# HP Graphics Server Blade VMware vDGA Quick Start

Implementation procedure overview on WS460c Graphics Server Blades

Version 1, December 2013

## Table of contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose of this document</td>
<td>2</td>
</tr>
<tr>
<td>Prerequisite Recommend Reading</td>
<td>2</td>
</tr>
<tr>
<td>HP WS460c Gen8 Graphics Server Blade VMware vDGA guidelines</td>
<td>2</td>
</tr>
<tr>
<td>VMware system configuration recommendations</td>
<td>2</td>
</tr>
<tr>
<td>VMware vDGA Considerations</td>
<td>3</td>
</tr>
<tr>
<td>Tips and Tricks you need to know - Tech Nuggets and Known issues</td>
<td>3</td>
</tr>
<tr>
<td>Setting up VMware for GPU Accelerated VDI vDGA</td>
<td>4</td>
</tr>
<tr>
<td>Configuring vDGA pass-through device via Virtual Center</td>
<td>4</td>
</tr>
<tr>
<td>Configuring vDGA pass-through device via Virtual Center Web Interface</td>
<td>7</td>
</tr>
<tr>
<td>Configure and Connect to vDGA configured VM</td>
<td>10</td>
</tr>
<tr>
<td>Appendix A – Graphics Server Blade Video Modes</td>
<td>11</td>
</tr>
<tr>
<td>Resources</td>
<td>13</td>
</tr>
</tbody>
</table>
Purpose of this document

Provides an overview of the configuration steps for setting up VMware vDGA technology on HP WS460c Graphics Server

Prerequisite Recommend Reading


- HP “Hardware Accelerated Graphics for Desktop Virtualization” for a detailed conceptual overview of the overall solution before proceeding. This document can be found at: http://h71028.www7.hp.com/enterprise/downloads/4AA4-1701ENW.pdf

HP WS460c Gen8 Graphics Server Blade VMware vDGA guidelines

VMware system configuration recommendations

<table>
<thead>
<tr>
<th>Table 1. Required BIOS setting for Graphics Server Blade</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Setting</strong></td>
</tr>
<tr>
<td>Power Management Options/HP Power Profile</td>
</tr>
<tr>
<td>Video Mode for Setup and Production</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2. VMware required minimum versions for vDGA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Components</strong></td>
</tr>
<tr>
<td>ESXi Host</td>
</tr>
<tr>
<td>VMware Horizon View</td>
</tr>
<tr>
<td>VMware Horizon View Agent</td>
</tr>
<tr>
<td>Horizon View Client</td>
</tr>
<tr>
<td>NVIDIA Graphics Driver for VM (vDGA only)</td>
</tr>
</tbody>
</table>
VMware vDGA Considerations

- Using VMware vDGA technology directly attaches a GPU to a VM for maximum graphics performance. For that reason, it is required that proper video driver be installed within the virtual machine.
- High availability and vMotion is not supported when host is in vDGA mode.
- V10 version hardware on a VM can no longer be managed via Virtual Center standard console, must use web access to manage.
- If using VMware vSGA and vDGA technologies, if you have a GPU configured for vDGA it cannot be used for vSGA.

Tips and Tricks you need to know - Tech Nuggets and Known issues

- When configuring VM with vDGA, only select the PCI GPU device itself and not the Audio Device on GPU's that support it.
- There is a known issue where after configuring VM for vDGA, when you try to start the VM it will shut down after a few seconds. To resolve, make sure the following is included as a VM configured parameter:
  - Right Click VM and click “Edit Settings”
  - Click “Options” tab.
  - Click Advanced/General and then click button on lower right “Configuration Parameters”
  - On Bottom right click “Add Row”
  - Manually add the following rows if they don’t exist

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>pciPassthru.msiEnabled</td>
<td>False</td>
</tr>
<tr>
<td>pciHole.start</td>
<td>2048</td>
</tr>
</tbody>
</table>

NOTE – Only the pcihole.start parameter is needed for this specific issues, however as seen in the configuration section later, we recommend both.

- The Graphics Server Blade support multiple video modes, when using VMware, only “SETUP” mode needs to be used.
  - If unfamiliar with video modes, please read Appendix A before continuing.
- In order to activate the NVIDIA display adapter in vDGA on a newly configured VM, you must connect via PCoIP, and at full screen from the endpoint (at native resolution), or the virtual machine will use the SVGA 3D display adapter.
  - To test, attempt to open the NVIDIA control panel. If it fails the NVIDIA adapter is not active.
- vDGA will not work through the vSphere console session, you will only see a blank screen in this mode.
- vDGA multi-monitor support is dynamic, it will try to automatically detect your endpoint client monitor configuration and display on all. All monitors must be the same and GPU must support multi-monitor.
  - Although not fully supported, quad display will only work on GPU’s that support it, for example NVIDIA K5000 and K6000.
  - If you try to connect from an endpoint configuration that is not supported and vDGA cannot dynamically configure session, you may get a single display output across all monitor as seen in this example.
    - To Resolve this, go back to a single monitor full screen or log out and then connect with a supported configuration.
Setting up VMware for GPU Accelerated VDI vDGA

Configuring vDGA pass-through device via Virtual Center

1. Update the WS460c Graphics Server Blade with the latest HP SPP firmware by booting to the SPP DVD. You can obtain the latest SPP at: [http://www.hp.com/go/spp](http://www.hp.com/go/spp)

2. Insure the System BIOS level is at least 9/20/2013

3. Setup system BIOS for installation
   a. Boot to the HP RBSU (BIOS configuration utility) by pressing F9 during boot
   b. Go to and select “System Default Options” > “Restore Default Options” and select to restore
   c. After reset boot back into HP RBSU (BIOS configuration utility) by pressing F9 during boot
   d. Set video mode to “SETUP”
      i. Select System Options
      ii. Select Remote Console Mode
      iii. Set to “SETUP” mode
   e. Set blade to static high performance mode
      i. Select Power Management Options > HP Power Profile
      ii. Select and enable “Maximum performance”

4. Install and complete basic configuration of vSphere host according to your specification

5. Ensure graphics cards are properly installed in workstation blade

6. Confirm GPU is seen by host, Enter Following Command:
   
   ```bash
   # esxcli hardware pci list --c 0x0300 --m 0xff
   ```

   You should see an output similar to the following:

   ```
   000:00:0:0
   Address: 000:00:0:0
   Segment: 0x0000
   Bus: 0x01
   Slot: 0x00
   Function: 0x00
   VMkernel Name:
   Vendor Name: NVIDIA Corporation
   Device Name: NVIDIA Quadro 6000
   ```

7. From vCenter, go to configuration tab for vDGA host server and select “Advanced Options”

8. On right side, click “Edit”
9. Once on edit screen you will see all the PCI devices on the systems. Scroll down to find the graphics cards you wish to connect to. In this example we have eight NVIDIA Quadro 3000M cards. Select every card you plan to use in pass-through then click “OK”

10. Back on the advanced configuration screen, select refresh and the newly selected pass-through device/s should be seen, and host message will appear stating a reboot is needed to complete configuration

11. Reboot Host

12. After reboot, go back to advanced configuration screen to confirm all cards are successfully set up for pass-through

13. Create and configure VM to be used with supported operating systems, VMware tools, operating system updates, and software

14. Shutdown VM

15. **IMPORTANT** - Edit the VM configuration on all vDGA VM’s.
   a. Right Click VM and click “Edit Settings”
   b. Click “Options” tab
   c. Click Advanced/General and then click button on lower right “Configuration Parameters”
   d. On Bottom right click “Add Row”
e. Manually add the following rows if they don't exist

<table>
<thead>
<tr>
<th>PCI Passthru 0.msiEnabled</th>
<th>PCI Passthru 0.msiEnabled</th>
</tr>
</thead>
<tbody>
<tr>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>PCIPassthru0.msiEnabled</td>
<td>PCIPassthru0.msiEnabled</td>
</tr>
<tr>
<td>PCI Hole /start</td>
<td>PCI Hole /start</td>
</tr>
<tr>
<td>2048</td>
<td>2048</td>
</tr>
</tbody>
</table>

16. Add the graphics device to you VM
   a. Edit VM settings and click “Add” to add new device
   b. From Add Hardware screen select “PCI Device” then “Next”
   c. From “Choose PCI Device” screen, select available PCI device not used in another VM

   ![Add Hardware Screen](image1)
   ![Choose PCI Device Screen](image2)

**IMPORTANT NOTE** – If using NVIDIA cards you may see audio card listed as well. Do not pass-through the audio controller if available on the card. In above example the audio controller is listed as 00.1.

17. Start VM
18. Once stated, log in and go to Device Manager
19. Under “Display Adapters” you should see a new “Standard VGA Graphics Adapter”, that is your graphics card without drivers loaded, as well as standard “VMware SVGA Adapter”

![Device Manager Screen](image3)

20. Load supported version of NVIDIA graphics driver from NVIDIA web site
21. Once you reboot the VM after installing the driver, the VCenter console will go black as the desktop is going to the NVIDIA card. To manage the VM, you will need to remove GPU or access remotely
22. Install View Horizon Agent on VM.
23. In some cases it may be necessary to enable NVIDIA API, usually it is not needed.
   C:\Program Files\Common Files\VMware\Teradici\PCoIP Server\MontereyEnable.exe” -enable
   Note: If MontereyEnable.exe is not found, use NvFBCEnable.exe. In the new SDK, MontereyEnable is replaced with NvFBCEnable.
24. After the process is complete, Restart the virtual machine.
25. Connect to VM, skip to section “Configure and Connect to vDGA configured VM”

**Configuring vDGA pass-through device via Virtual Center Web Interface**

1. Log into Virtual Center Web Access
2. Go to “Host and Clusters” view
3. Click on server host to be configured
4. Click on “Manage” tab on right
5. Click on Hardware/PCI Devices
6. In the “DirectPath I/O PCI Devices” screen there most likely is nothing shown as we have not configured GPUs.
7. Click on the pencil icon to edit
8. Scroll through the screen to find the GPU’s that you wish to pass-through and select
9. Click “Save” on DirectPath screen and then click refresh until the cards show up
10. Card will show up similar to the following example and on the server it will indicate a reboot is required

11. Reboot the host
12. After reboot, go back to DirectPath page and confirm cards are shown

13. Create and configure VM to be used with supported operating systems, VMware tools, operating system updates, and software
14. Shutdown VM
15. **IMPORTANT** - Edit the VM configuration on all vDGA VM’s.
   a. Right Click VM and click “Edit Settings”
   b. Click “Options” tab
   c. Click Advanced/General and then click button on lower right “Configuration Parameters”
   d. On Bottom right click “Add Row”
   e. Manually add the following rows if they don’t exist

16. Right click on VM and go to setting
17. At bottom of setting page, click on New Devices Drop down
18. Select “PCI Device”
19. Select an available GPU for the VM

**IMPORTANT NOTE**
– Only select the PCI GPU devices and not the Audio Device.

**IMPORTANT NOTE**
– Only one GPU per device is supported.

20. Start VM.
21. Once stated, log in and go to Device Manager.
22. Under Display Adapters you should see a new “Standard VGA Graphics Adapter”, that is your vDGA graphics card without drivers loaded, as well as standard “VMware SVGA Adapter”.
23. Load supported version of NVIDIA graphics driver from NVIDIA web site.
24. Once you reboot the VM after installing the driver, the vCenter console may go black as the desktop is going to the NVIDIA card. To manage the VM, you will need to remove GPU or access remotely.
25. Install View Horizon Agent on VM.
26. In some cases it may be necessary to enable NVIDIA API, usually it is not needed.

C:\Program Files\Common Files\VMware\Teradici PCoIP Server\MontereyEnable.exe –enable
Note: If MontereyEnable.exe is not found, use NvFBCEnable.exe. In the new SDK, MontereyEnable is replaced with NvFBCEnable.

27. After the process is complete, restart the virtual machine.

Configure and Connect to vDGA configured VM
1. On endpoint client, download and install latest version of Horizon View client if not done already.
   Latest clients can be download here: https://www.vmware.com/go/viewclients
2. Install/Configure Horizon View Server if not done.
3. Configure Horizon View for new desktop with vDGA configured VM’s. (see VMware Documentation)
4. From endpoint client, connect to View server via Horizon View client
5. A list of VM’s you have access to should be presented.
6. Right click on desktop icon to display mode
   a. All Monitors – Will attempt to connect full screen on all configured displays on endpoint client
   b. Fullscreen – Full Screen on primary monitor on endpoint client
   c. Window Large/Small – Opens in a single display monitor in windowed mode
   d. Custom – You can use to configure a custom size single display monitor window

7. Connect to desktop

   IMPORTANT NOTE – In order to activate the NVIDIA display adapter, you must connect via PCoIP, and at full screen from the endpoint or the virtual machine will use the SVGA 3D display adapter. vDGA will not work through the vSphere console session.

8. After the virtual machine has rebooted and you have connected via PCoIP in full screen, check to ensure that the GPU is active by viewing the display information in DXDiag.exe.
Appendix A – Graphics Server Blade Video Modes

Configuring Server Blade video mode

The WS460c G6/Gen8 Graphics Server Blade has four distinct graphics modes that may be available. Depending on what OS is installed, different modes are used. In a nutshell, the Graphics Server Blade has an embedded graphics card as well a high end graphics card, depending on configuration, it is necessary to disable one or the other card. Table 11 shows the supported modes for each environment and the following sections describe each.

Table 11. Proper use of video mode for each operating system

<table>
<thead>
<tr>
<th>Hardware Configuration</th>
<th>Bare Metal</th>
<th>Microsoft Server for RemoteFX and XenApp</th>
<th>Citrix XenServer</th>
<th>VMware vSphere</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS setup and configuration</td>
<td>Setup mode</td>
<td>Setup mode</td>
<td>Setup mode</td>
<td>Setup mode</td>
</tr>
<tr>
<td>Running in production</td>
<td>User mode</td>
<td>User mode</td>
<td>Setup mode</td>
<td>Setup mode</td>
</tr>
</tbody>
</table>

User mode

- This mode is used when it is necessary to disable the embedded graphics card because the installed operating system does not support using two different graphics architectures (Microsoft Server and Client OS's)
- In this mode the add-in card is enabled and the embedded is disabled. The OS only sees the add-in card
- This is the primary production mode when the following operating systems are installed:
  - Bare Metal Installations of Windows or Linux
  - Microsoft RemoteFX environments
  - Microsoft/XenApp environments
- In this mode the iLO and front SUV console is available in POST but when control is passed to the OS, it becomes inaccessible because it uses the embedded video card to generate its video. In this mode the server console will show a message indicating it is in User mode
- In this mode, once the operating system is booted, the system can be accessed via remote protocol only

Setup mode

- This mode has both video cards enabled but the add-in card is secondary and the embedded is primary
- This mode is used in the following configurations:
  - Used in full production mode when the installed operating does support using two different graphics architectures at the same time. Currently the operating systems that use this mode for production include the VMware and XenServer Hypervisors.
  - Used during system install and configuration when the installed operating does not support using two different graphics architectures at the same time. For example, this mode is used during install, configuration and driver install for Windows 7 and Server, as well as Linux (Bare Metal). The mode is switch to “User” before going into production.
- This is the primary production mode when the following operating systems are install:
  - Citrix XenServer
  - VMware vSphere
- In this mode the iLO console is accessible at all times

Admin mode

- This mode is used for troubleshooting only and is not used in production
- This mode disables the add-in card and only the ATI is active
Server mode (WS460c G6 only)
- This mode is deprecated and rarely used for troubleshooting only and is not used/supported in production
- This mode operates in the same way as Admin mode

Procedure to set Remote Console:

1. Using either iLO remote console or the Local I/O Connector, connect to server console
2. When prompted during boot, press the F9 key. The ROM-based Setup Utility appears
3. Select System Options > Remote Console Mode. The current Remote Console Mode appears. (figure 16)
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 Graphics Server Blade will reboot, and then the Remote Console Mode will be in effect

Figure 16. Remote Console Mode
Resources

To read more about HP ProLiant WS460c Gen8 Graphics Server Blade, go to hp.com/go/bladeworkstation
To learn more about HP Client Virtualization reference architectures, go to hp.com/go/cv
HP and Citrix, hp.com/go/citrix and citrix.com/hp
Citrix XenDesktop, citrix.com/xendesktop
Citrix XenApp, citrix.com/xenapp
Microsoft Hyper-V, microsoft.com/hyper-v
HP and VMware View, hp.com/go/vmware and vmware.com/view

Learn more at hp.com/go/cv