VIRTUAL GPU SOFTWARE R430 FOR CITRIX HYPERVERSOR

RN-06927-001_v9.0 through 9.1 Revision 02 | September 2019

Release Notes
**TABLE OF CONTENTS**

Chapter 1. Release Notes........................................................................................................ 1
1.1. Updates in Release 9.0.................................................................................................... 2
1.2. Updates in Release 9.1.................................................................................................. 2

Chapter 2. Validated Platforms............................................................................................. 3
2.1. Supported NVIDIA GPUs and Validated Server Platforms..................................... 3
2.2. Hypervisor Software Releases.................................................................................... 4
2.3. Guest OS Support........................................................................................................ 6
    2.3.1. Windows Guest OS Support.................................................................................. 7
    2.3.2. Linux Guest OS Support...................................................................................... 8
2.4. NVIDIA CUDA Toolkit Version Support.................................................................... 8

Chapter 3. Known Product Limitations.............................................................................. 10
3.1. Issues may occur with graphics-intensive OpenCL applications on vGPU types with limited
    frame buffer.................................................................................................................. 10
3.2. In pass through mode, all GPUs connected to each other through NVLink must be assigned
    to the same VM.............................................................................................................. 11
3.3. vGPU profiles with 512 Mbytes or less of frame buffer support only 1 virtual display head
    on Windows 10.............................................................................................................. 11
3.4. NVENC requires at least 1 Gbyte of frame buffer...................................................... 12
3.5. VM running older NVIDIA vGPU drivers fails to initialize vGPU when booted............. 12
3.6. Single vGPU benchmark scores are lower than pass-through GPU......................... 13
3.7. nvidia-smi fails to operate when all GPUs are assigned to GPU pass-through mode...... 14
3.8. Windows Aero is disabled on Citrix Virtual Apps and Desktops session using 3 or 4 monitors
    in 2560×1600 resolution.............................................................................................. 15
3.9. VMs configured with large memory fail to initialize vGPU when booted................... 16
3.10. vGPU host driver RPM upgrade fails....................................................................... 17

Chapter 4. Resolved Issues................................................................................................. 18

Chapter 5. Known Issues.................................................................................................... 20
5.1. Migrating a VM configured with NVIDIA vGPU software release 9.1 to a host running
    release 9.0 fails............................................................................................................. 20
5.2. ECC memory with NVIDIA vGPU is not supported on Tesla M60 and Tesla M6......... 21
5.3. Virtual GPU fails to start if ECC is enabled............................................................... 21
5.4. RAPIDS cuDF merge fails on NVIDIA vGPU............................................................. 23
5.5. 9.0 only: Users’ view sessions may become corrupted after migration.................... 24
5.6. Migration of VMs configured with vGPU stops before the migration is complete....... 24
5.7. 9.0 only: Incorrect ECC error counts are reported for vGPUs on some GPUs........... 25
5.8. ECC memory settings for a vGPU cannot be changed by using NVIDIA X Server Settings... 25
5.9. Changes to ECC memory settings for a Linux vGPU VM by nvidia-smi might be ignored..... 26
5.10. 9.0 only: VM crashes after the volatile ECC error count is reset.............................. 27
5.11. 9.0 only: On Linux VMs, the license directory is not deleted when the guest driver is
    uninstalled..................................................................................................................... 27
5.12. Incorrect GPU type shown for Quadro RTX 8000 GPUs in Citrix XenCenter................... 28
5.13. NVIDIA Notification Icon prevents log off of Citrix Published Application user sessions... 28
5.14. Vulkan applications crash in Windows 7 guest VMs configured with NVIDIA vGPU........ 29
5.15. Host core CPU utilization is higher than expected for moderate workloads.............. 29
5.16. Frame capture while the interactive logon message is displayed returns blank screen... 30
5.17. RDS sessions do not use the GPU with some Microsoft Windows Server releases......... 30
5.18. Even when the scheduling policy is equal share, unequal GPU utilization is reported.... 31
5.19. When the scheduling policy is fixed share, GPU utilization is reported as higher than
expected.............................................................................................................. 32
5.20. nvidia-smi reports that vGPU migration is supported on all hypervisors................. 33
5.21. NVIDIA Control Panel Crashes in a VM connected to two 4K displays..................... 33
5.22. vGPU guest VM driver not properly loaded on servers with more than 512 GB or 1 TB
or more of system memory................................................................................... 34
5.23. Luxmark causes a segmentation fault on an unlicensed Linux client....................... 35
5.24. Resolution is not updated after a VM acquires a license and is restarted................. 35
5.25. A segmentation fault in DBus code causes nvidia-gridd to exit on Red Hat Enterprise
Linux and CentOS.............................................................................................. 36
5.26. No Manage License option available in NVIDIA X Server Settings by default............ 36
5.27. Licenses remain checked out when VMs are forcibly powered off.......................... 37
5.28. Memory exhaustion can occur with vGPU profiles that have 512 Mbytes or less of frame
buffer.................................................................................................................... 38
5.29. VM bug checks after the guest VM driver for Windows 10 RS2 is installed................. 39
5.30. On Citrix Hypervisor 7.0, VMs unexpectedly reboot and Citrix Hypervisor crashes or
freezes..................................................................................................................... 40
5.31. With no NVIDIA driver installed, Citrix Hypervisor misidentifies Tesla M10 cards...... 41
5.32. GNOME Display Manager (GDM) fails to start on Red Hat Enterprise Linux 7.2 and CentOS
7.0............................................................................................................................ 41
5.33. Video goes blank when run in loop in Windows Media Player.................................. 42
5.34. Local VGA console is momentarily unblanked when Citrix Virtual Apps and Desktops
changes resolution of the VM desktop.................................................................. 43
5.35. VM bugchecks on shutdown/restart when Citrix Virtual Apps and Desktops is installed
and NVIDIA driver is uninstalled or upgraded....................................................... 43
5.36. Application frame rate may drop when running Citrix Virtual Apps and Desktops at
2560x1600 resolution.......................................................................................... 44
5.37. Windows VM BSOD...................................................................................... 44
5.38. Windows VM BSOD when upgrading NVIDIA drivers over a Citrix Virtual Apps and Desktops
session.................................................................................................................. 45
5.39. XenCenter does not allow vGPUs to be selected as a GPU type for Linux VMs........... 46
5.40. If X server is killed on a RHEL7 VM running vGPU, XenCenter console may not
automatically switch to text console..................................................................... 47
5.41. Citrix Virtual Apps and Desktops shows only a black screen when connected to a vGPU VM47
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 Citrix Hypervisor.

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:

<table>
<thead>
<tr>
<th>Software</th>
<th>9.0</th>
<th>9.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVIDIA Virtual GPU Manager for the Citrix Hypervisor releases listed in Hypervisor Software Releases</td>
<td>430.27</td>
<td>430.46</td>
</tr>
<tr>
<td>NVIDIA Windows driver</td>
<td>431.02</td>
<td>431.79</td>
</tr>
<tr>
<td>NVIDIA Linux driver</td>
<td>430.30</td>
<td>430.46</td>
</tr>
</tbody>
</table>

Caution

If you install the wrong NVIDIA vGPU software packages for the version of Citrix Hypervisor 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

- 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 Windows 10 May 2019 Update (1903) as a guest OS
- Support for Ubuntu 18.04 LTS as a guest OS on Citrix Hypervisor 7.1 cumulative update 2

1.2. Updates in Release 9.1

New Features in Release 9.1

- Security updates
- Miscellaneous bug fixes

Hardware and Software Support Introduced in Release 9.1

- Support for the following versions of Citrix Virtual Apps and Desktops:
  - Version 7 1906
  - Version 7 1909
This release family of NVIDIA vGPU software provides support for several NVIDIA GPUs on validated server hardware platforms, Citrix Hypervisor 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 Citrix Hypervisor, running on validated server hardware platforms:

- GPUs based on the NVIDIA Maxwell™ graphic architecture:
  - Tesla M6
  - Tesla M10
  - Tesla M60
- GPUs based on the NVIDIA Pascal™ architecture:
  - Tesla P4
  - Tesla P6
  - Tesla P40
  - Tesla P100 PCIe 16 GB (XenMotion with vGPU is not supported.)
  - Tesla P100 SXM2 16 GB (XenMotion with vGPU is not supported.)
  - Tesla P100 PCIe 12GB (XenMotion with vGPU is not supported.)
- GPUs based on the NVIDIA Volta™ architecture:
  - Tesla V100 SXM2
  - Tesla V100 SXM2 32GB
  - Tesla V100 PCIe
  - Tesla V100 PCIe 32GB
Validated Platforms

- Tesla V100 FHHL
- GPUs based on the NVIDIA Turing™ architecture:
  - Tesla T4
  - Quadro RTX 6000 in displayless mode (GRID Virtual PC and GRID Virtual Applications are not supported.)
  - Quadro RTX 8000 in displayless mode (GRID Virtual PC and GRID Virtual Applications are not supported.)

In displayless mode, local physical display connectors are disabled.

For a list of validated server platforms, refer to NVIDIA GRID Certified Servers.

---

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 vComputeServer vGPU types. Furthermore, vComputeServer is not supported on any GPU on Citrix Hypervisor.

---

2.2. Hypervisor Software Releases

Changes to Citrix Product Names and Release Numbers

Citrix has changed the product names of its hypervisor and virtual desktop software.

<table>
<thead>
<tr>
<th>Old Name</th>
<th>New Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>XenServer</td>
<td>Citrix Hypervisor</td>
</tr>
<tr>
<td>XenApp and XenDesktop</td>
<td>Citrix Virtual Apps and Desktops</td>
</tr>
<tr>
<td>Citrix Receiver</td>
<td>Citrix Workspace App</td>
</tr>
</tbody>
</table>

The release number of Citrix Virtual Apps and Desktops releases after 7.18 is in the form 7 yyymm, where yyymm denotes the year and month that the product was released. For example, for a product released in August 2018, yyymm is 1808

For more information, see these topics in the Citrix documentation:

- New product names
New product and component version numbers

Supported Citrix Hypervisor Releases

This release family of NVIDIA vGPU software is supported on the Citrix Hypervisor releases listed in the table.

Support for NVIDIA vGPU software requires the Premium Edition (previously Enterprise Edition) of Citrix Hypervisor. For details, see Licensing in the Citrix documentation. Cumulative update releases for a base release of Citrix Hypervisor are compatible with the base release and can also be used with this version of NVIDIA vGPU software unless expressly stated otherwise.

<table>
<thead>
<tr>
<th>Software</th>
<th>Releases Supported</th>
<th>Notes</th>
</tr>
</thead>
</table>
| Citrix Hypervisor 8.0           | RTM build and compatible cumulative update releases | In NVIDIA vGPU mode, all NVIDIA GPUs that support NVIDIA vGPU software are supported.  
vComputeServer is not supported.  
In GPU pass-through mode, the following GPUs are not supported:  
  ▪ Quadro RTX 6000  
  ▪ Quadro RTX 8000  
  This release supports XenMotion with vGPU on suitable GPUs as listed in Supported NVIDIA GPUs and Validated Server Platforms. |
| Citrix Hypervisor (XenServer) 7.6 | RTM build and compatible cumulative update releases | All NVIDIA GPUs that support NVIDIA vGPU software are supported.  
vComputeServer is not supported.  
This release supports XenMotion with vGPU on suitable GPUs as listed in Supported NVIDIA GPUs and Validated Server Platforms. |
Validated Platforms

**Software Releases Supported**

<table>
<thead>
<tr>
<th>Software</th>
<th>Releases Supported</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citrix Hypervisor (XenServer) 7.1</td>
<td>RTM build and compatible cumulative update releases</td>
<td>Not supported on Tesla T4. vComputeServer is <strong>not</strong> supported. XenMotion with vGPU is <strong>not</strong> supported.</td>
</tr>
<tr>
<td>Citrix Hypervisor (XenServer) 7.0</td>
<td>RTM build 125380 and compatible cumulative update releases</td>
<td>Not supported on Tesla T4. vComputeServer is <strong>not</strong> supported. XenMotion with vGPU is <strong>not</strong> supported.</td>
</tr>
</tbody>
</table>

**Supported Virtual Desktop Software Releases**

This release supports **only** the virtual desktop software releases listed in the table. Except where otherwise stated, HDX 3D Pro mode is supported but not required.

<table>
<thead>
<tr>
<th>Software</th>
<th>Releases Supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citrix Virtual Apps and Desktops</td>
<td>Since 9.1: Versions 7 1909, 7 1906, 7 1903, and 7 1808 9.0 only: Versions 7 1903 and 7 1808</td>
</tr>
<tr>
<td>XenApp/XenDesktop</td>
<td>Versions 7.18 and 7.17 7.15 in HDX 3D Pro mode (required to ensure that the NVIDIA GPU can be used unrestricted)</td>
</tr>
</tbody>
</table>

### 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.

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.

In GPU pass-through mode, the following GPUs are **not** supported on Citrix Hypervisor 8.0:
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 Citrix Hypervisor. The releases of Citrix Hypervisor for which a Windows release is supported depend on whether NVIDIA vGPU or pass-through GPU is used.

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

XenMotion with vGPU is supported on supported Windows guest OS releases.

<table>
<thead>
<tr>
<th>Guest OS</th>
<th>NVIDIA vGPU - Citrix Hypervisor Releases</th>
<th>Pass-Through GPU - Citrix Hypervisor Releases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows Server 2019</td>
<td>8.0, 7.1 cumulative update 2</td>
<td>8.0, 7.1 cumulative update 2</td>
</tr>
<tr>
<td>Windows Server 2016 1709, 1607</td>
<td>8.0, 7.6, 7.1, 7.0</td>
<td>8.0, 7.6, 7.1, 7.0</td>
</tr>
<tr>
<td>Windows Server 2012 R2</td>
<td>8.0, 7.6, 7.1, 7.0</td>
<td>8.0, 7.6, 7.1, 7.0</td>
</tr>
<tr>
<td>Windows Server 2008 R2</td>
<td>8.0, 7.6, 7.1, 7.0</td>
<td>8.0, 7.6, 7.1, 7.0</td>
</tr>
<tr>
<td>Windows 10:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‣ May 2019 Update (1903)</td>
<td>8.0, 7.6, 7.1, 7.0</td>
<td>8.0, 7.6, 7.1, 7.0</td>
</tr>
<tr>
<td>‣ October 2018 Update (1809)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‣ Spring Creators Update (1803)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‣ Fall Creators Update (1709)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‣ Creators Update (1703)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‣ Anniversary Update (1607)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‣ November Update (1511)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‣ RTM (1507)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Windows 8.1 Update</td>
<td>8.0, 7.6, 7.1, 7.0</td>
<td>8.0, 7.6, 7.1, 7.0</td>
</tr>
<tr>
<td>Windows 8.1</td>
<td>8.0, 7.6, 7.1, 7.0</td>
<td>-</td>
</tr>
<tr>
<td>Windows 8</td>
<td>8.0, 7.6, 7.1, 7.0</td>
<td>-</td>
</tr>
<tr>
<td>Windows 7</td>
<td>8.0, 7.6, 7.1, 7.0</td>
<td>8.0, 7.6, 7.1, 7.0</td>
</tr>
</tbody>
</table>

Supported only on GPUs based on the Maxwell architecture.
2.3.2. Linux Guest OS Support

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

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

XenMotion with vGPU is not supported on any Linux guest OS release.

<table>
<thead>
<tr>
<th>Guest OS</th>
<th>NVIDIA vGPU - Citrix Hypervisor Releases</th>
<th>Pass-Through GPU - Citrix Hypervisor Releases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red Hat Enterprise Linux 7.0-7.6</td>
<td>8.0, 7.6, 7.1, 7.0</td>
<td>8.0, 7.6, 7.1, 7.0</td>
</tr>
<tr>
<td>CentOS 7.0-7.6</td>
<td>8.0, 7.6, 7.1, 7.0</td>
<td>8.0, 7.6, 7.1, 7.0</td>
</tr>
<tr>
<td>Ubuntu 18.04 LTS</td>
<td>8.0, 7.6, 7.1 cumulative update 2</td>
<td>8.0, 7.6, 7.1 cumulative update 2</td>
</tr>
<tr>
<td>Ubuntu 16.04 LTS</td>
<td>8.0, 7.6, 7.1, 7.0</td>
<td>8.0, 7.6, 7.1, 7.0</td>
</tr>
<tr>
<td>Ubuntu 14.04 LTS</td>
<td>8.0, 7.6, 7.1, 7.0</td>
<td>8.0, 7.6, 7.1, 7.0</td>
</tr>
</tbody>
</table>

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.

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.

2.5. XenMotion with vGPU Support

XenMotion with vGPU is supported only on a subset of supported GPUs, Citrix Hypervisor releases, and guest operating systems.

Supported GPUs:
- Tesla M6
- Tesla M10
- Tesla M60
- Tesla P4
- Tesla P6
- Tesla P40
- Tesla V100 SXM2
- Tesla V100 SXM2 32GB
- Tesla V100 PCIe
- Tesla V100 PCIe 32GB
- Tesla V100 FHHL
- Tesla T4
- Quadro RTX 6000
- Quadro RTX 8000

Supported Citrix Hypervisor releases: 8.0 and 7.6 only.

Supported guest OS releases: Windows only. XenMotion with vGPU is **not** supported on 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 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.
3.2. 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.

3.3. 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.
3.4. 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.

3.5. 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 Citrix Hypervisor 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 Citrix Hypervisor VM’s `/var/log/messages` log file reports one of the following errors:

- An error message:
Known Product Limitations

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.

3.6. 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 specifying frame_rate_limiter=0 in the VM's platform:vgpu_extra_args parameter:

```
[root@xenserver ~]# xe vm-param-set uuid=e71afda4-53f4-3a1b-6c92-a364a7f619c2 platform:vgpu_extra_args="frame_rate_limiter=0"
[root@xenserver ~]#
```

The setting takes effect the next time the VM is started or rebooted.

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 in one of the following ways:
Removing the `vgpu_extra_args` key from the `platform` parameter

Removing `frame_rate_limiter=0` from the `vgpu_extra_args` key

Setting `frame_rate_limiter=1`. For example:

```
[root@xenserver ~]# xe vm-param-set uuid=e71afda4-53f4-3a1b-6c92-a364a7f619c2 platform:vgpu_extra_args="frame_rate_limiter=1"
[root@xenserver ~]#
```

3.7. **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@xenserver-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 Citrix Hypervisor dom0.

To confirm that all GPUs are operating in pass-through mode, use XenCenter’s GPU tab to review current GPU assignment:
3.8. Windows Aero is disabled on Citrix Virtual Apps and Desktops session using 3 or 4 monitors in 2560×1600 resolution

Description

Windows Aero may be disabled when Citrix Virtual Apps and Desktops is connected to a VM with a vGPU or passthrough GPU, with 3 or 4 monitors at 2560×1600 resolution.

This limitation is a limitation of Windows 7. For details, see the Microsoft knowledge base article Desktop background disappears with very large extended desktop on Windows 7.
3.9. VMs configured with large memory fail to initialize vGPU when booted

Description

When starting multiple VMs configured with large amounts of RAM (typically more than 32GB per VM), a VM may fail to initialize vGPU. In this scenario, the VM boots in standard VGA mode with reduced resolution and color depth. The NVIDIA vGPU software 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)

The Citrix Hypervisor VM’s /var/log/messages log file contains these error messages:

```plaintext
vmiop_log: error: NVOS status 0x29
vmiop_log: error: Assertion Failed at 0x7620fd4b:179
vmiop_log: error: 8 frames returned by backtrace
  ...
vmiop_log: error: VGPU message 12 failed, result code: 0x29
  ...
vmiop_log: error: NVOS status 0x8
vmiop_log: error: Assertion Failed at 0x7620c8df:280
vmiop_log: error: 8 frames returned by backtrace
  ...
vmiop_log: error: VGPU message 26 failed, result code: 0x8
```

Resolution

vGPU reserves a portion of the VM’s framebuffer for use in GPU mapping of VM system memory. The reservation is sufficient to support up to 32GB of system memory, and may be increased to accommodate up to 64GB by specifying enable_large_sys_mem=1 in the VM’s platform:vgpu_extra_args parameter:

```
[root@xenserver ~]# xe vm-param-set uuid=e71afda4-53f4-3a1b-6c92-a364a7f619c2
  platform:vgpu_extra_args="enable_large_sys_mem=1"
```

The setting takes effect the next time the VM is started or rebooted. With this setting in place, less GPU FB is available to applications running in the VM. To accommodate system memory larger than 64GB, the reservation can be further increased by specifying extra_fb_reservation in the VM’s platform:vgpu_extra_args parameter, and setting its value to the desired reservation size in megabytes. The default value of 64M is sufficient to support 64GB of RAM. We recommend adding 2M of reservation for each additional 1GB of system memory. For example, to support 96GB of RAM, set extra_fb_reservation to 128:

```
platform:vgpu_extra_args="enable_large_sys_mem=1, extra_fb_reservation=128"
```

The reservation can be reverted back to its default setting in one of the following ways:
Removing the `vgpu_extra_args` key from the `platform` parameter
Removing `enable_large_sys_mem` from the `vgpu_extra_args` key
Setting `enable_large_sys_mem=0`

3.10. vGPU host driver RPM upgrade fails

Description
Upgrading vGPU host driver RPM fails with an error message about failed dependencies on the console.

```
[root@xenserver ~]# rpm -U NVIDIA-vGPU-xenserver-6.5-352.46.x86_64.rpm
error: Failed dependencies:
NVIDIA-vgx-xenserver conflicts with NVIDIA-vGPU-xenserver-6.5-352.46.x86_64
```

Resolution
Uninstall the older vGPU RPM before installing the latest driver.

Use the following command to uninstall the older vGPU RPM:

```
[root@xenserver ~]# rpm -e NVIDIA-vgx-xenserver
```
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

<table>
<thead>
<tr>
<th>Bug ID</th>
<th>Summary and Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Virtual GPU fails to start if ECC is enabled</td>
</tr>
<tr>
<td></td>
<td>NVIDIA vGPU does not support error correcting code (ECC) memory. If ECC memory is enabled, NVIDIA vGPU fails to start.</td>
</tr>
<tr>
<td></td>
<td>Starting with NVIDIA vGPU software release 9.0, NVIDIA vGPU supports ECC memory on GPUs and hypervisor software versions that support ECC.</td>
</tr>
<tr>
<td>200269717</td>
<td>On Tesla P40, P6, and P4 GPUs, the default ECC setting prevents NVIDIA vGPU from starting</td>
</tr>
<tr>
<td></td>
<td>Starting with NVIDIA vGPU software release 9.0, NVIDIA vGPU supports ECC memory on GPUs and hypervisor software versions that support ECC.</td>
</tr>
<tr>
<td>2285306</td>
<td>Cloned VMs configured with a vGPU type different than the type in the master image fail to start</td>
</tr>
<tr>
<td></td>
<td>Cloned VMs configured with a vGPU type different than the type in the master image fail to start.</td>
</tr>
<tr>
<td></td>
<td>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>.</td>
</tr>
</tbody>
</table>
## Issues Resolved in Release 9.1

<table>
<thead>
<tr>
<th>Bug ID</th>
<th>Summary and Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>200534988</td>
<td>Error XID 47 followed by multiple XID 32 errors</td>
</tr>
<tr>
<td></td>
<td>After disconnecting Citrix Virtual Apps and Desksops 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.</td>
</tr>
<tr>
<td>200526633</td>
<td>9.0 only: VM crashes after the volatile ECC error count is reset</td>
</tr>
<tr>
<td></td>
<td>After the command <code>nvidia-smi -p 0</code> is run from a guest VM to reset the volatile ECC error count, the VM crashes.</td>
</tr>
<tr>
<td>200525006</td>
<td>9.0 only: Incorrect ECC error counts are reported for vGPUs on some GPUs</td>
</tr>
<tr>
<td></td>
<td>Incorrect ECC error counts are reported for vGPUs on some GPUs when the command <code>nvidia-smi -q</code> is run from a guest VM.</td>
</tr>
<tr>
<td>200524555</td>
<td>9.0 only: On Linux VMs, the license directory is not deleted when the guest driver is uninstalled</td>
</tr>
<tr>
<td></td>
<td>On Linux guest VMs, the license directory <code>/etc/nvidia/license</code> is not deleted when the NVIDIA vGPU software graphics driver is uninstalled.</td>
</tr>
<tr>
<td>200434909</td>
<td>9.0 only: Users' view sessions may become corrupted after migration</td>
</tr>
<tr>
<td></td>
<td>When a VM configured with vGPU under heavy load is migrated to another host, users' view sessions may become corrupted after the migration.</td>
</tr>
</tbody>
</table>
5.1. Migrating a VM configured with NVIDIA vGPU software release 9.1 to a host running release 9.0 fails

Description

If a VM configured with NVIDIA vGPU software 9.1 is migrated to a host running NVIDIA vGPU software 9.0, the migration fails and the VM shuts down.

This issue occurs only with the following combination of releases of guest VM graphics driver, vGPU manager on the source host, and vGPU manager on the destination host:

<table>
<thead>
<tr>
<th>Guest VM Graphics Driver</th>
<th>Source vGPU Manager</th>
<th>Destination vGPU Manager</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.1</td>
<td>9.1</td>
<td>9.0</td>
</tr>
</tbody>
</table>

Workaround

Tesla M10 GPUs do not support this workaround. Even after applying this workaround to a system on which this issue occurs, vGPU migration with Tesla M10 GPUs fails with the following error:

```
Unexpected migration data block encountered.
```

1. On the host that is running vGPU Manager 9.1, set the registry key `RMSetVGPUVersionMax` to `0x30001`.
2. Start the VM.
3. Confirm that the vGPU version in the log files is `0x30001`.

```
2019-09-10T10:19:05.420Z| vthread-2142280| I125: vmiop_log: vGPU version: 0x30001
```
The VM can now be migrated.

**Status**

Not a bug

**Ref. #**

200533827

### 5.2. 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.
- **Since 9.1:** 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**

Open

### 5.3. 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.

However, NVIDIA vGPU does not support ECC memory with the following GPUs:

- Tesla M60 GPUs
Tesla M6 GPUs

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

Citrix XenCenter displays the following error message:

An emulator required to run this VM failed to start

The following error is logged in the Citrix Hypervisor host’s /var/log/messages log file:

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

Resolution

If you are using Tesla M60 or Tesla M6 GPUs, 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.

```bash
# shutdown -r now
```

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

```bash
# nvidia-smi -q
```

```
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:
  ```bash
  # nvidia-smi -e 1
  ```

- If you want to change the ECC status to on for a specific GPU, run this command:
  ```bash
  # 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.
  ```

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

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

### 5.4. 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
5.5. 9.0 only: Users' view sessions may become corrupted after migration

Description
When a VM configured with vGPU under heavy load is migrated to another host, users' view sessions may become corrupted after the migration.

Workaround
Restart the VM.

Status
Resolved in NVIDIA vGPU software 9.1

Ref. #
200434909

5.6. Migration of VMs configured with vGPU stops before the migration is complete

Description
When a VM configured with vGPU is migrated to another host, the migration stops before it is complete. After the migration stops, the VM is no longer accessible.

This issue occurs if the ECC memory configuration (enabled or disabled) on the source and destination hosts are different. The ECC memory configuration on both the source and destination hosts must be identical.

Workaround
Reboot the hypervisor host to recover the VM. Before attempting to migrate the VM again, ensure that the ECC memory configuration on both the source and destination hosts are identical.

Status
Not an NVIDIA bug
5.7. 9.0 only: Incorrect ECC error counts are reported for vGPUs on some GPUs

Description
Incorrect ECC error counts are reported for vGPUs on some GPUs when the command `nvidia-smi -q` is run from a guest VM.

This issue affects only vGPUs that reside on physical GPUs based on the NVIDIA Volta GPU architecture. For vGPUs on GPUs based on other architectures, the ECC error count is correct.

Status
Resolved in NVIDIA vGPU software 9.1

5.8. ECC memory settings for a vGPU cannot be changed by using NVIDIA X Server Settings

Description
The ECC memory settings for a vGPU cannot be changed from a Linux guest VM by using NVIDIA X Server Settings. After the ECC memory state has been changed on the ECC Settings page and the VM has been rebooted, the ECC memory state remains unchanged.

Workaround
Use the `nvidia-smi` command in the guest VM to enable or disable ECC memory for the vGPU as explained in Virtual GPU Software User Guide.

If the ECC memory state remains unchanged even after you use the `nvidia-smi` command to change it, use the workaround in Changes to ECC memory settings for a Linux vGPU VM by nvidia-smi might be ignored.
5.9. Changes to ECC memory settings for a Linux vGPU VM by `nvidia-smi` might be ignored

**Description**

After the ECC memory state for a Linux vGPU VM has been changed by using the `nvidia-smi` command and the VM has been rebooted, the ECC memory state might remain unchanged.

This issue occurs when multiple NVIDIA configuration files in the system cause the kernel module option for setting the ECC memory state `RMGuestECCState` in `/etc/modprobe.d/nvidia.conf` to be ignored.

When the `nvidia-smi` command is used to enable ECC memory, the file `/etc/modprobe.d/nvidia.conf` is created or updated to set the kernel module option `RMGuestECCState`. Another configuration file in `/etc/modprobe.d/` that contains the keyword `NVreg_RegistryDwordsPerDevice` might cause the kernel module option `RMGuestECCState` to be ignored.

**Workaround**

This workaround requires administrator privileges.

1. Move the entry containing the keyword `NVreg_RegistryDwordsPerDevice` from the other configuration file to `/etc/modprobe.d/nvidia.conf`.
2. Reboot the VM.

**Status**

Open

**Ref. #**

200505777
5.10. 9.0 only: VM crashes after the volatile ECC error count is reset

Description
After the command `nvidia-smi -p 0` is run from a guest VM to reset the volatile ECC error count, the VM crashes.

This issue does not occur if the EEC state in the VM is set to off.

Status
Resolved in NVIDIA vGPU software 9.1

Ref. #
200526633

5.11. 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
5.12. Incorrect GPU type shown for Quadro RTX 8000 GPUs in Citrix XenCenter

Description
On the GPU tab in Citrix XenCenter, Quadro RTX 8000 GPUs are incorrectly shown as [QUADRO RTX 6000].

However, the virtual GPU types for a Quadro RTX 8000 GPU are listed correctly. NVIDIA vGPU software does not support systems with a mixture of GPU types. Therefore, you can assume that any GPU in a system that you know to contain Quadro RTX 8000 GPUs that is shown as [QUADRO RTX 6000] is, in fact, a Quadro RTX 8000 GPU.

Status
Not an NVIDIA bug

Ref. #
200492607

5.13. NVIDIA Notification Icon prevents log off of Citrix Published Application user sessions

Description
By default on Windows Server operating systems, the NVIDIA Notification Icon application is started with every Citrix Published Application user session. This application might prevent the Citrix Published Application user session from being logged off even after the user has quit all other applications.

Resolution
Disable the NVIDIA Notification Icon application for Citrix Published Application user sessions as explained in Virtual GPU Software User Guide.

Status
Resolved by the provision of Windows registry keys for disabling the NVIDIA Notification Icon application for Citrix Published Application user sessions.
5.14. 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. #
2206368

5.15. 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
5.16. 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
See Black Screen at Logon with Nvidia HDX 3D Pro enabled in the Citrix knowledge base.

Status
Not a bug

Ref. #
2115733

5.17. 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.

2. Set the **Use the hardware default graphics adapter for all Remote Desktop Services sessions** option.

### 5.18. 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

### 5.19. 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.
5.20. `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

5.21. NVIDIA Control Panel Crashes in a VM connected to two 4K displays

Description
When NVIDIA Control Panel is started in a VM connected through a pass-through GPU to two 4K displays, a Citrix HDX 3D Pro Warning pop-up window opens containing the following error message:

"Full screen applications are not supported on Citrix HDX 3D Pro. Press Ok to exit"

NVIDIA Control Panel then crashes.
After the VM is rebooted, NVIDIA Control Panel is available again.

Version
Citrix Virtual Apps and Desktops 7.16
**Status**

Not an NVIDIA bug

**Ref. #**

200393766

### 5.22. vGPU guest VM driver not properly loaded on servers with more than 512 GB or 1 TB or more of system memory

**Description**

If PV IOMMU is enabled, support for vGPU is limited to servers with a maximum of 512 GB of system memory. On servers with more than 512 GB of system memory and PV IOMMU enabled, the guest VM driver is not properly loaded. **Device Manager** marks the vGPU with a yellow exclamation point.

If PV IOMMU is disabled, support for vGPU is limited to servers with less than 1 TB of system memory. This limitation applies only to systems with supported GPUs based on the Maxwell architecture: Tesla M6, Tesla M10, and Tesla M60. On servers with 1 TB or more of system memory, VMs configured with vGPU fail to power on. However, support for GPU pass through is not affected by this limitation.

**Resolution**

1. If PV IOMMU is enabled, disable it.

   ```bash
   [root@xenserver ~]# /opt/xensource/libexec/xen-cmdline --set-xen iommu=dom0-passthrough
   ```

2. If the server has 1 TB or more of system memory, limit the amount of system memory on the server to 1 TB minus 16 GB.

   If the server has less than 1 TB or more of system memory, omit this step.

3. Reboot the server.

**Status**

Not an NVIDIA bug

**Ref. #**

1799582
5.23. Luxmark causes a segmentation fault on an unlicensed Linux client

Description
If the Luxmark application is run on a Linux guest VM configured with NVIDIA vGPU that is booted without acquiring a license, a segmentation fault occurs and the application core dumps. The fault occurs when the application cannot allocate a CUDA object on NVIDIA vGPUs where CUDA is disabled. On NVIDIA vGPUs that can support CUDA, CUDA is disabled in unlicensed mode.

Status
Not an NVIDIA bug.

Ref. #
200330956

5.24. 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
5.25. 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

5.26. 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.

*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 vComputeServer, 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

---

**5.27. 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

5.28. 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.

When memory exhaustion occurs, the NVIDIA host driver reports Xid error 31 and Xid error 43 in the Citrix Hypervisor /var/log/messages file.

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.

For more information, see also Windows 10 Optimization for XenDesktop on the Citrix blog.

**Status**

Open

**Ref. #**

- 200130864
- 1803861

### 5.29. VM bug checks after the guest VM driver for Windows 10 RS2 is installed

**Description**

When the VM is rebooted after the guest VM driver for Windows 10 RS2 is installed, the VM bug checks. When Windows boots, it selects one of the standard supported video modes. If Windows is booted directly with a display that is driven by an NVIDIA driver, for example a vGPU on Citrix Hypervisor, a blue screen crash occurs.

This issue occurs when the screen resolution is switched from VGA mode to a resolution that is higher than 1920×1200.
Fix
Download and install Microsoft Windows Update KB4020102 from the Microsoft Update Catalog.

Workaround
If you have applied the fix, ignore this workaround.
Otherwise, you can work around this issue until you are able to apply the fix by not using resolutions higher than 1920x1200.

1. Choose a GPU profile in Citrix XenCenter that does not allow resolutions higher than 1920x1200.
2. Before rebooting the VM, set the display resolution to 1920x1200 or lower.

Status
Not an NVIDIA bug

Ref. #
200310861

5.30. On Citrix Hypervisor 7.0, VMs unexpectedly reboot and Citrix Hypervisor crashes or freezes

Description
On Citrix Hypervisor 7.0, VMs to which a vGPU is attached unexpectedly reboot and Citrix Hypervisor crashes or freezes.

The event log in the Citrix Hypervisor /var/log/crash/xen.log file lists the following errors:

- A fatal bus error on a component at the slot where the GPU card is installed
- A fatal error on a component at bus 0, device 2, function 0

This issue occurs when page-modification logging (PML) is enabled on Intel Broadwell CPUs running Citrix Hypervisor 7.0. Citrix is aware of this issue and is working on a permanent fix.

Workaround
Disable page-modification logging (PML) as explained in XenServer 7 host crash while starting multiple virtual machines in the Citrix Support Knowledge Center.
### 5.31. With no NVIDIA driver installed, Citrix Hypervisor misidentifies Tesla M10 cards

**Description**
An erroneous entry in the `pci.ids` database causes Citrix Hypervisor to identify Tesla M10 cards as GRID M40 when no NVIDIA driver is installed.

**Version**
Citrix Hypervisor 6.5 and 7.0

**Workaround**
None

**Status**
Not an NVIDIA bug

**Ref. #**
1853248

### 5.32. 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

### 5.33. Video goes blank when run in loop in Windows Media Player

**Description**

When connected to a vGPU-enabled VM using Citrix Virtual Apps and Desktops, a video played back in looping mode on Windows Media Player goes blank or freezes after a few iterations.

**Workaround**

None

**Status**

Not an NVIDIA bug

**Ref. #**

1306623
5.34. Local VGA console is momentarily unblanked when Citrix Virtual Apps and Desktops changes resolution of the VM desktop

Description
When Citrix Virtual Apps and Desktops establishes a remote connection to a VM using vGPU, the VM's local VGA console display in XenCenter is blanked (assuming the VM local console has not been disabled by setting platform:vgpu_extra_args="disable_vnc=1"). If the Citrix Virtual Apps and Desktops session changes resolution of the VM's desktop, the local VGA console momentarily unblanks, allowing a XenCenter user to briefly view the desktop.

Workaround
Disable the VM's local VGA console

```
xe vm-param-set uuid=vm-uuid platform:vgpu_extra_args="disable_vnc=1"
```

Status
Open

Ref. #
NVIDIA-145/1375164

5.35. VM bugchecks on shutdown/restart when Citrix Virtual Apps and Desktops is installed and NVIDIA driver is uninstalled or upgraded.

Description
If the Citrix Virtual Apps and Desktops agent is installed in a VM before any NVIDIA GPU driver is installed, the VM will bugcheck (bluescreen) when the NVIDIA driver is subsequently upgraded or uninstalled. The bugcheck code is 0x7E, SYSTEM_THREAD_EXCEPTION_NOT_HANDLED.

Workaround
Use one of the following workarounds:
- Do a force shutdown of the VM and restart it.
- Install the NVIDIA driver in guest VMs before installing Citrix Virtual Apps and Desktops.

**Status**

Open

**Ref. #**

NVIDIA-295/200018125

---

**5.36. Application frame rate may drop when running Citrix Virtual Apps and Desktops at 2560×1600 resolution.**

**Description**

An application’s rendering frame rate may drop when running Citrix Virtual Apps and Desktops at 2560×1600 resolution, relative to the frame rate obtained at lower resolutions.

**Fix**

Using the Windows `regedit` utility within the VM, open the `HKLM\SOFTWARE\Citrix\Graphics` registry key and create a new `DWORD` value, `EncodeSpeed`, with a value of 2. Reboot the VM. This setting may improve the delivered frame rate at the expense of a reduction in image quality.

**Status**

Open

**Ref. #**

NVIDIA-190/1416336

---

**5.37. Windows VM BSOD**

**Description**

Windows VM bugchecks on Citrix Hypervisor when running a large number of vGPU based VMs.
The Citrix Hypervisor /var/log/messages file contains these error messages:

NVRM: Xid (PCI:0000:08:00): 31, Ch 0000001e, engmask 00000111, intr 10000000
NVRM: Xid (PCI:0000:08:00): 31, Ch 00000016, engmask 00000111, intr 10000000
...  
vmiop_log: error: Assertion Failed at 0xb5b898d8:4184
vmiop_log: error: 8 frames returned by backtrace
vmiop_log: error: /usr/lib/libnvidia-vgx.so(_nv000793vgx+0x69d) [0xb5b8064d]
vmiop_log: error: /usr/lib/libnvidia-vgx.so(_nv000479vgx+0x118) [0xb5b898d8]
vmiop_log: error: /usr/lib/libnvidia-vgx.so(_nv000782vgx+0x59) [0xb5b85f49]
vmiop_log: error: /usr/lib/libnvidia-vgx.so(_nv000347vgx+0x3db) [0xb5b932db]
vmiop_log: error: /usr/lib/libnvidia-vgx.so [0xb5b78e4a]
vmiop_log: error: /usr/lib/xen/bin/vgpu [0x80554be]
vmiop_log: error: /lib/libpthread.so.0 [0xb7612912]
vmiop_log: error: /lib/libc.so.6(clone+0x5e) [0xb76fc5ee]
vmiop_log: error: failed to initialize guest PTE entries
vmiop_log: error: failed to fill up guest PTE entries 3
vmiop_log: error: VGPU message 27 failed, result code: 0xff000003
vmiop_log: error: 0x1c000001, 0xff010000, 0x1a77ba000, 0x0, 0x1,
vmiop_log: error: 0x1, 0x1000, 0x10202, 0xc1d00001, 0xff010000,
vmiop_log: error: 0xc0f00004, 0x0
vmiop_log: error: Timeout occurred, reset initiated.

Version
Citrix Hypervisor 6.2

Fix
Ensure that you are running the latest OEM firmware for your NVIDIA vGPU software boards.

Status
Closed

Ref. #
NVIDIA-327/1632120

5.38. Windows VM BSOD when upgrading NVIDIA drivers over a Citrix Virtual Apps and Desktops session

Description
Windows VM bugchecks when NVIDIA guest drivers are upgraded over a Citrix Virtual Apps and Desktops session.

If the VM is restarted after the bugcheck, the upgraded driver loads correctly and full functionality is available.
5.39. XenCenter does not allow vGPUs to be selected as a GPU type for Linux VMs

Description
When creating a new Linux VM or editing the properties of an existing Linux VM, XenCenter does not allow vGPUs to be selected as a GPU type.

vGPU on Linux VMs is supported as a technical preview on Citrix Hypervisor 6.5, and does include XenCenter integration.

Version
Affects the XenCenter integration with Citrix Hypervisor 6.5 only.
Resolved in the XenCenter integration with Citrix Hypervisor 7.0.

Workaround
Refer to XenServer vGPU Management in Virtual GPU Software User Guide for how to configure vGPU by using the \texttt{xe} CLI.

Status
Closed

Ref. #
NVIDIA-360
5.40. If X server is killed on a RHEL7 VM running vGPU, XenCenter console may not automatically switch to text console

Description
If X server is killed on a RHEL7 VM running vGPU, XenCenter console may display a corrupted image and fail to switchover to text console.

The failure to switchover to text console is due to a bug in RHEL7, which causes X server to not start correctly under certain configurations.

Workaround
Use CTRL+ALT+F1, F2, or F3 to switch between Linux terminals.

Status
Closed

Ref. #
NVIDIA-350/200123378

5.41. Citrix Virtual Apps and Desktops shows only a black screen when connected to a vGPU VM

Description
Citrix Virtual Apps and Desktops sometimes displays only a black screen when it is connected to an NVIDIA vGPU VM. The probable cause is that the display that is connected to the NVIDIA vGPU is entering a lower power state.

Fix
Disable all display-related power management settings.

For detailed instructions, visit Microsoft power plans frequently asked questions and from the list, select your OS version.

Status
Not an NVIDIA bug
Ref. #
1719877
Notice

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 NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE.

Information furnished is believed to be accurate and reliable. However, NVIDIA Corporation assumes no responsibility for the consequences of use of such information or for any infringement of patents or other rights of third parties that may result from its use. No license is granted by implication of otherwise under any patent rights of NVIDIA Corporation. Specifications mentioned in this publication are subject to change without notice. This publication supersedes and replaces all other information previously supplied. NVIDIA Corporation products are not authorized as critical components in life support devices or systems without express written approval of NVIDIA Corporation.

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, vGPU, Pascal, 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

© 2013-2019 NVIDIA Corporation. All rights reserved.