



# Virtual GPU Software R430 for Nutanix AHV

Release Notes

# Table of Contents

<b>Chapter 1. Release Notes.....</b>	<b>1</b>
1.1. Updates in Release 9.0.....	2
1.2. Updates in Release 9.1.....	2
1.3. Updates in Release 9.2.....	2
1.4. Updates in Release 9.3.....	3
1.5. Updates in Release 9.4.....	3
<b>Chapter 2. Validated Platforms.....</b>	<b>4</b>
2.1. Supported NVIDIA GPUs and Validated Server Platforms.....	4
2.2. Hypervisor Software Releases.....	5
2.3. Guest OS Support.....	6
2.3.1. Windows Guest OS Support.....	6
2.3.2. Linux Guest OS Support.....	8
2.4. NVIDIA CUDA Toolkit Version Support.....	9
<b>Chapter 3. Known Product Limitations.....</b>	<b>10</b>
3.1. Issues occur when the channels allocated to a vGPU are exhausted.....	10
3.2. Total frame buffer for vGPUs is less than the total frame buffer on the physical GPU....	11
3.3. Issues may occur with graphics-intensive OpenCL applications on vGPU types with limited frame buffer.....	12
3.4. In pass through mode, all GPUs connected to each other through NVLink must be assigned to the same VM.....	12
3.5. vGPU profiles with 512 Mbytes or less of frame buffer support only 1 virtual display head on Windows 10.....	13
3.6. NVENC requires at least 1 Gbyte of frame buffer.....	13
3.7. VM running older NVIDIA vGPU drivers fails to initialize vGPU when booted.....	14
3.8. Single vGPU benchmark scores are lower than pass-through GPU.....	15
3.9. nvidia-smi fails to operate when all GPUs are assigned to GPU pass-through mode.....	16
<b>Chapter 4. Resolved Issues.....</b>	<b>17</b>
<b>Chapter 5. Known Issues.....</b>	<b>20</b>
5.1. NVIDIA Control Panel fails to start if launched too soon from a VM without licensing information.....	20
5.2. Since 9.2: NVIDIA vGPU runtime statistics gathering does not function.....	20
5.3. DWM crashes randomly occur in Windows VMs.....	21
5.4. Citrix Virtual Apps and Desktops session freezes when the desktop is unlocked.....	21
5.5. NVIDIA vGPU software graphics driver fails after Linux kernel upgrade with DKMS enabled.....	22

5.6. 9.0, 9.1 Only: VMs fail to boot with failed assertions.....	23
5.7. 9.0, 9.1 Only: ECC memory with NVIDIA vGPU is not supported on Tesla M60 and Tesla M6.....	24
5.8. Virtual GPU fails to start if ECC is enabled.....	25
5.9. 9.0 only: Hypervisor host hangs when persistence mode is set.....	27
5.10. RAPIDS cuDF merge fails on NVIDIA vGPU.....	27
5.11. 9.0 only: No vCS option available in NVIDIA X Server Settings.....	28
5.12. 9.0 only: On Linux VMs, the license directory is not deleted when the guest driver is uninstalled.....	29
5.13. Vulkan applications crash in Windows 7 guest VMs configured with NVIDIA vGPU.....	30
5.14. Host core CPU utilization is higher than expected for moderate workloads.....	30
5.15. Frame capture while the interactive logon message is displayed returns blank screen.....	31
5.16. RDS sessions do not use the GPU with some Microsoft Windows Server releases.....	32
5.17. Even when the scheduling policy is equal share, unequal GPU utilization is reported.....	32
5.18. When the scheduling policy is fixed share, GPU utilization is reported as higher than expected.....	33
5.19. Benign warnings during Virtual GPU Manager installation or uninstallation after hypervisor upgrade.....	34
5.20. Benign not in gzip format messages during Virtual GPU Manager installation or uninstallation.....	35
5.21. License is not acquired in Windows VMs.....	36
5.22. nvidia-smi reports that vGPU migration is supported on all hypervisors.....	36
5.23. Resolution is not updated after a VM acquires a license and is restarted.....	37
5.24. A segmentation fault in Dbus code causes nvidia-gridd to exit on Red Hat Enterprise Linux and CentOS.....	37
5.25. No Manage License option available in NVIDIA X Server Settings by default.....	38
5.26. Licenses remain checked out when VMs are forcibly powered off.....	39
5.27. Memory exhaustion can occur with vGPU profiles that have 512 Mbytes or less of frame buffer.....	40
5.28. GNOME Display Manager (GDM) fails to start on Red Hat Enterprise Linux 7.2 and CentOS 7.0.....	41



---

# Chapter 1. Release Notes

These *Release Notes* summarize current status, information on validated platforms, and known issues with NVIDIA vGPU software and associated hardware on Nutanix AHV.



**Note:** The most current version of the documentation for this release of NVIDIA vGPU software can be found online at [NVIDIA Virtual GPU Software Documentation](#).

The releases in this release family of NVIDIA vGPU software include the software listed in the following table:

Software	9.0	9.1	9.2	9.3	9.4
NVIDIA Virtual GPU Manager for the Nutanix AHV releases listed in <a href="#">Hypervisor Software Releases</a>	430.27	430.46	430.67	430.83	430.99
NVIDIA Windows driver	431.02	431.79	432.08	432.33	432.44
NVIDIA Linux driver	430.30	430.46	430.63	430.83	430.99



**CAUTION:**

If you install the wrong NVIDIA vGPU software packages for the version of Nutanix AHV you are using, NVIDIA Virtual GPU Manager will fail to load.

The releases of the vGPU Manager and guest VM drivers that you install must be compatible. Different versions of the vGPU Manager and guest VM driver from within the same main release branch can be used together. For example, you can use the vGPU Manager from release 9.1 with guest VM drivers from release 9.0. However, versions of the vGPU Manager and guest VM driver from different main release branches cannot be used together. For example, you cannot use the vGPU Manager from release 9.1 with guest VM drivers from release 7.2.

See [VM running older NVIDIA vGPU drivers fails to initialize vGPU when booted](#).

This requirement does not apply to the NVIDIA vGPU software license sever. All releases of NVIDIA vGPU software are compatible with **all** releases of the license server.

## 1.1. Updates in Release 9.0

### New Features in Release 9.0

- ▶ NVIDIA Virtual Compute Server (V) vGPUs for artificial intelligence, deep learning, and high-performance computing workloads
- ▶ Error correcting code (ECC) memory support
- ▶ Page retirement support
- ▶ Configurable times slices for equal share schedulers and fixed share schedulers
- ▶ New configuration parameter to specify host ID of a licensed client
- ▶ Miscellaneous bug fixes

### Hardware and Software Support Introduced in Release 9.0

- ▶ Support for the Tesla T4 GPU
- ▶ Support for Nutanix AHV 5.11
- ▶ Support for Windows 10 May 2019 Update (1903) as a guest OS

## 1.2. Updates in Release 9.1

### New Features in Release 9.1

- ▶ Security updates
- ▶ Miscellaneous bug fixes

## 1.3. Updates in Release 9.2

### New Features in Release 9.2

- ▶ Miscellaneous bug fixes
- ▶ Security updates
- ▶ Limitation on the maximum number of NVIDIA Virtual Compute Server vGPUs to eight vGPUs per physical GPU, irrespective of the available hardware resources of the physical GPU

## 1.4. Updates in Release 9.3

### New Features in Release 9.3

- ▶ Miscellaneous bug fixes
- ▶ Security updates (see [Security Bulletin: NVIDIA GPU Display Driver - February 2020](#))

### Feature Support Withdrawn in Release 9.3

- ▶ The following OS releases are no longer supported as a guest OS:
  - ▶ Windows Server 2008 R2
  - ▶ Red Hat Enterprise Linux 7.0-7.4
  - ▶ CentOS 7.0-7.4

## 1.5. Updates in Release 9.4

### New Features in Release 9.4

- ▶ Miscellaneous bug fixes
- ▶ Security updates - see [Security Bulletin: NVIDIA GPU Display Driver - June 2020](#)

### Hardware and Software Support Introduced in Release 9.4

- ▶ Support for Nutanix AHV 5.15
- ▶ Support for the following OS releases as a guest OS:
  - ▶ Red Hat Enterprise Linux 7.8
  - ▶ CentOS 7.8

### Feature Support Withdrawn in Release 9.4

- ▶ Nutanix AHV 5.9 and 5.8 are no longer supported
- ▶ The following OS releases are no longer supported as a guest OS:
  - ▶ Red Hat Enterprise Linux 7.5
  - ▶ CentOS 7.5

---

# Chapter 2. Validated Platforms

This release family of NVIDIA vGPU software provides support for several NVIDIA GPUs on validated server hardware platforms, Nutanix AHV hypervisor software versions, and guest operating systems. It also supports the version of NVIDIA CUDA Toolkit that is compatible with R430 drivers.

## 2.1. Supported NVIDIA GPUs and Validated Server Platforms

This release of NVIDIA vGPU software provides support for the following NVIDIA GPUs on Nutanix AHV, running on validated Nutanix NX series appliances and OEM server hardware platforms:

- ▶ GPUs based on the NVIDIA Maxwell™ graphic architecture:
  - ▶ Tesla M10 (vCS is **not** supported.)
  - ▶ Tesla M60 (vCS is **not** supported.)
- ▶ GPUs based on the NVIDIA Pascal™ architecture:
  - ▶ Tesla P4
  - ▶ Tesla P40
- ▶ GPUs based on the NVIDIA Volta architecture:
  - ▶ Tesla V100 PCIe
  - ▶ Tesla V100 PCIe 32GB
- ▶ GPUs based on the NVIDIA Turing™ architecture:
  - ▶ Tesla T4

For a list of validated server platforms, refer to [NVIDIA GRID Certified Nutanix Servers](#).



**Note:**

Tesla M60 and M6 GPUs support compute mode and graphics mode. NVIDIA vGPU requires GPUs that support both modes to operate in graphics mode.



Recent Tesla M60 GPUs and M6 GPUs are supplied in graphics mode. However, your GPU might be in compute mode if it is an older Tesla M60 GPU or M6 GPU, or if its mode has previously been changed.

To configure the mode of Tesla M60 and M6 GPUs, use the `gpumodeswitch` tool provided with NVIDIA vGPU software releases.

Even in compute mode, Tesla M60 and M6 GPUs do **not** support NVIDIA Virtual Compute Server vGPU types.

## 2.2. Hypervisor Software Releases

This release is supported on the Nutanix AHV releases listed in the table.



### Note:

Updates to a base release of Nutanix AHV are compatible with the base release and can also be used with this version of NVIDIA vGPU software unless expressly stated otherwise.

**Since 9.2:** In NVIDIA vGPU software release 9.2, an API incompatibility between the NVIDIA vGPU software drivers and Nutanix AHV was introduced that prevents the gathering of runtime vGPU statistics. However, NVIDIA vGPU software functionality is not affected and Nutanix AHV remains fully supported.

Software	Releases Supported	Notes
<b>Since 9.4:</b> Nutanix AOS Hypervisor (AHV) 5.15	5.15 RTM and compatible 5.15.x updates	
Nutanix AOS Hypervisor (AHV) 5.11	5.11 RTM and compatible 5.11.x updates	
Nutanix AOS Hypervisor (AHV) 5.10	5.10 RTM and compatible 5.10.x updates	<p>This release does <b>not</b> support error correcting code (ECC) memory.</p> <p>Support for the following GPUs is introduced in release 5.10.1:</p> <ul style="list-style-type: none"> <li>▶ Tesla P4</li> <li>▶ Tesla V100 PCIe 32GB</li> </ul> <p>Support for the Tesla T4 GPU is introduced in release 5.10.4.</p> <p>5.10 RTM does <b>not</b> support Tesla P4, Tesla T4 or Tesla V100 PCIe 32GB.</p>
<b>9.0-9.3 only:</b> Nutanix AOS Hypervisor (AHV) 5.9	5.9 RTM and compatible 5.9.x updates	<p>This release does <b>not</b> support ECC memory.</p> <p>Not supported on the following GPUs:</p>

Software	Releases Supported	Notes
		<ul style="list-style-type: none"> <li>▶ Tesla P4</li> <li>▶ Tesla T4</li> <li>▶ Tesla V100 PCIe 32GB</li> </ul>
<b>9.0-9.3 only:</b> Nutanix AOS Hypervisor (AHV) 5.8	5.8 RTM and compatible 5.8.x updates	<p>This release does <b>not</b> support ECC memory.</p> <p>Not supported on the following GPUs:</p> <ul style="list-style-type: none"> <li>▶ Tesla P4</li> <li>▶ Tesla T4</li> <li>▶ Tesla V100 PCIe 32GB</li> </ul>
Nutanix AOS Hypervisor (AHV) 5.5	5.5 RTM and compatible 5.5.x updates	<p>This release does <b>not</b> support ECC memory.</p> <p>Not supported on the following GPUs:</p> <ul style="list-style-type: none"> <li>▶ Tesla P4</li> <li>▶ Tesla T4</li> <li>▶ Tesla V100 PCIe 32GB</li> </ul>

## 2.3. Guest OS Support

NVIDIA vGPU software supports several Windows releases and Linux distributions as a guest OS. The supported guest operating systems depend on the hypervisor software version.



### Note:

Use only a guest OS release that is listed as supported by NVIDIA vGPU software with your virtualization software. To be listed as supported, a guest OS release must be supported not only by NVIDIA vGPU software, but also by your virtualization software. NVIDIA **cannot** support guest OS releases that your virtualization software does not support.

NVIDIA vGPU software supports **only** 64-bit guest operating systems. No 32-bit guest operating systems are supported.

### 2.3.1. Windows Guest OS Support

NVIDIA vGPU software supports **only** the 64-bit Windows releases listed in the table as a guest OS on Nutanix AHV. The releases of Nutanix AHV for which a Windows release is supported depend on whether NVIDIA vGPU or pass-through GPU is used.



### Note:

If a specific release, even an update release, is not listed, it's **not** supported.

Guest OS	NVIDIA vGPU - Nutanix AHV Releases	Pass-Through GPU - Nutanix AHV Releases
Windows Server 2019	<b>Since 9.4:</b> 5.15, 5.11, 5.10 <b>9.0-9.3 only:</b> 5.11, 5.10	<b>Since 9.4:</b> 5.15, 5.11, 5.10 <b>9.0-9.3 only:</b> 5.11, 5.10
Windows Server 2016 1709, 1607	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5
Windows Server 2012 R2	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5
<b>9.0-9.2 only:</b> Windows Server 2008 R2	5.11, 5.10, 5.9, 5.8, 5.5	5.11, 5.10, 5.9, 5.8, 5.5
Windows 10: <ul style="list-style-type: none"> <li>▶ May 2019 Update (1903)</li> <li>▶ October 2018 Update (1809)</li> <li>▶ Spring Creators Update (1803)</li> <li>▶ Fall Creators Update (1709)</li> <li>▶ Creators Update (1703)</li> <li>▶ Anniversary Update (1607)</li> <li>▶ November Update (1511)</li> <li>▶ RTM (1507)</li> </ul>	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5
Windows 8.1 Update	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5
Windows 8.1	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5
Windows 8	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5

Guest OS	NVIDIA vGPU - Nutanix AHV Releases	Pass-Through GPU - Nutanix AHV Releases
Windows 7	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5

## 2.3.2. Linux Guest OS Support

NVIDIA vGPU software supports **only** the Linux distributions listed in the table as a guest OS on Nutanix AHV. The releases of Nutanix AHV for which a Linux release is supported depend on whether NVIDIA vGPU or pass-through GPU is used.



### Note:

If a specific release, even an update release, is not listed, it's **not** supported.

Guest OS	NVIDIA vGPU - Nutanix AHV Releases	Pass-Through GPU - Nutanix AHV Releases
<b>Since 9.4:</b> Red Hat Enterprise Linux 7.6-7.8	5.15, 5.11, 5.10, 5.5	5.15, 5.11, 5.10, 5.5
<b>9.3 only:</b> Red Hat Enterprise Linux 7.5-7.7	5.11, 5.10, 5.9, 5.8, 5.5	5.11, 5.10, 5.9, 5.8, 5.5
<b>9.1, 9.2 only:</b> Red Hat Enterprise Linux 7.0-7.7	5.11, 5.10, 5.9, 5.8, 5.5	5.11, 5.10, 5.9, 5.8, 5.5
<b>9.0 only:</b> Red Hat Enterprise Linux 7.0-7.6	5.11, 5.10, 5.9, 5.8, 5.5	5.11, 5.10, 5.9, 5.8, 5.5
<b>Since 9.4:</b> CentOS 7.6-7.8	5.15, 5.11, 5.10, 5.5	5.15, 5.11, 5.10, 5.5
<b>9.3 only:</b> CentOS 7.5-7.7	5.11, 5.10, 5.9, 5.8, 5.5	5.11, 5.10, 5.9, 5.8, 5.5
<b>9.1, 9.2 only:</b> CentOS 7.0-7.7	5.11, 5.10, 5.9, 5.8, 5.5	5.11, 5.10, 5.9, 5.8, 5.5
<b>9.0 only:</b> CentOS 7.0-7.6	5.11, 5.10, 5.9, 5.8, 5.5	5.11, 5.10, 5.9, 5.8, 5.5
Ubuntu 16.04 LTS	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5
Ubuntu 14.04 LTS	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5	<b>Since 9.4:</b> 5.15, 5.11, 5.10, 5.5 <b>9.0-9.3 only:</b> 5.11, 5.10, 5.9, 5.8, 5.5

## 2.4. NVIDIA CUDA Toolkit Version Support

The releases in this release family of NVIDIA vGPU software support NVIDIA CUDA Toolkit 10.1 Update 1.

For more information about NVIDIA CUDA Toolkit, see [CUDA Toolkit 10.1 Documentation](#).

**Note:**

If you are using NVIDIA vGPU software with CUDA on Linux, avoid conflicting installation methods by installing CUDA from a distribution-independent runfile package. Do not install CUDA from distribution-specific RPM or Deb package.

To ensure that the NVIDIA vGPU software graphics driver is not overwritten when CUDA is installed, deselect the CUDA driver when selecting the CUDA components to install.

For more information, see [NVIDIA CUDA Installation Guide for Linux](#).

---

# Chapter 3. Known Product Limitations

Known product limitations for this release of NVIDIA vGPU software are described in the following sections.

## 3.1. Issues occur when the channels allocated to a vGPU are exhausted

### Description

Issues occur when the channels allocated to a vGPU are exhausted and the guest VM to which the vGPU is assigned fails to allocate a channel to the vGPU. A physical GPU has a fixed number of channels and the number of channels allocated to each vGPU is inversely proportional to the maximum number of vGPUs allowed on the physical GPU.

When the channels allocated to a vGPU are exhausted and the guest VM fails to allocate a channel, the following errors are reported on the hypervisor host or in an NVIDIA bug report:

```
Jun 26 08:01:25 srvxen06f vgpu-3[14276]: error: vmiop_log: (0x0): Guest attempted to
allocate channel above its max channel limit 0xfb
Jun 26 08:01:25 srvxen06f vgpu-3[14276]: error: vmiop_log: (0x0): vGPU message 6
failed, result code: 0x1a
Jun 26 08:01:25 srvxen06f vgpu-3[14276]: error: vmiop_log: (0x0):
0xc1d004a1, 0xff0e0000, 0xff0400fb, 0xc36f,
Jun 26 08:01:25 srvxen06f vgpu-3[14276]: error: vmiop_log: (0x0):          0x1,
0xff1fe314, 0xff1fe038, 0x100b6f000, 0x1000,
Jun 26 08:01:25 srvxen06f vgpu-3[14276]: error: vmiop_log: (0x0):
0x80000000, 0xff0e0200, 0x0, 0x0, (Not logged),
Jun 26 08:01:25 srvxen06f vgpu-3[14276]: error: vmiop_log: (0x0):          0x1, 0x0
Jun 26 08:01:25 srvxen06f vgpu-3[14276]: error: vmiop_log: (0x0): , 0x0
```

### Workaround

Use a vGPU type with more frame buffer, thereby reducing the maximum number of vGPUs allowed on the physical GPU. As a result, the number of channels allocated to each vGPU is increased.

## 3.2. Total frame buffer for vGPUs is less than the total frame buffer on the physical GPU

Some of the physical GPU's frame buffer is used by the hypervisor on behalf of the VM for allocations that the guest OS would otherwise have made in its own frame buffer. The frame buffer used by the hypervisor is not available for vGPUs on the physical GPU. In NVIDIA vGPU deployments, frame buffer for the guest OS is reserved in advance, whereas in bare-metal deployments, frame buffer for the guest OS is reserved on the basis of the runtime needs of applications.

If error-correcting code (ECC) memory is enabled on a physical GPU that does not have HBM2 memory, the amount of frame buffer that is usable by vGPUs is further reduced. All types of vGPU are affected, not just vGPUs that support ECC memory.

On all GPUs that support ECC memory and, therefore, dynamic page retirement, additional frame buffer is allocated for dynamic page retirement. The amount that is allocated is inversely proportional to the maximum number of vGPUs per physical GPU. All GPUs that support ECC memory are affected, even GPUs that have HBM2 memory or for which ECC memory is disabled.

The approximate amount of frame buffer that NVIDIA vGPU software reserves can be calculated from the following formula:

$$\text{max-reserved-fb} = \text{vgpu-profile-size-in-mb} \div 16 + 16 + \text{ecc-adjustments} + \text{page-retirement-allocation}$$

### **max-reserved-fb**

The maximum total amount of reserved frame buffer in Mbytes that is not available for vGPUs.

### **vgpu-profile-size-in-mb**

The amount of frame buffer in Mbytes allocated to a single vGPU. This amount depends on the vGPU type. For example, for the T4-16Q vGPU type, *vgpu-profile-size-in-mb* is 16384.

### **ecc-adjustments**

The amount of frame buffer in Mbytes that is not usable by vGPUs when ECC is enabled on a physical GPU that does not have HBM2 memory. If ECC is disabled or the GPU has HBM2 memory, *ecc-adjustments* is 0.

### **page-retirement-allocation**

The amount of frame buffer in Mbytes that is reserved for dynamic page retirement.

- ▶ On GPUs based on the NVIDIA Maxwell GPU architecture, *page-retirement-allocation* =  $4 \div \text{max-vgpus-per-gpu}$ .
- ▶ On GPUs based on NVIDIA GPU architectures **after** the Maxwell architecture, *page-retirement-allocation* =  $128 \div \text{max-vgpus-per-gpu}$

***max-vgpus-per-gpu***

The maximum number of vGPUs that can be created simultaneously on a physical GPU. This number varies according to the vGPU type. For example, for the T4-16Q vGPU type, *max-vgpus-per-gpu* is 1.



**Note:** In VMs running a Windows guest OS that supports Windows Display Driver Model (WDDM) 1.x, namely, Windows 7, Windows 8.1, Windows Server 2008, and Windows Server 2012, an additional 48 Mbytes of frame buffer are reserved and not available for vGPUs.

### 3.3. Issues may occur with graphics-intensive OpenCL applications on vGPU types with limited frame buffer

#### Description

Issues may occur when graphics-intensive OpenCL applications are used with vGPU types that have limited frame buffer. These issues occur when the applications demand more frame buffer than is allocated to the vGPU.

For example, these issues may occur with the Adobe Photoshop and LuxMark OpenCL Benchmark applications:

- ▶ When the image resolution and size are changed in Adobe Photoshop, a program error may occur or Photoshop may display a message about a problem with the graphics hardware and a suggestion to disable OpenCL.
- ▶ When the LuxMark OpenCL Benchmark application is run, XID error 31 may occur.

#### Workaround

For graphics-intensive OpenCL applications, use a vGPU type with more frame buffer.

### 3.4. In pass through mode, all GPUs connected to each other through NVLink must be assigned to the same VM

#### Description

In pass through mode, all GPUs connected to each other through NVLink must be assigned to the same VM. If a subset of GPUs connected to each other through NVLink is passed through



to a VM, unrecoverable error `XID_74` occurs when the VM is booted. This error corrupts the NVLink state on the physical GPUs and, as a result, the NVLink bridge between the GPUs is unusable.

### Workaround

Restore the NVLink state on the physical GPUs by resetting the GPUs or rebooting the hypervisor host.

## 3.5. vGPU profiles with 512 Mbytes or less of frame buffer support only 1 virtual display head on Windows 10

### Description

To reduce the possibility of memory exhaustion, vGPU profiles with 512 Mbytes or less of frame buffer support only 1 virtual display head on a Windows 10 guest OS.

The following vGPU profiles have 512 Mbytes or less of frame buffer:

- ▶ Tesla M6-0B, M6-0Q
- ▶ Tesla M10-0B, M10-0Q
- ▶ Tesla M60-0B, M60-0Q

### Workaround

Use a profile that supports more than 1 virtual display head and has at least 1 Gbyte of frame buffer.

## 3.6. NVENC requires at least 1 Gbyte of frame buffer

### Description

Using the frame buffer for the NVIDIA hardware-based H.264/HEVC video encoder (NVENC) may cause memory exhaustion with vGPU profiles that have 512 Mbytes or less of frame buffer. To reduce the possibility of memory exhaustion, NVENC is disabled on profiles that have 512 Mbytes or less of frame buffer. Application GPU acceleration remains fully supported and available for all profiles, including profiles with 512 Mbytes or less of frame buffer.

NVENC support from both Citrix and VMware is a recent feature and, if you are using an older version, you should experience no change in functionality.

The following vGPU profiles have 512 Mbytes or less of frame buffer:

- ▶ Tesla M6-0B, M6-0Q
- ▶ Tesla M10-0B, M10-0Q
- ▶ Tesla M60-0B, M60-0Q

### Workaround

If you require NVENC to be enabled, use a profile that has at least 1 Gbyte of frame buffer.

## 3.7. VM running older NVIDIA vGPU drivers fails to initialize vGPU when booted

### Description

A VM running a version of the NVIDIA guest VM drivers from a previous main release branch, for example release 4.4, will fail to initialize vGPU when booted on a Nutanix AHV platform running the current release of Virtual GPU Manager.

In this scenario, the VM boots in standard VGA mode with reduced resolution and color depth. The NVIDIA virtual GPU is present in **Windows Device Manager** but displays a warning sign, and the following device status:

```
Windows has stopped this device because it has reported problems. (Code 43)
```

Depending on the versions of drivers in use, the Nutanix AHV VM's `/var/log/messages` log file reports one of the following errors:

- ▶ An error message:  
`vmiop_log: error: Unable to fetch Guest NVIDIA driver information`
- ▶ A version mismatch between guest and host drivers:  
`vmiop_log: error: Guest VGX version(1.1) and Host VGX version(1.2) do not match`
- ▶ A signature mismatch:  
`vmiop_log: error: vGPU message signature mismatch.`

### Resolution

Install the current NVIDIA guest VM driver in the VM.

## 3.8. Single vGPU benchmark scores are lower than pass-through GPU

### Description

A single vGPU configured on a physical GPU produces lower benchmark scores than the physical GPU run in pass-through mode.

Aside from performance differences that may be attributed to a vGPU's smaller frame buffer size, vGPU incorporates a performance balancing feature known as Frame Rate Limiter (FRL). On vGPUs that use the best-effort scheduler, FRL is enabled. On vGPUs that use the fixed share or equal share scheduler, FRL is disabled.

FRL is used to ensure balanced performance across multiple vGPUs that are resident on the same physical GPU. The FRL setting is designed to give good interactive remote graphics experience but may reduce scores in benchmarks that depend on measuring frame rendering rates, as compared to the same benchmarks running on a pass-through GPU.

### Resolution

FRL is controlled by an internal vGPU setting. On vGPUs that use the best-effort scheduler, NVIDIA does not validate vGPU with FRL disabled, but for validation of benchmark performance, FRL can be temporarily disabled by setting the extra vGPU parameter `frame_rate_limiter=0` through the ACLI when a vGPU type is assigned to a VM.

For example:

```
vm.gpu_assign a7f24fc0-dd05-4595-a76f-8b1901942aae gpu=Nvidia_GRID_M60-4Q  
extra_param=frame_rate_limiter=0
```

The setting takes effect the next time any VM using the given vGPU type is started.

With this setting in place, the VM's vGPU will run without any frame rate limit.

The FRL can be reverted back to its default setting by setting the extra vGPU parameter `frame_rate_limiter=1`

## 3.9. `nvidia-smi` fails to operate when all GPUs are assigned to GPU pass-through mode

### Description

If all GPUs in the platform are assigned to VMs in pass-through mode, `nvidia-smi` will return an error:

```
[root@vgx-test ~]# nvidia-smi  
Failed to initialize NVML: Unknown Error
```

This is because GPUs operating in pass-through mode are not visible to `nvidia-smi` and the NVIDIA kernel driver operating in the Nutanix AHV host.

### Resolution

N/A

---

# Chapter 4. Resolved Issues

Only resolved issues that have been previously noted as known issues or had a noticeable user impact are listed. The summary and description for each resolved issue indicate the effect of the issue on NVIDIA vGPU software **before the issue was resolved**.

## Issues Resolved in Release 9.0

Bug ID	Summary and Description
-	<p><b>Virtual GPU fails to start if ECC is enabled</b></p> <p>NVIDIA vGPU does not support error correcting code (ECC) memory. If ECC memory is enabled, NVIDIA vGPU fails to start.</p> <p>Starting with NVIDIA vGPU software release 9.0, NVIDIA vGPU supports ECC memory on GPUs and hypervisor software versions that support ECC.</p>
2285306	<p><b>Cloned VMs configured with a vGPU type different than the type in the master image fail to start</b></p> <p>Cloned VMs configured with a vGPU type different than the type in the master image fail to start.</p> <p>When a Windows 10 VM is booted, the VM becomes stuck in a loop and alternately displays <code>Getting devices ready: 50%</code> and <code>Preparation in progress</code>.</p>

## Issues Resolved in Release 9.1

Bug ID	Summary and Description
200534988	<p><b>Error XID 47 followed by multiple XID 32 errors</b></p> <p>After disconnecting Citrix Virtual Apps and Desktops and clicking the power button in the VM, error XID 47 occurs followed by multiple XID 32 errors. When these errors occur, the hypervisor host becomes unusable.</p>
200537078	<p><b><u>9.0 only: Hypervisor host hangs when persistence mode is set</u></b></p>

Bug ID	Summary and Description
	When persistence mode is set through the <code>nvidia-smi</code> command, a very large number of <code>nv_queue</code> worker threads is spawned, causing the hypervisor host to hang.
200524555	<p><b><u>9.0 only: On Linux VMs, the license directory is not deleted when the guest driver is uninstalled</u></b></p> <p>On Linux guest VMs, the license directory <code>/etc/nvidia/license</code> is not deleted when the NVIDIA vGPU software graphics driver is uninstalled.</p>
200522255	<p><b><u>9.0 only: No vCS option available in NVIDIA X Server Settings</u></b></p> <p>The <b>vCS</b> option is missing from the <b>Manage License</b> section in the <b>NVIDIA X Server Settings</b> window.</p>

## Issues Resolved in Release 9.2

Bug ID	Summary and Description
2644858	<p><b><u>9.0, 9.1 Only: VMs fail to boot with failed assertions</u></b></p> <p>In some scenarios with heavy workloads running on multiple VMs configured with NVIDIA vGPUs on a single physical GPU, additional VMs configured with NVIDIA vGPU on the same GPU fail to boot. The failure of the VM to boot is followed by failed assertions. This issue affects GPUs based on the NVIDIA Volta GPU architecture and later architectures.</p>
-	<p><b><u>9.0, 9.1 Only: ECC memory with NVIDIA vGPU is not supported on Tesla M60 and Tesla M6</u></b></p> <p>Error-correcting code (ECC) memory with NVIDIA vGPU is not supported on Tesla M60 and Tesla M6 GPUs. The effect of starting NVIDIA vGPU when it is configured on a Tesla M60 or Tesla M6 GPU on which ECC memory is enabled depends on your NVIDIA vGPU software release.</p> <ul style="list-style-type: none"> <li>▶ <b>9.0 only:</b> The hypervisor host fails.</li> <li>▶ <b>9.1 only:</b> The VM fails to start.</li> </ul>
-	<p><b><u>Virtual GPU fails to start if ECC is enabled</u></b></p> <p>NVIDIA vGPU does not support ECC memory with the following GPUs and hypervisor software releases:</p> <ul style="list-style-type: none"> <li>▶ <b>9.0, 9.1 only:</b> Tesla M60 GPUs</li> <li>▶ <b>9.0, 9.1 only:</b> Tesla M6 GPUs</li> <li>▶ Nutanix AHV releases earlier than 5.11</li> </ul>

Bug ID	Summary and Description
	<p>If ECC memory is enabled and your GPU or Nutanix AHV release does not support ECC, NVIDIA vGPU fails to start.</p> <p>This issue is resolved <b>only</b> for ECC support on Tesla M60 and Tesla M6 GPUs. ECC remains unsupported on Nutanix AHV releases earlier than 5.11.</p>

### Issues Resolved in Release 9.3

No resolved issues are reported in this release for Nutanix AHV.

### Issues Resolved in Release 9.4

No resolved issues are reported in this release for Nutanix AHV.

---

## Chapter 5. Known Issues

### 5.1. NVIDIA Control Panel fails to start if launched too soon from a VM without licensing information

#### Description

If NVIDIA licensing information is not configured on the system, any attempt to start **NVIDIA Control Panel** by right-clicking on the desktop within 30 seconds of the VM being started fails.

#### Workaround

Wait at least 30 seconds before trying to launch **NVIDIA Control Panel**.

#### Status

Open

#### Ref. #

200623179

### 5.2. Since 9.2: NVIDIA vGPU runtime statistics gathering does not function

#### Description

In NVIDIA vGPU software release 9.2, an API incompatibility between the NVIDIA vGPU software drivers and Nutanix AHV was introduced that prevents the gathering of runtime vGPU statistics. However, NVIDIA vGPU software functionality is not affected and Nutanix AHV remains fully supported.



### Version

This issue affects NVIDIA vGPU software 9.x releases since 9.2.

### Status

Open

### Ref. #

2699736

## 5.3. DWM crashes randomly occur in Windows VMs

### Description

Desktop Windows Manager (DWM) crashes randomly occur in Windows VMs, causing a blue-screen crash and the bug check `CRITICAL_PROCESS_DIED`. Computer Management shows problems with the primary display device.

### Version

This issue affects Windows 10 1809, 1903 and 1909 VMs.

### Status

Not an NVIDIA bug

### Ref. #

2730037

## 5.4. Citrix Virtual Apps and Desktops session freezes when the desktop is unlocked

### Description

When a Citrix Virtual Apps and Desktops session that is locked is unlocked by pressing **Ctrl+Alt+Del**, the session freezes. This issue affects only VMs that are running Microsoft Windows 10 1809 as a guest OS.

## Version

Microsoft Windows 10 1809 guest OS

## Workaround

Restart the VM.

## Status

Not an NVIDIA bug

## Ref. #

2767012

# 5.5. NVIDIA vGPU software graphics driver fails after Linux kernel upgrade with DKMS enabled

## Description

After the Linux kernel is upgraded (for example by running `sudo apt full-upgrade`) with Dynamic Kernel Module Support (DKMS) enabled, the `nvidia-smi` command fails to run. If DKMS is enabled, an upgrade to the Linux kernel triggers a rebuild of the NVIDIA vGPU software graphics driver. The rebuild of the driver fails because the compiler version is incorrect. Any attempt to reinstall the driver fails because the kernel fails to build.

When the failure occurs, the following messages are displayed:

```
-> Installing DKMS kernel module:
    ERROR: Failed to run `/usr/sbin/dkms build -m nvidia -v 430.30 -k 5.3.0-28-generic`:
    Kernel preparation unnecessary for this kernel. Skipping...
    Building module:
    cleaning build area...
    'make' -j8 NV_EXCLUDE_BUILD_MODULES='' KERNEL_UNAME=5.3.0-28-generic
    IGNORE_CC_MISMATCH='T' modules... (bad exit status: 2)
    ERROR (dkms apport): binary package for nvidia: 430.30 not found
    Error! Bad return status for module build on kernel: 5.3.0-28-generic
    (x86_64)
    Consult /var/lib/dkms/nvidia/430.30/build/make.log for more information.
    -> error.
    ERROR: Failed to install the kernel module through DKMS. No kernel module
    was installed;
    please try installing again without DKMS, or check the DKMS logs for more
    information.
    ERROR: Installation has failed. Please see the file '/var/log/nvidia-
    installer.log' for details.
    You may find suggestions on fixing installation problems in the README
    available on the Linux driver download page at www.nvidia.com.
```

## Workaround

When installing the NVIDIA vGPU software graphics driver with DKMS enabled, specify the `--no-cc-version-check` option.

## Status

Not a bug.

## Ref. #

2836271

# 5.6. 9.0, 9.1 Only: VMs fail to boot with failed assertions

## Description

In some scenarios with heavy workloads running on multiple VMs configured with NVIDIA vGPUs on a single physical GPU, additional VMs configured with NVIDIA vGPU on the same GPU fail to boot. The failure of the VM to boot is followed by failed assertions. This issue affects GPUs based on the NVIDIA Volta GPU architecture and later architectures.

When this error occurs, error messages similar to the following examples are logged to the Nutanix AHV log file:

```
nvidia-vgpu-mgr[31526]: error: vmiop_log: NVOS status 0x1e
nvidia-vgpu-mgr[31526]: error: vmiop_log: Assertion Failed at 0xb2d3e4d7:96
nvidia-vgpu-mgr[31526]: error: vmiop_log: 12 frames returned by backtrace
nvidia-vgpu-mgr[31526]: error: vmiop_log: /usr/lib64/libnvidia-vgpu.so(_nv003956vgpu
+0x18) [0x7f4bb2cfb338] vmiop_dump_stack
nvidia-vgpu-mgr[31526]: error: vmiop_log: /usr/lib64/libnvidia-vgpu.so(_nv004018vgpu
+0xd4) [0x7f4bb2d09ce4] vmiopd_alloc_pb_channel
nvidia-vgpu-mgr[31526]: error: vmiop_log: /usr/lib64/libnvidia-vgpu.so(_nv002878vgpu
+0x137) [0x7f4bb2d3e4d7] vgpufceInitCopyEngine_GK104
nvidia-vgpu-mgr[31526]: error: vmiop_log: /usr/lib64/libnvidia-vgpu.so(+0x80e27)
[0x7f4bb2cd0e27]
nvidia-vgpu-mgr[31526]: error: vmiop_log: /usr/lib64/libnvidia-vgpu.so(+0x816a7)
[0x7f4bb2cd16a7]
nvidia-vgpu-mgr[31526]: error: vmiop_log: vgpu() [0x413820]
nvidia-vgpu-mgr[31526]: error: vmiop_log: vgpu() [0x413a8d]
nvidia-vgpu-mgr[31526]: error: vmiop_log: vgpu() [0x40e11f]
nvidia-vgpu-mgr[31526]: error: vmiop_log: vgpu() [0x40bb69]
nvidia-vgpu-mgr[31526]: error: vmiop_log: vgpu() [0x40b51c]
nvidia-vgpu-mgr[31526]: error: vmiop_log: /lib64/libc.so.6(__libc_start_main+0x100)
[0x7f4bb2fed20]
nvidia-vgpu-mgr[31526]: error: vmiop_log: vgpu() [0x4033ea]
nvidia-vgpu-mgr[31526]: error: vmiop_log: (0x0): Alloc Channel(Gpfifo) for device
failed error: 0x1e
nvidia-vgpu-mgr[31526]: error: vmiop_log: (0x0): Failed to allocate FCE channel
nvidia-vgpu-mgr[31526]: error: vmiop_log: (0x0): init_device_instance failed for
inst 0 with error 2 (init frame copy engine)
nvidia-vgpu-mgr[31526]: error: vmiop_log: (0x0): Initialization:
init_device_instance failed error 2
```

```
nvidia-vgpu-mgr[31526]: error: vmiop_log: display_init failed for inst: 0
nvidia-vgpu-mgr[31526]: error: vmiop_env_log: (0x0): vmiope_process_configuration:
plugin registration error
nvidia-vgpu-mgr[31526]: error: vmiop_env_log: (0x0): vmiope_process_configuration
failed with 0x1a
kernel: [858113.083773] [nvidia-vgpu-vfio] ace3f3bb-17d8-4587-920e-199b8fed532d:
start failed. status: 0x1
```

## Status

Resolved in NVIDIA vGPU software 9.2.

## Ref. #

2644858

# 5.7. 9.0, 9.1 Only: ECC memory with NVIDIA vGPU is not supported on Tesla M60 and Tesla M6

## Description

Error-correcting code (ECC) memory with NVIDIA vGPU is not supported on Tesla M60 and Tesla M6 GPUs. The effect of starting NVIDIA vGPU when it is configured on a Tesla M60 or Tesla M6 GPU on which ECC memory is enabled depends on your NVIDIA vGPU software release.

- ▶ **9.0 only:** The hypervisor host fails.
- ▶ **9.1 only:** The VM fails to start.

## Workaround

Ensure that ECC memory is disabled on Tesla M60 and Tesla M6 GPUs. For more information, see [Virtual GPU fails to start if ECC is enabled](#).

## Status

Resolved in NVIDIA vGPU software 9.2

## 5.8. Virtual GPU fails to start if ECC is enabled

### Description

Tesla M60, Tesla M6, and GPUs based on the Pascal GPU architecture, for example Tesla P100 or Tesla P4, support error correcting code (ECC) memory for improved data integrity. Tesla M60 and M6 GPUs in graphics mode are supplied with ECC memory disabled by default, but it may subsequently be enabled using `nvidia-smi`. GPUs based on the Pascal GPU architecture are supplied with ECC memory enabled.

NVIDIA vGPU does not support ECC memory with the following GPUs and hypervisor software releases:

- ▶ **9.0, 9.1 only:** Tesla M60 GPUs
- ▶ **9.0, 9.1 only:** Tesla M6 GPUs
- ▶ Nutanix AHV releases earlier than 5.11

If ECC memory is enabled and your GPU or Nutanix AHV release does not support ECC, NVIDIA vGPU fails to start.

The following error is logged in the Nutanix AHV host's `/var/log/messages` log file:

```
vmiop_log: error: Initialization: VGX not supported with ECC Enabled.
```

### Workaround

If you are using Tesla M60 or Tesla M6 GPUs, or a release of Nutanix AHV that does not support ECC, ensure that ECC is disabled on all GPUs.

Before you begin, ensure that NVIDIA Virtual GPU Manager is installed on your hypervisor.

1. Use `nvidia-smi` to list the status of all GPUs, and check for ECC noted as enabled on GPUs.

```
# nvidia-smi -q
=====NVSMI LOG=====

Timestamp                : Tue Dec 19 18:36:45 2017
Driver Version           : 384.99

Attached GPUs            : 1
GPU 0000:02:00.0

[...]

  Ecc Mode
    Current                : Enabled
    Pending                : Enabled

[...]
```

2. Change the ECC status to off on each GPU for which ECC is enabled.

- ▶ If you want to change the ECC status to off for all GPUs on your host machine, run this command:

```
# nvidia-smi -e 0
```

- ▶ If you want to change the ECC status to off for a specific GPU, run this command:

```
# nvidia-smi -i id -e 0
```

*id* is the index of the GPU as reported by `nvidia-smi`.

This example disables ECC for the GPU with index `0000:02:00.0`.

```
# nvidia-smi -i 0000:02:00.0 -e 0
```

3. Reboot the host.

4. Confirm that ECC is now disabled for the GPU.

```
# nvidia-smi -q
=====NVSMI LOG=====
Timestamp                : Tue Dec 19 18:37:53 2017
Driver Version           : 384.99
Attached GPUs            : 1
GPU 0000:02:00.0
[...]

  Ecc Mode
  Current                : Disabled
  Pending                : Disabled
[...]
```

If you later need to enable ECC on your GPUs, run one of the following commands:

- ▶ If you want to change the ECC status to on for all GPUs on your host machine, run this command:

```
# nvidia-smi -e 1
```

- ▶ If you want to change the ECC status to on for a specific GPU, run this command:

```
# nvidia-smi -i id -e 1
```

*id* is the index of the GPU as reported by `nvidia-smi`.

This example enables ECC for the GPU with index `0000:02:00.0`.

```
# nvidia-smi -i 0000:02:00.0 -e 1
```

After changing the ECC status to on, reboot the host.

## Status

Resolved in NVIDIA vGPU software 9.2

This issue is resolved **only** for ECC support on Tesla M60 and Tesla M6 GPUs. ECC remains unsupported on Nutanix AHV releases earlier than 5.11.

## 5.9. 9.0 only: Hypervisor host hangs when persistence mode is set

### Description

When persistence mode is set through the `nvidia-smi` command, a very large number of `nv_queue` worker threads is spawned, causing the hypervisor host to hang.

When this issue occurs, the hypervisor host creates a kernel stack trace similar to the following example.

```
Jun 24 04:42:16 example-02 kernel: kernel BUG at /build/linux-Ue9GXV/linux-4.15.0/
drivers/pci/msi.c:352!
...
Jun 24 04:42:16 example-02 kernel: Call Trace:
Jun 24 04:42:16 example-02 kernel: pci_disable_msi+0x100/0x120
Jun 24 04:42:16 example-02 kernel: nv_open_device+0x411/0x8e0 [nvidia]
Jun 24 04:42:16 example-02 kernel: ? kmem_cache_alloc+0x158/0x1b0
Jun 24 04:42:16 example-02 kernel: nvidia_open+0x339/0x550 [nvidia]
Jun 24 04:42:16 example-02 kernel: ? lookup_fast+0xcc/0x320
Jun 24 04:42:16 example-02 kernel: nvidia_frontend_open+0x58/0xa0 [nvidia]
Jun 24 04:42:16 example-02 kernel: chrdev_open+0xc4/0x1b0
```

### Status

Resolved in NVIDIA vGPU software 9.1

### Ref. #

200537078

## 5.10. RAPIDS cuDF merge fails on NVIDIA vGPU

### Description

The `merge` function of the RAPIDS cuDF GPU data frame library fails on NVIDIA vGPU. This function fails because RAPIDS uses the Unified Memory feature of CUDA, which NVIDIA vGPU does not support.

### Status

Open

### Ref. #

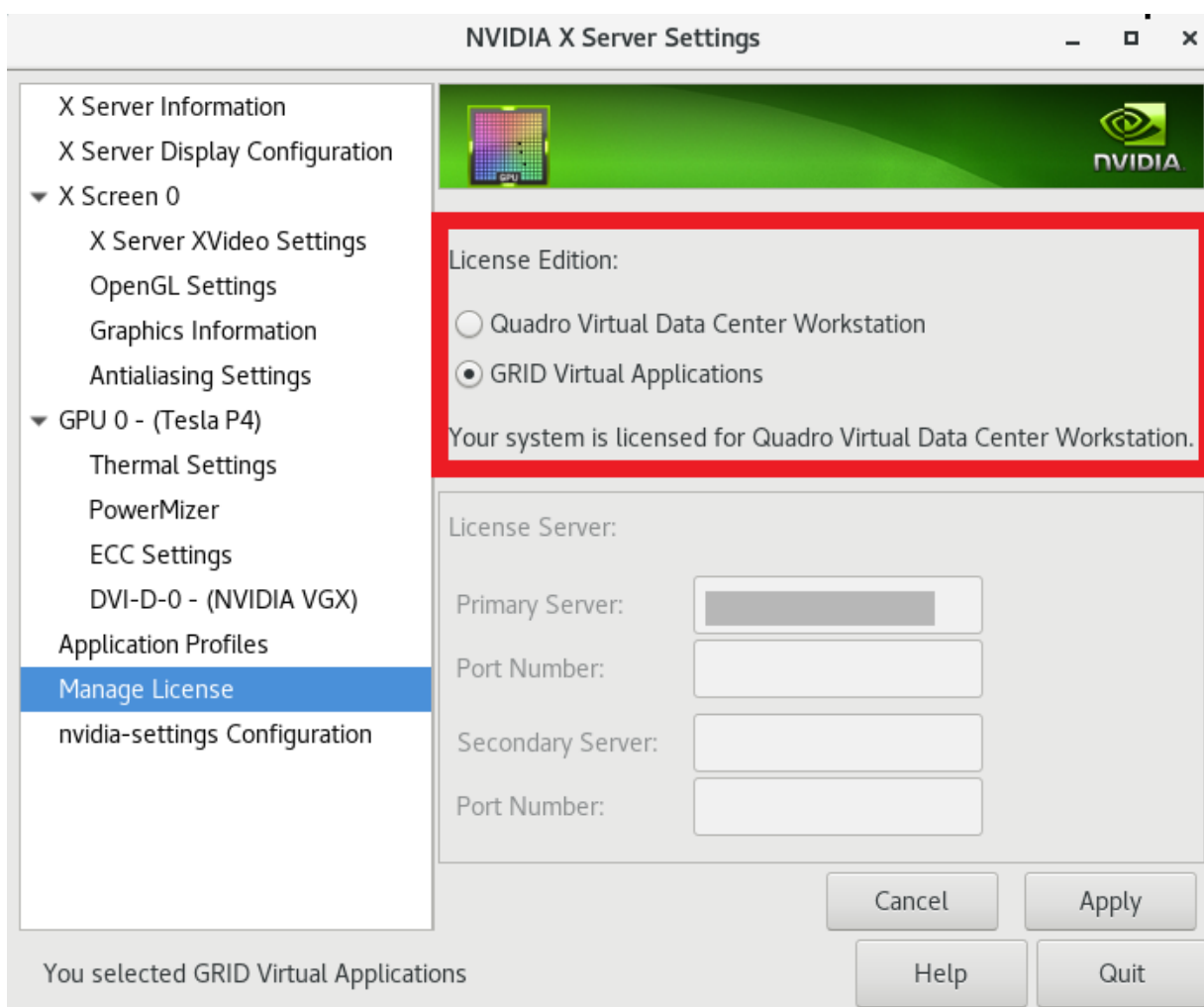
2642134

## 5.11. 9.0 only: No vCS option available in NVIDIA X Server Settings

### Description

The **vCS** option is missing from the **Manage License** section in the **NVIDIA X Server Settings** window.

As a result of this missing option, the **NVIDIA X Server Settings** window incorrectly states that the system is licensed for Quadro vDWS when, in fact, the system is licensed for vCS.





## Workaround

If you are licensing a physical GPU for vCS, you **must** use the configuration file `/etc/nvidia/gridd.conf`. See [Virtual GPU Client Licensing User Guide](#).

## Status

Resolved in NVIDIA vGPU software 9.1

## Ref. #

200522255

# 5.12. 9.0 only: On Linux VMs, the license directory is not deleted when the guest driver is uninstalled

## Description

On Linux guest VMs, the license directory `/etc/nvidia/license` is not deleted when the NVIDIA vGPU software graphics driver is uninstalled.

The following error message is written to the `nvidia-uninstaller` log file:

```
Failed to delete the directory '/etc/nvidia' (Directory not empty).
```

## Workaround

As root, remove the `/etc/nvidia/license` directory after the NVIDIA vGPU software graphics driver is uninstalled.

## Status

Resolved in NVIDIA vGPU software 9.1

## Ref. #

200524555

## 5.13. Vulkan applications crash in Windows 7 guest VMs configured with NVIDIA vGPU

### Description

In Windows 7 guest VMs configured with NVIDIA vGPU, applications developed with Vulkan APIs crash or throw errors when they are launched. Vulkan APIs require sparse texture support, but in Windows 7 guest VMs configured with NVIDIA vGPU, sparse textures are not enabled.

In Windows 10 guest VMs configured with NVIDIA vGPU, sparse textures are enabled and applications developed with Vulkan APIs run correctly in these VMs.

### Status

Open

### Ref. #

200381348

## 5.14. Host core CPU utilization is higher than expected for moderate workloads

### Description

When GPU performance is being monitored, host core CPU utilization is higher than expected for moderate workloads. For example, host CPU utilization when only a small number of VMs are running is as high as when several times as many VMs are running.

### Workaround

Disable monitoring of the following GPU performance statistics:

- ▶ vGPU engine usage by applications across multiple vGPUs
- ▶ Encoder session statistics
- ▶ Frame buffer capture (FBC) session statistics
- ▶ Statistics gathered by performance counters in guest VMs

## Status

Open

## Ref. #

2414897

# 5.15. Frame capture while the interactive logon message is displayed returns blank screen

## Description

Because of a known limitation with NvFBC, a frame capture while the interactive logon message is displayed returns a blank screen.

An NvFBC session can capture screen updates that occur after the session is created. Before the logon message appears, there is no screen update after the message is shown and, therefore, a black screen is returned instead. If the NvFBC session is created after this update has occurred, NvFBC cannot get a frame to capture.

## Workaround

Press **Enter** or wait for the screen to update for NvFBC to capture the frame.

## Status

Not a bug

## Ref. #

2115733

## 5.16. RDS sessions do not use the GPU with some Microsoft Windows Server releases

### Description

When some releases of Windows Server are used as a guest OS, Remote Desktop Services (RDS) sessions do not use the GPU. With these releases, the RDS sessions by default use the Microsoft Basic Render Driver instead of the GPU. This default setting enables 2D DirectX applications such as Microsoft Office to use software rendering, which can be more efficient than using the GPU for rendering. However, as a result, 3D applications that use DirectX are prevented from using the GPU.

### Version

- ▶ Windows Server 2016
- ▶ Windows Server 2012

### Solution

Change the local computer policy to use the hardware graphics adapter for all RDS sessions.

1. Choose **Local Computer Policy > Computer Configuration > Administrative Templates > Windows Components > Remote Desktop Services > Remote Desktop Session Host > Remote Session Environment**.
2. Set the **Use the hardware default graphics adapter for all Remote Desktop Services sessions** option.

## 5.17. Even when the scheduling policy is equal share, unequal GPU utilization is reported

### Description

When the scheduling policy is equal share, unequal GPU engine utilization can be reported for the vGPUs on the same physical GPU.

For example, GPU engine usage for three P40-8Q vGPUs on a Tesla P40 GPU might be reported as follows:

```
[root@localhost:~] nvidia-smi vgpu
```

Wed Jun 27 10:33:18 2018

```

+-----+-----+
| NVIDIA-SMI 390.59 | Driver Version: 390.59 |
+-----+-----+
| GPU  Name | Bus-Id | GPU-Util |
| vGPU ID  Name | VM ID  VM Name | vGPU-Util |
+-----+-----+
| 0  Tesla P40 | 00000000:81:00.0 | 52% |
| 2122661  GRID P40-8Q | 2122682  centos7.4-xmpl-211... | 19% |
| 2122663  GRID P40-8Q | 2122692  centos7.4-xmpl-211... | 0% |
| 2122659  GRID P40-8Q | 2122664  centos7.4-xmpl-211... | 25% |
+-----+-----+
| 1  Tesla P40 | 00000000:85:00.0 | 58% |
| 2122662  GRID P40-8Q | 2122689  centos7.4-xmpl-211... | 0% |
| 2122658  GRID P40-8Q | 2122667  centos7.4-xmpl-211... | 59% |
| 2122660  GRID P40-8Q | 2122670  centos7.4-xmpl-211... | 0% |
+-----+-----+

```

The vGPU utilization of the vGPU 2122658 is reported as 59%. However, the expected vGPU utilization should not exceed 33%.

This behavior is a result of the mechanism that is used to measure GPU engine utilization.

## Status

Open

## Ref. #

2175888

# 5.18. When the scheduling policy is fixed share, GPU utilization is reported as higher than expected

## Description

When the scheduling policy is fixed share, GPU engine utilization can be reported as higher than expected for a vGPU.

For example, GPU engine usage for six P40-4Q vGPUs on a Tesla P40 GPU might be reported as follows:

```

[root@localhost:~] # nvidia-smi vgpu
Mon Aug 20 10:33:18 2018
+-----+-----+
| NVIDIA-SMI 390.42 | Driver Version: 390.42 |
+-----+-----+
| GPU  Name | Bus-Id | GPU-Util |
| vGPU ID  Name | VM ID  VM Name | vGPU-Util |
+-----+-----+
| 0  Tesla P40 | 00000000:81:00.0 | 99% |
| 85109  GRID P40-4Q | 85110  win7-xmpl-146048-1 | 32% |
| 87195  GRID P40-4Q | 87196  win7-xmpl-146048-2 | 39% |
| 88095  GRID P40-4Q | 88096  win7-xmpl-146048-3 | 26% |
+-----+-----+

```

	89170	GRID P40-4Q		89171	win7-xmpl-146048-4		0%	
	90475	GRID P40-4Q		90476	win7-xmpl-146048-5		0%	
	93363	GRID P40-4Q		93364	win7-xmpl-146048-6		0%	
+-----+-----+-----+-----+-----+-----+-----+-----+-----+								
	1	Tesla P40		00000000:85:00.0			0%	
+-----+-----+-----+-----+-----+-----+-----+-----+-----+								

The vGPU utilization of vGPU 85109 is reported as 32%. For vGPU 87195, vGPU utilization is reported as 39%. And for 88095, it is reported as 26%. However, the expected vGPU utilization of any vGPU should not exceed approximately 16.7%.

This behavior is a result of the mechanism that is used to measure GPU engine utilization.

## Status

Open

## Ref. #

2227591

# 5.19. Benign warnings during Virtual GPU Manager installation or uninstallation after hypervisor upgrade

## Description

If the Virtual GPU Manager is installed or uninstalled after the hypervisor is upgraded, several warning messages about missing files are displayed.

The messages are similar to the following examples:

```
WARNING: Can't read module /lib/modules/4.4.77-1.el6.nutanix.20170830.124.x86_64/
weak-updates/ixgbevf.ko: No such file or directory
WARNING: Can't read module /lib/modules/4.4.77-1.el6.nutanix.20170830.124.x86_64/
weak-updates/i40evf.ko: No such file or directory
...
egrep: /lib/modules/4.4.77-1.el6.nutanix.20170830.124.x86_64//weak-updates/
ixgbevf.ko: No such file or directory
egrep: /lib/modules/4.4.77-1.el6.nutanix.20170830.124.x86_64//weak-updates/
i40evf.ko: No such file or directory
...
gzip: /boot/initramfs-4.4.77-1.el6.nutanix.20170830.124.x86_64.tmp: not in gzip
format
WARNING: Can't read module /lib/modules/4.4.77-1.el6.nutanix.20170830.124.x86_64/
weak-updates/ixgbevf.ko: No such file or directory
WARNING: Can't read module /lib/modules/4.4.77-1.el6.nutanix.20170830.124.x86_64/
weak-updates/i40evf.ko: No such file or directory
...
```

After a hypervisor upgrade, files that are the targets of some symbolic links no longer exist. Although these missing files cause warning messages to be displayed during the installation

or uninstallation of the Virtual GPU Manager, the installation or uninstallation is completed without errors.

### Workaround

Ignore these messages as they are benign.

### Status

Open

### Ref. #

200423757

## 5.20. Benign not in gzip format messages during Virtual GPU Manager installation or uninstallation

### Description

During the installation or uninstallation of the Virtual GPU Manager, the warning messages are displayed:

```
gzip: /boot/initramfs-4.4.77-1.el6.nutanix.20170830.100726.x86_64.img: not in gzip format
gzip: /boot/initramfs-4.4.77-1.el6.nutanix.20170830.100726.x86_64.tmp: not in gzip format
```

### Workaround

Ignore these messages as they are benign.

### Status

Open

### Ref. #

200405700

## 5.21. License is not acquired in Windows VMs

### Description

When a windows VM configured with a licensed vGPU is started, the VM fails to acquire a license.

Error messages in the following format are written to the NVIDIA service logs:

```
[000000020.860152600 sec] - [Logging.lib] ERROR: [nvGridLicensing.FlexUtility]
353@FlexUtility::LogFneError : Error: Failed to add trusted storage. Server
URL : license-server-url -
[1,7E2,2,1[7000003F,0,9B00A7]]
```

```
System machine type does not match expected machine type..
```

### Workaround

This workaround requires administrator privileges.

1. Stop the **NVIDIA Display Container LS** service.
2. Delete the contents of the folder %SystemDrive%\Program Files\NVIDIA Corporation\Grid Licensing.
3. Start the **NVIDIA Display Container LS** service.

### Status

Closed

### Ref. #

200407287

## 5.22. nvidia-smi reports that vGPU migration is supported on all hypervisors

### Description

The command `nvidia-smi vgpu -m` shows that vGPU migration is supported on all hypervisors, even hypervisors or hypervisor versions that do not support vGPU migration.



## Status

Closed

## Ref. #

200407230

# 5.23. Resolution is not updated after a VM acquires a license and is restarted

## Description

In a Red Enterprise Linux 7.3 guest VM, an increase in resolution from 1024×768 to 2560×1600 is not applied after a license is acquired and the `gridd` service is restarted. This issue occurs if the `multimonitor` parameter is added to the `xorg.conf` file.

## Version

Red Enterprise Linux 7.3

## Status

Open

## Ref. #

200275925

# 5.24. A segmentation fault in Dbus code causes `nvidia-gridd` to exit on Red Hat Enterprise Linux and CentOS

## Description

On Red Hat Enterprise Linux 6.8 and 6.9, and CentOS 6.8 and 6.9, a segmentation fault in Dbus code causes the `nvidia-gridd` service to exit.

The `nvidia-gridd` service uses Dbus for communication with **NVIDIA X Server Settings** to display licensing information through the **Manage License** page. Disabling the GUI for licensing resolves this issue.

To prevent this issue, the GUI for licensing is disabled by default. You might encounter this issue if you have enabled the GUI for licensing and are using Red Hat Enterprise Linux 6.8 or 6.9, or CentOS 6.8 and 6.9.

## Version

Red Hat Enterprise Linux 6.8 and 6.9

CentOS 6.8 and 6.9

## Status

Open

## Ref. #

- ▶ 200358191
- ▶ 200319854
- ▶ 1895945

# 5.25. No Manage License option available in NVIDIA X Server Settings by default

## Description

By default, the **Manage License** option is not available in **NVIDIA X Server Settings**. This option is missing because the GUI for licensing on Linux is disabled by default to work around the issue that is described in [A segmentation fault in Dbus code causes nvidia-gridd to exit on Red Hat Enterprise Linux and CentOS](#).

## Workaround

This workaround requires `sudo` privileges.



**Note:** Do **not** use this workaround with Red Hat Enterprise Linux 6.8 and 6.9 or CentOS 6.8 and 6.9. To prevent a segmentation fault in Dbus code from causing the `nvidia-gridd` service from exiting, the GUI for licensing must be disabled with these OS versions.

If you are licensing a physical GPU for vCS, you **must** use the configuration file `/etc/nvidia/gridd.conf`.

1. If **NVIDIA X Server Settings** is running, shut it down.

2. If the `/etc/nvidia/gridd.conf` file does not already exist, create it by copying the supplied template file `/etc/nvidia/gridd.conf.template`.
3. As root, edit the `/etc/nvidia/gridd.conf` file to set the `EnableUI` option to `TRUE`.
4. Start the `nvidia-gridd` service.

```
# sudo service nvidia-gridd start
```

When **NVIDIA X Server Settings** is restarted, the **Manage License** option is now available.

## Status

Open

# 5.26. Licenses remain checked out when VMs are forcibly powered off

## Description

NVIDIA vGPU software licenses remain checked out on the license server when non-persistent VMs are forcibly powered off.

The NVIDIA service running in a VM returns checked out licenses when the VM is shut down. In environments where non-persistent licensed VMs are not cleanly shut down, licenses on the license server can become exhausted. For example, this issue can occur in automated test environments where VMs are frequently changing and are not guaranteed to be cleanly shut down. The licenses from such VMs remain checked out against their MAC address for seven days before they time out and become available to other VMs.

## Resolution

If VMs are routinely being powered off without clean shutdown in your environment, you can avoid this issue by shortening the license borrow period. To shorten the license borrow period, set the `LicenseInterval` configuration setting in your VM image. For details, refer to [Virtual GPU Client Licensing User Guide](#).

## Status

Closed

## Ref. #

1694975

## 5.27. Memory exhaustion can occur with vGPU profiles that have 512 Mbytes or less of frame buffer

### Description

Memory exhaustion can occur with vGPU profiles that have 512 Mbytes or less of frame buffer.

This issue typically occurs in the following situations:

- ▶ Full screen 1080p video content is playing in a browser. In this situation, the session hangs and session reconnection fails.
- ▶ Multiple display heads are used with Citrix Virtual Apps and Desktops or VMware Horizon on a Windows 10 guest VM.
- ▶ Higher resolution monitors are used.
- ▶ Applications that are frame-buffer intensive are used.
- ▶ NVENC is in use.

To reduce the possibility of memory exhaustion, NVENC is disabled on profiles that have 512 Mbytes or less of frame buffer.

The following vGPU profiles have 512 Mbytes or less of frame buffer:

- ▶ Tesla M6-0B, M6-0Q
- ▶ Tesla M10-0B, M10-0Q
- ▶ Tesla M60-0B, M60-0Q

The root cause is a known issue associated with changes to the way that recent Microsoft operating systems handle and allow access to overprovisioning messages and errors. If your systems are provisioned with enough frame buffer to support your use cases, you should not encounter these issues.

### Workaround

- ▶ Use an appropriately sized vGPU to ensure that the frame buffer supplied to a VM through the vGPU is adequate for your workloads.
- ▶ Monitor your frame buffer usage.
- ▶ If you are using Windows 10, consider these workarounds and solutions:
  - ▶ Use a profile that has 1 Gbyte of frame buffer.
  - ▶ Optimize your Windows 10 resource usage.

To obtain information about best practices for improved user experience using Windows 10 in virtual environments, complete the [NVIDIA GRID vGPU Profile Sizing Guide for Windows 10 download request form](#).

## Status

Open

## Ref. #

- ▶ 200130864
- ▶ 1803861

# 5.28. GNOME Display Manager (GDM) fails to start on Red Hat Enterprise Linux 7.2 and CentOS 7.0

## Description

GDM fails to start on Red Hat Enterprise Linux 7.2 and CentOS 7.0 with the following error:

```
Oh no! Something has gone wrong!
```

## Workaround

Permanently enable permissive mode for Security Enhanced Linux (SELinux).

1. As root, edit the `/etc/selinux/config` file to set `SELINUX` to `permissive`.

```
SELINUX=permissive
```

2. Reboot the system.

```
~]# reboot
```

For more information, see [Permissive Mode](#) in *Red Hat Enterprise Linux 7 SELinux User's and Administrator's Guide*.

## Status

Not an NVIDIA bug

## Ref. #

200167868

## Notice

This document is provided for information purposes only and shall not be regarded as a warranty of a certain functionality, condition, or quality of a product. NVIDIA Corporation ("NVIDIA") makes no representations or warranties, expressed or implied, as to the accuracy or completeness of the information contained in this document and assumes no responsibility for any errors contained herein. NVIDIA shall have no liability for the consequences or use of such information or for any infringement of patents or other rights of third parties that may result from its use. This document is not a commitment to develop, release, or deliver any Material (defined below), code, or functionality.

NVIDIA reserves the right to make corrections, modifications, enhancements, improvements, and any other changes to this document, at any time without notice.

Customer should obtain the latest relevant information before placing orders and should verify that such information is current and complete.

NVIDIA products are sold subject to the NVIDIA standard terms and conditions of sale supplied at the time of order acknowledgement, unless otherwise agreed in an individual sales agreement signed by authorized representatives of NVIDIA and customer ("Terms of Sale"). NVIDIA hereby expressly objects to applying any customer general terms and conditions with regards to the purchase of the NVIDIA product referenced in this document. No contractual obligations are formed either directly or indirectly by this document.

NVIDIA products are not designed, authorized, or warranted to be suitable for use in medical, military, aircraft, space, or life support equipment, nor in applications where failure or malfunction of the NVIDIA product can reasonably be expected to result in personal injury, death, or property or environmental damage. NVIDIA accepts no liability for inclusion and/or use of NVIDIA products in such equipment or applications and therefore such inclusion and/or use is at customer's own risk.

NVIDIA makes no representation or warranty that products based on this document will be suitable for any specified use. Testing of all parameters of each product is not necessarily performed by NVIDIA. It is customer's sole responsibility to evaluate and determine the applicability of any information contained in this document, ensure the product is suitable and fit for the application planned by customer, and perform the necessary testing for the application in order to avoid a default of the application or the product. Weaknesses in customer's product designs may affect the quality and reliability of the NVIDIA product and may result in additional or different conditions and/or requirements beyond those contained in this document. NVIDIA accepts no liability related to any default, damage, costs, or problem which may be based on or attributable to: (i) the use of the NVIDIA product in any manner that is contrary to this document or (ii) customer product designs.

No license, either expressed or implied, is granted under any NVIDIA patent right, copyright, or other NVIDIA intellectual property right under this document. Information published by NVIDIA regarding third-party products or services does not constitute a license from NVIDIA to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property rights of the third party, or a license from NVIDIA under the patents or other intellectual property rights of NVIDIA.

Reproduction of information in this document is permissible only if approved in advance by NVIDIA in writing, reproduced without alteration and in full compliance with all applicable export laws and regulations, and accompanied by all associated conditions, limitations, and notices.

THIS DOCUMENT AND ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, "MATERIALS") ARE BEING PROVIDED "AS IS." NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE. TO THE EXTENT NOT PROHIBITED BY LAW, IN NO EVENT WILL NVIDIA BE LIABLE FOR ANY DAMAGES, INCLUDING WITHOUT LIMITATION ANY DIRECT, INDIRECT, SPECIAL, INCIDENTAL, PUNITIVE, OR CONSEQUENTIAL DAMAGES, HOWEVER CAUSED AND REGARDLESS OF THE THEORY OF LIABILITY, ARISING OUT OF ANY USE OF THIS DOCUMENT, EVEN IF NVIDIA HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. Notwithstanding any damages that customer might incur for any reason whatsoever, NVIDIA's aggregate and cumulative liability towards customer for the products described herein shall be limited in accordance with the Terms of Sale for the product.

## VESA DisplayPort

DisplayPort and DisplayPort Compliance Logo, DisplayPort Compliance Logo for Dual-mode Sources, and DisplayPort Compliance Logo for Active Cables are trademarks owned by the Video Electronics Standards Association in the United States and other countries.

## HDMI

HDMI, the HDMI logo, and High-Definition Multimedia Interface are trademarks or registered trademarks of HDMI Licensing LLC.

## OpenCL

OpenCL is a trademark of Apple Inc. used under license to the Khronos Group Inc.

## Trademarks

NVIDIA, the NVIDIA logo, NVIDIA GRID, NVIDIA GRID vGPU, NVIDIA Maxwell, NVIDIA Pascal, NVIDIA Turing, NVIDIA Volta, Quadro, and Tesla are trademarks or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

## Copyright

© 2013-2020 NVIDIA Corporation. All rights reserved.