot process XenServer dynamically allocates a vGPU to the VM.

2. Follow the Windows operating system installation screens.
3. Once the operating system installation completes, reboot the VM.
4. Install the appropriate driver for the GPU inside the guest. The following example shows the specific case for in-guest installation of the NVIDIA GRID drivers.
5. Copy the 32 or 64-bit NVIDIA Windows driver package to the VM, open the zip file, and run setup.exe.
6. Follow the installation wizard to install the driver.
7. Once the driver installation has completed, you may be prompted to reboot the VM. Select **Restart Now** to reboot the VM immediately, alternatively, exit the installer package, and reboot the VM when ready. When the VM starts it will boot to a Windows desktop.
8. To verify that the NVIDIA driver is running, right-click on the desktop and select **NVIDIA Control Panel**.
9. In the NVIDIA Control Panel, select **System Information**. This will show the GPU Type in use by the VM, its capabilities, and the NVIDIA driver version in use:



Note:

Depending on the NVIDIA graphics card used, customers may need NVIDIA subscription or a license. Refer to the [NVIDIA product information](#) for details.

This completes the process for setting up a single VM to use GPU. The VM is now ready to run the full range of DirectX and OpenGL graphics applications supported by the GPU.

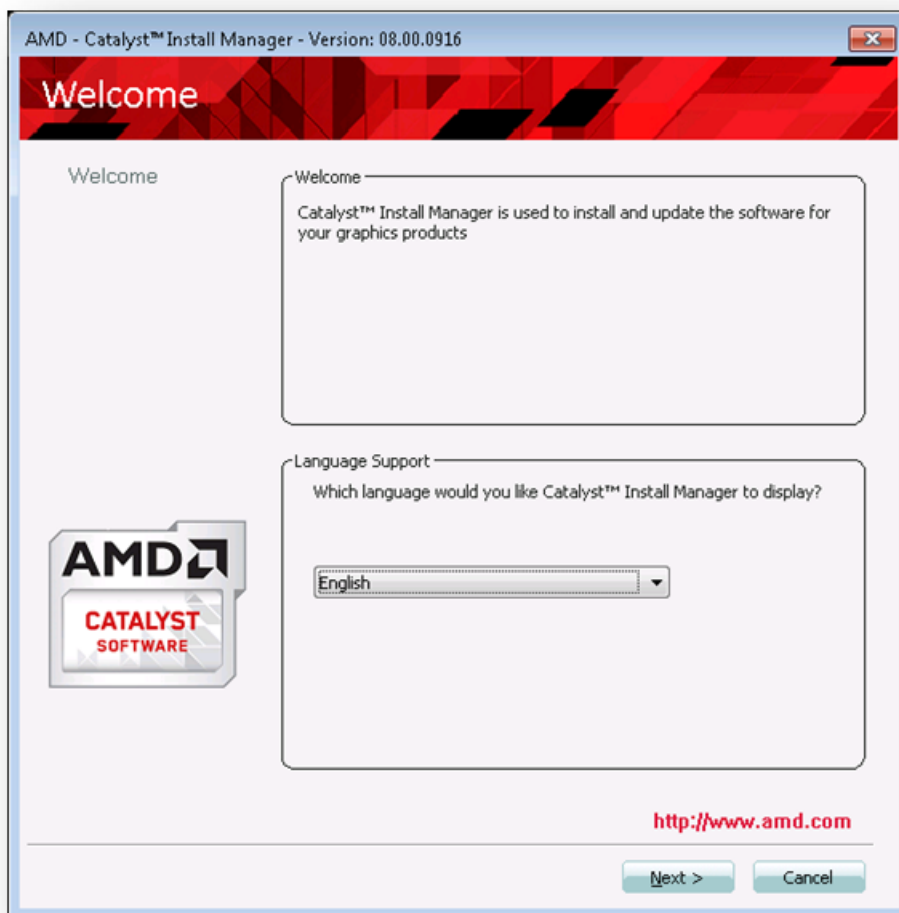
5.2.2. Install the AMD Drivers

To enable GPU operation, AMD drivers must be installed into the VM.

1. Start the VM. In the **Resources** pane, right-click on the VM, and click **Start**.

During this boot process XenServer dynamically allocates a GPU to the VM.

2. Follow the Windows operating system installation screens.
3. Once the operating system installation completes, reboot the VM.
4. Copy the 32 or 64-bit AMD Windows drivers (AMD Catalyst Install Manager) to the VM.
5. Run the AMD Catalyst Install Manager; select your **Destination Folder** and then click **Install**.

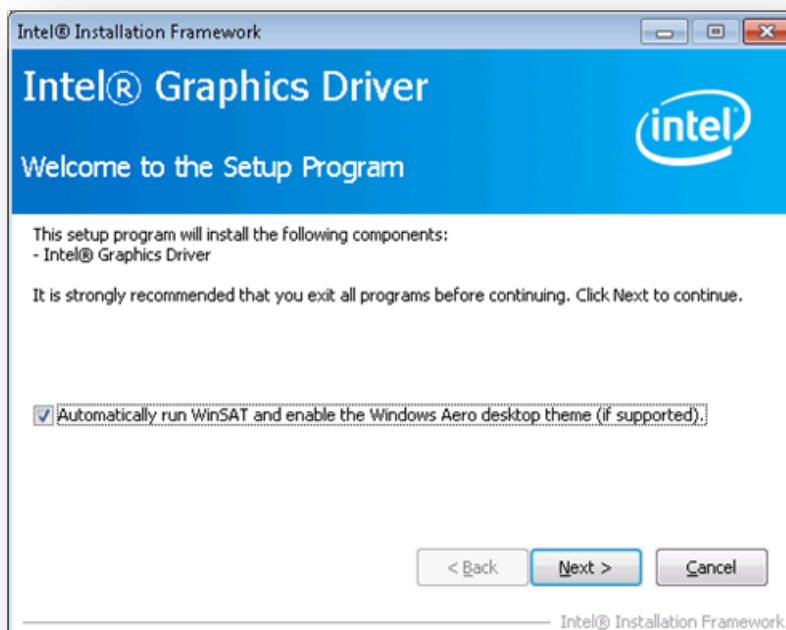


6. Follow the Installation wizard to install the driver.
7. To complete the installation, reboot your VM.
8. When the VM restarts, check that graphics are working correctly; open the Windows Device Manager, expand **Display adapters**, and ensure that the AMD Graphics Adapter does not have any warning symbols.

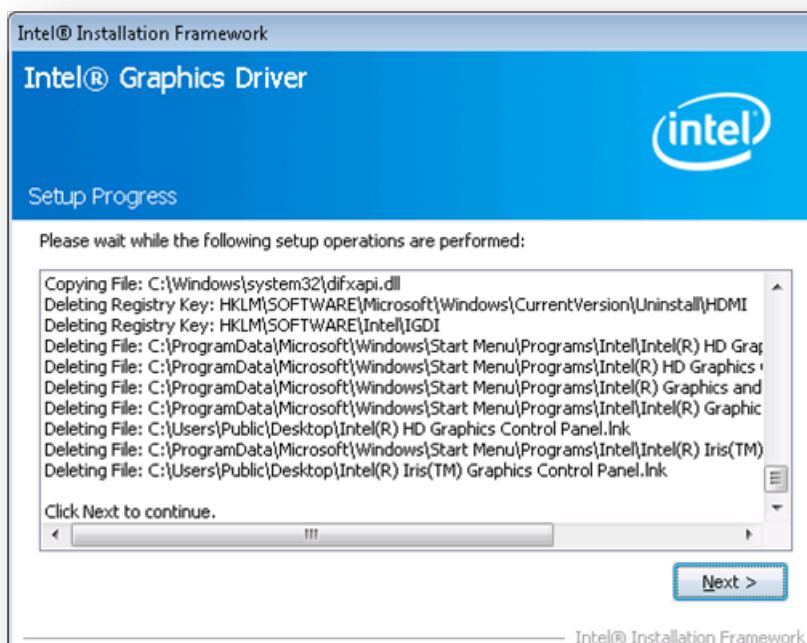
5.2.3. Install the Intel Drivers

To enable GPU operation, Intel drivers must be installed into the VM.

1. Start the VM. In the **Resources** pane, right-click on the VM, and click **Start**.
 During this boot process XenServer dynamically allocates a GPU to the VM.
2. Follow the Windows operating system installation screens.
3. Once the operating system installation completes, reboot the VM.
4. Copy the 32 or 64-bit Intel Windows driver (Intel Graphics Driver) to the VM.
5. Run the **Intel Graphics Driver** setup program
6. Select **Automatically run WinSAT**, and then click **Next**.



7. To accept the License Agreement, click **Yes**, and on the Readme File Information screen, click **Next**.
8. Wait while setup operations complete. When you are prompted, click **Next**.



9. In order to complete the installation, you will be prompted to restart the VM. Select **Yes, I want to restart this computer now**, and click **Finish**.



10. When the VM restarts, check that graphics are working correctly; open the Windows Device Manager, expand **Display adapters**, and ensure that the Intel Graphics Adapter does not have any warning symbols.

Note:

You can obtain the latest drivers from the [Intel website](#).

5.3. Configure VDA for HDX 3D Pro

1. Mount the XenDesktop installation media. If Autoplay is not enabled, navigate to and run AutoSelect.exe on the installation media.
2. In the XenDesktop installation wizard, select Virtual Delivery Agent for Windows Desktop OS:
3. On the Configuration screen, select Create A Master Image, then click Next.
4. On the HDX 3D Pro screen, click Yes, install the VDA for HDX 3D Pro, and click Next.
5. Work through the remainder of the Installation wizard.