Hardware Accelerated Graphics for VDI

Implementing Workstation Class Graphics Performance on HP VDI Architecture using HP Hardware Accelerated VDI Technology.

Technical White Paper

Document Version 1.0
# Table of contents

HP, Client Virtualization, and Hardware Accelerated Graphics for VDI ............................................ 4
The Enterprise Client Virtualization Reference Architecture for HP VirtualSystem ................................ 5
Blade Workstation ................................................................................................................................. 6
Hardware Accelerated Graphics for VDI Technology Concepts ............................................................... 6
  Shared GPU ........................................................................................................................................ 6
  Direct Mapped Graphics ..................................................................................................................... 7
  How many virtual machines per host ................................................................................................... 8
Hardware Accelerated Graphics for Server Virtualization ....................................................................... 9
Configuring Hardware Accelerated Graphics for VDI on WS460c Blade Workstation ............................. 9
  WS460c Blade Workstation Hardware Configuration ........................................................................... 10
  Citrix XenServer Direct Map System Configuration Recommendations ........................................ 12
  Microsoft Hyper-V / RemoteFX System Configuration Recommendations .................................. 14
Sample Hardware Accelerated Graphics for VDI Configuration Process .................................................... 15
  Installing Graphics cards in Blade Workstation ............................................................................... 15
  Setting up XenServer and HDX 3D Pro for GPU Accelerated VDI .................................................... 15
  Setting up Microsoft HYPER-V and RemoteFX for GPU Accelerated VDI .................................... 18
  Configuring Virtual Monitors ............................................................................................................ 21
For more information ............................................................................................................................. 24
**Purpose of this document**
- Gives IT decision makers, architects and implementation specialists an overview of how HP and its virtualization partners approach and implement Hardware Accelerated Graphics for VDI.
- Outlines the steps required to configure and deploy Hardware Accelerated Graphics for VDI.
- This document does not discuss the in-depth install and configuration of virtualization platforms except that which is unique to deploying hardware accelerated graphics for VDI solution on HP systems.

**Target audience**
This document is targeted at IT architects and engineers that plan on implementing virtualization platforms including Hardware Accelerated Graphics for VDI.

**Skill Set**
It is expected that the person utilizing this document will be familiar with servers, networking, storage principles and virtualization concepts. The installer should also be familiar with HP BladeSystem.

**Abbreviations and naming conventions**
Table 1 is a list of abbreviations and names used throughout this document and their intended meaning.

<table>
<thead>
<tr>
<th>Convention</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV</td>
<td>Client Virtualization</td>
</tr>
<tr>
<td>DAS</td>
<td>Direct attached storage</td>
</tr>
<tr>
<td>GPU</td>
<td>Graphical Processing Unit (Graphics Card)</td>
</tr>
<tr>
<td>HDX 3D</td>
<td>Citrix XenApp HDX 3D</td>
</tr>
<tr>
<td>HDX 3D Pro</td>
<td>HDX 3D Pro Graphics</td>
</tr>
<tr>
<td>ICA</td>
<td>Independent Computing Architecture protocol (used by Citrix HDX)</td>
</tr>
<tr>
<td>OA</td>
<td>Onboard Administrator</td>
</tr>
<tr>
<td>RBSU</td>
<td>ROM Based Setup Utility (BIOS)</td>
</tr>
<tr>
<td>RDP</td>
<td>Microsoft Remote Desktop Protocol</td>
</tr>
<tr>
<td>RFX</td>
<td>Microsoft RemoteFX</td>
</tr>
<tr>
<td>RGS</td>
<td>HP Remote Graphics Software</td>
</tr>
<tr>
<td>VDI</td>
<td>Virtual Desktop Infrastructure</td>
</tr>
<tr>
<td>vGPU</td>
<td>Virtual GPU, Virtualized Graphics</td>
</tr>
<tr>
<td>VM</td>
<td>Virtual Machine</td>
</tr>
</tbody>
</table>
What this document produces

Utilizing Citrix XenDesktop HDX 3D Pro, Citrix XenApp HDX 3D and/or Microsoft RemoteFX technologies, this document will help construct a blade system capable of supporting these new technologies from Citrix and Microsoft allowing for GPU accelerated virtual machines.

HP, Client Virtualization, and Hardware Accelerated Graphics for VDI

HP and its virtualization partners have introduced new technology that expands the capabilities of standard VDI to deliver full workstation class performance in VDI environments. This technology allows for true hardware graphics acceleration in virtual machines. This acceleration enables offloading of the graphical and remote protocol overhead of the virtual machines instead of relying on CPU intensive software rendered graphics. This technology enables not only the sharing of a GPU resource between many virtual machines, but also direct attaching dedicated GPUs to individual virtual machines, thus creating a workstation class virtual machine with all the features of a desk side workstation. This solution can be purchased separately (WS460c Blade) or integrated into the HP BladeSystem portfolio.

HP Client Virtualization with BladeSystem can help customers achieve the goals of IT and workforce support, without compromising performance, operating costs, information security, and user experience with HP Client Virtualization Reference Architectures. These reference architectures provide:

**Simplicity**: with an integrated data center solution for rapid installation and startup and easy ongoing operations
- Self-contained and modular server, storage, and networking architecture
- 3x improvement in IT productivity

**Optimization**: a tested solution with the right combination of compute, storage, networking, and system management tuned for Client Virtualization efficiency
- Scalable performance, enhanced security, always available
- 60% less rack space compared to competitors
- 95% fewer NICs, HBAs, and switches; 65% lower cost; 40% less power for LAN/SAN connections

**Flexibility**: with options to scale up and/or scale out to meet precise customer requirements
- Flexible solution for all workers in an organization from task workers to PC power users
- Support for up to 7,800 VDI users and 6,400 Citrix XenApp connections in three racks using the different desktop delivery methods offered by Citrix XenDesktop with FlexCast technology and leveraging Microsoft Hyper-V Dynamic Memory.
- Unmatched price/performance with both DAS and SAS tiered storage in a single rack (50% cheaper than SAN)

**Workstation Performance**: with options that use hardware accelerated graphics to meet the demands of customers high end user needs
- Flexible solution with options that give customers a full range portfolio of VDI solutions meeting the requirements from your least demanding task worker to your most demanding high end knowledge worker.
• Current support for up to two virtual machines per host with directly attached (1:1) GPU for graphics acceleration. Remaining resources on each host can be used for standard VDI users (not hardware accelerated.)

• Current support for theoretical max of up to 32 virtual machines per GPU with shared (many:1) GPU for graphics acceleration. Remaining resources on each host can be used for standard VDI users (not hardware accelerated.)

• Support for offloading the video rendering and protocol overhead to GPU improving graphical and host CPU performance.

The Enterprise Client Virtualization Reference Architecture for HP VirtualSystem

VDI is one possible implementation of the Client Virtualization (CV) Reference Architecture. Blade Workstations and server based computing also fit the CV model. With VDI, a desktop is created as a virtual machine. Applications and user personality are injected into the core desktop virtual machine and a brokering mechanism manages connecting end users to the virtual machine through a connection protocol.

Figure 1: Architecture of a VDI implementation.

More than simply a virtual machine, the runtime VDI instance is the real time compilation of the end user’s data, personal settings and application settings with a core operating system instance and shared generic profile where the applications are either installed locally as a fully installed or packaged instance or streamed from outside the virtual machine. While seemingly complex at first glance, the proper management of these individual pieces yields a far more flexible end user computing environment that is simpler to manage. This is accomplished in part by the decoupling of the user and application from the core operating system. Once the user and application are segmented from the device, the device itself
starts to matter less as well. With the right approach, an end user can receive the same experience regardless of what core compute resource they are logged onto at the moment.


**Blade Workstation**

HP has long enjoyed a strong market position deploying Client Virtualization to end users that require a robust compute environment. These users might include graphics designers, financial traders, or any number of heavy compute users. An HP ProLiant WS460c Workstation Blade provides a local workstation experience to end users over the network using the HP Remote Graphics Software protocol, Citrix ICA (Used by XenApp), HDX 3D, XenDesktop HDX 3D Pro and Microsoft Remote Desktop Protocol 7.1 (Standard and RemoteFX). Historically this has been a 1:1 solution, meaning one full blade is dedicated to a single user.

HP has been working for several years to develop a way to bring high end graphical performance of the Blade Workstation into the virtual environment. The wait is over with the introduction of WS460 G6 Blade Workstation support for Xen Server multi GPU pass-through and Microsoft RemoteFX (RFX) technologies. These new technologies enable the user to get full fidelity “just like desktop” functionality with RemoteFX or a workstation class performance in a virtual environment with XenServer and HDX 3D Pro.

**Hardware Accelerated Graphics for VDI Technology Concepts**

**Shared GPU**

Also referred to as “Virtual GPU”, this method allows the GPU to be managed by the hypervisor and multiple virtual machines can share the resource of the GPU (see figure 2). Similar in concept to the way that hypervisors share CPU cycles via vCPU’s, they also can share the resources of the GPU through a custom video driver in each virtual machine. In this model, the GPU driver is loaded within the hypervisor and it manages access to the GPU by the virtual machine.

Enterprise Hypervisors and Servers using this technology include:

- Microsoft Server 2008r2 Enterprise SP1 or later with Hyper-V RemoteFX.
- Citrix XenApp (with Microsoft Server 2008r2 Enterprise R2 SP1)

**Note** – When you use XenApp on XenServer, each VM is an instance of the multi-user XenApp product. Each of these virtual machines running XenApp has access to one server-side GPU(via XenServer GPU Pass-through) which can be shared by multiple users for hardware acceleration of DirectX-based graphics.
Advantages:

- Lower cost of ownership in that many virtual machines benefit from the resource of a single GPU.
- Theoretical Max of up to 32 virtual machines supported by a single GPU (Dependent on card used, monitor layout and resolution).
- Allows each user to have a power user performance with support for some 3D, Windows Aero etc.
- Can designate users that have access to GPU while others still run in a software graphics mode.

Disadvantages:

- GPU can become a performance bottleneck as many users draw on resource of one card.
- Some limitation on 3D content that can be used. OpenGL not fully supported.
- Unacceptable performance for high end 3D knowledge or workstation user. Extremely Limited support for OpenGL applications.

Direct Mapped Graphics

Also referred to as “Mulit-GPU Pass-Through” or just “GPU Pass-Through”, this method allows the GPU to be directly mapped to a virtual machine for dedicated 1:1 use (see Figure 3). This method enables the virtual machine to have full access to GPU allowing for full workstation class graphics performance in a virtual machine. This method works well for high end 3D and GPU Compute users. HP provides a full range of graphics cards from mid-range MXM form factor to top end full length PCIe cards.

Enterprise Hypervisors using this technology include:

- Citrix Xen Server 6.0 or newer with Citrix XenDesktop HDX 3D Pro.
Figure 3: Virtual GPU Model

Advantages:
- Workstation class graphics experience in a virtual machine
- Support for all 3D technologies and application.
- Can designate users that have access to GPU while others still run in a software GPU mode.

Disadvantages:
- Higher cost of ownership per connection as it has a dedicated GPU per virtual machine.
- Some limitations on running 3D applications in full screen mode at this time with HDX 3D Pro.

How many virtual machines per host

In order to determine how many virtual machines per host, you use the same calculation you would use in a non-hardware accelerated environment but simply compensating for the amount of resources you give to the graphics accelerated virtual machines as you will give them more resources capabilities.

XenServer Direct Mapped Considerations
- When you direct map graphics to a virtual machine, giving it a workstation class graphics performance, you will also need to give it workstation class resources.
  - Recommendation for best performance
    - Memory - 2GB or more.
    - CPU - 2 virtual cores per virtual machine or more.
- Once you have defined the resource needs for your "virtual workstation", you can define the remaining resources available for virtual machines running standard VDI session to maximize the resource utilization of each host server.

Hyper-V RemoteFX Considerations
- The GPU has a dedicated amount of video RAM, for example, an NVIDIA Quadro 6000 has 6GB of video RAM. Microsoft RemoteFX virtual machines consume a specific amount of video RAM based on the max number of monitors and resolution set for each virtual machine. This will
dictate the maximum number of virtual machine per GPU. The following table gives you the
memory allocation based on monitor and resolution configuration.

<table>
<thead>
<tr>
<th>Maximum Resolution</th>
<th>Host System Memory reserved by max number of monitors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1024x768</td>
<td>85MB</td>
</tr>
<tr>
<td>1280x1024</td>
<td>98MB</td>
</tr>
<tr>
<td>1600x1200</td>
<td>114MB</td>
</tr>
<tr>
<td>1920x1200</td>
<td>124MB</td>
</tr>
</tbody>
</table>

- RemoteFX technology shares the resources of the GPU, allowing multiple virtual machines to use the recourse of that GPU. The more virtual machines you configure with RFX per host, the lower the potential performance will be.
- Adding multiple graphics cards enhances performance as Hyper-V will load balance between cards as virtual machines start up.
- Recommendation for memory and vCPU allocation is the same as recommended by Microsoft for RemoteFX.
- Once you have defined the resource needs for your "virtual workstation", you can define the remaining resources available for virtual machines running standard VDI session to maximize the resource utilization of each host server.

Hardware Accelerated Graphics for Server Virtualization

Although not thoroughly covered in this paper, GPU accelerated VDI is also available for server virtualization. GPUs can be attached in the same way as client operation systems to server virtual machines when supported. Possible use cases include: XenApp and Hyper-V for RemoteFX.

Configuring Hardware Accelerated Graphics for VDI on WS460c Blade Workstation

The following section covers only configuration requirements specific to the HP WS460c G6 and HP BladeSystem infrastructure, it does not cover the full hypervisor configuration. See the documentation specific to vendor for their specific configuration steps.

- Citrix XenServer and XenDesktop Documentation - [http://support.citrix.com/edocs](http://support.citrix.com/edocs)
WS460c Blade Workstation Hardware Configuration


The following information complements the Blade Workstation documentation with configuration requirements specific to Blade Workstation WS460c G6 being used for Hardware Accelerated Graphics for VDI.

### HP BladeSystem Hardware / Firmware recommended minimum revisions

<table>
<thead>
<tr>
<th>Components</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP Onboard Administrator</td>
<td>3.30</td>
</tr>
<tr>
<td>HP Virtual Connect</td>
<td>3.18</td>
</tr>
<tr>
<td>HP ProLiant Server System ROM</td>
<td>Blade Workstation WS460c 05/05/2011</td>
</tr>
<tr>
<td>HP Integrated Lights-Out 2 (iLO 2)</td>
<td>2.05</td>
</tr>
</tbody>
</table>

### Supported Video cards for HP Hardware Accelerated Graphics for VDI solution

- NVIDIA Quadro 4000, 5000, 6000
- Quadro MXM Fermi-based form factor cards (not yet available at this writing).
- Quadro Tesla M2070Q

### Understanding Blade Workstation video subsystem modes

The WS460c G6 blade workstation has four distinct graphics modes (see Figure 3).

#### User Mode

- This is the primary production mode when used in Microsoft RemoteFX environments.
- In this mode the NVIDIA card is enabled and the embedded ATI is disabled.
- In this mode the iLO console is inaccessible because it uses the ATI to generate its video.
- System can be accessed via remote protocol only.
- This is the primary Production mode when used in all non-virtualized environments.

#### Setup Mode

- This is the primary production mode when used in Citrix XenServer environments.
- This mode has both video cards enabled but the NVIDIA card is secondary.
- In this mode the iLO console is accessible.
Admin Mode

- This mode disables the NVIDIA card and only the ATI is active.
- Add in graphics cards are disabled but system still registers as a WS460c Blade Workstation.

Server Mode

- This operates the same as Admin mode but switches the BIOS to run as ProLiant BL server and not a workstation.

Configuring blade workstation video mode.

As shown in Figure 4, Remote Console Mode can be viewed or changed during POST. To view or set Remote Console Mode during boot (POST):

1. Using either iLO remote console or the Local I/O Connector, view the boot console using iLO 2, which provides direct control of the Blade Workstation.
2. When prompted during boot, press the F9 key. The ROM-based Setup Utility appears.
4. To change the Remote Console Mode, press Enter. The Remote Console Mode menu appears. Use the Up and Down arrow keys to select the desired mode. When done, press Enter and then perform the steps indicated to exit the ROM-based Setup Utility.
5. The workstation blade performs a reboot, and then the Remote Console Mode appears.

Setting the HP Graphics Expander BIOS settings

The current generation of Blade Workstation (ws460c G6) supports up to two full size graphics cards in the graphics expansion bay. The PCIe graphics expansion bay slots can be configured for one card at x16 or two cards at x8.

A BIOS setting dictates if the system can see one or two cards. This setting must be manually set.

Setting “HP Graphics Expander x16”

Enabled
- This mode enables one slot at x16

Disabled
Configuring the blade workstation PCIe mode

PCIe mode can be viewed or changed during POST. To view or set PCIe Mode during boot (POST):

1. Using either iLO remote console or the Local I/O Connector, view the boot console using iLO 2, which provides direct control of the Blade Workstation.
2. When prompted during boot, press the F9 key. The ROM-based Setup Utility appears.
3. As shown in Figure 5, select System Options > HP Graphics Expander x16. The current Expander mode appears.
4. To change mode, press Enter. The mode menu appears. Use the Up and Down arrow keys to select the desired mode.
   a. Enable – Turns on one card at x16
   b. Disable – Turns on two cards at x8
5. When done, press Enter, and then perform the steps indicated to exit the ROM-based Setup Utility.
6. The workstation blade performs a reboot, and then the Remote Console Mode appears.

Citrix XenServer Direct Map System Configuration Recommendations

<table>
<thead>
<tr>
<th>Required BIOS setting for Xen Server 6 on Blade Workstation</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced Options / Option ROM Loading Sequence</td>
<td>Load Option Card Devices First</td>
</tr>
<tr>
<td>Power Management Options / Adv. Pow. Mgnt Options / Min. Processor Idle Power Core State</td>
<td>No C-States</td>
</tr>
<tr>
<td>Video Mode (see above section)</td>
<td>Setup Mode</td>
</tr>
</tbody>
</table>
### Citrix required minimum versions

<table>
<thead>
<tr>
<th>Components</th>
<th>Software description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citrix XenServer</td>
<td>Citrix XenServer 6</td>
</tr>
<tr>
<td>Citrix XenDesktop Controller</td>
<td>XenDesktop 4, 5 or 5.5</td>
</tr>
<tr>
<td>Citrix XenDesktop Virtual Desktop Agent</td>
<td>XenDesktop 5.5 VDA</td>
</tr>
<tr>
<td>Citrix Receiver</td>
<td>Citrix Receiver 3.0 with online plug-in 13.0.0.6684, Receiver for Linux 12.0 Works with any other Citrix Receiver, too, using standard codecs</td>
</tr>
<tr>
<td>XenDesktop Setup Wizard</td>
<td>Citrix XenDesktop 5.5 Setup Wizard</td>
</tr>
</tbody>
</table>

### End user virtual machines

<table>
<thead>
<tr>
<th>Components</th>
<th>Software description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supported Operating Systems for HDX 3D Pro.</td>
<td>Microsoft Windows 7 (Recommended) Windows XP (Performance Limitations)</td>
</tr>
<tr>
<td>High Performance Connection Protocols</td>
<td>Citrix ICA with HDX 3D Pro</td>
</tr>
<tr>
<td>Standard Performance Connection Protocols</td>
<td>Microsoft RDP, Citrix HDX(JCA)</td>
</tr>
</tbody>
</table>

### XenServer 6 Important notes
- Use the latest available driver for virtual machine client OS from NVIDIA web site.
- No driver for NVIDIA card is installed on XenServer, the driver is loaded directly on the virtual machine just like a desk side unit after the GPU is assigned to the VM.
- Because XenServer does not load drivers for the NVIDIA cards, the systems can run in “Setup” video mode to allow access through iLO remote consoles.
- XenServer only allows one GPU to be attached to any virtual machine.
- For best performance, run HDX 3D Citrix ICA and HDX 3D Pro on client devices with a dual core 3GHz processor are recommended. (Minimum requirement for a single monitor configuration is a single core 2 GHz processor.) For more information, see the Citrix documentation.
- Need to configure virtual monitors in the GPU control panel of the VM to support multiple monitors with Citrix XenDesktop HDX 3D Pro.

### Configure HDX3D Pro for best performance
- 2 vCPU per virtual machine or more.
  - In order to set more than two cores per socket (allowing for 4 cores or more) on a virtual machine, you must configure the virtual machine to enable multiple cores per socket. Use the following command for XenServer console to enable up to 4 cores per socket and set the virtual machine to use 4 cores at next boot.
xe vm-param-set uuid=<VM UUID> platform:cores-per-socket=4 VCPUs-max=8 VCPUs-at-startup=4

- 2GB or more of virtual machine memory.
- Turn off Aero support unless needed. Aero will consume one CPU core.
- Set FRAMECAPTURERATE on the HDX3D Pro configuration tool to at least “26” on the virtual machine. For more information, see HDX 3D Pro documentation on Admin Control Tool.
- If local LAN has good network bandwidth, “2D Drawing” setting can be used on the client. This setting is also used for complex 2D drawing applications in a WAN environment.
- Smaller screen resolutions perform better.

Microsoft Hyper-V / RemoteFX System Configuration Recommendations

### Microsoft required minimum versions

<table>
<thead>
<tr>
<th>Components</th>
<th>Software description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microsoft RFX</td>
<td>Microsoft Windows Server 2008 R2 SP1</td>
</tr>
<tr>
<td>Microsoft RDP</td>
<td>Microsoft RDP 7.1 (Included in Windows 7 SP1)</td>
</tr>
</tbody>
</table>

### Required BIOS setting for Xen Server 6 on Blade Workstation WS460c G6

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced Options / Option ROM Loading Sequence</td>
<td>Load Option Card Devices First</td>
</tr>
<tr>
<td>Video Mode for Server Setup</td>
<td>Setup mode</td>
</tr>
<tr>
<td>Video Mode for Production</td>
<td>User Mode</td>
</tr>
</tbody>
</table>

### End user virtual machines

<table>
<thead>
<tr>
<th>Components</th>
<th>Software description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supported Operating System for RemoteFX</td>
<td>Microsoft Windows 7 SP1</td>
</tr>
<tr>
<td>High Performance Connection Protocols</td>
<td>RemoteFX (RFX)</td>
</tr>
<tr>
<td>Standard Performance Connection Protocols</td>
<td>Microsoft RDP</td>
</tr>
</tbody>
</table>
Microsoft RemoteFX Important notes

- Use the latest available driver for Windows Server 2008 R2 from NVIDIA web site.
- Because Windows 2008 does load drivers for the NVIDIA cards, but does not support two types of video cards, the following modes must be used.
  - Setup mode – for systems install, setup, and NVIDIA driver loading.
  - User Mode – For production RFX mode.
    - In this mode it is not possible to reach iLO console through iLO, OA, or front i/o dongle. You must enable remote RDP console for systems management.
- You can install more than one graphics card of same type and Windows will load balance between them at virtual machine startup.
- GPU assignments are not dynamically managed after virtual machine startup for load balancing.
- At this time this solution only has minimal support of Open-GL applications.

Configure RemoteFX for best performance

- Set the following group policy setting to “Highest (Best Quality)”
  - Policy: Optimize visual experience when using RemoteFX
  - Location: Computer Configuration\Administrative Templates\Windows Components\Remote Desktop Services\Remote Desktop Session Host \Remote Session Environment
- When connecting to RFX session from RDP client, ensure the following settings are configured:
  - Display Tab
    - Colors setting set to “Highest Quality (32bit)”.
  - Experience Tab
    - Connection Speed must be set to “LAN (10Mbps or higher)”.
    - All check boxes should be checked.
- Recommend 2 vCPU per virtual machine or more.
- Recommend 2GB or more of virtual machine memory.
- Turn off Aero support unless needed.
- Fewer virtual machines per GPU improves performance potential.

Sample Hardware Accelerated Graphics for VDI Configuration Process

Installing Graphics cards in Blade Workstation

To Install Blade workstation with graphics card, Refer to Maintenance guide for instructions on how to install graphics card at:

Setting up XenServer and HDX 3D Pro for GPU Accelerated VDI

1. Install XenServer 6.
2. Create or Import new virtual machine and install operating system and XenServer tools.
3. Configure and update operating system to your specifications.
4. Install XenDesktop Agent
   a. Select “Install Virtual Desktop Agent”.
   b. Select Advanced.
   c. Select HDX 3D Pro and point to key file.

![Select Components to Install](image)

e. Enter address for the desktop controller and validate if needed.

![Controller Location](image)

f. Finalize any remaining VM configuration you need to do.

g. Shut down VM.

h. From XenCenter, Right click VM and Select properties.

i. Under properties, select GPU on menu.

![GPU](image)

j. Select the NVIDIA card installed in your system.
IMPORTANT NOTES

- Once a graphics card is installed and seen by XenServer, the card will remain in list even if it is not physically installed anymore.
- DO NOT choose the embedded card, ATI in this case.
- You can only add as many cards to VMs as you have actual cards. Even if they are not turned on.
- You can only add one direct attached GPU to any virtual machine.

k. Start Virtual Machine
l. Once started, you can go to device manager and see an extra video card showing up as a VGA generic card.
m. Install NVIDIA driver on Virtual Machine. Use latest NVIDIA driver from their web site for the card and client operating system used.
n. Reboot
o. After reboot, if successful, when the virtual machine and the NVIDIA card initialize the virtual machine default console in XenCenter will seem to be locked up (See Below). It is not locked up; the video output has been redirected to NVIDIA card. If you need to do maintenance, you need to log in from HDX or RDP remote console or remove the NVIDIA card.
p. You are now ready to create XenDesktop group and connect via HDX 3D. See Citrix documentation for further instructions.

Setting up Microsoft HYPER-V and RemoteFX for GPU Accelerated VDI

1. Install Server 2008 R2 SP1.
2. Download and install latest driver from NVIDIA for Server 2008 R2 x64
3. Reboot and confirm driver loaded by going to device manager.
5. Reboot
6. Create and setup VM according to Microsoft documentation and your company guidelines.
7. Once created shutdown virtual machine.
8. Right click on virtual machine and select “Settings”.
9. On right side of “Settings” screen, select “RemoteFX 3D Video Adapter”.

![RemoteFX 3D Video Adapter](image)

10. Click “Add”
11. Select max number and resolution size of displays.
   **Note** – Your choice will dictate the max amount of RemoteFX virtual machines you can run per GPU. See “Microsoft Hyper-V / RemoteFX System Configuration Recommendations” section of this document for more information.

![Settings for VM](image)

12. Boot VM and allow time to fully start. During this boot the RemoteFX driver will be loaded in the virtual machine and enabled on next boot.
   **Note** – As longs as you are using Win7 for your virtual machine, the virtualization tools and drivers are already included and do not need to be installed separately.
13. Restart virtual machine.
14. On next boot, if all is configured correctly you should get the following screen on the virtual machine you have configured stating that “Video remoting was disconnected” and software console will not be displayed.

![Video remoting was disconnected](image)
15. On the end client, start remote desktop connection client.
16. For full RemoteFX performance, set the following settings on these tabs before connecting.
   a. Set display colors to “Highest Quality (32 bit)”
      **Note** – Not setting 32 bit quality will disable RemoteFX.
   b. Enable all on Experience tab and select LAN connection type.
      **Note** – Not selecting LAN connection type will disable RemoteFX.

17. Connect to VM.
**Note** – Quick way to determine you are in RemoteFX mode is you have “Shutdown” option on start menu (Assuming you don’t set policy in GPO). Standard RDP session does not allow this and you will only get Logoff.

**Configuring Virtual Monitors**

When running in GPU pass-through mode the virtual machine has direct access to graphics card. However in most cases there is not an actual physical display attached to the card on the host. NVIDIA drivers by default will not initialize if there is not a monitor physically connected. Although it is an option to physically connect monitors to all graphics cards on host system, it’s not practical as they are not actually used.
To resolve this, you can create a virtual monitor by connecting physical monitors to any host running a NVIDIA Quadro card and export the monitor information, this information is called EDID information. Once exported, you can import it in the NVIDIA control panel on the virtual machine running pass-through graphics, thus creating a virtual monitor.

To Create and export an EDID file:

1. On a system with an NVIDIA Quadro card installed and connected to physical monitors you wish to capture information, open the NVIDIA control panel and select "View System Topology", then click on the DVI “**Connected: HP….. (Manage EDID..)”**.

2. The following screen comes up, click “Export EDID”.

3. Save to file.
To import an EDID file.

1. On a virtual machine configured for pass-through and with NVIDIA card driver installed, open the NVIDIA control panel and select “View System Topology”, then click on the DVI “Not Connected: (Manage EDID.).”.

2. Under “Load EDID” select connector type “DVI”.
3. Click “Load EDID” button, and browse to the EDID file you with it import and load.

4. **Important Note** If loading in HDX3D while in a session, the screen may become unresponsive to mouse, simply use your arrow key to select to save the NVIDIA configuration before the timeout is reached. After save, disconnect session and reconnect, the mouse control should be back.

5. At this point the VM NVIDIA driver will detect a monitor connected even though there is none.
For more information

To read more about HP and Client Virtualization, go to www.hp.com/go/cv
Other documents in the Client Virtualization reference architecture series can be found at the same URL.

Microsoft Hyper-V, http://www.microsoft.com/hyper-v

To help us improve our documents, please provide feedback at http://h20219.www2.hp.com/ActiveAnswers/us/en/solutions/technical_tools_feedback.html.