TABLE OF CONTENTS

Chapter 1. Release Notes........................................................................................................ 1
  1.1. Updates in Release 10.0.......................................................................................... 2
  1.2. Updates in Release 10.1...................................................................................... 3
  1.3. Updates in Release 10.2...................................................................................... 3

Chapter 2. Validated Platforms............................................................................................... 5
  2.1. Supported NVIDIA GPUs and Validated Server Platforms................................. 5
  2.2. Hypervisor Software Releases............................................................................. 6
  2.3. Guest OS Support............................................................................................... 9
    2.3.1. Windows Guest OS Support...................................................................... 10
    2.3.2. Linux Guest OS Support.......................................................................... 11
  2.4. NVIDIA CUDA Toolkit Version Support............................................................ 12
  2.5. Multiple vGPU Support..................................................................................... 13
  2.6. Peer-to-Peer CUDA Transfers over NVLink Support........................................... 15

Chapter 3. Known Product Limitations............................................................................... 17
  3.1. Total frame buffer for vGPUs is less than the total frame buffer on the physical GPU... 17
  3.2. Issues may occur with graphics-intensive OpenCL applications on vGPU types with limited
       frame buffer...................................................................................................... 18
  3.3. In pass through mode, all GPUs connected to each other through NVLink must be assigned
       to the same VM............................................................................................... 19
  3.4. vGPU profiles with 512 Mbytes or less of frame buffer support only 1 virtual display head
       on Windows 10............................................................................................... 19
  3.5. NVENC requires at least 1 Gbyte of frame buffer................................................. 20
  3.6. VM running older NVIDIA vGPU drivers fails to initialize vGPU when booted........... 20
  3.7. Single vGPU benchmark scores are lower than pass-through GPU........................ 21
  3.8. nvidia-smi fails to operate when all GPUs are assigned to GPU pass-through mode... 22
  3.9. Windows Aero is disabled on Citrix Virtual Apps and Desktops session using 3 or 4 monitors
       in 2560×1600 resolution.................................................................................. 23
  3.10. VMs configured with large memory fail to initialize vGPU when booted................. 24
  3.11. vGPU host driver RPM upgrade fails.................................................................. 25

Chapter 4. Resolved Issues................................................................................................. 26

Chapter 5. Known Issues..................................................................................................... 27
  5.1. Out of memory errors cause VM crashes on Citrix Hypervisor............................. 27
  5.2. On Linux, the frame rate might drop to 1 after several minutes............................ 28
  5.3. 10.0, 10.1 Only: NVIDIA Control Panel cannot be used to change the display resolution...29
  5.4. Citrix XenCenter displays misleading information about vGPU types....................... 29
  5.5. Citrix Virtual Apps and Desktops connection freezes initially............................... 30
  5.6. Uncorrectable machine check exception occurs after initial reboot of the hypervisor host.31
  5.7. DWM crashes randomly occur in Windows VMs.................................................. 32
  5.8. NVIDIA Control Panel fails to launch in a platform layer or published image....... 32
  5.9. Remote desktop session freezes with assertion failure and XID error 43 after migration.. 32
5.10. 10.0, 10.1 Only: The Desktop color depth list is empty.................................33
5.11. Citrix Virtual Apps and Desktops session freezes when the desktop is unlocked........33
5.12. NVIDIA vGPU software graphics driver fails after Linux kernel upgrade with DKMS enabled34
5.13. A PCIe bus error occurs in iLO after Tesla M10 installation..................................35
5.14. 10.0 Only: XID errors and TDR occur after vGPU migration.................................36
5.15. On Citrix Hypervisor, all vGPUs in a VM must be of the same type....................36
5.16. Console VGA cannot be disabled................................................................40
5.17. Console VNC is unusable with Xorg on multiple vGPUs in a VM.........................38
5.18. Publisher not verified warning during Windows 7 driver installation....................38
5.19. RAPIDS cuDF merge fails on NVIDIA vGPU.......................................................39
5.20. Migration of VMs configured with vGPU stops before the migration is complete.....39
5.21. ECC memory settings for a vGPU cannot be changed by using NVIDIA X Server Settings.. 40
5.22. Changes to ECC memory settings for a Linux vGPU VM by nvidia-smi might be ignored.. 41
5.23. Incorrect GPU type shown for Quadro RTX 8000 GPUs in Citrix XenCenter..............41
5.24. NVIDIA Notification Icon prevents log off of Citrix Published Application user sessions...42
5.25. Vulkan applications crash in Windows 7 guest VMs configured with NVIDIA vGPU. 42
5.26. Host core CPU utilization is higher than expected for moderate workloads............... 43
5.27. Frame capture while the interactive logon message is displayed returns blank screen... 44
5.28. RDS sessions do not use the GPU with some Microsoft Windows Server releases......... 44
5.29. Even when the scheduling policy is equal share, unequal GPU utilization is reported.....45
5.30. When the scheduling policy is fixed share, GPU utilization is reported as higher than  
     expected....................................................................................................46
5.31. nvidia-smi reports that vGPU migration is supported on all hypervisors................47
5.32. NVIDIA Control Panel Crashes in a VM connected to two 4K displays....................47
5.33. vGPU guest VM driver not properly loaded on servers with more than 512 GB or 1 TB  
     or more of system memory.........................................................................48
5.34. Luxmark causes a segmentation fault on an unlicensed Linux client......................49
5.35. Resolution is not updated after a VM acquires a license and is restarted............... 49
5.36. A segmentation fault in DBus code causes nvidia-gridd to exit on Red Hat Enterprise  
     Linux and CentOS.........................................................................................50
5.37. No Manage License option available in NVIDIA X Server Settings by default............50
5.38. Licenses remain checked out when VMs are forcibly powered off..........................51
5.39. Memory exhaustion can occur with vGPU profiles that have 512 Mbytes or less of frame  
     buffer........................................................................................................52
5.40. VM bug checks after the guest VM driver for Windows 10 RS2 is installed...............53
5.41. On Citrix Hypervisor 7.0, VMs unexpectedly reboot and Citrix Hypervisor crashes or 
     freezes.......................................................................................................54
5.42. With no NVIDIA driver installed, Citrix Hypervisor misidentifies Tesla M10 cards.......55
5.43. GNOME Display Manager (GDM) fails to start on Red Hat Enterprise Linux 7.2 and CentOS 
     7.0.............................................................................................................55
5.44. Video goes blank when run in loop in Windows Media Player.................................56
5.45. Local VGA console is momentarily unblanked when Citrix Virtual Apps and Desktops  
     changes resolution of the VM desktop......................................................57
5.46. VM bugchecks on shutdown/restart when Citrix Virtual Apps and Desktops is installed and NVIDIA driver is uninstalled or upgraded................................................................. 57
5.47. Application frame rate may drop when running Citrix Virtual Apps and Desktops at 2560×1600 resolution........................................................................................................ 58
5.48. Windows VM BSOD.......................................................................................................................... 58
5.49. Windows VM BSOD when upgrading NVIDIA drivers over a Citrix Virtual Apps and Desktops session........................................................................................................ 59
5.50. XenCenter does not allow vGPUs to be selected as a GPU type for Linux VMs................. 60
5.51. If X server is killed on a RHEL7 VM running vGPU, XenCenter console may not automatically switch to text console................................................................. 61
5.52. Citrix Virtual Apps and Desktops shows only a black screen when connected to a vGPU VM
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>10.0</th>
<th>10.1</th>
<th>10.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVIDIA Virtual GPU Manager for the Citrix Hypervisor releases listed in Hypervisor Software Releases</td>
<td>440.43</td>
<td>440.53</td>
<td>440.87</td>
</tr>
<tr>
<td>NVIDIA Windows driver</td>
<td>441.66</td>
<td>442.06</td>
<td>443.05</td>
</tr>
<tr>
<td>NVIDIA Linux driver</td>
<td>440.43</td>
<td>440.56</td>
<td>440.87</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 10.1 with guest VM drivers from release 10.0. 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 10.1 with guest VM drivers from release 9.2.

See VM running older NVIDIA vGPU drivers fails to initialize vGPU when booted.
This requirement does not apply to the NVIDIA vGPU software license server. All releases of NVIDIA vGPU software are compatible with all releases of the license server.

1.1. Updates in Release 10.0

New Features in Release 10.0

- Support for NVIDIA® GRID™ Virtual PC and GRID Virtual Applications on Quadro RTX 6000 and Quadro RTX 8000 GPUs
- Increase in the maximum number of virtual display heads supported by -1Q, -2B, and -1B4 vGPUs:
  - All -1Q vGPUs now support 4 heads instead of 2 heads.
  - All -2B vGPUs now support 4 heads instead of 2 heads.
  - All -1B4 vGPUs now support 4 heads instead of 1 head.
- Flexible virtual display resolutions
  Instead of a fixed maximum resolution per head, vGPUs now support a maximum combined resolution based on their frame buffer size. This behavior allows the same number of lower resolution displays to be used as before, but alternatively allows a smaller number of higher resolution displays to be used.
- Virtual display resolutions greater than 4096×2160
- 10-bit color
- Support for multiple vGPUs in a single VM
- Changes to allow cross-branch driver support in future main release branches

This feature cannot be used until the next NVIDIA vGPU software main release branch is available.

The purpose of this change is to allow a release of the Virtual GPU Manager from a later main release branch to be used with the NVIDIA vGPU software graphics drivers for the guest VMs from the previous branch.

- Miscellaneous bug fixes

Hardware and Software Support Introduced in Release 10.0

- Support for passively cooled Quadro RTX 6000 and Quadro RTX 8000 GPUs
- Support for Tesla V100S PCIe 32GB GPUs
- Support for Citrix Hypervisor 8.1
- Support for Red Hat Enterprise Linux 8.1 as a guest OS
- Support for Red Hat Enterprise Linux 8.0 and CentOS 8.0 as a guest OS
- Support for Windows 10 November 2019 Update (1909) as a guest OS
- Support for Citrix Virtual Apps and Desktops version 7 1912
Features Deprecated in Release 10.0

The following table lists features that are deprecated in this release of NVIDIA vGPU software. Although the features remain available in this release, they might be withdrawn in a future release. In preparation for the possible removal of these features, use the preferred alternative listed in the table.

<table>
<thead>
<tr>
<th>Deprecated Feature</th>
<th>Preferred Alternative</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1B4 vGPU types</td>
<td>-1B vGPU types</td>
</tr>
<tr>
<td>-2B4 vGPU types</td>
<td>-2B vGPU types</td>
</tr>
</tbody>
</table>

1.2. Updates in Release 10.1

New Features in Release 10.1

- Miscellaneous bug fixes

Feature Support Withdrawn in Release 10.1

- Citrix Hypervisor 7.6 is no longer supported.
- Red Hat Enterprise Linux 8.0 is no longer supported as a guest OS.

1.3. Updates in Release 10.2

New Features in Release 10.2

- Miscellaneous bug fixes

Hardware and Software Support Introduced in Release 10.2

- Support for the following OS releases as a guest OS:
  - Red Hat Enterprise Linux 8.2
  - CentOS Linux 8 (1911)
  - Red Hat Enterprise Linux 7.8
  - CentOS 7.8
- Support for Citrix Virtual Apps and Desktops version 7 2003

Feature Support Withdrawn in Release 10.2

- The following guest OS releases are no longer supported:
- Red Hat Enterprise Linux 7.0-7.5
- CentOS 7.0-7.5

The following releases of Citrix virtual desktop software are no longer supported:

- Citrix Virtual Apps and Desktops version 7 1808
- Citrix XenApp/XenDesktop versions 7.18 and 7.17
Chapter 2.
VALIDATED PLATFORMS

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 R440 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 V100S PCIe 32GB
- Tesla V100 FHHL
- GPUs based on the NVIDIA Turing™ architecture:
  - Tesla T4
  - Quadro RTX 6000 in displayless mode
  - Quadro RTX 6000 passive in displayless mode
  - Quadro RTX 8000 in displayless mode
  - Quadro RTX 8000 passive in displayless mode

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 yymm, where yymm denotes the year and month that the product was released. For example, for a product released in August 2018, yymm is **1808**.

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

- 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 8.1 | 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 6000 passive  
  - Quadro RTX 8000  
  - Quadro RTX 8000 passive  
  This release supports XenMotion with vGPU on suitable GPUs as listed in Supported NVIDIA GPUs and Validated Server Platforms. |
<p>| 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. Multiple vGPUs on a single VM are not supported. vComputeServer is not supported. |</p>
<table>
<thead>
<tr>
<th>Software</th>
<th>Releases Supported</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>In GPU pass-through mode, the following GPUs are not supported:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‣ Quadro RTX 6000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‣ Quadro RTX 6000 passive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‣ Quadro RTX 8000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‣ Quadro RTX 8000 passive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>This release supports XenMotion with vGPU on suitable GPUs as listed in Supported NVIDIA GPUs and Validated Server Platforms.</td>
</tr>
<tr>
<td>10.0 only: Citrix Hypervisor (XenServer) 7.6</td>
<td>RTM build and compatible cumulative update releases</td>
<td>All NVIDIA GPUs that support NVIDIA vGPU software are supported.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple vGPUs on a single VM are not supported.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>vComputeServer is not supported.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>This release supports XenMotion with vGPU on suitable GPUs as listed in Supported NVIDIA GPUs and Validated Server Platforms.</td>
</tr>
<tr>
<td>Citrix Hypervisor (XenServer) 7.1</td>
<td>RTM build and compatible cumulative update releases</td>
<td>Not supported on the following GPUs:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‣ Tesla T4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‣ Quadro RTX 6000 passive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‣ Quadro RTX 8000 passive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple vGPUs on a single VM are not supported.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>vComputeServer is not supported.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>XenMotion with vGPU is not 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 the following GPUs:</td>
</tr>
</tbody>
</table>
Validated Platforms

Software Releases Supported Notes

‣ Tesla T4
‣ Quadro RTX 6000 passive
‣ Quadro RTX 8000 passive

Multiple vGPUs on a single VM are not supported.
vComputeServer is not supported.
XenMotion with vGPU is not supported.

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 10.2: Versions 7 2003, Versions 7 1912, 7 1909, 7 1906, and 7 1903</td>
</tr>
<tr>
<td></td>
<td>10.0, 10.1 only: Versions 7 1912, 7 1909, 7 1906, 7 1903, and 7 1808</td>
</tr>
<tr>
<td>XenApp/XenDesktop</td>
<td>Since 10.2: Version 7.15</td>
</tr>
<tr>
<td></td>
<td>10.0, 10.1 only: Versions 7.18, 7.17, and 7.15</td>
</tr>
<tr>
<td></td>
<td>Version 7.15 is supported only in HDX 3D Pro mode. HDX 3D Pro mode is 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:

- Quadro RTX 6000
- Quadro RTX 8000

### 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>8.1, 8.0, 7.1 cumulative update 2</td>
<td>8.1, 8.0, 7.1 cumulative update 2</td>
</tr>
<tr>
<td>Windows Server 2016 1709, 1607</td>
<td>Since 10.1: 8.1, 8.0, 7.1, 7.0</td>
<td>Since 10.1: 8.1, 8.0, 7.1, 7.0</td>
</tr>
<tr>
<td></td>
<td>10.0 only: 8.1, 8.0, 7.6, 7.1, 7.0</td>
<td>10.0 only: 8.1, 8.0, 7.6, 7.1, 7.0</td>
</tr>
<tr>
<td>Windows Server 2012 R2</td>
<td>Since 10.1: 8.1, 8.0, 7.1, 7.0</td>
<td>Since 10.1: 8.1, 8.0, 7.1, 7.0</td>
</tr>
<tr>
<td></td>
<td>10.0 only: 8.1, 8.0, 7.6, 7.1, 7.0</td>
<td>10.0 only: 8.1, 8.0, 7.6, 7.1, 7.0</td>
</tr>
<tr>
<td>Windows Server 2008 R2</td>
<td>Since 10.1: 8.1, 8.0, 7.1, 7.0</td>
<td>Since 10.1: 8.1, 8.0, 7.1, 7.0</td>
</tr>
<tr>
<td></td>
<td>10.0 only: 8.1, 8.0, 7.6, 7.1, 7.0</td>
<td>10.0 only: 8.1, 8.0, 7.6, 7.1, 7.0</td>
</tr>
<tr>
<td>Windows 10:</td>
<td>Since 10.1: 8.1, 8.0, 7.1, 7.0</td>
<td>Since 10.1: 8.1, 8.0, 7.1, 7.0</td>
</tr>
<tr>
<td>‣ November 2019 Update (1909)</td>
<td>10.0 only: 8.1, 8.0, 7.6, 7.1, 7.0</td>
<td>10.0 only: 8.1, 8.0, 7.6, 7.1, 7.0</td>
</tr>
<tr>
<td>‣ May 2019 Update (1903)</td>
<td></td>
<td></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>
</tbody>
</table>
Validated Platforms

www.nvidia.com
Virtual GPU Software R440 for Citrix Hypervisor

<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>RTM (1507)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Windows 8.1 Update</td>
<td>Since 10.1: 8.1, 8.0, 7.1, 7.0</td>
<td>Since 10.1: 8.1, 8.0, 7.1, 7.0</td>
</tr>
<tr>
<td></td>
<td>10.0 only: 8.1, 8.0, 7.6, 7.1, 7.0</td>
<td>10.0 only: 8.1, 8.0, 7.6, 7.1, 7.0</td>
</tr>
<tr>
<td>Windows 8.1</td>
<td>Since 10.1: 8.1, 8.0, 7.1, 7.0</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>10.0 only: 8.1, 8.0, 7.6, 7.1, 7.0</td>
<td></td>
</tr>
<tr>
<td>Windows 8</td>
<td>Since 10.1: 8.1, 8.0, 7.1, 7.0</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>10.0 only: 8.1, 8.0, 7.6, 7.1, 7.0</td>
<td></td>
</tr>
<tr>
<td>Windows 7</td>
<td>Since 10.1: 8.1, 8.0, 7.1, 7.0</td>
<td>Since 10.1: 8.1, 8.0, 7.1, 7.0</td>
</tr>
<tr>
<td></td>
<td>10.0 only: 8.1, 8.0, 7.6, 7.1, 7.0</td>
<td>10.0 only: 8.1, 8.0, 7.6, 7.1, 7.0</td>
</tr>
<tr>
<td></td>
<td>Supported only on GPUs based on the Maxwell architecture</td>
<td></td>
</tr>
</tbody>
</table>

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>Since 10.2: Red Hat Enterprise Linux 8.2</td>
<td>8.1</td>
<td>8.1</td>
</tr>
<tr>
<td>Red Hat Enterprise Linux 8.1</td>
<td>8.1</td>
<td>8.1</td>
</tr>
<tr>
<td>Since 10.2: CentOS Linux 8 (1911)</td>
<td>8.1</td>
<td>8.1</td>
</tr>
<tr>
<td>10.0 only: Red Hat Enterprise Linux 8.0</td>
<td>8.1</td>
<td>8.1</td>
</tr>
<tr>
<td>CentOS 8.0</td>
<td>8.1</td>
<td>8.1</td>
</tr>
<tr>
<td>Since 10.2: Red Hat Enterprise Linux 7.6-7.8</td>
<td>8.1, 8.0, 7.1, 7.0</td>
<td>8.1, 8.0, 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 10.2.

For more information about NVIDIA CUDA Toolkit, see CUDA Toolkit 10.2 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 V100S PCIe 32GB
- Tesla V100 FHHL
- Tesla T4
- Quadro RTX 6000
- Quadro RTX 6000 passive
- Quadro RTX 8000
- Quadro RTX 8000 passive

Supported Citrix Hypervisor releases:
- Since 10.1: 8.1 and 8.0 only
- 10.0 only: 8.1, 8.0, and 7.6 only

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

2.6. Multiple vGPU Support

To support applications and workloads that are compute or graphics intensive, multiple vGPUs can be added to a single VM. The assignment of more than one vGPU to a VM is supported only on a subset of vGPUs and Citrix Hypervisor releases.

Supported vGPUs

Only Q-series and C-series vGPUs that are allocated all of the physical GPU’s frame buffer are supported.
<table>
<thead>
<tr>
<th>GPU Architecture</th>
<th>Board</th>
<th>vGPU</th>
</tr>
</thead>
</table>
| Turing           | Tesla T4 | T4-16Q  
|                  |        | T4-16C  |
|                  | Quadro RTX 6000 | RTX6000-24Q  
|                  |        | RTX6000-24C  |
|                  | Quadro RTX 6000 passive | RTX6000P-24Q  
|                  |        | RTX6000P-24C  |
|                  | Quadro RTX 8000 | RTX8000-48Q  
|                  |        | RTX8000-48C  |
|                  | Quadro RTX 8000 passive | RTX8000P-48Q  
|                  |        | RTX8000P-48C  |
| Volta            | Tesla V100 SXM2 32GB | V100DX-32Q  
|                  |        | V100D-32C  |
|                  | Tesla V100 PCIe 32GB | V100D-32Q  
|                  |        | V100D-32C  |
|                  | Tesla V100S PCIe 32GB | V100S-32Q  
|                  |        | V100S-32C  |
|                  | Tesla V100 SXM2 | V100X-16Q  
|                  |        | V100X-16C  |
|                  | Tesla V100 PCIe | V100-16Q  
|                  |        | V100-16C  |
|                  | Tesla V100 FHHL | V100L-16Q  
|                  |        | V100L-16C  |
| Pascal           | Tesla P100 SXM2 | P100X-16Q  
|                  |        | P100X-16C  |
|                  | Tesla P100 PCIe 16GB | P100-16Q  
|                  |        | P100-16C  |
|                  | Tesla P100 PCIe 12GB | P100C-12Q  
|                  |        | P100C-12C  |
## Validated Platforms

**www.nvidia.com**

**Virtual GPU Software R440 for Citrix Hypervisor**

**RN-06927-001 _v10.0 through 10.2 | 15**

### GPU Architecture

<table>
<thead>
<tr>
<th>Board</th>
<th>vGPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>P40-24Q</td>
<td>P40-24C</td>
</tr>
<tr>
<td>P6-16Q</td>
<td>P6-16C</td>
</tr>
<tr>
<td>P4-8Q</td>
<td>P4-8C</td>
</tr>
<tr>
<td>M60-8Q</td>
<td>M10-8Q</td>
</tr>
<tr>
<td>M6-8Q</td>
<td></td>
</tr>
</tbody>
</table>

### Maximum vGPUs per VM

NVIDIA vGPU software supports up to a maximum of 16 vGPUs per VM on Citrix Hypervisor.

### Supported Hypervisor Releases

Citrix Hypervisor 8.1 only.

### 2.7. Peer-to-Peer CUDA Transfers over NVLink Support

Peer-to-peer CUDA transfers enable device memory between vGPUs on different GPUs that are assigned to the same VM to be accessed from within the CUDA kernels. NVLink is a high-bandwidth interconnect that enables fast communication between such vGPUs. Peer-to-Peer CUDA Transfers over NVLink is supported only on a subset of vGPUs, Citrix Hypervisor releases, and guest OS releases.

### Supported vGPUs

Only Q-series and C-series vGPUs that are allocated all of the physical GPU's frame buffer on physical GPUs that support NVLink are supported.

<table>
<thead>
<tr>
<th>GPU Architecture</th>
<th>Board</th>
<th>vGPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turing</td>
<td>Quadro RTX 6000</td>
<td>RTX6000-24Q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RTX6000-24C</td>
</tr>
<tr>
<td></td>
<td>Quadro RTX 6000 passive</td>
<td>RTX6000P-24Q</td>
</tr>
<tr>
<td>GPU Architecture</td>
<td>Board</td>
<td>vGPU</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RTX6000P-24C</td>
</tr>
<tr>
<td>Quadro RTX 8000</td>
<td>RTX8000-48Q</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RTX8000-48C</td>
<td></td>
</tr>
<tr>
<td>Quadro RTX 8000 passive</td>
<td>RTX8000P-48Q</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RTX8000P-48C</td>
<td></td>
</tr>
<tr>
<td>Volta</td>
<td>Tesla V100 SXM2 32GB</td>
<td>V100DX-32Q</td>
</tr>
<tr>
<td></td>
<td>V100DX-32C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tesla V100 SXM2</td>
<td>V100X-16Q</td>
</tr>
<tr>
<td></td>
<td>V100X-16C</td>
<td></td>
</tr>
<tr>
<td>Pascal</td>
<td>Tesla P100 SXM2</td>
<td>P100X-16Q</td>
</tr>
<tr>
<td></td>
<td>P100X-16C</td>
<td></td>
</tr>
</tbody>
</table>

**Supported Hypervisor Releases**

Peer-to-Peer CUDA Transfers over NVLink are supported on all hypervisor releases that support the assignment of more than one vGPU to a VM. For details, see Multiple vGPU Support.

**Supported Guest OS Releases**

Linux only. Peer-to-Peer CUDA Transfers over NVLink are not supported on Windows.

**Limitations**

- Only direct connections are supported. NVSwitch is not supported.
- PCIe is not supported.
- SLI is not supported.
Chapter 3.
KNOWN PRODUCT LIMITATIONS

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

3.1. 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:

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

\[
\text{max-reserved-fb}
\]

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 that support dynamic page retirement, page-retirement-allocation = \(128 \div \max\text{-vgpus-per-gpu}\), where \(\max\text{-vgpus-per-gpu}\) is 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, \(\max\text{-vgpus-per-gpu}\) is 1.
- On GPUs that don't support dynamic page retirement, page-retirement-allocation is 0.

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.

### 3.2. 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.
3.3. 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.

3.4. 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.
3.5. 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.

3.6. 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.

3.7. 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:
Known Product Limitations

- 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"
```

3.8. `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:
Resolution
N/A

3.9. 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.
3.10. 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`

3.11. 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
[root@xenserver ~]#

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

Use the following older 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 10.0

No resolved issues are reported in this release for Citrix Hypervisor.

Issues Resolved in Release 10.1

<table>
<thead>
<tr>
<th>Bug ID</th>
<th>Summary and Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>200556412</td>
<td>10.0 Only: XID errors and TDR occur after vGPU migration</td>
</tr>
<tr>
<td></td>
<td>After several vGPU migrations, Direct X 12 applications might show corruption or crash with XID errors 13 or 32 and a TDR followed by XID error 43. This issue affects only GPUs based on the NVIDIA Turing™ architecture.</td>
</tr>
</tbody>
</table>

Issues Resolved in Release 10.2

<table>
<thead>
<tr>
<th>Bug ID</th>
<th>Summary and Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2920224</td>
<td>10.0, 10.1 Only: NVIDIA Control Panel cannot be used to change the display resolution</td>
</tr>
<tr>
<td></td>
<td>After the user selects a new display resolution in NVIDIA Control Panel and clicks Apply, the resolution is not changed and the selection on the list is reset to the previous value.</td>
</tr>
<tr>
<td>200555917</td>
<td>10.0, 10.1 Only: The Desktop color depth list is empty</td>
</tr>
<tr>
<td></td>
<td>The Desktop color depth list on the Change resolution page in NVIDIA Control Panel for the VM display NVIDIA VGX is empty. This list should include options such as SDR 24 bit and SDR 30 bit.</td>
</tr>
</tbody>
</table>
5.1. Out of memory errors cause VM crashes on Citrix Hypervisor

Description

Out of memory errors caused by extreme memory fragmentation cause VM crashes on Citrix Hypervisor.

When this issue occurs, error messages similar to the following examples are logged:

Feb 4 10:45:53 copic-xs-4 kernel: [3067294.076077] Out of memory: Kill process 6989 (xapi) score 7 or sacrifice child
Feb 4 10:45:53 copic-xs-4 kernel: [3067294.076399] Killed process 6989 (xapi)
    total-vm:638288kB, anon-rss:53824kB, file-rss:9292kB, shmem-rss:4kB
Feb 4 10:45:53 copic-xs-4 kernel: [3067294.138425] oom_reaper: reaped process 6989 (xapi), now anon-rss:0kB, file-rss:8kB, shmem-rss:4kB

The kill process in the messages might be different than xapi.

Workaround

Increase the memory for the Citrix Hypervisor dom0 domain to at least 16 GB.

Status

Open

Ref. #

2870295
5.2. On Linux, the frame rate might drop to 1 after several minutes

Description
On Linux, the frame rate might drop to 1 frame per second (FPS) after NVIDIA vGPU software has been running for several minutes. Only some applications are affected, for example, glxgears. Other applications, such as Unigine Heaven, are not affected. This behavior occurs because Display Power Management Signaling (DPMS) for the Xorg server is enabled by default and the display is detected to be inactive even when the application is running. When DPMS is enabled, it enables power saving behavior of the display after several minutes of inactivity by setting the frame rate to 1 FPS.

Workaround
1. If necessary, stop the Xorg server.
   
   # /etc/init.d/xorg stop

2. In a plain text editor, edit the /etc/X11/xorg.conf file to set the options to disable DPMS and disable the screen saver.
   
   a. In the Monitor section, set the DPMS option to false.
      
      Option "DPMS" "false"

   b. At the end of the file, add a ServerFlags section that contains option to disable the screen saver.
      
      Section "ServerFlags"
      Option "BlankTime" "0"
      EndSection

   c. Save your changes to /etc/X11/xorg.conf file and quit the editor.

3. Start the Xorg server.
   
   # etc/init.d/xorg start

Status
Open

Ref. #
200605900
5.3. 10.0, 10.1 Only: **NVIDIA Control Panel cannot be used to change the display resolution**

**Description**
After the user selects a new display resolution in **NVIDIA Control Panel** and clicks **Apply**, the resolution is not changed and the selection on the list is reset to the previous value.

**Workaround**
Use Microsoft **Display settings** in **System settings** to change the display resolution.

**Status**
Resolved in NVIDIA vGPU software 10.2

**Ref. #**
2920224

5.4. Citrix XenCenter displays misleading information about vGPU types

**Description**
The **GPU type** drop-down list in the Citrix XenCenter management GUI displays misleading information about the maximum combined resolution supported by Q-series and B-series vGPU types. The GUI lists the maximum number of displays supported and the highest resolution supported in a manner that implies a maximum combined resolution higher than is actually supported. These vGPU types support the number of displays listed only at resolutions lower than the resolution listed and support the resolution listed only with fewer displays than the number listed.

For example, as shown in the following screen capture, the GUI lists a resolution of 5120×2880 and four displays for the V100-1Q vGPU type. However, this vGPU type supports only one display at a resolution of 5120×2880 and supports a maximum of four displays only at resolutions of 2560×1600 and lower.
Version
All supported Citrix Hypervisor versions.

Workaround
For information about display configurations supported by Q-series and B-series vGPU types, see Virtual GPU Software User Guide.

Status
Not an NVIDIA bug

Ref. #
200556224

5.5. Citrix Virtual Apps and Desktops connection freezes initially

Description
When -0B and -0Q vGPU types are used with Citrix Virtual Apps and Desktops version 7 1903 and later versions, the session freezes or a black screen is seen when the connection is first made. When this issue occurs, the error message DXGI_ERROR_DEVICE_REMOVED is displayed. The affected versions of Citrix Virtual Apps and Desktops use Microsoft DDAPI.

Version
Citrix Virtual Apps and Desktops version 7 1903 and later versions
5.6. Uncorrectable machine check exception occurs after initial reboot of the hypervisor host

Description

When the hypervisor host is first rebooted after the Virtual GPU Manager is installed, an uncorrectable machine check exception occurs during POST immediately after the message SYSTEM CHIPSET INITIALIZATION UPI LINK INITIALIZATION - START. This issue affects HP systems that are running a hypervisor release that has a newer kernel, for example, Citrix Hypervisor 8.0 or 8.1.

After the uncorrectable machine check exception, the hypervisor host is automatically reset and returns to a working state.

When this issue occurs, the following error messages are logged:

Uncorrectable Machine Check Exception (Processor 1, APIC ID 0x00000000, Bank 0x00000006, Status 0xFB800000'00000E0B, Address 0x00000000'00000000, Misc 0x00000000'5B000000). 
Uncorrectable PCI Express Error Detected. Slot 255 (Segment 0x0, Bus 0x5D, Device 0x48, Function 0x0) 
Uncorrectable PCI Express Error Detected. Slot 255 (Segment 0x0, Bus 0x5D, Device 0x88, Function 0x0) 
Uncorrectable PCI Express Error Detected. Slot 255 (Segment 0x0, Bus 0x5D, Device 0x80, Function 0x0) 
Uncorrectable PCI Express Error Detected. Slot 255 (Segment 0x0, Bus 0x5D, Device 0x40, Function 0x0)

Workaround

Wait for the host to return to a working state. No other action is required.

Status

Open
5.7. 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

5.8. NVIDIA Control Panel fails to launch in a platform layer or published image

**Description**
When NVIDIA vGPU software is used with Citrix App Layering and a platform layer for a vGPU is being configured, **NVIDIA Control Panel** might fail to launch. The driver might be working normally or it might fail with code 31 or code 43.

For more information and a workaround for this issue, see [NVIDIA Control Panel fails to launch in platform layer or published image](https://www.nvidia.com) in the Citrix Support Knowledge Center.

5.9. Remote desktop session freezes with assertion failure and XID error 43 after migration

**Description**
After multiple VMs configured with vGPU on a single hypervisor host are migrated simultaneously, the remote desktop session freezes with an assertion failure and XID
error 43. This issue affects only GPUs that are based on the Volta GPU architecture. It does not occur if only a single VM is migrated.

When this error occurs, the following error messages are logged to the Citrix Hypervisor log file /var/log/messages:

```
Jan  3 14:35:48 ch81-m1 vgpu-12[8050]: error: vmiop_log: NVOS status 0x1f
Jan  3 14:35:48 ch81-m1 vgpu-12[8050]: error: vmiop_log: Assertion Failed at 0x4b8cacf6:286
Jan  3 14:35:59 ch81-m1 vgpu-12[8050]: error: vmiop_log: (0x0): XID 43 detected on physical_chid:0x174, guest_chid:0x14
```

**Status**

Open

**Ref. #**

200581703

### 5.10. 10.0, 10.1 Only: The Desktop color depth list is empty

**Description**

The Desktop color depth list on the Change resolution page in NVIDIA Control Panel for the VM display NVIDIA VGX is empty. This list should include options such as SDR 24 bit and SDR 30 bit.

**Status**

Resolved in NVIDIA vGPU software 10.2

**Ref. #**

200555917

### 5.11. 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

5.12. 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:

```
$ sudo dkms install nvidia

-> Installing DKMS kernel module:
   ERROR: Failed to run `/usr/sbin/dkms build -m nvidia -v 440.43 -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='' modules...(bad exit status: 2)
   ERROR (dkms apport): binary package for nvidia: 440.43 not found
   Error! Bad return status for module build on kernel: 5.3.0-28-generic
   (x86_64)
   Consult /var/lib/dkms/nvidia/ 440.43/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.
```
5.13. A PCIe bus error occurs in iLO after Tesla M10 installation

Description
When a Tesla M10 GPU is installed in an HPE ProLiant DL380 Gen9 or Gen10 server that is running Citrix Hypervisor 8.0, a PCIe bus error occurs in the Integrated Lights-Out (iLO) remote server management processor. When this error occurs, the server can no longer be shut down, the error message CPU Panic is issued, and a reboot of the server is forced.

This issue is specific to the PLX chip on the Tesla M10 GPU.

Version
This issue affects only the following combination of software and hardware:

- **Hypervisor software:** Citrix Hypervisor 8.0
- **Server hardware:** HPE ProLiant DL380 Gen9 or Gen10
- **GPU hardware:** Tesla M10

Status
Open

Ref. #
200588187
5.14. 10.0 Only: XID errors and TDR occur after vGPU migration

Description
After several vGPU migrations, Direct X 12 applications might show corruption or crash with XID errors 13 or 32 and a TDR followed by XID error 43. This issue affects only GPUs based on the NVIDIA Turing™ architecture.

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

```
WARNING: NVRM: Xid (PCI:0000:05:00): 13, pid=2103457, Graphics Exception on GPC 0: 3D-Z KIND Violation. Coordinates: (0x0, 0x0)
WARNING: NVRM: Xid (PCI:0000:05:00): 13, pid=2103457, Graphics Exception: ESR 0x500420=0x80000080 0x500434=0x0 0x500438=0xd0000 0x50043c=0x10006
WARNING: NVRM: Xid (PCI:0000:05:00): 32, pid=2103457, Channel ID 00000036 intr0 00040000
WARNING: NVRM: Xid (PCI:0000:05:00): 43, pid=2103457, Ch 00000015
```

Status
Resolved in NVIDIA vGPU software 10.1

Ref. #
200556412

5.15. On Citrix Hypervisor, all vGPUs in a VM must be of the same type

Description
When a VM on Citrix Hypervisor is configured with multiple vGPUs, all vGPUs must be of the same type. Citrix Hypervisor does not support the assignment multiple vGPUs of different types to a single VM.

What happens when you try to assign a different type of vGPU than the vGPUs already assigned to a VM depends on the Citrix utility that you are using.

- If you are using Citrix XenCenter, only vGPUs of the same type as the vGPUs already assigned are available for selection.
- If you are using the xe command, the command fails and the following error message is displayed:

```
Cannot create a virtual GPU that is incompatible with the existing types on the VM.
```
5.16. Console VGA cannot be disabled

Description
In Citrix Hypervisor releases since 8.1, a VM's console VGA interface cannot be disabled by using the command `xe vm-param-set uuid=vm-uuid platform:vgpu_extra_args="disable_vnc=1"`. This behavior is the result of the following changes that Citrix introduced in Citrix Hypervisor release 8.1:

- The command `xe vm-param-set` for assigning plugin configuration parameters has been withdrawn.
- Plugin parameters can no longer be set for vGPUs that were created by using Citrix XenCenter. They can be set only for vGPUs that were created from the command line by using the `xe` command.

Version
Since Citrix Hypervisor release 8.1

Workaround
Create the vGPU by using the `xe` command and specify plugin parameters for the group to which the vGPU belongs as explained in Virtual GPU Software User Guide.

Status
Not a bug

Ref. #
200569785
5.17. Console VNC is unusable with Xorg on multiple vGPUs in a VM

Description
If a Linux VM is configured with multiple vGPUs and the Xorg display server is running on more than one vGPU in the VM, severe corruption in console VNC is observed.

Workaround
Ensure that the Xorg display server is running on only one vGPU in the VM.

Status
Not an NVIDIA bug.

Ref. #
200568261

5.18. Publisher not verified warning during Windows 7 driver installation

Description
During installation of the NVIDIA vGPU software graphics driver for Windows on Windows 7, Windows warns that it can’t verify the publisher of the driver software. If Device Manager is used to install the driver, Device Manager warns that the driver is not digitally signed. If you install the driver, error 52 (CM_PROB_UNSIGNED_DRIVER) occurs.

This issue occurs because Microsoft is no longer dual signing WHQL-tested software binary files by using the SHA-1 and SHA-2 hash algorithms. Instead, WHQL-tested software binary files are signed only by using the SHA-2 hash algorithm. All NVIDIA vGPU software graphics drivers for Windows are WHQL tested.

By default, Windows 7 systems cannot recognize signatures that were created by using the SHA-2 hash algorithm. As a result, software binary files that are signed only by using the SHA-2 hash algorithm are considered unsigned.

For more information, see 2019 SHA-2 Code Signing Support requirement for Windows and WSUS on the Microsoft Windows support website.
Version
Windows 7

Workaround
If you experience this issue, install the following updates and restart the VM or host before installing the driver:

- Servicing stack update (SSU) (KB4490628)
- SHA-2 update (KB4474419)

Status
Not a bug

5.19. 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

5.20. 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

Ref. #

200520027

5.21. 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.

Status

Open

Ref. #

200523086
5.22. Changes to ECC memory settings for a Linux vGPU VM by \texttt{nvidia-smi} might be ignored

\textbf{Description}

After the ECC memory state for a Linux vGPU VM has been changed by using the \texttt{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 \texttt{RMGuestECCState} in \texttt{/etc/modprobe.d/nvidia.conf} to be ignored.

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

\textbf{Workaround}

This workaround requires administrator privileges.

1. Move the entry containing the keyword \texttt{NVreg\_RegistryDwordsPerDevice} from the other configuration file to \texttt{/etc/modprobe.d/nvidia.conf}.
2. Reboot the VM.

\textbf{Status}

Open

\textbf{Ref. #}

200505777

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

\textbf{Description}

On the \textbf{GPU} tab in \textbf{Citrix XenCenter}, Quadro RTX 8000 GPUs are incorrectly shown as \texttt{[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

### 5.24. 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.

**Ref. #**

2206368

### 5.25. 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

### 5.26. 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
5.27. 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

5.28. 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.

5.29. 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
5.30. 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%       |
|-------------------------------+--------------------------------+------------|
|   2  Tesla P40                | 00000000:86:00.0               |   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
5.31. **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

5.32. **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
5.33. 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.

   `[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.

3. Reboot the server.

Status

Not an NVIDIA bug

Ref. #

1799582
5.34. 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

5.35. 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
5.36. 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

5.37. 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

5.38. 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

**5.39. 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

**5.40. 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 1920×1200.

1. Choose a GPU profile in Citrix XenCenter that does not allow resolutions higher than 1920×1200.
2. Before rebooting the VM, set the display resolution to 1920×1200 or lower.

Status
Not an NVIDIA bug

Ref. #
200310861

5.41. 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.
5.42. 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

5.43. 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`.
   ```bash
   SELINUX=permissive
   ```
2. Reboot the system.
   ```bash
   ~]# 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

### 5.44. 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
5.45. 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

5.46. 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

### 5.47. 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

### 5.48. 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 00000001, engmask 00000001, intr 10000000
NVRM: Xid (PCI:0000:08:00): 31, Ch 0000001e, engmask 00000000, 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
vmiop_log: error: VGPU message 27 failed, result code: 0xff000003
vmiop_log: error: Oxc1d00001, 0xff010000, 0x1a77ba000, 0x0, 0x1,
vmiop_log: error: Oxl, 0x1000, 0x10202, Oxc1d00001, 0xff010000,
vmiop_log: error: 0xcaf00004, 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

5.49. 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.
Known Issues

Fix
Upgrade Citrix Virtual Apps and Desktops to 7.6 Feature Pack 3

Status
Closed

Ref. #
NVIDIA-370/200130780

5.50. 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 xe CLI.

Status
Closed

Ref. #
NVIDIA-360
5.51. 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

5.52. 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

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.

www.nvidia.com
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 NONINFRINGEMENT, 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.