Technical white paper

Dell EMC PowerEdge Servers with NVIDIA GPUs and VMware vSphere

How to configure Dell EMC PowerEdge Servers with NVIDIA GPUs and VMware vSphere

Abstract

This white paper describes how to configure Dell EMC PowerEdge servers with Nvidia GPUs and VMware vSphere. Also included are a support matrix of GPUs supported on Dell EMC PowerEdge servers, as well as GPU-specific troubleshooting information.

March 2020

Revisions

Date	Description
March 2020	Initial release

Acknowledgements

This paper was produced by the following:

Author: Hypervisor Engineering

Support: Shiva Katta

Other: Sherry Keller, and Ramya D R, IDD team

The information in this publication is provided “as is.” Dell Inc. makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Table of contents

Executive summary

Virtualization technology removes hardware from the network equation, and allows you to host multiple and varied workloads that share the same hardware. In its infancy, virtualization technology was limited to CPU, memory, storage, and network applications. Now, however, virtualization can also benefit graphic workload balancing. The same common set of IT resources can be used to host multiple graphic workloads, or to provide virtual desktop infrastructure (VDI) for loads as low as simple document editing and as large as gaming design.

Selecting hardware combinations and configurations can significantly impact the successful deployment of VDI over VMware. The following provides guidance on selecting VMware vSphere features for VDI, and discusses troubleshooting solutions for issues you might encounter during setup and deployment.

1 Use cases for NVIDIA GPUs in vSphere

The use cases discussed here are divided into VDI and non-VDI. NVIDIA GPUs are further classified into those used by multiple users and those operating in dedicated mode. This technical white paper briefs about configuring the NVIDIA GPUs with vSphere for various.

The features available in vSphere with NVIDIA GPUs on the Dell EMC PowerEdge servers are:

- vDGA (Virtual Dedicated Graphics Accelerator)
- vGPU (Virtual GPU)
- VM DirectPath-I/O GPU

1.1 VMware vDGA

vDGA provides direct pass-through to a physical GPU. This method provides the following:

- Unrestricted and dedicated access to the GPU.
- Best performance to the user as the GPU device is dedicated to a single VM which accesses the GPU directly.
- Limits the GPU usage to a single VM and prevents the use of Motion feature.

In this method, the GPU device is passed through to the VM. The relevant GPU driver must be installed inside the VM guest operating system. No special drivers are required to be installed in ESXi.

1.2 NVIDIA grid vGPU

NVIDIA vGPU allows sharing one or more NVIDIA GPUs to multiple VMs. This method provides the following:

- Direct access to the physical GPU on the ESXi host across multiple VMs.
- Options for multiple vGPU assignments to a single VM.
- GPU-enabled VMs can be migrated to remote hosts with GPUs.

In this method, GPU profiles are created based on the physical GPU and those profiles are mapped to the VMs. This method requires software components installed in both the ESXi host and the VM. The VM that has the GPU profiles attached requires a GPU driver. The ESXi host also requires that the vGPU manager software is installed. The vGPU option is used for VDI and virtual workstations.

1.3 VM DirectPath I/O GPU

In this approach, a GPU is assigned as a PCIe pass-through device to the VM. The guest operating system deployed in a VM can access the GPU directly and can offload all the relevant computational or graphical operations to the GPU. The vGPU is not shared across the VMs. Performance is expected to be closer to a bare-metal deployment. When VM DirectPath is used, other vSphere functions such as vMotion, DRS, and cloning, or snapshots are not supported. This feature is targeted for machine learning, HPC, and other AI-related workloads in virtualized environments.

2 Hardware and software requirements

The hardware and software requirements for configuring the GPUs are:

- Hardware:
  - PowerEdge servers must be certified for VMware vSphere ESXi. See VMware HCL.
  - Ensure that criteria are met for PCIe device pass-through, which is also known as VM DirectPath-IO as listed in VMware KB 2142307
- Software:
  - VMware vSphere Hypervisor, ESXi versions
  - GPU drivers in guest operating system and host operating system
  - Specific CUDA software libraries
  - VMware Horizon client

2.1 Configuring vDGA

Configuring vDGA involves configuring Windows VM with direct access to the GPU, and then configuring the vDGA feature. The step-by-step procedure is provided in the following sections.

2.1.1 Configuring Windows VM with direct access to the GPU

To configure a Windows VM with direct access to the GPU, complete the following steps:

1. Update the server with the supported BIOS or firmware and NVIDIA GPU.
2. Install vSphere ESXi and enable NVIDIA GPU for pass-through, or Virtual DirectPath I/O.
3. Configure and deploy the virtual machine with a supported version of the Windows operating system.
4. Assign the GPU to the VM.
5. Install the relevant driver or software within the VM.

2.1.2 Configuring vDGA feature with vSphere

To configure the vDGA feature with vSphere on a Dell EMC PowerEdge server, complete the following steps:

1. See the support matrix to select the supported and certified GPU for your PowerEdge server.
2. Ensure that the appropriate PSUs are added to the server supplying power to the GPUs.
3. Turn off the system and install the NVIDIA GPU graphics card on the PowerEdge server.
4. Verify that VT-d or AMD IOMMU is enabled in the server BIOS.
5. Ensure that the minimum BIOS version is installed on the server. See the VMware HCL to verify the certified BIOS version for vDGA support on the installed GPU.
6. Install the supported, certified ESXi version on the PowerEdge server.
7. After the successful installation of ESXi, enable pass-through for the GPU in the ESXi host configuration and reboot the host.
8. Create the VM and deploy the supported guest operating system.
9. Ensure that ESXi host has adequate memory to create the VM.
10. Add a PCI device to the VM and select the appropriate PCIe function to enable GPU pass-through on the virtual machine.
11. Configure the VM video card 3D capabilities.
12. Obtain the GPU drivers from the GPU vendor and install the GPU device drivers in the guest operating system of the VM.
13. Install VMware Tools and Horizon Agent in the guest operating system and reboot the VM.
14. After the successful reboot of the VM, add the VM to the manual desktop pool, so that the guest operating system can be accessed using PCoIP or VMware Blast Extreme. In PCoIP or VMware Blast session, activate the NVIDIA display adapter.

For more information, see the *VMware Horizon 7 Documentation* at docs.vmware.com.

2.2 Configuring vGPU

To configure the Windows VM with direct access to a GPU, complete the following steps:

1. Update the server with the supported BIOS or firmware and NVIDIA GPU.
2. Install vSphere ESXi and enable the NVIDIA GPU.
3. Download and install the supported vGPU Manager on the ESXi host.
4. Build the Horizon infrastructure.
5. Configure and deploy VM with guest operating system.
6. For the VM, select the appropriate GPU profile and assign the GPU to VM.
7. Install the relevant driver or software on the VM.

To configure the vDGA feature with vSphere on a Dell EMC PowerEdge server, complete the following steps:

1. See the support matrix included in this white paper to select the supported and certified GPU for your PowerEdge server.
2. Ensure that the appropriate PSUs are added to the server that supply power to the GPUs.
3. Verify that VT-d or AMD IOMMU is enabled in the server BIOS.
4. Ensure that the minimum BIOS version is installed on the server. See the VMware HCL to get the certified BIOS version for the vDGA.
5. Install the NVIDIA GPU graphics card on the PowerEdge server.
6. Install a supported and certified ESXi version on the PowerEdge server.
7. Download the NVIDIA vGPU Manager vSphere Installation Bundle (VIB) for the appropriate version of ESXi. Verify the compatibility of this VIB with the ESXi version.
8. Install vGPU Manager in ESXi.
9. Update VMware Tools and Virtual Hardware (vSphere Compatibility) for the template of each VM that will use vGPU. See the vSphere Compatibility matrixes for information on compatible virtual hardware.
10. In the vSphere Web Client, edit the VM settings and add a shared PCI device. PCI devices require reserving guest memory. Expand **New PCI Device** and click **Reserve all guest memory**. You can also modify this setting in the **VM Memory settings**.
11. Select the appropriate GPU profile for your use case. For sizing guidelines, see *NVIDIA vGPU™ GRID Deployment Guide for VMware Horizon 7.x* on VMware vSphere 6.7.
12. Download the NVIDIA guest driver installer package to the VM. Ensure that it matches the version of the installed NVIDIA VIB on ESXi.
13. Choose one of the following methods to install the NVIDIA guest driver. After the NVIDIA driver is installed, vCenter Server console for VM displays a blank screen.

    a. Desktop Pool
    b. View Agent Direct-Connection Plugin
    c. RDP

After the base VM is configured and licensed for vGPU, this VM can be configured as template. From this templated, designed VMs can be deployed.

For more information about the configuration, see the VMware Horizon 7 Documentation or Ready Solutions.

2.3 Configuring VM DirectPath I/O GPU

To configure a Windows VM with direct access to the GPU, complete the following steps:

1. Update the server with the supported BIOS or firmware and NVIDIA GPU.
2. Install vSphere ESXi and enable NVIDIA GPU for pass-through, or Virtual DirectPath I/O.
3. Configure and Deploy VM with Linux operating systems preferably for HPC and machine learning workloads.
4. Assign the GPU to the VM.
5. Install the relevant driver or software in the VM for executing the workloads.

To configure the vDGA feature with vSphere on a PowerEdge server, complete the following steps:

1. See the support matrix included in this white paper to select the supported and certified GPU for your PowerEdge server.
2. Ensure that the PSUs are added to the server, and that they can adequately power the GPUs.
3. Verify that VT-d or AMD IOMMU is enabled in the server BIOS.
4. Ensure that the minimum BIOS version is installed on the Server. See the VMware HCL to get the certified BIOS version for the vDGA.
5. Install the NVIDIA GPU Graphics Card on the PowerEdge server.
6. Install a supported and certified ESXi version on the PowerEdge server.
7. After successful installation of ESXi, enable pass-through for the GPU in the ESXi host configuration and reboot.
8. Create a VM and deploy a supported Linux OS as guest operating system.
9. Ensure that ESXi host has adequate memory to create the VM.
10. Add a PCI device to the VM and select the appropriate PCI device to enable GPU pass-through.
11. Obtain the GPU drivers from the GPU vendor and install the GPU device drivers in the guest operating system.
12. Install VMware Tools in the guest operating system and reboot the VM.
13. After the successful reboot of the VM, install relevant libraries for executing the workloads related to HPC, machine learning, and so on.

3 GPU support matrix with Dell EMC PowerEdge servers

Note: The below section lists the various support matrixes for the GPU features as they relate to the supported server and ESXi versions.

3.1 PowerEdge yx5x servers supporting NVIDIA GPU

The following tables list out the PowerEdge yx5x severs and supported NVIDIA GPU:

Table 1 yx5x PowerEdge servers and NVIDIA GPU support

	NVIDIA Tesla M10	NVIDIA Tesla V100	NVIDIA Tesla T4	NVIDIA Quadro RTX6000	NVIDIA Quadro RTX8000	NVIDIA Tesla V100S
PowerEdge R7525	Y	Y	Y	Y	Y	Y
PowerEdge R7515	N	N	Y	N	N	N
PowerEdge R6525	N	N	Y	N	N	N
PowerEdge R6515	N	N	Y	N	N	N
PowerEdge C6525	N	N	Y	N	N	N

3.2 PowerEdge yx4x servers supporting NVIDIA GPU

Table 2 yx4x PowerEdge servers and NVIDIA GPU support

	NVIDIA Tesla K80	NVIDIA Quadro P4000	NVIDIA Tesla P100	NVIDIA Tesla M60	NVIDIA Tesla P4	NVIDIA Tesla V100	NVIDIA Tesla P40	NVIDIA Tesla M10	NVIDIA Tesla T4
PowerEdge R940xa	N	N	Y	N	N	Y	Y	N	N
PowerEdge R840	N	N	Y	N	N	Y	Y	Y	N
PowerEdge R740	Y	Y	Y	Y	Y	Y	Y	Y	Y
PowerEdge R740xd	Y	Y	Y	Y	Y	Y	Y	Y	Y
Dell XC740xd-24	N	N	N	Y	N	N	N	N	N
PowerEdge R7425	N	N	N	N	Y	Y	Y	Y	Y
PowerEdge R640	N	N	N	N	N	N	N	N	Y
PowerEdge T640	Y	N	Y	Y	N	Y	Y	Y	N
PowerEdge T440	N	Y	N	N	N	N	N	N	N
PowerEdge C4140	N	N	Y	N	N	Y	Y	N	N
PowerEdge C6420	N	N	N	N	N	N	N	N	Y

3.3 PowerEdge yx3x and yx2x servers supporting NVIDIA GPU

Table 3 yx3x and yx2x PowerEdge servers and NVIDIA GPU support

	Grid K1	Grid K2	Quadro K2000	Quadro K2200	Quadro K400	Quadro K4200	Quadro K5200	Quadro K6000	Quadro M2000	Quadro P5000	Tesla P6000	Tesla M4000	Tesla M60	K40m	K20X	K20m	K20c	K10 CA
PowerEdge R720	Y	Y	Y	N	Y	N	N	N	N	N	N	N	N	Y	N	N	N	N
Dell Precision Rack 7910	N	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	N	N
PowerEdge T620	N	Y	Y	N	Y	N	N	N	N	N	N	N	N	N	N	N	Y	N
PowerEdge R730	Y	Y	N	N	N	N	N	N	N	N	N	N	Y	Y	N	N	N	N
PowerEdge T630	N	Y	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	Y	N
Dell XC730-16G	Y	Y	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
PowerEdge C4130	Y	Y	N	N	N	N	N	N	N	N	N	N	N	Y	Y	Y	N	Y
VRTX	N	Y	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N

3.4 vGPU support matrix

3.4.1 vGPU supported servers and GPUs matrix

Table 4 vGPU supported PowerEdge servers and GPUs

	Tesla M10	Tesla M60	Tesla P4	Tesla P40	Tesla V100	Tesla T4
PowerEdge C4140	N	N	N	Y	Y	N
PowerEdge C4130	Y	Y	N	Y	Y	N
PowerEdge R640	N	N	N	N	N	Y
Power Edge XR2	N	N	N	N	N	Y
PowerEdge R740	Y	Y	Y	Y	Y	Y
PowerEdge R740xd	Y	Y	Y	Y	Y	Y
PowerEdge R7425	Y	N	Y	Y	Y	Y
PowerEdge R730	Y	Y	Y	Y	N	N
Dell XC740xd	Y	Y	Y	Y	Y	N
Dell XC730-16G	Y	Y	N	N	N	N
PowerEdge R940xa	N	N	N	Y	Y	N
PowerEdge R840	Y	N	N	Y	Y	N
PowerEdge T640	Y	Y	N	Y	Y	N
PowerEdge T630	N	Y	N	Y	Y	N

3.5 vDGA support matrix

3.5.1 vDGA certified on PowerEdge servers with NVIDIA GPUs

Table 5 vDGA certified on PowerEdge servers with NVIDIA GPUs

	Grid K1	Grid K2	Quadro K2000	Quadro K2200	Quadro K400	Quadro K4200	Quadro K5200	Quadro K6000	Quadro M2000	Quadro P5000
PowerEdge R720	Y	Y	Y	N	Y	N	N	N	N	N
Dell Precision Rack 7910	N	Y	N	Y	N	Y	Y	Y	Y	Y
PowerEdge T620	N	Y	Y	N	Y	N	N	N	N	N
PowerEdge C8220x	N	N	N	N	N	N	N	N	N	N
PowerEdge R730	Y	Y	N	N	N	N	N	N	N	N
PowerEdge T630	N	Y	N	N	N	N	N	N	N	N
Dell XC730-16G	Y	Y	N	N	N	N	N	N	N	N
PowerEdge C4130	Y	Y	N	N	N	N	N	N	N	N
PowerEdge R740	N	N	N	N	N	N	N	N	N	N
PowerEdge R740xd	N	N	N	N	N	N	N	N	N	N
PowerEdge R840	N	N	N	N	N	N	N	N	N	N
PowerEdge R940xa	N	N	N	N	N	N	N	N	N	N
PowerEdge T640	N	N	N	N	N	N	N	N	N	N
PowerEdge C4140	N	N	N	N	N	N	N	N	N	N
Dell XC740xd-24	N	N	N	N	N	N	N	N	N	N
PowerEdge R7425	N	N	N	N	N	N	N	N	N	N
PowerEdge R7515	N	N	N	N	N	N	N	N	N	N
PowerEdge R6515	N	N	N	N	N	N	N	N	N	N

Table 6 vDGA certified on PowerEdge servers with NVIDIA GPUs

	Tesla P6000	Tesla M4000	Tesla M60	P100 12 GB	P100 16 GB	V100 16 GB	V100 32 GB	P40	M10	V4	P4
PowerEdge R720	N	N	N	N	N	N	N	N	N	N	N
Dell Precision Rack 7910	Y	Y	N	N	N	N	N	N	N	N	N
PowerEdge T620	N	N	N	N	N	N	N	N	N	N	N
PowerEdge C8220x	N	N	N	N	N	N	N	N	N	N	N
PowerEdge R730	N	N	Y	N	N	N	N	N	N	N	N
PowerEdge T630	N	N	Y	N	N	N	N	N	N	N	N
Dell XC730-16G	N	N	N	N	N	N	N	N	N	N	N
PowerEdge C4130	N	N	Y	N	N	N	N	N	N	N	N
PowerEdge R740	N	N	Y	Y	Y	Y	Y	N	N	N	N
PowerEdge R740xd	N	N	Y	Y	Y	Y	Y	N	N	N	N
PowerEdge R840	N	N	N	Y	Y	Y	Y	Y	Y	N	N
PowerEdge R940xa	N	N	N	Y	Y	Y	N	Y	Y	N	N
PowerEdge T640	N	N	Y	Y	Y	Y	Y	Y	Y	N	N

	Tesla P6000	Tesla M4000	Tesla M60	P100 12 GB	P100 16 GB	V100 16 GB	V100 32 GB	P40	M10	V4	P4
PowerEdge C4140	N	N	Y	Y	Y	Y	Y	Y	N	N	N
Dell XC740xd-24	N	N	Y	N	N	N	N	N	N	N	N
PowerEdge R7425	N	N	N	N	N	Y	N	Y	Y	Y	Y
PowerEdge R7515	N	N	N	N	N	N	N	N	N	Y	N
PowerEdge R6515	N	N	N	N	N	N	N	N	N	Y	N

3.5.2 vDGA certified on Dell EMC VxRail servers with NVIDIA GPUs

Table 7 vDGA certified on Dell EMC VxRail servers with NVIDIA GPUs

	Tesla M10	Tesla M60	Tesla P4	Tesla P6	Tesla P60	Tesla V100	Tesla V100S	Tesla T4	RTX6000	RTX8000
Dell EMC VxRail V470	Y	Y	N	N	N	N	N	N	N	N
Dell EMC VxRail V470F	Y	Y	N	N	N	N	N	N	N	N
Dell EMC VxRail V570	Y	Y	N	N	Y	Y	Y	Y	Y	Y
Dell EMC VxRail V570F	Y	Y	N	N	Y	Y	Y	Y	Y	Y
Dell EMC VxRail E560	N	N	N	N	N	N	N	Y	N	N
Dell EMC VxRail E560F	N	N	N	N	N	N	N	Y	N	N
Dell EMC VxRail E560N	N	N	N	N	N	N	N	Y	N	N

4 Known issues and resolution

This section focuses on the known issues for configuring the GPU features described in this document.

1. PowerEdge R730 with NVIDIA Grid K2 and ESXi 6.x, a Windows 7 64-bit VM configured with vDGA fails to boot and display BSOD.

- **Resolution**:
  This is a known issue and to overcome the VM crash, set **pciPassthru0.msiEnabled** is set to **False** in the VMs VMX file. By default, **pciPassthru0.msiEnabled** is set to **True.**

2. VM configured with vGPU fails to start with the following error:
   The available memory resources in the parent resource pool are insufficient for the operation.

- **Resolution:**
  Verify the memory assigned to the VM. Ensure that it does not exceed or result in a memory overcommit.

3. VMs configured with vGPU cannot utilize vMotion and DRS functionalities.

- **Resolution:**
  With versions of ESXi 6.0.x and 6.5.x, the vMotion or similar live operations on VM are not supported. With ESXi 6.7.x, VM configured with vGPU can use vMotion, provided the destination host has the required, supported, and compatible hardware.

4. VM configured with vGPU fails to power on.

- **Resolution:**
  Ensure that the service X.Org is in a running state on the ESXi host. Operations such as start and stop can be performed either from vSphere Web Client or through SSH to the ESXi host.

5. On the PowerEdge R740 server, after installing the vGPU VIB in ESXi, the command `nvidia-smi` fails to display the GPU statistics with following error message:
   Failed to initialize NVML: Unknown Error

- **Resolution:**
  The above error can occur for many reasons, including misconfiguration. To resolve the issue:
  1. Ensure that the VGPU VIB installed successfully without any errors.
  2. Verify that the NVIDIA GPUs in the ESXi host are not configured as pass-through devices for VM DirectPath IO or vDGA.
  3. Run the command `lspci | grep -i nvida` on ESXi shell and ensure that there are entries related to NVIDIA GPUs present in the server.
  4. On Dell EMC PowerEdge yx4x servers, ensure the below settings in System BIOS are set:
     - **Memory Mapped I/O above 4 GB** is set to Enable
     - **Memory Mapped I/O Base** is set to 512 GB

6. On the PowerEdge R740 server with NVIDIA Tesla T4, an attempt to configure a VM with an assigned vGPU or to perform a GPU pass-through fails.

- **Resolution:**
  When the above failure is encountered, verify if Tesla T4 is enumerated as 32 separate GPUs in ESXi. If it is, ensure that the SR-IOV capability is enabled in the server BIOS and retry.

7. On the PowerEdge R740 with NVIDIA Tesla T4 and installed with ESXi 6.7 and attempt to list the NVIDIA GPU in Ispci command fails.

- **Resolution**:
  This behavior may be due to multiple reasons. Check the following:
  - Ensure that both the PSUs are plugged-in and working.
  - Ensure that the correct wattage of PSU is used for the GPU configuration.
  - Ensure that the GPU power cables are connected.
  - Ensure that the GPU is not configured as a pass-through device in the ESXi host.

8. On the PowerEdge R730 server with Tesla M60, the VM configured with Tesla M60 for vDGA displays a blank screen or fails to power-on.

- **Resolution:**
  The Tesla M60 can work in both graphics mode and computation mode. Ensure that, the NVIDIA M60 GPU is configured in graphics mode and not in computation mode. For toggling the modes, use the tool gpumodeswitch. For more information, see the document *GPUMODESWITCH User Guide* at NVIDIA support site.

9. On the PowerEdge R740 with Tesla T4 and ESXi 6.5, VM configured with multi vGPU fails to power on or fails to initialize vGPU after boot.

- **Resolution:**
  Assigning multiple vGPUs is not supported in ESXi 6.5. Either update the host to ESXi 6.7 U3 or ensure that the VM has only one vGPU associated to it.

10. On the PowerEdge R740 server with Tesla T4 and ESXi 6.5 U1, the VM with vGPU associated to it fails to boot.

- **Resolution:**
  The VM associated with vGPU power-on fails with the error message “The amount of graphics resource available in the parent resource pool is insufficient for the operation.” This behavior is seen if the VM **Graphics Type** is set to **Shared**. Change the **Graphics Type** to the **Shared Direct** option. Note that the default option is set to **Shared**.

11. On a PowerEdge R730 with M60 GPU and ESXi 6.5 U3, the VM failed to power-on. The VM log file contains the following entry:

```
2019-10-07T06:57:51.499Z| vmx| I120: PCIPassthru: total number of pages needed (2097186) exceeds limit (917504), failing
2019-10-07T06:57:51.499Z| vmx| I120: Module DevicePowerOn power on failed
```

- **Resolution:**
  In order to resolve this issue, either reduce the memory assigned to the VM and power on, or perform the followings:
  - Ensure that the BIOS configuration setting **Memory Mapped I/O above 4GB** is set to **Enable**.
  - Add the below command in the VMX file of VM:
    - `pciPassthru.use64bitMMIO=”TRUE”`
    - `pciPassthru.64bitMMIOSizeGB = “64”`