

NVIDIA Grid vGPU on Nutanix

Nutanix Solution Note

Version 1.1 • February 2017 • SN-2046

Copyright

Copyright 2017 Nutanix, Inc.

Nutanix, Inc.
1740 Technology Drive, Suite 150
San Jose, CA 95110

All rights reserved. This product is protected by U.S. and international copyright and intellectual property laws.

Nutanix is a trademark of Nutanix, Inc. in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies.

Contents

- 1. Executive Summary..... 4**
- 2. Introduction.....5**
 - 2.1. Audience..... 5
 - 2.2. Purpose.....5
- 3. Solution Overview..... 6**
 - 3.1. The Nutanix Enterprise Cloud Platform.....6
- 4. NVIDIA Grid GPU.....10**
 - 4.1. Dedicated GPU.....10
 - 4.2. Shared GPU..... 11
- 5. Hypervisor Implementations on Nutanix..... 15**
 - 5.1. VMware vSphere..... 15
 - 5.2. XenServer..... 15
 - 5.3. Microsoft Hyper-V..... 15
 - 5.4. AHV.....16
- 6. Nutanix vGPU Options.....17**
- 7. Conclusion..... 18**
- 8. Appendix..... 19**
 - 8.1. Installation: vGPU Architecture.....19
 - 8.2. Troubleshooting..... 20
 - 8.3. About Nutanix..... 20
- List of Figures.....21**
- List of Tables..... 22**

1. Executive Summary

This solution note discusses the advantages of running the Nutanix Enterprise Cloud Platform with NVIDIA Grid vGPU for virtual desktop infrastructure (VDI) or server-based computing (SBC) projects. NVIDIA Grid vGPU provides a desktop delivery method that supports GPU-intensive use cases well beyond the obvious examples of CAD and CAM. The industry is increasingly deploying GPU-enhanced workloads, as hardware acceleration benefits even commonplace applications such as browsers (Internet Explorer, Firefox, and Chrome) and Office 2016. Nutanix delivers a powerful, flexible, and reliable platform for the full spectrum of desktop virtualization requirements, including unrivaled uptime and the freedom to mix and match workloads to suit your enterprise needs and end user objectives, from task workers to power users. With VMware vSphere 6, Nutanix offers vGPU capabilities for our NX-3175-G4, NX-3175-G5, NX-3155-G4, and NX-3155-G5 nodes, all of which can be ordered with NVIDIA Grid cards.

2. Introduction

2.1. Audience

This solution note is a part of the Nutanix solutions library and provides an overview of the combination of the Nutanix Enterprise Cloud Platform and NVIDIA Grid technologies. It is intended for IT architects and administrators as a technical introduction to the solution.

2.2. Purpose

This document covers the following subject areas:

- Overview of the Nutanix solution.
- Overview of NVIDIA Grid on different hypervisors.
- The benefits of implementing NVIDIA vGPU on Nutanix.

Table 1: Document Version History

Version Number	Published	Notes
1.0	April 2016	Original publication.
1.1	February 2017	Added material on: XenServer Support M10/M60 G5 Models

3. Solution Overview

3.1. The Nutanix Enterprise Cloud Platform

The Nutanix Enterprise Cloud Platform delivers a hyperconverged infrastructure solution purpose-built to provide the flexibility of the public cloud alongside the predictability and affordability of on-premises hardware. The Nutanix Enterprise Cloud Platform is composed of two product families—Nutanix Acropolis, a complete data storage fabric, and Nutanix Prism, a simple and efficient management suite.

Highlights of Nutanix Acropolis and Prism include:

- Storage and compute resources hyperconverged on industry-standard x86 servers.
- 100 percent software-based intelligence, allowing the solutions to constantly evolve.
- Unlimited scale, as data, metadata, and operations are distributed across the entire cluster.
- Automatic self-healing, returning the cluster to a redundant state without intervention.
- API-based automation and rich analytics, including capacity management predictions.
- One-click operations, including firmware management and software upgrades.
- Single pane of glass management for the entire storage and virtualization stack.

The hardware for the Nutanix Enterprise Cloud Platform consists of highly dense storage and server compute (CPU and RAM) building blocks. Each building block is based on industry-standard Intel processor technology and delivers a unified, scale-out, shared-nothing architecture with no single points of failure. These blocks comprise a turnkey solution that generates significant OPEX savings on rack space, power, and cooling costs.

The Acropolis Distributed Storage Fabric

Acropolis does not rely on traditional SAN or NAS storage or on expensive storage network interconnects. All storage management is VM-centric, and the Acropolis Distributed Storage Fabric (DSF) optimizes I/O at the VM virtual disk level. As seen in the figure below, the file system automatically tiers data across different types of storage devices using intelligent data placement algorithms. These algorithms make sure the most frequently used data is available in memory or in flash for the fastest possible performance. The web-scale engineering behind the DSF allows you to start small and scale as you grow, while providing all of the benefits you expect from a storage platform, such as built-in disaster recovery and data protection. The DSF also has self-healing mechanisms that return the system to a fully redundant state after any component failure.

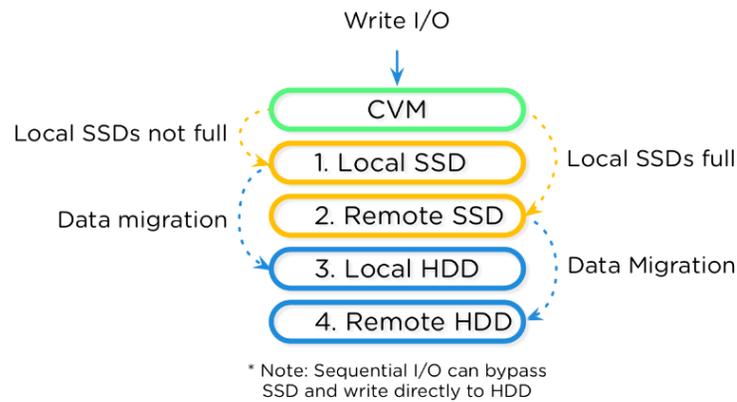


Figure 1: Information Life Cycle Management

Each Nutanix storage Controller VM (CVM) in the DSF connects directly to the local storage controller and its associated disks, as the following figure illustrates. Using local storage controllers on each host localizes access to data through the DSF, thereby reducing storage I/O latency and chatter on the network. The DSF replicates writes synchronously to at least one other node in the system, distributing data throughout the cluster for resiliency and availability. Having a local storage controller on each host ensures that storage performance as well as storage capacity linearly scale when adding nodes.

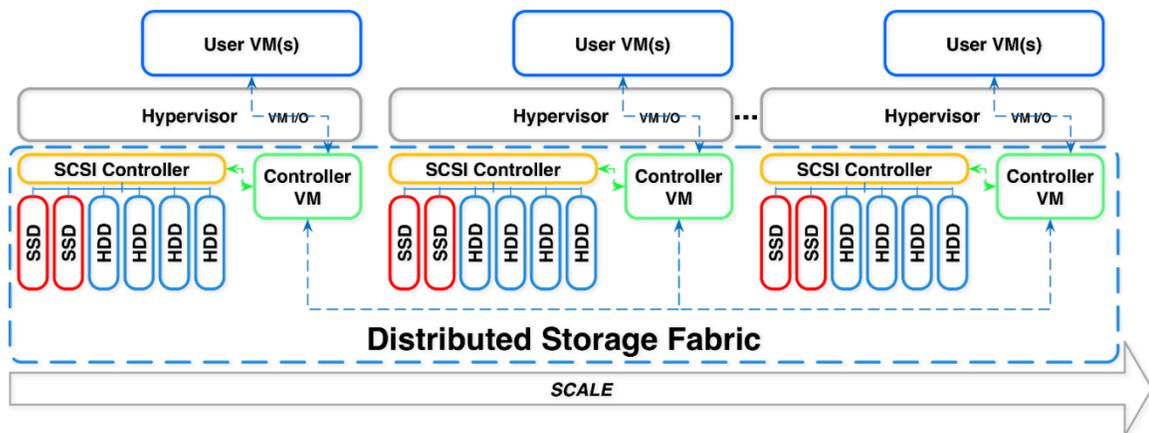


Figure 2: Overview of the Acropolis Distributed Storage Fabric

The Acropolis App Mobility Fabric

The Acropolis App Mobility Fabric (AMF) enables unprecedented workload flexibility, allowing workloads to move between different hypervisors (ESXi, Hyper-V, AHV), and even to the public cloud. This design supports a range of powerful use cases, such as cross-hypervisor disaster recovery. No matter what hypervisor you choose to run in each environment (dev/

test, production, or disaster recovery), the AMF allows you to quickly move workloads across hypervisors without complex troubleshooting or stability concerns.

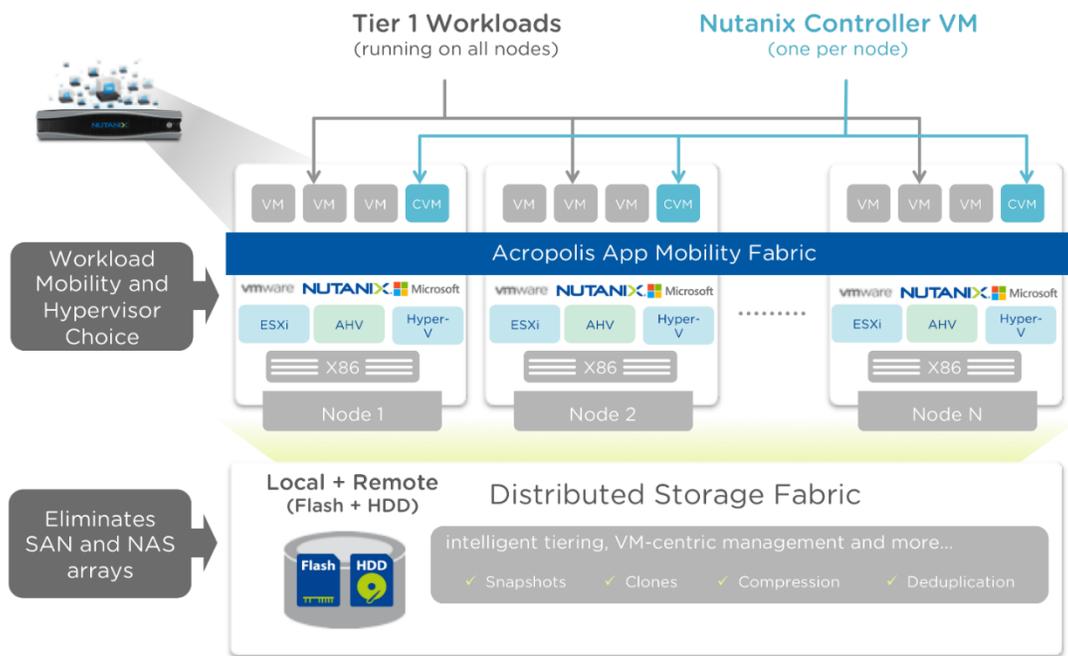


Figure 3: The Nutanix Architecture

AHV

AHV integrates tightly with the DSF and provides cost-effective virtualization with elegant and responsive management. A full type-1 virtualization solution, AHV includes a distributed VM management service responsible for storing VM configuration, making scheduling decisions, and exposing a management interface. The Prism interface, a robust REST API, and an interactive command line interface called aCLI (Acropolis CLI) combine to eliminate the complex management associated with open source hypervisors.

AHV provides the following capabilities:

- Virtual machine storage. Storage devices for the VM, such as SCSI and IDE devices.
- Crash-consistent snapshots. Includes VM configuration and disk contents.
- Virtual networks (L2). Layer-2 network communication between VMs and to the external network, with support for multiple vSwitches and VLANS.
- Managed networks (L3). IP address management (IPAM) to provide layer-3 addresses for VMs.

Nutanix Prism

Nutanix Prism lets you manage the entire stack—storage and virtualization—at any scale from a single pane of glass. Designed with simplicity as a preeminent concern, Prism allows one-click management of typically complex operations, such as hypervisor upgrades, firmware management, and cluster expansion. Prism also provides advanced data analytics and heuristics, giving you powerful and actionable insight into the system state.

4. NVIDIA Grid GPU

NVIDIA is the best-known manufacturer of graphic cards designed for desktop virtualization. AMD's graphic cards work only in certain use cases and do not deliver the same optimizations that NVIDIA cards offer.

NVIDIA Grid boards (which include GPUs based on the Maxwell GPU architecture) allow GPU virtualization, enabling multiple users to share a single graphics card. GPU virtualization not only supports higher user densities, but also delivers native performance while accessing a virtual desktop. NVIDIA GPUs also have an engine for H.264 encoding that offloads processes from the CPU, which further increases user density on your hardware.

NVIDIA cards have multiple GPUs, which improves scaling. The M10 has four entry-level Maxwell GPUs and 32 GB of RAM (8 GB per GPU). The M60 has two high-end Maxwell GPUs and 16 GB of RAM (8 GB per GPU). The M60 can deliver far more CUDA cores than the M10—4,096 to 2,560, respectively. NVIDIA recently introduced the M6, M60 and M10 cards; these cards are based on the newer Maxwell GPUs and deliver either additional performance or a higher desktop density than the Kepler-based cards. The M6 is meant for the blade server form factor, which isn't directly applicable to the current line of Nutanix hardware; Nutanix supports both the M10 and the M60, which are based on a full PCI card.

In the following sections, we discuss the ways in which vGPU improves desktop virtualization user experience.

4.1. Dedicated GPU

With GPU passthrough, you can create a VM with a dedicated Kepler GPU. This configuration provides user experience comparable to using a fat client with a high-end graphics card. However, assigning a GPU core to a single VM, either a hosted shared desktop (SBC) or a hosted private desktop (VDI), limits scalability.

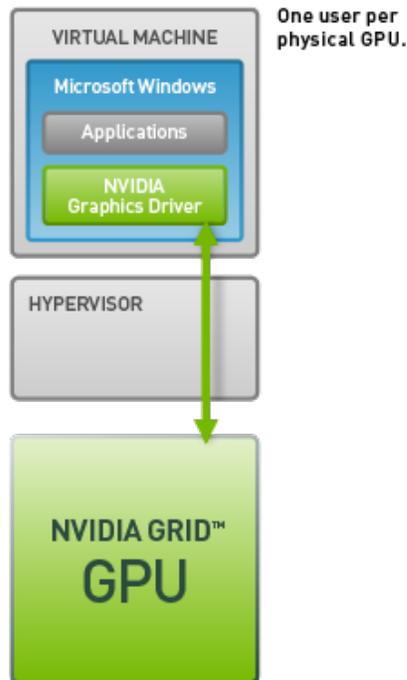


Figure 4: Dedicated GPU. Source: NVIDIA

4.2. Shared GPU

Grid technology lets multiple virtual desktops share a GPU, while offering the same user experience as native GPUs. An NVIDIA Grid M10 card, for example, has four physical GPU cores that can host up to sixteen users per core, resulting in 64 users, each with a vGPU-enabled desktop, per M10 card.

The GPU processes VM graphics commands directly, which means that users get high-end graphics without a performance penalty due to hypervisor interference. vGPU is more scalable than passthrough, as we assign vGPU profiles to our users and thus get more users on the same card.

vGPU profiles deliver dedicated graphics memory through the vGPU Manager, which assigns the configured memory for each desktop. A vSphere installation bundle (VIB) installs the vGPU Manager on the hypervisor. Each VDI instance has preset resources based on the needs of its applications.

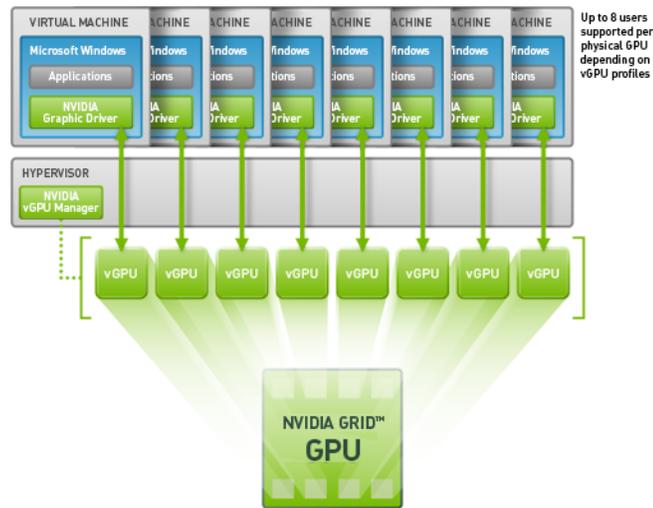


Figure 5: Shared GPU. Source: NVIDIA

The following options are available for our GPU-enhanced solutions.

Table 2: Nutanix GPU Options

Node Type	Expansion Slots	Options
NX-3175-G4	1x PCIe	1x M10 1x M60
NX-3155-G4	3x PCIe	2x M10
NX-3175-G5	1x PCIe	1x M10 1x M60
NX-3155-G5	3x PCIe	2x M10 2x M60

The vGPU profiles can be mixed per card but not per GPU core; account for this restriction when sizing the NVIDIA Grid cards for your environment. Admins can use “Depth” or “Breadth” mode to load balance the system. Depth mode lets you maximize the number of users on a single card. Using this approach, you could reach resource limits sooner than with Breadth mode, which employs a round-robin load balancing mechanism. Breadth mode’s drawback is that it may use the NVIDIA Grid cards less efficiently.

vGPU can also deliver adjusted performance based on profiles. You can configure a vGPU profile based on a VM in a way that accounts equally for user performance and scalability needs. Note that, as the table below shows, high powered VMs take more resources and thus fit on fewer desktops than less powerful profiles, which you can place on more desktops.

Table 3: vGPU Profiles

NVIDIA Grid Graphics Board	Virtual GPU Profile	Graphics Memory	Max Displays Per User	Max Resolution Per Display	Max Users Per Graphics Board	High End Designer
Grid M10	M10-8Q	8,192 MB	4	3840x2160 (4k)	4	High End Designer
	M10-4Q	4,096 MB	4	3840x2160 (4k)	8	Designer
	M10-2Q	2,048 MB	4	2560x1600	16	Designer
	M10-1Q	1,024 MB	2	2560x1600	32	Designer
	M10-0Q	512 MB	2	2560x1600	64	Power User/Designer
	M10-2B	2,048 MB	2	2560x1600	16	Power User
	M10-1B	1,024 MB	2	2560x1600	32	Power User
	M10-0B	512 MB	2	2560x1600	64	Power User
Grid M60	M60-8Q	8,192 MB	4	3840x2160 (4k)	2	High End Designer
	M60-4Q	4,096 MB	4	3840x2160 (4k)	4	Designer
	M60-2Q	2,048 MB	4	2560x1600	8	Designer
	M60-1Q	1,024 MB	2	2560x1600	16	Designer
	M60-0Q	512 MB	2	2560x1600	32	Power User/Designer

NVIDIA Grid Graphics Board	Virtual GPU Profile	Graphics Memory	Max Displays Per User	Max Resolution Per Display	Max Users Per Graphics Board	High End Designer
	M60-2B	2,048 MB	2	2560x1600	8	Power User
	M60-1B	1,024 MB	2	2560x1600	16	Power User
	M60-0B	512 MB	2	2560x1600	32	Power User

All the GPU profiles with a Q go through the same certification process as the current NVIDIA workstation processors, which means that these profiles should perform at least as well. Note that the B models are the non-Quadro “Business” profiles; these profiles are replacing the non-Q profiles that NVIDIA offers with the K1 and the K2 Grid Cards (Grid 1.x).

The 0B, 1B, and 2B profiles are for knowledge workers. These profiles offer less graphical performance, but greater scalability. However, because the 0Q, 1Q, and 2Q profiles deliver comparable scalability and are certified at the same level as the workstation processors, the Q profiles are a good choice for low-end graphical enhancements.

vGPU usage is becoming increasingly commonplace, in part because the graphical richness of even commodity applications is growing rapidly. Office 2013, Flash/HTML, Windows 7/8.1, or even Windows 10 with Aero are now all good use cases for the 1Q and 2Q vGPU profiles.

With the introduction of Grid 2.0 and the newer Maxwell-based cards, NVIDIA has added more performance to the GPU cores. This performance upgrade improves user experience for everyone, although these products target developers with high graphics demands.

5. Hypervisor Implementations on Nutanix

Although Nutanix supports vSphere (ESXi), AHV, and Hyper-V, only vSphere and Hyper-V currently accommodate GPU workloads.

5.1. VMware vSphere

The vSGA is the Windows Display Driver Model driver, which is included when you install VMware tools on the virtual desktop. This driver adds support for DirectX and OpenGL. The VMware SVGA 3D graphics driver provides support for DirectX 9.0c and OpenGL 2.1. One of the benefits of VMware SVGA 3D for software 2D, 3D, and vSGA implementations is that a virtual machine can dynamically switch between software or hardware acceleration, with no reconfiguration.

Another option is vDGA, which requires a native graphics card driver that must be installed directly into the virtual desktop. vDGA supports OpenGL 4.3, DirectX 11, and CUDA. The recently added vGPU capability in vSphere 6 delivers the vGPU profiles, which provides much better performance than the software-emulated SVGA and vDGA.

5.2. XenServer

The XenServer 5.6 release introduced multi-GPU passthrough, which lets you map multiple GPUs to the same number of virtual machines (GPU passthrough). XenServer also offers vGPU, based on shared GPU technology, including Intel GVT-g, AMD cards, or NVIDIA. Nutanix delivers models based on NVIDIA Grid technology only. XenServer with GPUs based on NVIDIA Grid supports OpenGL, DirectX 11, and CUDA.

5.3. Microsoft Hyper-V

Microsoft offers Windows Server 2012 and Hyper-V 3.0 as a virtualization platform. This platform includes a suite of improvements similar to Citrix HDX called RemoteFX. It improves user experience on high-latency networks, as well as access to peripheral devices attached to the client, such as USB.

RemoteFX offers a vGPU feature too, which allows you to share a GPU core to multiple VMs and further improve user density for low-end and medium graphic performance. RemoteFX vGPU is technically different than the vGPU mentioned earlier. Microsoft's RemoteFX vGPU implementation is comparable to vSGA from vSphere, with a high DirectX compatibility, but

without clear focus on OpenGL; this difference can be decisive, as the majority of applications use OpenGL. Hyper-V supports DirectX 11.

5.4. AHV

At this point in time, AHV does not support shared or dedicated GPU, but shared and dedicated GPU support are in development.

6. Nutanix vGPU Options

Prior to vGPU, Nutanix GPU options were limited by hypervisor capabilities. XenServer, Hyper-V, and vSphere allow for software emulation and, with vSphere and XenServer, we can dedicate a physical GPU. However, software emulation doesn't generate performance levels equal to physical GPUs, and assigning a physical GPU to a VM reduces the number of users who get the enhanced features for graphical content.

With vSphere and XenServer, we offer vGPU capabilities for our NX-3175-G4, NX-3175-G5, NX-3155-G4, and NX-3155-G5 nodes, all of which can be ordered with NVIDIA Grid cards.

7. Conclusion

NVIDIA Grid cards are now the standard for providing enhanced graphical performance for any desktop virtualization platform. Prior to vGPU, Nutanix supported XenServer, Hyper-V, and vSphere implementations, which used software-based emulation and passthrough. Each of these implementations has limitations, however. Although software-based emulation works for low-end graphical content, it taxes the system's CPUs. Passthrough limits scalability because it binds one GPU to a VM, resulting in only two or four VMs per card. vGPU, in contrast, provides both flexibility and great user experience. vGPU profiles can determine the number of GPU cycles a user gets and, thus, the quality of their user experience. Moreover, with vGPU on Nutanix, high-quality user experience is possible without concessions on scalability, as you can scale up to 64 users per card.

8. Appendix

8.1. Installation: vGPU Architecture

This is a representation of all the components needed for vGPU.

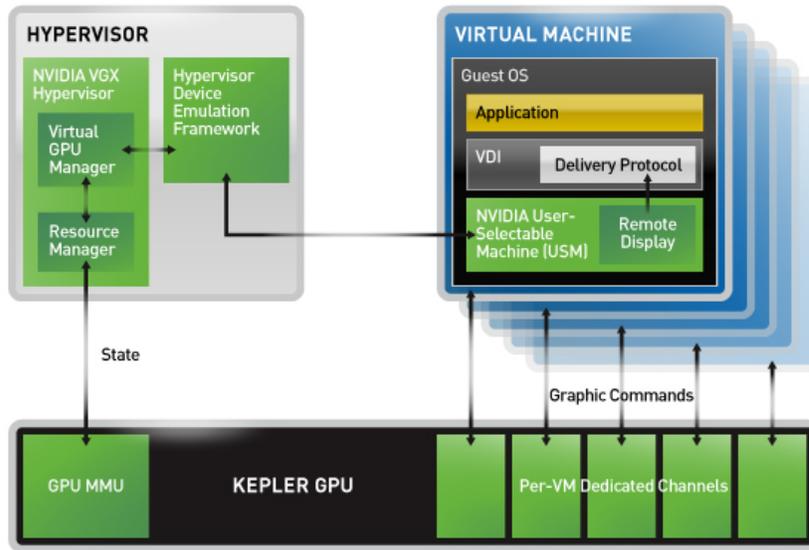


Figure 6: GPU Architecture. Source: NVIDIA

The following components are necessary to configure vGPU.

Table 4: Architectural Components

Component	Implementation
Maxwell GPU	M10 or M60 cards
Virtual GPU Manager/Resource Manager	NVIDIA VIB
Hypervisor	vSphere 6
NVIDIA Selectable Machine	Windows drivers for the vGPU Profile
Delivery Protocol	PCoIP/ICA

8.2. Troubleshooting

The following commands are useful when troubleshooting vGPU-related issues:

Table 5: Troubleshooting Commands

Goal	Command	Component
Confirming GPU installation	<code>lspci grep -i display</code>	Hypervisor
Confirming GPU configuration	<code>esxcli hardware pci list -c 0x0300 -m 0xff</code>	Hypervisor
Verifying VIB installation	<code>esxcli software vib list grep Nvidia</code>	Virtual GPU Manager/ Resource Manager
Confirming VIB loading	<code>esxcfg-module -l grep Nvidia</code>	Virtual GPU Manager/ Resource Manager
Manually loading VIB	<code>esxcli system module load -m Nvidia</code>	Virtual GPU Manager/ Resource Manager
Verifying if module is loading	<code>cat /var/log/vmkernel.log grep NVRM</code>	Virtual GPU Manager/ Resource Manager
Checking Xorg status	<code>/etc/init.d/xorg status</code>	Hypervisor
Manually starting Xorg	<code>/etc/init.d/xorg start</code>	Hypervisor
Checking Xorg logging	<code>cat /var/log/Xorg.log grep -E "GPU nv"</code>	Hypervisor
Checking vGPU management	<code>Nvidia-smi</code>	Virtual GPU Manager/ Resource Manager

8.3. About Nutanix

Nutanix makes infrastructure invisible, elevating IT to focus on the applications and services that power their business. The Nutanix Enterprise Cloud Platform leverages web-scale engineering and consumer-grade design to natively converge compute, virtualization and storage into a resilient, software-defined solution with rich machine intelligence. The result is predictable performance, cloud-like infrastructure consumption, robust security, and seamless application mobility for a broad range of enterprise applications. Learn more at www.nutanix.com or follow up on Twitter [@nutanix](https://twitter.com/nutanix).

List of Figures

Figure 1: Information Life Cycle Management.....	7
Figure 2: Overview of the Acropolis Distributed Storage Fabric.....	7
Figure 3: The Nutanix Architecture.....	8
Figure 4: Dedicated GPU. Source: NVIDIA.....	11
Figure 5: Shared GPU. Source: NVIDIA.....	12
Figure 6: GPU Architecture. Source: NVIDIA.....	19

List of Tables

Table 1: Document Version History.....	5
Table 2: Nutanix GPU Options.....	12
Table 3: vGPU Profiles.....	13
Table 4: Architectural Components.....	19
Table 5: Troubleshooting Commands.....	20