

# Virtual GPU Software R470 for Citrix Hypervisor

Release Notes

# **Table of Contents**

Cr	napter I. Release Notes	I
	1.1. NVIDIA vGPU Software Driver Versions	1
	1.2. Compatibility Requirements for the NVIDIA vGPU Manager and Guest VM Driver	2
	1.3. Updates in Release 13.10	3
	1.4. Updates in Release 13.9	4
	1.5. Updates in Release 13.8	4
	1.6. Updates in Release 13.7	5
	1.7. Updates in Release 13.6	5
	1.8. Updates in Release 13.5	5
	1.9. Updates in Release 13.4	6
	1.10. Updates in Release 13.3	6
	1.11. Updates in Release 13.2	7
	1.12. Updates in Release 13.1	7
	1.13. Updates in Release 13.0	8
Ch	napter 2. Validated Platforms	9
	2.1. Supported NVIDIA GPUs and Validated Server Platforms	9
	2.1.1. Switching the Mode of a GPU that Supports Multiple Display Modes	. 13
	2.1.2. Switching the Mode of a Tesla M60 or M6 GPU	. 14
	2.2. Hypervisor Software Releases	15
	2.3. Guest OS Support	. 17
	2.3.1. Windows Guest OS Support	. 18
	2.3.2. Linux Guest OS Support	.19
	2.4. NVIDIA CUDA Toolkit Version Support	.20
	2.5. XenMotion with vGPU Support	21
	2.6. Multiple vGPU Support	. 23
	2.7. Peer-to-Peer CUDA Transfers over NVLink Support	. 25
	2.8. Unified Memory Support	. 26
	2.9. Since 13.1: NVIDIA Deep Learning Super Sampling (DLSS) Support	. 27
Ch	napter 3. Known Product Limitations	29
	3.1. NVENC does not support resolutions greater than 4096×4096	29
	3.2. Nested Virtualization Is Not Supported by NVIDIA vGPU	.30
	3.3. Issues occur when the channels allocated to a vGPU are exhausted	30
	3.4. Total frame buffer for vGPUs is less than the total frame buffer on the physical	
	GPU	31

3.5. Issues may occur with graphics-intensive OpenCL applications on vGPU types v limited frame buffer	
3.6. In pass through mode, all GPUs connected to each other through NVLink must assigned to the same VM	
3.7. vGPU profiles with 512 Mbytes or less of frame buffer support only 1 virtual displaced on Windows 10	
3.8. NVENC requires at least 1 Gbyte of frame buffer	35
3.9. VM running an incompatible NVIDIA vGPU guest driver fails to initialize vGPU w booted	
3.10. Single vGPU benchmark scores are lower than pass-through GPU	37
3.11. nvidia-smi fails to operate when all GPUs are assigned to GPU pass-through mode	•
3.12. Windows Aero is disabled on Citrix Virtual Apps and Desktops session using 3 4 monitors in 2560×1600 resolution	
3.13. VMs configured with large memory fail to initialize vGPU when booted	40
3.14. vGPU host driver RPM upgrade fails	41
Chapter 4. Resolved Issues	43
Chapter 5. Known Issues	
5.1. Frame buffer seems to be missing from GPUs	
5.2. Graphics applications are corrupted on some Windows vGPU VMs	
5.3. NVIDIA Control Panel crashes if a user session is disconnected and reconnected	
5.4. 13.0-13.7 Only: Remote desktop connection is lost and the NVIDIA vGPU softw graphics driver is unloaded	
5.5. Pixelation occurs on a Windows VM configured with a Tesla T4 vGPU	50
5.6. 13.4-13.6 Only: Destination host becomes unstable after migration of a VM runr an NVIDIA vGPU software 11 release	0
5.7. 13.0-13.6 Only: Windows VMs fail to acquire a license in environments with mult active desktop sessions	•
5.8. 13.0-13.6 Only: Some graphics applications can cause a TDR on a Windows configured with an NVIDIA A16 vGPU	VM
5.9. 13.0-13.4 Only: VMs configured with a vGPU based on the NVIDIA Amp architecture can become slow to respond	
5.10. NLS client fails to acquire a license with the error The allowed time to procresponse has expired	
5.11. With multiple active sessions, NVIDIA Control Panel incorrectly shows that system is unlicensed	the
5.12. VP9 and AV1 decoding with web browsers are not supported on Micros Windows Server 2019	soft
5.13. 13.0-13.2 Only: Linux VM might fail to return a license after shutdown if the lice server is specified by its name	ense

5.14. Ubuntu guest driver initialization fails with vGPUs and GPUs that support SR-IO. 1.5.	56
5.15. 13.0-13.6 Only: NVIDIA Control Panel is started only for the RDP user that logs on first	56
5.16. 13.0-13.2 Only: Windows vGPU VM sometimes crashes after guest OS upgrade5	
5.17. 13.0-13.2 Only: Memory leaks in the vGPU manager plugin cause the VM to hang	
5.18. nvidia-smi ignores the second NVIDIA vGPU device added to a Microsoft Windows	
Server 2016 VM	59
5.19. Desktop session freezes when a VM is migrated to or from a host running an NVIDIA vGPU software 14 release	60
5.20. The reported NVENC frame rate is double the actual frame rate	.61
5.21. VM hangs after vGPU migration from a host running a vGPU manager 11 release to a host running a vGPU manager 13 release	61
5.22. Windows 2012 R2 licensed clients cannot acquire licenses from a DLS instance 6	32
5.23. 13.0 Only: Windows 2012 R2 licensed clients cannot acquire licenses from a CLS or DLS instance	63
5.24. Desktop session freezes when a VM is migrated to or from a host running an NVIDIA vGPU software 11 release	63
5.25. Hypervisor host reboots when multiple cloned VMs are simultaneously powered on or migrated	64
5.26. NVENC does not work with Teradici Cloud Access Software on Windows	. 65
5.27. A licensed client might fail to acquire a license if a proxy is set	. 65
5.28. Session connection fails with four 4K displays and NVENC enabled on a 2Q, 3Q, or 4Q vGPU	66
5.29. Disconnected sessions cannot be reconnected or might be reconnected very slowly with NVWMI installed	67
5.30. Migrating a VM with a Tesla T4 vGPU between a host running NVIDIA vGPU software 11.3 and a host running a different release fails	. 68
5.31. Linux VM hangs after vGPU migration to a host running a newer vGPU manager	
version	69
5.32. Idle Teradici Cloud Access Software session disconnects from Linux VM	
5.33. No virtual GPU types are listed in Citrix XenCenter	70
5.34. NVIDIA vGPU software graphics driver fails to load after upgrade on Citrix Hypervisor	71
5.35. Windows guest VMs with vGPUs or GPUs with large BAR memory settings fail to boot to the desktop in UEFI mode	72
5.36. Migrated VM with more than two vGPUs crashes on destination host	73
5.37. Driver upgrade in a Linux guest VM with multiple vGPUs might fail	73
5.38. NVIDIA Control Panel fails to start if launched too soon from a VM without licensing information.	74

5.39. VNC client session goes blank and console VNC is corrupted when the guest driver is uninstalled	
5.41. Suspend and resume between hosts running different versions of the vGPU	
manager fails	
5.42. On Linux, the frame rate might drop to 1 after several minutes	
5.43. Citrix XenCenter displays misleading information about vGPU types	
5.44. Citrix Virtual Apps and Desktops connection freezes initially	78
5.45. DWM crashes randomly occur in Windows VMs	
<ul><li>5.46. NVIDIA Control Panel fails to launch in a platform layer or published image</li><li>5.47. Remote desktop session freezes with assertion failure and XID error 43 after migration</li></ul>	
5.48. Citrix Virtual Apps and Desktops session freezes when the desktop is unlocked	
5.49. NVIDIA vGPU software graphics driver fails after Linux kernel upgrade with DKMS enabled	
5.50. On Citrix Hypervisor, all vGPUs in a VM must be of the same type	82
5.51. Console VGA cannot be disabled	83
5.52. Console VNC is unusable with Xorg on multiple vGPUs in a VM	84
5.53. Migration of VMs configured with vGPU stops before the migration is complete	84
5.54. ECC memory settings for a vGPU cannot be changed by using NVIDIA X Server Settings	85
5.55. Changes to ECC memory settings for a Linux vGPU VM by nvidia-smi might be ignored	85
5.56. Incorrect GPU type shown for Quadro RTX 8000 GPUs in Citrix XenCenter	86
5.57. NVIDIA Notification Icon prevents log off of Citrix Published Application user sessions	87
5.58. Host core CPU utilization is higher than expected for moderate workloads	87
5.59. Frame capture while the interactive logon message is displayed returns blank	
screen	
5.60. RDS sessions do not use the GPU with some Microsoft Windows Server releases	89
5.61. When the scheduling policy is fixed share, GPU utilization is reported as higher than expected	89
5.62. nvidia-smi reports that vGPU migration is supported on all hypervisors	90
5.63. NVIDIA Control Panel Crashes in a VM connected to two 4K displays	91
5.64. vGPU guest VM driver not properly loaded on servers with more than 512 GB or 1 TB or more of system memory	91
5.65. Luxmark causes a segmentation fault on an unlicensed Linux client	
5.66. A segmentation fault in DBus code causes nvidia-gridd to exit on Red Hat	
Enterprise Linux and CentOS	93
5.67. No Manage License option available in NVIDIA X Server Settings by default	94
5.68. Licenses remain checked out when VMs are forcibly powered off	95

5.69. Memory exhaustion can occur with vGPU profiles that have 512 Mbytes or les of frame buffer	
5.70. VM bug checks after the guest VM driver for Windows 10 RS2 is installed	97
5.71. On Citrix Hypervisor 7.0, VMs unexpectedly reboot and Citrix Hypervisor crashe or freezes	
5.72. With no NVIDIA driver installed, Citrix Hypervisor misidentifies Tesla M10 cards.	98
5.73. GNOME Display Manager (GDM) fails to start on Red Hat Enterprise Linux 7.2 an CentOS 7.0	
5.74. Video goes blank when run in loop in Windows Media Player	100
5.75. Local VGA console is momentarily unblanked when Citrix Virtual Apps an Desktops changes resolution of the VM desktop	
5.76. VM bugchecks on shutdown/restart when Citrix Virtual Apps and Desktops installed and NVIDIA driver is uninstalled or upgraded	
5.77. Application frame rate may drop when running Citrix Virtual Apps and Desktop at 2560×1600 resolution	
5.78. Windows VM BSOD	102
5.79. Windows VM BSOD when upgrading NVIDIA drivers over a Citrix Virtual Apps an Desktops session	
5.80. XenCenter does not allow vGPUs to be selected as a GPU type for Linux VMs	. 104
5.81. If X server is killed on a RHEL7 VM running vGPU, XenCenter console may no automatically switch to text console	
5.82. Citrix Virtual Apps and Desktops shows only a black screen when connected t 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.



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

# 1.1. NVIDIA vGPU Software Driver Versions

Each release in this release family of NVIDIA vGPU software includes a specific version of the NVIDIA Virtual GPU Manager, NVIDIA Windows driver, and NVIDIA Linux driver.

NVIDIA vGPU Software Version	NVIDIA Virtual GPU Manager Version	NVIDIA Windows Driver Version	NVIDIA Linux Driver Version
13.10	470.239.01	474.82	470.239.06
13.9	470.223.02	474.64	470.223.02
13.8	470.199.03	474.44	470.199.02
13.7	470.182.02	474.30	470.182.03
13.6	470.161.02	474.14	470.161.03
13.5	470.161.02	474.04	470.161.03
13.4	470.141.05	473.81	470.141.03
13.3	470.129.04	473.47	470.129.06
13.2	470.103.02	472.98	470.103.01
13.1	470.82	472.39	470.82.01
13.0	470.63	471.68	470.63.01

For details of which Citrix Hypervisor releases are supported, see <u>Hypervisor Software</u> Releases.

# 1.2. Compatibility Requirements for the NVIDIA vGPU Manager and Guest **VM** Driver

The releases of the NVIDIA vGPU Manager and guest VM drivers that you install must be compatible. If you install an incompatible guest VM driver release for the release of the vGPU Manager that you are using, the NVIDIA vGPU fails to load. See VM running an incompatible NVIDIA vGPU quest driver fails to initialize vGPU when booted.



Note: You must use NVIDIA License System with every release in this release family of NVIDIA vGPU software. The legacy NVIDIA vGPU software license server has reached end of life (EOL) and is no longer supported.

## Compatible NVIDIA vGPU Manager and Guest VM Driver Releases

The following combinations of NVIDIA vGPU Manager and guest VM driver releases are compatible with each other.

- NVIDIA vGPU Manager with guest VM drivers from the same release
- NVIDIA vGPU Manager with guest VM drivers from different releases within the same major release branch
- NVIDIA vGPU Manager from a later major release branch with guest VM drivers from the previous branch
- NVIDIA vGPU Manager from a later long-term support branch with guest VM drivers from the previous long-term support branch



#### Note:

When NVIDIA vGPU Manager is used with guest VM drivers from a different release within the same branch or from the previous branch, the combination supports only the features, hardware, and software (including guest OSes) that are supported on both releases.

For example, if vGPU Manager from release 13.10 is used with guest drivers from release 11.2, the combination does not support Red Hat Enterprise Linux 7.6 because NVIDIA vGPU software release 13.10 does not support Red Hat Enterprise Linux 7.6.

The following table lists the specific software releases that are compatible with the components in the NVIDIA vGPU software 13 major release branch.

NVIDIA vGPU Software Component	Releases	Compatible Software Releases
NVIDIA vGPU Manager	13.0 through 13.10	<ul> <li>Guest VM driver releases 13.0 through 13.10</li> <li>All guest VM driver 12.x releases</li> <li>All guest VM driver 11.x releases</li> </ul>
Guest VM drivers	13.0 through 13.10	NVIDIA vGPU Manager releases 13.0 through 13.10

## Incompatible NVIDIA vGPU Manager and Guest VM Driver Releases

The following combinations of NVIDIA vGPU Manager and guest VM driver releases are incompatible with each other.

- NVIDIA vGPU Manager from a later major release branch with guest VM drivers from a production branch two or more major releases before the release of the vGPU Manager
- NVIDIA vGPU Manager from an earlier major release branch with guest VM drivers from a later branch

The following table lists the specific software releases that are incompatible with the components in the NVIDIA vGPU software 13 major release branch.

NVIDIA vGPU Software Component	Releases	Incompatible Software Releases
NVIDIA vGPU Manager	13.0 through 13.10	All guest VM driver releases 10.x and earlier
Guest VM drivers	13.0 through 13.10	All NVIDIA vGPU Manager releases 12.x and earlier

# Updates in Release 13.10

### New Features in Release 13.10

- Security updates see Security Bulletin: NVIDIA GPU Display Driver February 2024, which is posted shortly after the release date of this software and is listed on the **NVIDIA Product Security page**
- Miscellaneous bug fixes

## Newly Supported Hardware and Software in Release 13.10

- Newly supported guest OS releases:
  - Red Hat Enterprise Linux 8.9
- Newly supported remoting solutions:
  - Citrix Virtual Apps and Desktops version 7 2311

### **Updates in Release 13.9** 1.4.

### New Features in Release 13.9

- Security updates see Security Bulletin: NVIDIA GPU Display Driver October 2023, which is posted shortly after the release date of this software and is listed on the **NVIDIA Product Security page**
- Miscellaneous bug fixes

### Newly Supported Hardware and Software in Release 13.9

- Newly supported remoting solutions:
  - Citrix Virtual Apps and Desktops version 7 2308

### Feature Support Withdrawn in Release 13.9

- Guest OSes no longer supported:
  - Windows Server 2012 R2

# Updates in Release 13.8

### New Features in Release 13.8

- Security updates see Security Bulletin: NVIDIA GPU Display Driver June 2023, which is posted shortly after the release date of this software and is listed on the NVIDIA **Product Security** page
- Miscellaneous bug fixes

## Hardware and Software Support Introduced in Release 13.8

- Support for Red Hat Enterprise Linux 8.8 as a guest OS
- Support for Citrix Virtual Apps and Desktops version 7 2305

### Feature Support Withdrawn in Release 13.8

▶ Red Hat Enterprise Linux 8.7 and 8.4 are no longer supported as a guest OS.

### Updates in Release 13.7 1.6.

### New Features in Release 13.7

- Security updates see Security Bulletin: NVIDIA GPU Display Driver March 2023, which is posted shortly after the release date of this software and is listed on the NVIDIA Product Security page
- Miscellaneous bug fixes

### Hardware and Software Support Introduced in Release 13.7

- Support for Windows 10 2022 Update (22H2) as a guest OS
- Support for Citrix Virtual Apps and Desktops version 7 2303

# Updates in Release 13.6

#### New Features in Release 13.6

Security updates - see Security Bulletin: NVIDIA GPU Display Driver - November 2022, which is updated shortly after the release date of this software and is listed on the **NVIDIA Product Security page** 

#### **Updates in Release 13.5** 1.8.

#### New Features in Release 13.5

- Security updates see Security Bulletin: NVIDIA GPU Display Driver November 2022, which is posted shortly after the release date of this software and is listed on the **NVIDIA** Product Security page
- Support for non-transparent local proxy servers when NVIDIA vGPU software is served licenses by a Cloud License Service (CLS) instance
- Miscellaneous bug fixes

### Hardware and Software Support Introduced in Release 13.5

Support for Red Hat Enterprise Linux 8.7 as a guest OS

Support for Citrix Virtual Apps and Desktops version 7 2209

### Features Deprecated in Release 13.5



Note: All versions of Microsoft Windows Server 2016 remain supported as a guest OS. The earlier statement that Microsoft Windows Server 2016 is deprecated is incorrect.

#### Updates in Release 13.4 1.9.

### New Features in Release 13.4

- Security updates see Security Bulletin: NVIDIA GPU Display Driver August 2022, which is posted shortly after the release date of this software and is listed on the NVIDIA **Product Security** page
- Miscellaneous bug fixes

### Hardware and Software Support Introduced in Release 13.4

Support for Citrix Virtual Apps and Desktops version 7 2206

### Feature Support Withdrawn in Release 13.4

Citrix XenServer 7.1 is no longer supported.

# 1.10. Updates in Release 13.3

#### New Features in Release 13.3

- Security updates see Security Bulletin: NVIDIA GPU Display Driver May 2022, which is posted shortly after the release date of this software and is listed on the NVIDIA **Product Security page**
- Miscellaneous bug fixes

### Hardware and Software Support Introduced in Release 13.3

- Support for Red Hat Enterprise Linux 8.6 as a guest OS
- Support for Citrix Virtual Apps and Desktops version 7 2203

## Feature Support Withdrawn in Release 13.3

Red Hat Enterprise Linux 8.5 and 8.2 are no longer supported as a guest OS.

# 1.11. Updates in Release 13.2

### New Features in Release 13.2

- Security updates see Security Bulletin: NVIDIA GPU Display Driver February 2022, which is posted shortly after the release date of this software and is listed on the **NVIDIA Product Security page**
- Miscellaneous bug fixes

### Hardware and Software Support Introduced in Release 13.2

- Support for Red Hat Enterprise Linux 8.5 as a quest OS
- Support for Citrix Virtual Apps and Desktops version 7 2112

### Feature Support Withdrawn in Release 13.2

- Red Hat Enterprise Linux 8.1 is no longer supported as a guest OS.
- Red Hat Enterprise Linux 7.8 and 7.7 are no longer supported as a guest OS.

# 1.12. Updates in Release 13.1

### New Features in Release 13.1

- Support for CUDA profilers on vGPUs on the following GPUs:
  - NVIDIA A40
  - NVIDIA A16
  - NVIDIA A10
  - NVIDIA RTX A6000
  - NVIDIA RTX A5000
- NVIDIA Deep Learning Super Sampling (DLSS) support on NVIDIA RTX Virtual Workstation
- Security updates see Security Bulletin: NVIDIA GPU Display Driver October 2021, which is available on the release date of this software and is listed on the NVIDIA **Product Security** page
- Miscellaneous bug fixes

## Hardware and Software Support Introduced in Release 13.1

Support for Citrix Virtual Apps and Desktops version 7 2109

# 1.13. Updates in Release 13.0

#### New Features in Release 13.0

- Support for the following NVIDIA CUDA Toolkit features on NVIDIA vGPU:
  - Development tools such as IDEs, debuggers, profilers, and utilities as listed under CUDA Toolkit Major Components in NVIDIA CUDA Toolkit Release Notes for CUDA 11.4
  - Tracing and profiling through the CUDA Profiling Tools Interface (CUPTI)
- Compatibility with the guest VM drivers from the previous long-term support branch (11)
- NVIDIA License System support
- Miscellaneous bug fixes

### Hardware and Software Support Introduced in Release 13.0

- Support for the following GPUs:
  - NVIDIA A100 PCIe 80GB
  - NVIDIA A30
  - NVIDIA A16
- Support for Citrix Virtual Apps and Desktops version 7 2106

# 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 R470 drivers.

# Supported NVIDIA GPUs and Validated Server Platforms

This release of NVIDIA vGPU software on Citrix Hypervisor provides support for several NVIDIA GPUs running on validated server hardware platforms. For a list of validated server platforms, refer to NVIDIA Virtual GPU Certified Servers.

The supported products for each type of NVIDIA vGPU software deployment depend on the GPU.

### GPUs Based on the NVIDIA Ampere Architecture

	Supported NVIDIA vGPU Software Products 1'2'3'4			
GPU	Time-Sliced NVIDIA vGPU	MIG-Backed NVIDIA vGPU	GPU Pass Through	
NVIDIA A100 PCIe 80GB	N/A	N/A	vCS	
NVIDIA A100 HGX 80GB	N/A	N/A	vCS	
NVIDIA A100 PCIe 40GB	N/A	N/A	vCS	
NVIDIA A100 HGX 40GB	N/A	N/A	vCS	
NVIDIA A40 <u>5</u>	► vWS ► vPC	N/A	► vCS ► vWS	

	Supported NVIDIA vGPU Software Products 1'2'3'4		
GPU	Time-Sliced NVIDIA vGPU	MIG-Backed NVIDIA vGPU	GPU Pass Through
	▶ vApps		► vApps
NVIDIA A30	N/A	N/A	vCS
NVIDIA A16	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>
NVIDIA A10	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>
NVIDIA RTX A6000 <u>5</u>	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>
NVIDIA RTX A5000 <u>5</u>	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>

# GPUs Based on the NVIDIA Turing $^{\mathsf{TM}}$ Architecture

	Supported NVIDIA vGPU Software Products 1'2'3'4			
GPU	Time-Sliced NVIDIA vGPU	MIG-Backed NVIDIA vGPU	GPU Pass Through	
Tesla T4	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>	
Quadro RTX 6000 <u>5</u>	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>	
Quadro RTX 6000 passive <u>5</u>	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>	

	Supported NVIDIA vGPU Software Products 1'2'3'4			
GPU	Time-Sliced NVIDIA vGPU	MIG-Backed NVIDIA vGPU	GPU Pass Through	
Quadro RTX 8000 <u>5</u>	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>	
Quadro RTX 8000 passive <u>5</u>	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>	

# GPUs Based on the NVIDIA Volta Architecture

	Supported NVIDIA vGPU Software Products 1'2'3'4		
GPU	Time-Sliced NVIDIA vGPU	MIG-Backed NVIDIA vGPU	GPU Pass Through
Tesla V100 SXM2	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>
Tesla V100 SXM2 32GB	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>
Tesla V100 PCIe	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>
Tesla V100 PCIe 32GB	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>
Tesla V100S PCIe 32GB	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>
Tesla V100 FHHL	<ul><li>VWS</li><li>VPC</li></ul>	N/A	► vCS ► vWS

	Supported NVIDIA vGPU Software Products 1'2'3'4		
GPU	Time-Sliced NVIDIA vGPU	MIG-Backed NVIDIA vGPU	GPU Pass Through
	vApps		► vApps

# GPUs Based on the NVIDIA Pascal<sup>™</sup> Architecture

**Note:** XenMotion with vGPU is **not** supported on any variant of the Tesla P100 GPU.

	Supported NVIDIA vGPU Software Products 1'2'3'4		
GPU	Time-Sliced NVIDIA vGPU	MIG-Backed NVIDIA vGPU	GPU Pass Through
Tesla P4	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>
Tesla P6	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>
Tesla P40	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>
Tesla P100 PCle 16 GB	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>
Tesla P100 SXM2 16 GB	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>
Tesla P100 PCIe 12GB	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vCS</li><li>vWS</li><li>vApps</li></ul>

# GPUs Based on the NVIDIA Maxwell<sup>™</sup> Graphic Architecture



Note: NVIDIA Virtual Compute Server (vCS) is not supported on GPUs based on the NVIDIA Maxwell graphic architecture.

	Supported NVIDIA vGPU Software Products 1'2'3'4		
GPU	Time-Sliced NVIDIA vGPU	MIG-Backed NVIDIA vGPU	GPU Pass Through
Tesla M6	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vWS</li><li>vApps</li></ul>
Tesla M10	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vWS</li><li>vApps</li></ul>
Tesla M60	<ul><li>vWS</li><li>vPC</li><li>vApps</li></ul>	N/A	<ul><li>vWS</li><li>vApps</li></ul>

# Switching the Mode of a GPU that Supports Multiple Display Modes

Some GPUs support displayless and display-enabled modes but must be used in NVIDIA vGPU software deployments in displayless mode.

The GPUs listed in the following table support multiple display modes. As shown in the table, some GPUs are supplied from the factory in displayless mode, but other GPUs are supplied in a display-enabled mode.

- vCS: NVIDIA Virtual Compute Server
- vWS: NVIDIA RTX Virtual Workstation
- vPC: NVIDIA Virtual PC
- vApps: NVIDIA Virtual Applications

The supported products are as follows:

<sup>&</sup>lt;sup>2</sup> N/A indicates that the deployment is not supported.

<sup>&</sup>lt;sup>3</sup> vCS is supported only on Linux operating systems.

<sup>&</sup>lt;sup>4</sup> vApps is supported only on Windows operating systems.

<sup>&</sup>lt;sup>5</sup> This GPU is supported only in displayless mode. In displayless mode, local physical display connectors are disabled.

GPU	Mode as Supplied from the Factory	
NVIDIA A40	Displayless	
NVIDIA RTX A5000	Display enabled	
NVIDIA RTX A6000	Display enabled	

A GPU that is supplied from the factory in displayless mode, such as the NVIDIA A40 GPU. might be in a display-enabled mode if its mode has previously been changed.

To change the mode of a GPU that supports multiple display modes, use the displaymodeselector tool, which you can request from the <a href="NVIDIA Display Mode">NVIDIA Display Mode</a> Selector Tool page on the NVIDIA Developer website.



#### Note:

Only the following GPUs support the displaymodeselector tool:

- NVIDIA A40
- NVIDIA RTX A5000
- NVIDIA RTX A6000

Other GPUs that support NVIDIA vGPU software do not support the displaymodeselector tool and, unless otherwise stated, do not require display mode switching.

# 2.1.2. Switching the Mode of a Tesla M60 or M6 **GPU**

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. If you are unsure which mode your GPU is in, use the gpumodeswitch tool to find out the mode.



#### Note:

Only Tesla M60 and M6 GPUs support the gpumodeswitch tool. Other GPUs that support NVIDIA vGPU do not support the qpumodeswitch tool and, except as stated in Switching the Mode of a GPU that Supports Multiple Display Modes, do not require mode switching.

Even in compute mode, Tesla M60 and M6 GPUs do not support NVIDIA Virtual Compute Server vGPU types. Furthermore, vCS is not supported on any GPU on Citrix Hypervisor.

For more information, refer to gpumodeswitch User Guide.

# **Hypervisor Software Releases**

### Changes to Citrix Product Names and Release Numbers

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

Old Name	New Name
XenServer	Citrix Hypervisor
XenApp and XenDesktop	Citrix Virtual Apps and Desktops
Citrix Receiver	Citrix Workspace App

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:

- 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.



#### Note:

Support for NVIDIA vGPU software requires the Premium Edition (previously Enterprise Edition) of Citrix Hypervisor. For details, see <u>Licensing</u> 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.

Software	Releases Supported	Notes
Citrix Hypervisor 8.2	RTM build and compatible cumulative update releases	This release supports XenMotion with vGPU on suitable GPUs as listed in Supported NVIDIA GPUs and Validated Server Platforms. The following GPUs are supported in GPU pass through mode only:

Software	Releases Supported	Notes
		<ul> <li>NVIDIA A100 PCIe 80GB</li> <li>NVIDIA A100 HGX 80GB</li> <li>NVIDIA A100 PCIe 40GB</li> <li>NVIDIA A100 HGX 40GB</li> <li>NVIDIA A30</li> <li>vCS is <b>not</b> supported.</li> </ul>
13.0-13.3 only: Citrix XenServer 7.1	RTM build and compatible cumulative update releases	Not supported on the following GPUs:  Tesla T4  Quadro RTX 6000 passive  Quadro RTX 8000 passive  NVIDIA RTX A6000  NVIDIA RTX A5000  NVIDIA A40  NVIDIA A40  NVIDIA A16  NVIDIA A10  NVIDIA A10  NVIDIA A100 PCIe 80GB  NVIDIA A100 PCIe 40GB  NVIDIA A100 PCIE 40GB

# 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.

Software	Releases Supported
Citrix Virtual Apps and Desktops	Note: All versions supported by earlier NVIDIA vGPU software 13 releases are also supported.

Software	Releases Supported		
	Since 13.10: Version 7 2311		
	<b>Since 13.9:</b> Version 7 2308		
	Since 13.8: Version 7 2305		
	Since 13.7: Version 7 2303		
	Since 13.5: Version 7 2209		
	Since 13.4: Version 7 2206		
	Since 13.3: Version 7 2203		
	Since 13.2: Version 7 2112		
	Since 13.1: Version 7 2109		
	Versions 7 2106, 7 2103, 7 2012, 7 2009, 7 2006, 7 2003, 7 1912, 7 1909, 7 1906, and 7 1903		
XenApp/XenDesktop	Version 7.15		
	Note: 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.		

### 2.3. **Guest OS Support**

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



#### Note:

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

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

#### 2.3.1. Windows Guest OS Support

NVIDIA vGPU software supports only the 64-bit Windows releases listed in the table as a guest OS on 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.



### Note:

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.

Guest OS	NVIDIA vGPU - Citrix Hypervisor Releases	Pass-Through GPU - Citrix Hypervisor Releases
Windows Server 2022	8.2	Since 13.4: 8.2
		<b>13.0-13.3 only:</b> 8.2, 7.1 cumulative update 2
Windows Server 2019	Since 13.4: 8.2	Since 13.4: 8.2
	<b>13.0-13.3 only:</b> 8.2, 7.1 cumulative update 2	<b>13.0-13.3 only:</b> 8.2, 7.1 cumulative update 2
Windows Server 2016 1709, 1607	Since 13.4: 8.2	Since 13.4: 8.2
	<b>13.0-13.3 only:</b> 8.2, 7.1	<b>13.0-13.3 only:</b> 8.2, 7.1
<b>13.0-13.8 only:</b> Windows Server 2012 R2 (not supported on GPUs based on architectures	<b>13.4-13.8</b> only: 8.2	<b>13.4-13.8 only:</b> 8.2
after the NVIDIA Turing™ architecture)	<b>13.0-13.3 only:</b> 8.2, 7.1	<b>13.0-13.3 only:</b> 8.2, 7.1
Since 13.7: Windows 10 2022 Update (22H2) and all Windows 10 releases supported by Microsoft up to and including this release	8.2	8.2
See Note (1)		
<b>13.1-13.6 only:</b> Windows 10 November 2021	<b>13.4-13.6</b> only: 8.2	<b>13.4-13.6 only:</b> 8.2
Update (21H2) and all Windows 10 releases supported by Microsoft up to and including this release	<b>13.1-13.3 only:</b> 8.2, 7.1	<b>13.1-13.3 only:</b> 8.2, 7.1 cumulative update 2
See Note (1)		
13.0 only: Windows 10 May 2021 Update (21H1) and all Windows 10 releases supported by Microsoft up to and including this release See Note (1)	8.2, 7.1	8.2, 7.1



Note:



1. The hardware-accelerated GPU scheduling feature introduced in Windows 10 May 2020 Update (2004) is not supported on GPUs based on the Maxwell architecture and is supported only in pass-through mode on GPUs based on later architectures.

#### 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.



#### Note:

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.

Guest OS	NVIDIA vGPU - Citrix Hypervisor Releases	Pass-Through GPU - Citrix Hypervisor Releases
Since 13.10: Red Hat Enterprise Linux 8.9	8.2	8.2
Since 13.8: Red Hat Enterprise Linux 8.8	8.2	8.2
13.5-13.7 only: Red Hat Enterprise Linux 8.7	8.2	8.2
Since 13.3: Red Hat Enterprise Linux 8.6	8.2	8.2
13.2 only: Red Hat Enterprise Linux 8.5	8.2	8.2
13.0-13.7 only: Red Hat Enterprise Linux 8.4	8.2	8.2
13.0-13.2 only: Red Hat Enterprise Linux 8.2	8.2	8.2
13.0, 13.1 only: Red Hat Enterprise Linux 8.1	8.2	8.2
Since 13.2: CentOS Linux 8 (2111)	8.2	8.2
CentOS Linux 8 (2105)	8.2	8.2
13.0-13.2 only: CentOS Linux 8 (2004)	8.2	8.2
<b>13.0, 13.1 only:</b> CentOS Linux 8 (1911)	8.2	8.2
Since 13.2: Red Hat Enterprise Linux 7.9	Since 13.4: 8.2	Since 13.4: 8.2
	<b>13.2, 13.3 only:</b> 8.2, 7.1	<b>13.2, 13.3 only:</b> 8.2, 7.1
<b>13.0, 13.1 only:</b> Red Hat Enterprise Linux 7.7-7.9	8.2, 7.1	8.2, 7.1
CentOS 7.6-7.8	Since 13.4: 8.2	Since 13.4: 8.2
	<b>13.0-13.3 only:</b> 8.2, 7.1	<b>13.0-13.3 only:</b> 8.2, 7.1
Ubuntu 20.04 LTS	8.2	8.2

Guest OS	NVIDIA vGPU - Citrix Hypervisor Releases	Pass-Through GPU - Citrix Hypervisor Releases
Ubuntu 18.04 LTS	Since 13.4: 8.2 13.0-13.3 only: 8.2, 7.1	Since 13.4: 8.2 13.0-13.3 only: 8.2, 7.1
Ubuntu 16.04 LTS	cumulative update 2  Since 13.4: 8.2  13.0-13.3 only: 8.2, 7.1	cumulative update 2  Since 13.4: 8.2  13.0-13.3 only: 8.2, 7.1
Ubuntu 14.04 LTS	Since 13.4: 8.2 13.0-13.3 only: 8.2, 7.1	Since 13.4: 8.2 13.0-13.3 only: 8.2, 7.1

# 2.4. NVIDIA CUDA Toolkit Version Support

The releases in this release family of NVIDIA vGPU software support NVIDIA CUDA Toolkit 11.4.

To build a CUDA application, the system must have the NVIDIA CUDA Toolkit and the libraries required for linking. For details of the components of NVIDIA CUDA Toolkit, refer to NVIDIA CUDA Toolkit Release Notes for CUDA 11.4.

To run a CUDA application, the system must have a CUDA-enabled GPU and an NVIDIA display driver that is compatible with the NVIDIA CUDA Toolkit release that was used to build the application. If the application relies on dynamic linking for libraries, the system must also have the correct version of these libraries.

For more information about NVIDIA CUDA Toolkit, refer to CUDA Toolkit 11.4 Documentation.



#### Note:

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



Note: vGPU migration is disabled for a VM for which any of the following NVIDIA CUDA Toolkit features is enabled:

- Unified memory
- Debuggers
- Profilers

## Supported GPUs

- Tesla M6
- Tesla M10
- Tesla M60
- Tesla P4
- Tesla P6
- Tesla P40
- Tesla V100 SXM2
- Tesla V100 SXM2 32GB
- Tesla V100 PCle
- Tesla V100 PCIe 32GB
- Tesla V100S PCle 32GB
- Tesla V100 FHHL
- Tesla T4
- Quadro RTX 6000
- Quadro RTX 6000 passive
- Quadro RTX 8000
- Quadro RTX 8000 passive
- NVIDIA A10
- NVIDIA A16
- NVIDIA A40
- NVIDIA RTX A5000
- NVIDIA RTX A6000

## Supported Citrix Hypervisor Releases

Citrix Hypervisor 8.2 only.

## Supported Guest OS Releases

Windows only. XenMotion with vGPU is **not** supported on Linux.

## Known Issues with XenMotion with vGPU Support

Use Case	Affected GPUs	Issue
Migration to or from a host running an NVIDIA vGPU software 11 release before 11.6	Tesla T4	VM hangs after vGPU migration from a host running a vGPU manager 11 release to a host running a vGPU manager 13 release
<ul> <li>Migration from         a host running         NVIDIA vGPU         software 11.3 to         a host running a         different release</li> <li>Migration to a host         running NVIDIA         vGPU software         11.3 from a host         running a different         release</li> </ul>	Tesla T4	Migrating a VM with a Tesla T4 vGPU between a host running NVIDIA vGPU software 11.3 and a host running a different release fails
Migration from a host that is running a vGPU manager 11 release to a host that is running a vGPU manager 13 release.	<ul><li>Tesla T4</li><li>Tesla V100</li></ul>	Linux VM hangs after vGPU migration to a host running a newer vGPU manager version
Migration to or from a host running an NVIDIA vGPU software 11 release	GPUs based on the NVIDIA Volta <sup>™</sup> architecture	Desktop session freezes when a VM is migrated to or from a host running an NVIDIA vGPU software 11 release
Migration to or from a host running an NVIDIA vGPU software 14 release	<ul><li>Tesla T4</li><li>Tesla V100</li></ul>	Desktop session freezes when a VM is migrated to or from a host running an NVIDIA vGPU software 14 release

Use Case	Affected GPUs	Issue
Simultaneous migration of multiple VMs	All GPUs that support XenMotion with vGPU	Hypervisor host reboots when multiple cloned VMs are simultaneously powered on or migrated
Migration between hosts with different ECC memory configuration	All GPUs that support XenMotion with vGPU	Migration of VMs configured with vGPU stops before the migration is complete

### 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 vGPUs that are allocated all of the physical GPU's frame buffer are supported.

GPU Architecture	Board	vGPU
Ampere	NVIDIA A40	A40-48Q See Note ( <u>1</u> ).
		A40-48C See Note ( <u>1</u> ).
	NVIDIA A16	A16-16Q See Note ( <u>1</u> ).
		A16-16C See Note ( <u>1</u> ).
	NVIDIA A10	A10-24Q See Note ( <u>1</u> ).
		A10-24C See Note ( <u>1</u> ).
	NVIDIA RTX A6000	A6000-48Q See Note ( <u>1</u> ).
		A6000-48C See Note ( <u>1</u> ).
	NVIDIA RTX A5000	A5000-24Q See Note ( <u>1</u> ).
		A5000-24C See Note (1).
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

GPU Architecture	Board	vGPU
		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
	Tesla P40	P40-24Q
		P40-24C
	Tesla P6	P6-16Q
		P6-16C
	Tesla P4	P4-8Q
		P4-8C
Maxwell	Tesla M60	M60-8Q
	Tesla M10	M10-8Q
	Tesla M6	M6-8Q



Note:

1. This type of vGPU cannot be assigned with other types of vGPU to the same VM.

### 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 8.2 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 are supported only on a subset of vGPUs, Citrix Hypervisor releases, and quest 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.

GPU Architecture	Board	vGPU
Ampere	NVIDIA A40	A40-48Q
		A40-48C
	NVIDIA A10	A10-24Q
		A10-24C
	NVIDIA RTX A6000	A6000-48Q
		A6000-48C
	NVIDIA RTX A5000	A5000-24Q
		A5000-24C
Turing	Quadro RTX 6000	RTX6000-24Q
		RTX6000-24C
	Quadro RTX 6000 passive	RTX6000P-24Q
		RTX6000P-24C
	Quadro RTX 8000	RTX8000-48Q
		RTX8000-48C

GPU Architecture	Board	vGPU
	Quadro RTX 8000 passive	RTX8000P-48Q
		RTX8000P-48C
Volta	Tesla V100 SXM2 32GB	V100DX-32Q
		V100DX-32C
	Tesla V100 SXM2	V100X-16Q
		V100X-16C
Pascal	scal Tesla P100 SXM2	P100X-16Q
		P100X-16C

### 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.
- PCle is not supported.
- SLI is not supported.

### **Unified Memory Support** 2.8.

Unified memory is a single memory address space that is accessible from any CPU or GPU in a system. It creates a pool of managed memory that is shared between the CPU and GPU to provide a simple way to allocate and access data that can be used by code running on any CPU or GPU in the system. Unified memory is supported only on a subset of vGPUs and guest OS releases.



Note: Unified memory is disabled by default. If used, you must enable unified memory individually for each vGPU that requires it by setting a vGPU plugin parameter.

## 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 unified memory are supported.

GPU Architecture	Board	vGPU
Ampere	NVIDIA A40	A40-48Q
		A40-48C
	NVIDIA A16	A16-16Q
		A16-16C
	NVIDIA A10	A10-24Q
		A10-24C
	NVIDIA RTX A6000	A6000-48Q
		A6000-48C
	NVIDIA RTX A5000	A5000-24Q
		A5000-24C

### Supported Guest OS Releases

Linux only. Unified memory is **not** supported on Windows.

### Limitations

- ▶ When unified memory is enabled for a VM, XenMotion with vGPU is disabled for the
- ▶ When unified memory is enabled for a VM, NVIDIA CUDA Toolkit profilers are disabled.

# Since 13.1: NVIDIA Deep Learning 2.9. Super Sampling (DLSS) Support

NVIDIA vGPU software supports NVIDIA DLSS on NVIDIA RTX Virtual Workstation.

**Supported DLSS versions:** 2.0. Version 1.0 is **not** supported.

### Supported GPUs:

- ▶ NVIDIA A40
- NVIDIA A16
- NVIDIA A10
- NVIDIA RTX A6000
- NVIDIA RTX A5000
- ▶ Tesla T4
- Quadro RTX 8000
- Quadro RTX 8000 passive

- Quadro RTX 6000
- Quadro RTX 6000 passive



Note: NVIDIA graphics driver components that DLSS requires are installed only if a supported GPU is detected during installation of the driver. Therefore, if the creation of VM templates includes driver installation, the template should be created from a VM that is configured with a supported GPU while the driver is being installed.

Supported applications: only applications that use nvngx\_dlss.dll version 2.0.18 or newer

# Chapter 3. Known Product Limitations

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

# **NVENC** does not support 3.1. resolutions greater than 4096×4096

### Description

The NVIDIA hardware-based H.264 video encoder (NVENC) does not support resolutions greater than 4096×4096. This restriction applies to all NVIDIA GPU architectures and is imposed by the GPU encoder hardware itself, not by NVIDIA vGPU software. The maximum supported resolution for each encoding scheme is listed in the documentation for NVIDIA Video Codec SDK. This limitation affects any remoting tool where H.264 encoding is used with a resolution greater than 4096×4096. Most supported remoting tools fall back to software encoding in such scenarios.

#### Workaround

If your GPU is based on a GPU architecture later than the NVIDIA Maxwell architecture, use H.265 encoding. H.265 is more efficient than H.264 encoding and has a maximum resolution of 8192×8192. On GPUs based on the NVIDIA Maxwell architecture, H.265 has the same maximum resolution as H.264, namely 4096×4096.



Note: Resolutions greater than 4096×4096 are supported only by the H.265 decoder that 64-bit client applications use. The H.265 decoder that 32-bit applications use supports a maximum resolution of 4096×4096.

Because the client-side Workspace App on Windows is a 32-bit application, resolutions greater than 4096×4096 are not supported for Windows clients of Citrix Virtual Apps

and Desktops. Therefore, if you are using a Windows client with Citrix Virtual Apps and Desktops, ensure that you are using H.264 hardware encoding with the default <u>Use video</u> codec for compression Citrix graphics policy setting, namely Actively Changing Regions. This policy setting encodes only actively changing regions of the screen (for example, a window in which a video is playing). Provided that the number of pixels along any edge of the actively changing region does not exceed 4096, H.264 encoding is offloaded to the NVENC hardware encoder.

# **Nested Virtualization Is Not** 3.2. Supported by NVIDIA vGPU

NVIDIA vGPU deployments do not support nested virtualization, that is, running a hypervisor in a guest VM. For example, enabling the Hyper-V role in a guest VM running the Windows Server OS is **not** supported because it entails enabling nested virtualization. Similarly, enabling Windows Hypervisor Platform is not supported because it requires the Hyper-V role to be enabled.

## Issues occur when the channels 3.3. allocated to a vGPU are exhausted

## Description

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

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

```
Jun 26 08:01:25 srvxen06f vgpu-3[14276]: error: vmiop log: (0x0): Guest attempted to
allocate channel above its max channel limit 0xfb
Jun 26 08:01:25 srvxen06f vgpu-3[14276]: error: vmiop log: (0x0): VGPU message 6
failed, result code: 0x1a
Jun 26 08:01:25 srvxen06f vgpu-3[14276]: error: vmiop log: (0x0):
0xcld004a1, 0xff0e0000, 0xff0400fb, 0xc36f,
Jun 26 08:01:25 srvxen06f vgpu-3[14276]: error: vmiop log: (0x0):
                                                                          0x1,
0xff1fe314, 0xff1fe038, 0x100b6f000, 0x1000,
Jun 26 08:01:25 srvxen06f vgpu-3[14276]: error: vmiop_log: (0x0):
0x80000000, 0xff0e0200, 0x0, 0x0, (Not logged),
Jun 26 08:01:25 srvxen06f vgpu-3[14276]: error: vmiop_log: (0x0):
                                                                          0x1, 0x0
Jun 26 08:01:25 srvxen06f vgpu-3[14276]: error: vmiop log: (0x0): , 0x0
```

### Workaround

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

## Total frame buffer for vGPUs is less 3.4. 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 quest 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:

max-reserved-fb = vgpu-profile-size-in-mb÷16 + 16 + ecc-adjustments + page-retirementallocation + compression-adjustment

#### 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-sizein-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 enabled on a physical GPU that does not have HBM2 memory eccadjustments is fb-without-ecc/16, which is equivalent to 64 Mbytes for every Gbyte of frame buffer assigned to the vGPU. fb-without-ecc is total amount of frame buffer with FCC disabled.

▶ If ECC is disabled or the GPU has HBM2 memory, ecc-adjustments is 0.

## page-retirement-allocation

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

- ▶ On GPUs based on the NVIDIA Maxwell GPU architecture, page-retirementallocation = 4÷max-vgpus-per-gpu.
- On GPUs based on NVIDIA GPU architectures after the Maxwell architecture, pageretirement-allocation = 128÷max-vgpus-per-gpu

### max-vgpus-per-gpu

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

## compression-adjustment

The amount of frame buffer in Mbytes that is reserved for the higher compression overhead in vGPU types with 12 Gbytes or more of frame buffer on GPUs based on the Turing architecture.

compression-adjustment depends on the vGPU type as shown in the following table.

vGPU Type	Compression Adjustment (MB)
T4-16Q	28
T4-16C	
T4-16A	
RTX6000-12Q	32
RTX6000-12C	
RTX6000-12A	
RTX6000-24Q	104
RTX6000-24C	
RTX6000-24A	
RTX6000P-12Q	32
RTX6000P-12C	
RTX6000P-12A	
RTX6000P-24Q	104
RTX6000P-24C	
RTX6000P-24A	
RTX8000-12Q	32
RTX8000-12C	
RTX8000-12A	
RTX8000-16Q	64

vGPU Type	Compression Adjustment (MB)
RTX8000-16C	
RTX8000-16A	
RTX8000-24Q	96
RTX8000-24C	
RTX8000-24A	
RTX8000-48Q	238
RTX8000-48C	
RTX8000-48A	
RTX8000P-12Q	32
RTX8000P-12C	
RTX8000P-12A	
RTX8000P-16Q	64
RTX8000P-16C	
RTX8000P-16A	
RTX8000P-24Q	96
RTX8000P-24C	
RTX8000P-24A	
RTX8000P-48Q	238
RTX8000P-48C	
RTX8000P-48A	

For all other vGPU types, compression-adjustment is 0.



Note: In VMs running Windows Server 2012 R2, which supports Windows Display Driver Model (WDDM) 1.x, an additional 48 Mbytes of frame buffer are reserved and not available for vGPUs.

## 3.5. Issues may occur with graphicsintensive 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.

## In pass through mode, all GPUs 3.6. 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.

# 3.7. 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.

## 3.8. 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

## VM running an incompatible NVIDIA vGPU guest driver fails to initialize vGPU when booted

## Description

A VM running a version of the NVIDIA quest VM driver that is incompatible with the current release of Virtual GPU Manager will fail to initialize vGPU when booted on a Citrix Hypervisor platform running that release of Virtual GPU Manager.

A guest VM driver is incompatible with the current release of Virtual GPU Manager in either of the following situations:

- The guest driver is from a release in a branch two or more major releases before the current release, for example release 9.4.
  - In this situation, the Citrix Hypervisor VM's /var/log/messages log file reports the following error:
  - vmiop log: (0x0): Incompatible Guest/Host drivers: Guest VGX version is older than the minimum version supported by the Host. Disabling vGPU.
- ▶ The guest driver is from a later release than the Virtual GPU Manager.
  - In this situation, the Citrix Hypervisor VM's /var/log/messages log file reports the following error:

```
vmiop\_log: (0x0): Incompatible Guest/Host drivers: Guest VGX version is newer
than the maximum version supported by the Host. Disabling vGPU.
```

In either situation, 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)

#### Resolution

Install a release of the NVIDIA quest VM driver that is compatible with current release of Virtual GPU Manager.

# 3.10. 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 vGPU's plugin parameters.

- Citrix Hypervisor 8.1 or later: Specify frame rate limiter=0 in the extra args parameter for the group to which the vGPU belongs:
  - [root@xenserver ~] # xe vgpu-param-set uuid=vgpu-uuid extra\_args=frame\_rate\_limiter=0
- ▶ Citrix Hypervisor earlier than 8.1: Specify frame rate limiter=0 in the VM's platform:vgpu extra args parameter:

```
[root@xenserver ~] # xe vm-param-set uuid=vm-uuid
platform:vgpu_extra_args="frame_rate_limiter=0"
```

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

With this setting in place, the VM's vGPU will run without any frame rate limit. The FRL can be reverted back to its default setting in one of the following ways:

- Removing the vGPU plugin's parameters:
  - Citrix Hypervisor 8.1 or later: Removing the extra args key the from group to which the vGPU belongs
  - Citrix Hypervisor earlier than 8.1: Removing the vgpu extra args key from the platform parameter

- Removing frame rate limiter=0 from the extra args or vgpu extra args key
- Setting frame rate limiter=1. For example:
  - Citrix Hypervisor 8.1 or later:

```
[root@xenserver ~] # xe vgpu-param-set uuid=vgpu-uuid
extra_args=frame_rate_limiter=1
```

Citrix Hypervisor earlier than 8.1:

```
[root@xenserver ~] # xe vm-param-set uuid=vm-uuid
platform:vgpu_extra_args="frame_rate_limiter=1"
```

# 3.11. nvidia-smi fails to operate when all GPUs are assigned to GPU passthrough mode

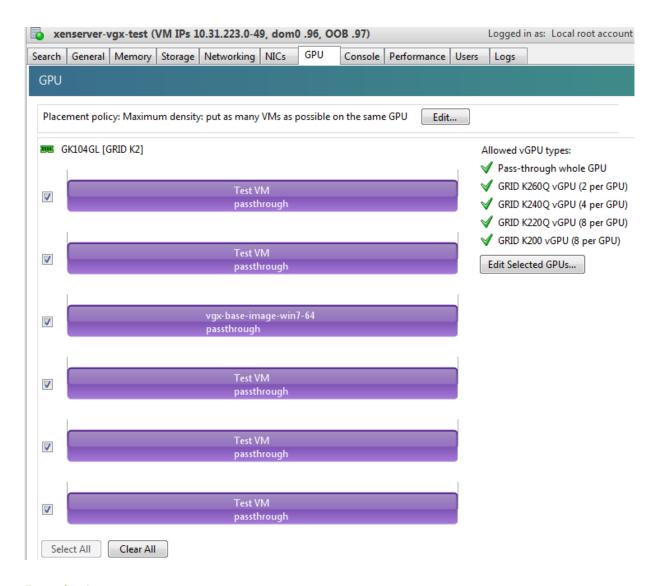
## Description

If all GPUs in the platform are assigned to VMs in pass-through mode, nvidia-smi will

```
[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

## 3.12. 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.

## 3.13. 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

```
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:

```
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 vGPU's plugin parameters:

- ► Citrix Hypervisor 8.1 or later: Specify enable large sys mem=1 in the extra args parameter for the group to which the vGPU belongs:
  - [root@xenserver ~] # xe vgpu-param-set uuid=vgpu-uuid extra\_args=enable\_large\_sys\_mem=1
- Citrix Hypervisor earlier than 8.1: Specify enable large sys mem=1 in the VM's platform:vgpu extra args parameter:

```
[root@xenserver ~] # xe vm-param-set uuid=vm-uuid
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:

Citrix Hypervisor 8.1 or later:

```
extra args="enable large sys mem=1,extra fb reservation=128"
```

Citrix Hypervisor earlier than 8.1:

```
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 plugin's parameters:
  - ▶ Citrix Hypervisor 8.1 or later: Removing the extra args key the from group to which the vGPU belongs
  - Citrix Hypervisor earlier than 8.1: Removing the vgpu extra args key from the platform parameter
- Removing enable large sys mem from the extra args or vgpu extra args key
- Setting enable large sys mem=0

## 3.14. 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 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 13.10

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

## Issues Resolved in Release 13.9

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

## Issues Resolved in Release 13.8

Bug ID	Summary and Description
3596327	13.0-13.7 Only: Remote desktop connection is lost and the NVIDIA vGPU software graphics driver is unloaded
	The remote desktop connection is lost and the NVIDIA vGPU software graphics driver is unloaded after an attempt to access a VM over RDP and VMware Horizon agent direct connect. After an attempt to log in again, a black screen is displayed.

## Issues Resolved in Release 13.7

Bug ID	Summary and Description
3334310	13.0-13.6 Only: NVIDIA Control Panel is started only for the RDP user that logs on first
	On all supported Windows Server guest OS releases, <b>NVIDIA Control Panel</b> is started only for the RDP user that logs on first. Other users cannot start <b>NVIDIA Control Panel</b> . If more than one RDP user is logged on when <b>NVIDIA Control Panel</b> is started, it always opens in the session of the RDP user that logged on first, irrespective of which user started <b>NVIDIA Control Panel</b> . Furthermore, on Windows Server 2016, <b>NVIDIA Control</b>

Bug ID	Summary and Description
	<b>Panel</b> crashes if a user session is disconnected and then reconnected while <b>NVIDIA Control Panel</b> is open.
3835855	13.0-13.6 Only: Windows VMs fail to acquire a license in environments with multiple active desktop sessions
	A race condition in the NVIDIA vGPU softwaregraphics driver for Windows can cause Windows VMs to fail to acquire a license. This issue occurs in environments where multiple active desktop sessions are trying to acquire a license simultaneously. When this issue occurs, the following error message is written to licensing event log on the client:  Mismatch between client and server with respect to licenses held. Returning the licenses
3966654	13.0-13.6 Only: Some graphics applications can cause a TDR on a Windows VM configured with an NVIDIA A16 vGPU
	Some graphics applications can cause a timeout detection and recovery (TDR) error on a Windows VM that is configured with a vGPU on an NVIDIA A16 GPU. The TDR might cause the application to crash in the VM.
3975182	13.4-13.6 Only: Destination host becomes unstable after migration of a VM running an NVIDIA vGPU software 11 release
	When a VM running an NVIDIA vGPU software 11 release is migrated between two hosts that are running NVIDIA vGPU software 13 releases, the destination host becomes unstable and might crash. When this issue occurs, error message XID error 38 might appear on the destination host. This issue occurs only if <b>all</b> the following conditions are met:
	The VM is running the graphics driver from any NVIDIA vGPU software 11 release.
	► The VM is configured with a vGPU that resides on a GPU that is based on the NVIDIA Turing architecture.
	The source host is running the Virtual GPU Manager from any NVIDIA vGPU software 13 release.
	The destination host is running the Virtual GPU Manager from a NVIDIA vGPU software release in the range 13.4 through 13.6.

## Issues Resolved in Release 13.6

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

## Issues Resolved in Release 13.5

Bug ID	Summary and Description
3658686	13.0-13.4 Only: VMs configured with a vGPU based on the NVIDIA  Ampere architecture can become slow to respond
	VMs configured with a vGPU on a GPU that is based on the NVIDIA Ampere GPU architecture can become slow to respond. When this error occurs, multiple XID error 62 and XID error 45 messages are written to the log file on the hypervisor host.

## Issues Resolved in Release 13.4

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

## Issues Resolved in Release 13.3

Bug ID	Summary and Description
200756399	13.0-13.2 Only: Linux VM might fail to return a license after shutdown if the license server is specified by its name
	If the license server is specified by its fully qualified domain name, a Linux VM might fail to return its license when the VM is shut down. This issue occurs if the nvidia-gridd service cannot resolve the fully qualified domain name of the license server because systemd-resolved.service is not available when the service attempts to return the license. When this issue occurs, the nvidia-gridd service writes the following message to the systemd journal:  General data transfer failure. Couldn't resolve host name
3465448	13.0-13.2 Only: Windows vGPU VM sometimes crashes after guest OS upgrade
	When a VM that is configured with NVIDIA vGPU is rebooted after an OS upgrade from Windows 10 1909 to Windows 10 20H2, the VM sometimes crashes. This issue is caused by a NULL pointer exception in the Virtual GPU Manager plugin (libnvidia-vgx.so). This NULL pointer exception might also cause the VM to crash in other situations. When this issue occurs, error messages that indicate that the Virtual GPU Manager process crashed are written to the standard activity log /var/log/messages on the hypervisor host.
200724807	13.0-13.2 Only: Memory leaks in the vGPU manager plugin cause the VM to hang

Bug ID	Summary and Description
	Applications running in a VM request memory to be allocated and freed
	by the vGPU manager plugin, which runs on the hypervisor host. When an
	application requests the vGPU manager plugin to free previously allocated
	memory, some of the memory is not freed. Some applications request
	memory more frequently than other applications. If such applications run
	for a long period of time, for example for two or more days, the failure to
	free all allocated memory might cause the hypervisor host to run out of
	memory. As a result, memory allocation for applications running in the VM
	might fail, causing the applications and, sometimes, the VM to hang.

## Issues Resolved in Release 13.2

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

## Issues Resolved in Release 13.1

Bug ID	Summary and Description
3392680	13.0 Only: Windows 2012 R2 licensed clients cannot acquire licenses from a CLS or DLS instance
	NVIDIA vGPU software licensed clients running in a Windows 2012 R2 VM cannot acquire licenses from a Cloud License Service (CLS) instance or a Delegated License Service (DLS) instance. During the license acquisition process, the vGPU licensing service compares the size of the message to be sent to the CLS or DLS instance before and after encryption. On Windows 2012 R2, the MSDN API for encrypting the message outputs a shorter encrypted message than the plain text message. As a result, the validation check in the service fails, which causes the client to fail to acquire a license.

## Issues Resolved in Release 13.0

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

## Chapter 5. Known Issues

## 5.1. Frame buffer seems to be missing from GPUs

## Description

On a host on which the Virtual GPU Manager is installed, GPU management tools, such as the nvidia-smi command, give the impression that some portion of a GPU's frame buffer is missing. For example, the NVIDIA A16 GPU has 16 GB of frame buffer, but total frame buffer is shown as 15.745 GB. This issue occurs because the Virtual GPU Manager does not report frame buffer that it has reserved for its own purposes, only the frame buffer that is available for applications.

## Version

This issue affects only releases in the NVIDIA vGPU software 13 branch.

### **Status**

Closed

### Ref. #

## 5.2. Graphics applications are corrupted on some Windows vGPU VMs

## Description

Graphics applications are corrupted on Windows VMs that are configured with one or more vGPUs that are based on the NVIDIA Ampere or NVIDIA Ada Lovelace GPU architecture.

## **Status**

Open

### Ref. #

3641947

## 5.3. **NVIDIA Control Panel** crashes if a user session is disconnected and reconnected

## Description

On all supported Windows Server guest OS releases, NVIDIA Control Panel crashes if a user session is disconnected and then reconnected while NVIDIA Control Panel is open.

### Version

This issue affects all supported Windows Server guest OS releases.

### **Status**

Open

### Ref. #

## 5.4. 13.0-13.7 Only: Remote desktop connection is lost and the NVIDIA vGPU software graphics driver is unloaded

## Description

The remote desktop connection is lost and the NVIDIA vGPU software graphics driver is unloaded after an attempt to access a VM over RDP and VMware Horizon agent direct connect. After an attempt to log in again, a black screen is displayed.

When this issue occurs, the following errors are written to the log files on the guest VM:

► A timeout detection and recovery (TDR) error:

```
vmiop_log: (0x0): Timeout occurred, reset initiated.
vmiop log: (0x0): TDR DUMP:0x52445456 0x006907d0 0x0000001cc 0x00000001
```

XID error 43:

```
vmiop log: (0x0): XID 43 detected on physical chid
```

vGPU error 22:

```
vmiop_log: (0x0): VGPU message 22 failed
```

Guest driver unloaded error:

```
vmiop log: (0x0): Guest driver unloaded!
```

## Workaround

To recover from this issue, reboot the VM.

Since 13.7: To prevent this issue from occurring, disable translation lookaside buffer (TLB) invalidation by setting the vGPU plugin parameter tlb invalidate enabled to 0.

### **Status**

Resolved in NVIDIA vGPU software 13.8

### Ref. #

## 5.5. Pixelation occurs on a Windows VM configured with a Tesla T4 vGPU

## Description

Users might experience poor graphics quality on a Windows VM that is configured with a vGPU on a Tesla T4 GPU. This issue can cause random pixelation on the entire screen, or only on some patches of the screen. No errors are reported or written to the log files when this issue occurs.

### Workaround

Contact NVIDIA Enterprise Support for assistance with a workaround for this issue.

### **Status**

Open

## Ref.#

3973158

## 13.4-13.6 Only: Destination host 5.6. becomes unstable after migration of a VM running an NVIDIA vGPU software 11 release

## Description

When a VM running an NVIDIA vGPU software 11 release is migrated between two hosts that are running NVIDIA vGPU software 13 releases, the destination host becomes unstable and might crash. When this issue occurs, error message XID error 38 might appear on the destination host. This issue occurs only if all the following conditions are met:

- The VM is running the graphics driver from any NVIDIA vGPU software 11 release.
- The VM is configured with a vGPU that resides on a GPU that is based on the NVIDIA Turing architecture.
- The source host is running the Virtual GPU Manager from any NVIDIA vGPU software 13 release.

▶ The destination host is running the Virtual GPU Manager from a NVIDIA vGPU software release in the range 13.4 through 13.6.

### Version

### **Status**

Resolved in NVIDIA vGPU software 13.7

### Ref. #

3975182

## 5.7. 13.0-13.6 Only: Windows VMs fail to acquire a license in environments with multiple active desktop sessions

## Description

A race condition in the NVIDIA vGPU softwaregraphics driver for Windows can cause Windows VMs to fail to acquire a license. This issue occurs in environments where multiple active desktop sessions are trying to acquire a license simultaneously. When this issue occurs, the following error message is written to licensing event log on the client: Mismatch between client and server with respect to licenses held. Returning the licenses

## Version

This issue affects only Windows guest VMs.

#### **Status**

Resolved in NVIDIA vGPU software 13.7

### Ref. #

## 5.8. 13.0-13.6 Only: Some graphics applications can cause a TDR on a Windows VM configured with an **NVIDIA A16 vGPU**

## Description

Some graphics applications can cause a timeout detection and recovery (TDR) error on a Windows VM that is configured with a vGPU on an NVIDIA A16 GPU. The TDR might cause the application to crash in the VM.

### **Status**

Resolved in NVIDIA vGPU software 13.7

#### Ref. #

3966654

13.0-13.4 Only: VMs configured 5.9. with a vGPU based on the NVIDIA Ampere architecture can become slow to respond

## Description

VMs configured with a vGPU on a GPU that is based on the NVIDIA Ampere GPU architecture can become slow to respond. When this error occurs, multiple XID error 62 and XID error 45 messages are written to the log file on the hypervisor host.

#### **Status**

Resolved in NVIDIA vGPU software 13.5

### Ref.#

# 5.10. NLS client fails to acquire a license with the error The allowed time to process response has expired

## Description

A licensed client of NVIDIA License System (NLS) fails to acquire a license with the error The allowed time to process response has expired. This error can affect clients of a Cloud License Service (CLS) instance or a Delegated License Service (DLS) instance.

This error occurs when the time difference between the system clocks on the client and the server that hosts the CLS or DLS instance is greater than 10 minutes. A common cause of this error is the failure of either the client or the server to adjust its system clock when daylight savings time begins or ends. The failure to acquire a license is expected to prevent clock windback from causing licensing errors.

### Workaround

Ensure that system clock time of the client and any server that hosts a DLS instance match the current time in the time zone where they are located.

To prevent this error from occurring when daylight savings time begins or ends, enable the option to automatically adjust the system clock for daylight savings time:

- Windows: Set the Adjust for daylight saving time automatically option.
- Linux: Use the hwclock command.

#### Status

Not a bug

#### Ref. #

## 5.11. With multiple active sessions, **NVIDIA Control Panel** incorrectly shows that the system is unlicensed

## Description

In an environment with multiple active desktop sessions, the Manage License page of NVIDIA Control Panel shows that a licensed system is unlicensed. However, the nvidiasmi command and the management interface of the NVIDIA vGPU software license server correctly show that the system is licensed. When an active session is disconnected and reconnected, the **NVIDIA Display Container** service crashes.

The Manage License page incorrectly shows that the system is unlicensed because of stale data in NVIDIA Control Panel in an environment with multiple sessions. The data is stale because NVIDIA Control Panel fails to get and update the settings for remote sessions when multiple sessions or no sessions are active in the VM. The NVIDIA Display Container service crashes when a session is reconnected because the session is not active at the moment of reconnection.

### **Status**

Open

## Ref. #

3761243

## 5.12. VP9 and AV1 decoding with web browsers are not supported on Microsoft Windows Server 2019

## Description

VP9 and AV1 decoding with web browsers are not supported on Microsoft Windows Server 2019. This issue occurs because starting with Windows Server 2019, the required codecs are not included with the OS and are not available through the Microsoft Store app. As a result, hardware decoding is not available for viewing YouTube videos or using collaboration tools such as Google Meet in a web browser.

### Version

This issue affects Microsoft Windows Server releases starting with Windows Server 2019.

### **Status**

Not an NVIDIA bug

### Ref. #

200756564

5.13. 13.0-13.2 Only: Linux VM might fail to return a license after shutdown if the license server is specified by its name

## Description

If the license server is specified by its fully qualified domain name, a Linux VM might fail to return its license when the VM is shut down. This issue occurs if the nvidia-gridd service cannot resolve the fully qualified domain name of the license server because systemd-resolved.service is not available when the service attempts to return the license. When this issue occurs, the nvidia-gridd service writes the following message to the systemd journal:

General data transfer failure. Couldn't resolve host name

#### **Status**

Resolved in NVIDIA vGPU software 13.3

### Ref. #

# 5.14. Ubuntu guest driver initialization fails with vGPUs and GPUs that support SR-IOV

## Description

Initialization of the NVIDIA vGPU software graphics driver fails for Ubuntu quest VMs configured with NVIDIA vGPUs or pass-through GPUs that support SR-IOV. An NVIDIA vGPU that supports SR-IOV resides on a physical GPU that supports SR-IOV with NVIDIA vGPU software, such as a GPU based on the NVIDIA Ampere architecture. This issue affects only guest VMs running a version of the Ubuntu OS with kernel version 5.13 or later, such as Ubuntu 20.04.

The failure of the driver to initialize is caused by an issue with extended Message Signaled Interrupts (MSI-X). When this issue occurs, the following error message is written to the kernel log file /var/log/kern.log in the guest VM:

RmInitAdapter failed!

#### **Status**

Open

#### Ref. #

3660894

## 5.15. 13.0-13.6 Only: **NVIDIA Control** Panel is started only for the RDP user that logs on first

## Description

On all supported Windows Server guest OS releases, NVIDIA Control Panel is started only for the RDP user that logs on first. Other users cannot start NVIDIA Control Panel. If more than one RDP user is logged on when NVIDIA Control Panel is started, it always opens in the session of the RDP user that logged on first, irrespective of which user started NVIDIA Control Panel. Furthermore, on Windows Server 2016, NVIDIA Control Panel crashes if a user session is disconnected and then reconnected while NVIDIA Control Panel is open.

### Version

This issue affects all supported Windows Server guest OS releases.

#### **Status**

Resolved in NVIDIA vGPU software 13.7

## Ref.#

3334310

# 5.16. 13.0-13.2 Only: Windows vGPU VM sometimes crashes after guest OS upgrade

## Description

When a VM that is configured with NVIDIA vGPU is rebooted after an OS upgrade from Windows 10 1909 to Windows 10 20H2, the VM sometimes crashes. This issue is caused by a NULL pointer exception in the Virtual GPU Manager plugin (libnvidia-vgx.so). This NULL pointer exception might also cause the VM to crash in other situations. When this issue occurs, error messages that indicate that the Virtual GPU Manager process crashed are written to the standard activity log /var/log/messages on the hypervisor host.

#### **Status**

Resolved in NVIDIA vGPU software 13.3

#### Ref. #

3465448

# 5.17. 13.0-13.2 Only: Memory leaks in the vGPU manager plugin cause the VM to hang

## Description

Applications running in a VM request memory to be allocated and freed by the vGPU manager plugin, which runs on the hypervisor host. When an application requests the vGPU manager plugin to free previously allocated memory, some of the memory is not freed. Some applications request memory more frequently than other applications. If such applications run for a long period of time, for example for two or more days, the failure to free all allocated memory might cause the hypervisor host to run out of memory. As a result, memory allocation for applications running in the VM might fail, causing the applications and, sometimes, the VM to hang.

When memory allocation fails, the error messages that are written to the log file on the hypervisor host depend on the hypervisor.

- For VMware vSphere ESXi, the following error messages are written to vmware.log: 2021-10-05T04:57:35.547Z| vthread-2329002| E110: vmiop log: Fail to create the buffer for translate pte rpc node 2021-06-05T10:48:33.007Z| vcpu-3| E105: PANIC: Unrecoverable memory allocation failure
- For Citrix Hypervisor and hypervisors based on Linux KVM, the following messages are written to the standard activity log in the /var/log directory (/var/log/messages or /var/log/syslog):

```
Feb 15 09:27:48 bkrzxen1 kernel: [1278743.170072] Out of memory: Kill process
20464 (vgpu) score 9 or sacrifice child
Feb 15 09:27:48 bkrzxen1 kernel: [1278743.170111] Killed process 20464 (vgpu)
total-vm:305288kB, anon-rss:56508kB, file-rss:30828kB, shmem-rss:0kB
Feb 15 09:27:48 bkrzxen1 kernel: [1278743.190484] oom reaper: reaped process
20464 (vgpu), now anon-rss:0kB, file-rss:27748kB, shmem-rss:4kB".
```

### Workaround

If an application or a VM hangs after a long period of usage, restart the VM every couple of days to prevent the hypervisor host from running out of memory.

#### **Status**

Resolved in NVIDIA vGPU software 13.3

### Ref. #

# 5.18. nvidia-smi ignores the second NVIDIA vGPU device added to a Microsoft Windows Server 2016

## Description

After a second NVIDIA vGPU device is added to a Microsoft Windows Server 2016 VM, the device does not appear in the output from the nvidia-smi command. This issue occurs only if the VM is already running NVIDIA vGPU software for the existing NVIDIA vGPU device when the second device is added to the VM.

The nvidia-smi command cannot retrieve the guest driver version, license status, and accounting mode of the second NVIDIA vGPU device.

```
nvidia-smi vgpu --query
GPU 00000000:37:00.0
   Active vGPUs
                                    : 3251695793
   vGPU ID
       VM ID
                                     : 3575923
                                    : SVR-Reg-W(P)-KuIn
       VM Name
                                    : GRID V100D-32Q
       vGPU Name
       vGPU Type
vGPU UUID
                                    : 185
                                    : 29097249-2359-11b2-8a5b-8e896866496b
       Guest Driver Version
                                    : 473.47
: Licensed
       License Status
                                     : Disabled
       Accounting Mode
GPU 00000000:86:00.0
   Active vGPUs
                                    : 1
                                    : 3251695797
: 3575923
   vGPU ID
       VM TD
       VM Name
                                    : SVR-Reg-W(P)-KuIn
       vGPU Name
                                    : GRID V100D-32Q
                                    : 185
       vGPU Type
                                    : 2926dd83-2359-11b2-8b13-5f22f0f74801
       vGPU UUID
       Guest Driver Version
License Status
                                    : Not Available
                                     : N/A
      Accounting Mode
                                  : N/A
```

#### Version

This issue affects only VMs that are running Microsoft Windows Server 2016 as a quest OS.

#### Workaround

To avoid this issue, configure the guest VM with both NVIDIA vGPU devices before installing the NVIDIA vGPU software graphics driver.

If you encounter this issue after the VM is configured, use one of the following workarounds:

- ▶ Reinstall the NVIDIA vGPU software graphics driver.
- Forcibly uninstall the Microsoft Basic Display Adapter and reboot the VM.
- Upgrade the guest OS on the VM to Microsoft Windows Server 2019.

#### **Status**

Not an NVIDIA bug

## Ref. #

3562801

## 5.19. Desktop session freezes when a VM is migrated to or from a host running an NVIDIA vGPU software 14 release

## Description

When a VM configured with a Tesla V100 or Tesla T4 vGPU is migrated between a host running an NVIDIA vGPU software 14 release and a host running a an NVIDIA vGPU software 13 release, the remote desktop session freezes. After the session freezes, the VM must be rebooted to recover the session. This issue occurs only when the NVIDIA hardware-based H.264/HEVC video encoder (NVENC) is enabled.

### Version

The issue affects migrations between a host running an NVIDIA vGPU software 14 release and a host running an NVIDIA vGPU software 13 release.

### Workaround

Disable NVENC.

#### Status

Open

#### Ref #

## 5.20. The reported NVENC frame rate is double the actual frame rate

## Description

The frame rate in frames per second (FPS) for the NVIDIA hardware-based H.264/HEVC video encoder (NVENC) reported by the nvidia-smi encodersessions command and NVWMI is double the actual frame rate. Only the reported frame rate is incorrect. The actual encoding of frames is **not** affected.

This issue affects only Windows VMs that are configured with NVIDIA vGPU.

### **Status**

Open

### Ref. #

2997564

# 5.21. VM hangs after vGPU migration from a host running a vGPU manager 11 release to a host running a vGPU manager 13 release

## Description

When a VM configured with a Tesla T4 vGPU is migrated from a host that is running a vGPU manager 11 release **before 11.6** to a host that is running a vGPU manager 13 release, the VM hangs. When this issue occurs, XID error 31 is written to the log files on the destination hypervisor host.

### Version

This issue affects migration from a host that is running a vGPU manager 11 release before 11.6, such as 11.4 or 11.5, to a host that is running a vGPU manager 13 release, such as 13.0 or 13.1.

### Workaround

Upgrade the host that is running a vGPU manager 11 release to release 11.6 before attempting the migration.

### **Status**

Open

## 5.22. Windows 2012 R2 licensed clients cannot acquire licenses from a DLS instance

## Description

NVIDIA vGPU software licensed clients running in a Windows 2012 R2 VM cannot acquire licenses from a Delegated License Service (DLS) instance. This issue occurs because the TLS handshake between the client VM and DLS instance is failing with schannel error code 0x80090326 (SEC E ILLEGAL MESSAGE), indicating that the client has encountered an unrecoverable error during the TLS handshake.

### Workaround

Use the legacy NVIDIA vGPU software license server instead of NVIDIA License System (NLS).

### **Status**

Open

### Ref. #

## 5.23. 13.0 Only: Windows 2012 R2 licensed clients cannot acquire licenses from a CLS or DLS instance

## Description

NVIDIA vGPU software licensed clients running in a Windows 2012 R2 VM cannot acquire licenses from a Cloud License Service (CLS) instance or a Delegated License Service (DLS) instance. During the license acquisition process, the vGPU licensing service compares the size of the message to be sent to the CLS or DLS instance before and after encryption. On Windows 2012 R2, the MSDN API for encrypting the message outputs a shorter encrypted message than the plain text message. As a result, the validation check in the service fails, which causes the client to fail to acquire a license.

#### Status

Resolved in NVIDIA vGPU software 13.1

#### Ref. #

3392680

## 5.24. Desktop session freezes when a VM is migrated to or from a host running an NVIDIA vGPU software 11 release

## Description

The remote desktop session freezes when a VM is migrated to or from a host running an NVIDIA vGPU software 11 release. This issue affects only VMs configured with vGPUs on GPUs based on the NVIDIA Volta<sup>™</sup> architecture.

When this issue occurs, the following error messages are written to the Citrix Hypervisor VM's log file /var/log/daemon.log on the destination host:

XID error 13

▶ XID error 43

### Version

The issue affects migrations to and from a host that is running an NVIDIA vGPU software 11 release.

## **Status**

Open

### Ref. #

200707632

# 5.25. Hypervisor host reboots when multiple cloned VMs are simultaneously powered on or migrated

## Description

When multiple cloned VMs are simultaneously powered on or migrated, the hypervisor host reboots. No crash dump is generated and no error messages related to the reboot are written to the log files daemon.log or kern.log.

### Version

This issue affects only Citrix Hypervisor 8.2.

### **Status**

Open

### Ref. #

## 5.26. NVENC does not work with Teradici Cloud Access Software on Windows

## Description

The NVIDIA hardware-based H.264/HEVC video encoder (NVENC) does not work with Teradici Cloud Access Software on Windows. This issue affects NVIDIA vGPU and GPU pass through deployments.

This issue occurs because the check that Teradici Cloud Access Software performs on the DLL signer name is case sensitive and NVIDIA recently changed the case of the company name in the signature certificate.

#### **Status**

Not an NVIDIA bug

This issue is resolved in the latest 21.07 and 21.03 Teradici Cloud Access Software releases.

### Ref. #

200749065

## 5.27. A licensed client might fail to acquire a license if a proxy is set

## Description

If a proxy is set with a system environment variable such as HTTP PROXY OR HTTPS PROXY, a licensed client might fail to acquire a license.

### Workaround

Perform this workaround on each affected licensed client.

1. Add the address of the NVIDIA vGPU software license server to the system environment variable NO PROXY.

The address must be specified exactly as it is specified in the client's license server settings either as a fully-qualified domain name or an IP address. If the NO PROXY environment variable contains multiple entries, separate the entries with a comma (,). If high availability is configured for the license server, add the addresses of the primary license server and the secondary license server to the system environment variable no proxy.

- 2. Restart the NVIDIA driver service that runs the core NVIDIA vGPU software logic.
  - On Windows, restart the NVIDIA Display Container service.
  - On Linux, restart the nvidia-gridd service.

### **Status**

Closed

#### Ref. #

200704733

## 5.28. Session connection fails with four 4K displays and NVENC enabled on a 2Q, 3Q, or 4Q vGPU

## Description

Desktop session connections fail for a 2Q, 3Q, or 4Q vGPU that is configured with four 4K displays and for which the NVIDIA hardware-based H.264/HEVC video encoder (NVENC) is enabled. This issue affects only Teradici Cloud Access Software sessions on Linux quest VMs.

This issue is accompanied by the following error message:

This Desktop has no resources available or it has timed out

This issue is caused by insufficient frame buffer.

### Workaround

Ensure that sufficient frame buffer is available for all the virtual displays that are connected to a vGPU by changing the configuration in one of the following ways:

Reducing the number of virtual displays. The number of 4K displays supported with NVENC enabled depends on the vGPU.

vGPU	4K Displays Supported with NVENC Enabled
2Q	1
3Q	2
4Q	3

Disabling NVENC. The number of 4K displays supported with NVENC disabled depends on the vGPU.

vGPU	4K Displays Supported with NVENC Disabled		
2Q	2		
3Q	2		
4Q	4		

▶ Using a vGPU type with more frame buffer. Four 4K displays with NVENC enabled on any Q-series vGPU with at least 6144 MB of frame buffer are supported.

#### **Status**

Not an NVIDIA bug

# Ref. #

200701959

# 5.29. Disconnected sessions cannot be reconnected or might be reconnected very slowly with **NVWMI** installed

# Description

Disconnected sessions cannot be reconnected or might be reconnected very slowly when the NVIDIA Enterprise Management Toolkit (NVWMI) is installed. This issue affects Citrix Virtual Apps and Desktops and VMware Horizon sessions on Windows guest VMs.

# Workaround

Uninstall NVWMI.

# **Status**

Open

### Ref. #

# 5.30. Migrating a VM with a Tesla T4 vGPU between a host running NVIDIA vGPU software 11.3 and a host running a different release fails

# Description

Migrating a VM with a Tesla T4 vGPU between hosts where one host is running NVIDIA vGPU software 11.3 and the other host is running a different release fails. After the migration, the destination host and VM become unstable.

When this issue occurs, error message XID error 38 is written to the Citrix Hypervisor VM's log file /var/log/daemon.log. Depending on the host configurations, the following messages might also be written to the log file:

- ▶ XID error 43
- ▶ VGPU message 58
- ▶ VGPU message 4

#### Version

This issue affects migrations between a host that is running NVIDIA vGPU software release 11.3 and a host that is running a different release. The issue affects migrations to and from the host that is running NVIDIA vGPU software release 11.3.

This issue does **not** affect migrations between two hosts that are both running NVIDIA vGPU software 11.3.

## Workaround

If you are migrating a VM between a host that is running NVIDIA vGPU software 11.3 and a host that is running a another NVIDIA vGPU software 11 release, contact NVIDIA Enterprise Support for assistance.

Otherwise, avoid migrating a VM between a host that is running NVIDIA vGPU software 11.3 and a host that is running a different NVIDIA vGPU software release.

#### **Status**

Closed

## Ref.#

200691763

200735219

# 5.31. Linux VM hangs after vGPU migration to a host running a newer vGPU manager version

# Description

When a Linux VM configured with a Tesla V100 or Tesla T4 vGPU is migrated from a host that is running a vGPU manager 11 release before 11.6 to a host that is running a vGPU manager 13 release, the VM hangs. After the migration, the destination host and VM become unstable. When this issue occurs, XID error 31 is written to the log files on the destination hypervisor host.

### Version

This issue affects migration from a host that is running a vGPU manager 11 release before 11.6 to a host that is running a vGPU manager 13 release.

## Workaround

If the VM is configured with a Tesla T4 vGPU, perform the following sequence of steps before attempting the migration:

- 1. Upgrade the host that is running a vGPU manager 11 release to release 11.6 or a later vGPU manager 11 release.
- 2. Disconnect any remoting tool that is using NVENC.



Note: You cannot use this workaround for a VM that is configured with a Tesla V100 vGPU.

## **Status**

Open

# Ref. #

# 5.32. Idle Teradici Cloud Access Software session disconnects from Linux VM

# Description

After a Teradici Cloud Access Software session has been idle for a short period of time, the session disconnects from the VM. When this issue occurs, the error messages NVOS status 0x19 and vGPU Message 21 failed are written to the log files on the hypervisor host. This issue affects only Linux guest VMs.

## **Status**

Open

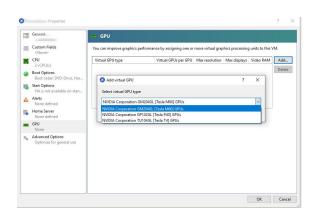
## Ref. #

200689126

# 5.33. No virtual GPU types are listed in Citrix XenCenter

# Description

No virtual GPU (vGPU) types are listed on the Select virtual GPU type drop-down list in Citrix XenCenter. However, the command nvidia-smi vgpu -s lists all supported vGPU types. This issue occurs because the driver version number of the Virtual GPU Manager contains multiple decimal points and is misinterpreted by Citrix XenCenter.



## Version

Citrix Hypervisor 8.2

## Resolution

Apply Hotfix XS82E002 - For Citrix Hypervisor 8.2 from Citrix.

## **Status**

Not an NVIDIA bug. Resolved by Citrix in hotfix XS82E002 for Citrix Hypervisor 8.2

#### Ref. #

200653755

# 5.34. NVIDIA vGPU software graphics driver fails to load after upgrade on Citrix Hypervisor

# Description

NVIDIA vGPU software graphics driver fails to load after an upgrade of NVIDIA vGPU software from an 11.x release to a 12.x release on Citrix Hypervisor. This issue occurs because the hypervisor name in the RPM package name has changed from xenserver to CitrixHypervisor.

The attempted upgrade fails its dependency checks as follows:

```
[root@localhost ~] # rpm -Uvh NVIDIA-vGPU-CitrixHypervisor-8.2-470.239.01.x86 64.rpm
error: Failed dependencies:
       NVIDIA-vGPU-xenserver conflicts with NVIDIA-vGPU-
CitrixHypervisor-1:8.2-470.239.01.x86_64
```

## Workaround

1. Determine the full name of the installed RPM package for the Virtual GPU Manager.

```
[root@localhost ~]# rpm -qa | grep NV
NVIDIA-vGPU-xenserver-8.2-460.107.x86 64
```

2. Remove the installed RPM package for the Virtual GPU Manager.

```
[root@localhost ~] # rpm -ev NVIDIA-vGPU-xenserver-8.2-460.107.x86 64
Preparing packages...
NVIDIA-vGPU-xenserver-1:8.2-460.107.x86 64
```

3. Perform a fresh installation of the latest RPM package for the Virtual GPU Manager.

```
[root@localhost ~] # rpm -ivh NVIDIA-vGPU-CitrixHypervisor-8.2-470.239.01.x86_64.rpm
Preparing...
Updating / installing...
   1:NVIDIA-vGPU-CitrixHypervisor-1:8.############################## [100%]
```

4. Reboot the hypervisor host.

[root@localhost ~] # reboot

#### Status

Open

# Ref. #

200682984

# 5.35. Windows guest VMs with vGPUs or GPUs with large BAR memory settings fail to boot to the desktop in UEFI mode

# Description

Windows quest VMs configured with vGPUs or physical GPUs that have large BAR memory settings fail to boot to the desktop in UEFI mode. The VM can be reached through its IP address but the desktop session is blank after the connection to the VM is established.

By default, Citrix Hypervisor creates a VM in UEFI boot mode and the boot mode of a VM cannot be changed after the VM is created.

# Version

Citrix Hypervisor 8.2

## Workaround

Delete the VM and re-create it, selecting BIOS Boot mode when you specify the installation method for the OS software on the new VM.



## **Status**

Open

## Ref. #

200676622

# 5.36. Migrated VM with more than two vGPUs crashes on destination host

# Description

When a VM that is configured with more than two vGPUs is migrated, the VM crashes on the destination host. When this issue occurs, the log file on the hypervisor host is flooded with NVOS status messages 0x23, 0x33, 0x5, and 0x21 and vGPU failure messages 4, 6, 7, 9, 12, 14, 19, 21, 23, 26, 32, 43, 54, and 56.

## **Status**

Open

#### Ref. #

200660221

# 5.37. Driver upgrade in a Linux guest VM with multiple vGPUs might fail

# Description

Upgrading the NVIDIA vGPU software graphics driver in a Linux guest VM with multiple vGPUs might fail. This issue occurs if the driver is upgraded by overinstalling the new release of the driver on the current release of the driver while the nvidia-gridd service is running in the VM.

# Workaround

- 1. Stop the nvidia-gridd service.
- 2. Try again to upgrade the driver.

## **Status**

Open

## Ref. #

200633548

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

# Description

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

## Workaround

Restart the VM and wait at least 30 seconds before trying to launch NVIDIA Control Panel.

## **Status**

Open

## Ref. #

200623179

# 5.39. VNC client session goes blank and console VNC is corrupted when the guest driver is uninstalled

# Description

When the NVIDIA vGPU software graphics driver is uninstalled from a Windows 10 quest VM for which the boot mode is UEFI, the VNC client session goes blank and console VNC is corrupted.

This issue affects only Windows 10 quest VMs for which the boot mode is UEFI. It does not affect Windows 10 guest VMs for which the boot mode is BIOS.

## Workaround

Reboot the VM by using Citrix XenCenter.

## Status

Not an NVIDIA bug

## Ref. #

200609669

# 5.40. Citrix Virtual Apps and Desktops session corruption occurs in the form of residual window borders

# Description

When a window is dragged across the desktop in a Citrix Virtual Apps and Desktops session, corruption of the session in the form of residual window borders occurs.

## Version

This issue affects only Citrix Virtual Apps and Desktops version 7 2003

# Workaround

Use Citrix Virtual Apps and Desktops version 7 1912 or 2006.

#### **Status**

Not an NVIDIA bug

### Ref. #

# 5.41. Suspend and resume between hosts running different versions of the vGPU manager fails

# Description

Suspending a VM configured with vGPU on a host running one version of the vGPU manager and resuming the VM on a host running a version from an older main release branch fails. For example, suspending a VM on a host that is running the vGPU manager from release 13.10 and resuming the VM on a host running the vGPU manager from release 12.4 fails. When this issue occurs, the error There were no servers available to complete the specified operation is reported on Citrix XenCenter.

### **Status**

Not an NVIDIA bug

## Ref. #

200602087

# 5.42. 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"
```

- c). Save your changes to /etc/X11/xorg.conf file and guit the editor.
- 3. Start the Xorg server.
  - # etc/init.d/xorg start

#### Status

Open

## Ref. #

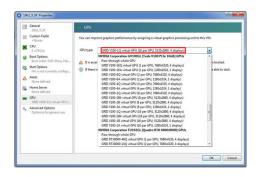
200605900

# 5.43. 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

# 5.44. Citrix Virtual Apps and Desktops connection freezes initially

# Description

When -OB and -OQ 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

# Workaround

Disable the use hardware encoding for video codec Citrix policy, which is enabled by default.

## **Status**

Open

## Ref. #

200494400

# 5.45. DWM crashes randomly occur in Windows VMs

# Description

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

# Version

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

## **Status**

Not an NVIDIA bug

## Ref. #

# 5.46. **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 NVDIA Control Panel fails to launch in platform layer or published image in the Citrix Support Knowledge Center.

# 5.47. 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-ml 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. #

# 5.48. 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

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

# Description

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

When the failure occurs, the following messages are displayed:

-> Installing DKMS kernel module:

```
ERROR: Failed to run \u00e7/usr/sbin/dkms build -m nvidia -v 470.63.01 -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: 470.63.01 not found
       Error! Bad return status for module build on kernel: 5.3.0-28-generic
       Consult /var/lib/dkms/nvidia/ 470.63.01/build/make.log for more information.
        -> error.
       ERROR: Failed to install the kernel module through DKMS. No kernel module
was installed;
       please try installing again without DKMS, or check the DKMS logs for more
information.
       ERROR: Installation has failed. Please see the file '/var/log/nvidia-
installer.log' for details.
        You may find suggestions on fixing installation problems in the README
available on the Linux driver download page at www.nvidia.com.
```

## Workaround

When installing the NVIDIA vGPU software graphics driver with DKMS enabled, use one of the following workarounds:

- Before running the driver installer, install the dkms package, then run the driver installer with the -dkms option.
- ▶ Run the driver installer with the --no-cc-version-check option.

## **Status**

Not a bug.

#### Ref #

2836271

# 5.50. 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

# **Status**

Not an NVIDIA bug.

#### Ref #

200568154

# 5.51. 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 <u>Virtual GPU Software User Guide</u>.

#### **Status**

Not a bug

# Ref. #

# 5.52. 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

# 5.53. 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. When this issue occurs, the message The VGPU is not compatible with any PGPU in the destination. appears in the Citrix XenCenter management GUI.

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

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

# 5.54. 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

# 5.55. Changes to ECC memory settings for a Linux vGPU VM by nvidia-smi might be ignored

# Description

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

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

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

# Workaround

This workaround requires administrator privileges.

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

### Status

Open

# Ref. #

200505777

# 5.56. Incorrect GPU type shown for Quadro RTX 8000 GPUs in Citrix **XenCenter**

# Description

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

However, the virtual GPU types for a Quadro RTX 8000 GPU are listed correctly.

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

## **Status**

Not an NVIDIA bug

## Ref. #

200492607

# 5.57. **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 guit 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

# 5.58. 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

# 5.59. 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. #

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

# Description

When some releases of Windows Server are used as a quest 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 2019
- Windows Server 2016
- Windows Server 2012

#### Solution

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

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

# 5.61. 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:

-	localhost: ~ g 20 10:33:	7] <b>nvidia-smi vg</b> 18 2018	ou 			
NVIDIA-SMI 390.42			Drive	Driver Version: 390.42		
GPU	Name vGPU ID	Name	Bus-Id   VM ID	VM Name	GPU-Util   vGPU-Util	
0	Tesla P40 85109 87195 88095 89170 90475 93363	- ~		0:81:00.0 win7-xmpl-146048-1 win7-xmpl-146048-2 win7-xmpl-146048-3 win7-xmpl-146048-4 win7-xmpl-146048-5 win7-xmpl-146048-6	99%   32%   39%   26%   0%   0%	
1	Tesla P40		0000000	00000000:85:00.0		

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

# 5.62. 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.#

# 5.63. 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

# 5.64. 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.
  - If the server has less than 1 TB or more of system memory, omit this step.
- 3. Reboot the server.

#### **Status**

Not an NVIDIA bug

#### Ref.#

1799582

# 5.65. Luxmark causes a segmentation fault on an unlicensed Linux client

# Description

If the Luxmark application is run on a Linux quest 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.#

# 5.66. 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**

# 5.67. 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 nvidiagridd to exit on Red Hat Enterprise Linux and CentOS.

#### Workaround

This workaround requires sudo privileges.



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

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

- 1. If **NVIDIA X Server Settings** is running, shut it down.
- 2. If the /etc/nvidia/gridd.conf file does not already exist, create it by copying the supplied template file /etc/nvidia/gridd.conf.template.
- 3. As root, edit the /etc/nvidia/gridd.conf file to set the EnableUI option to TRUE.
- 4. Start the nvidia-gridd service.
  - # sudo service nvidia-gridd start

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

#### **Status**

Open

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

# Description

NVIDIA vGPU software licenses remain checked out on the license server when nonpersistent 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

# 5.69. 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 quest 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**

# 5.70. 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. #

# 5.71. 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

- 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.

# **Status**

Not an NVIDIA bug

## Ref. #

1853248

# 5.72. 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.#

NVIDIA-420/1792341

# 5.73. 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

Oh no! Something has gone wrong!

## Workaround

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

- 1. As root, edit the /etc/selinux/config file to set SELINUX to permissive. SELINUX=permissive
- 2. Reboot the system.

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. #

# 5.74. 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

# 5.75. 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 disable vnc=1 in the vGPU plugin's parameters). 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.

Citrix Hypervisor 8.1 or later:

[root@xenserver ~] # xe vgpu-param-set uuid=vgpu-uuid extra\_args=disable\_vnc=1

Citrix Hypervisor earlier than 8.1:

xe vm-param-set uuid=vm-uuid platform:vgpu extra args="disable vnc=1"

#### **Status**

Open

## Ref. #

NVIDIA-145/1375164

5.76. 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

# 5.77. 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

# 5.78. Windows VM BSOD

# Description

Windows VM bugchecks on Citrix Hypervisor when running a large number of vGPU based VMs.

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

```
NVRM: Xid (PCI:0000:08:00): 31, Ch 0000001e, engmask 00000111, intr 10000000
NVRM: Xid (PCI:0000:08:00): 31, Ch 00000016, engmask 00000111, intr 10000000
vmiop log: error: Assertion Failed at 0xb5b898d8:4184
vmiop log: error: 8 frames returned by backtrace
vmiop_log: error: /usr/lib/libnvidia-vgx.so( nv000793vgx+0x69d) [0xb5b8064d]
vmiop log: error: /usr/lib/libnvidia-vgx.so( nv000479vgx+0x118) [0xb5b898d8]
vmiop_log: error: /usr/lib/libnvidia-vgx.so(_nv000782vgx+0x59) [0xb5b85f49]
vmiop_log: error: /usr/lib/libnvidia-vgx.so(_nv000347vgx+0x3db) [0xb5b932db]
vmiop_log: error: /usr/lib/libnvidia-vgx.so [0xb5b78e4a]
```

```
vmiop log: error: /usr/lib/xen/bin/vgpu [0x80554be]
vmiop_log: error: /lib/libpthread.so.0 [0xb7612912]
vmiop_log: error: /lib/libc.so.6(clone+0x5e) [0xb76fc5ee]
vmiop_log: error: failed to initialize guest PTE entries
vmiop_log: error: failed to fill up guest PTE entries 3
vmiop log: error: VGPU message 27 failed, result code: 0xff000003
vmiop log: error: 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

# 5.79. 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.

## Fix

Upgrade Citrix Virtual Apps and Desktops to 7.6 Feature Pack 3

#### **Status**

Closed

# Ref.#

NVIDIA-370/200130780

# 5.80. 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

# 5.81. 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

# 5.82. 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.#

#### Notice

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

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

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

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

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

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

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

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

THIS DOCUMENT AND ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, "MATERIALS") ARE BEING PROVIDED "AS IS." NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF 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.

#### Trademarks

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

#### Copyright

© 2013-2024 NVIDIA Corporation & affiliates. All rights reserved.

