# GRID VGPU FOR CITRIX XENSERVER VERSION 367.43/369.17

RN-06927-001 _v4.0 (GRID) | August 2016

## Release Notes

# TABLE OF CONTENTS

# Chapter 1. RELEASE NOTES

These *Release Notes* summarize current status, information on validated platforms, and known issues with NVIDIA GRID™ vGPU™ software and hardware on Citrix XenServer.

This release includes the following software:

- NVIDIA GRID Virtual GPU Manager version 367.43 for the following Citrix XenServer releases:
  - Citrix XenServer 7.0
  - Citrix XenServer 6.5 SP1
- NVIDIA Windows drivers for vGPU version 369.17
- NVIDIA Linux drivers for vGPU version 367.43

**Caution**

If you install the wrong package for the version of Citrix XenServer you are using, GRID vGPU Manager will fail to load.

The releases of GRID vGPU Manager and Windows guest VM drivers that you install must be compatible. Older VM drivers will not function correctly with this release of GRID vGPU Manager. However, some older GRID vGPU Managers may function correctly with this release of Windows guest drivers. See VM running older NVIDIA vGPU drivers fails to initialize vGPU when booted.

Updates in this release:

- Support for Tesla M10 cards
- Support for Windows Server 2016 as a guest OS
- Enhanced monitoring of GPU performance:
  - Monitoring of individual vGPUs
  - Monitoring from within a guest VM
- Miscellaneous bug fixes

# Chapter 2. VALIDATED PLATFORMS

This release of virtual GPU provides support for the following NVIDIA GPUs on Citrix XenServer, running on validated server hardware platforms:

- GRID K1
- GRID K2
- Tesla M6
- Tesla M10
- Tesla M60

For a list of validated server platforms, refer to Buy NVIDIA GRID Solutions.

## 2.1. Hypervisor Software Versions

This release has been tested with the following hypervisor software versions:

| Software | Version tested |
|---|---|
| Citrix XenServer 6.5 | Version 6.5 with XS65ESP1. The GRID vGPU Manager included in this release will not install without XenServer 6.5 SP1. |
| Citrix XenServer 7.0 | RTM build 125380 is supported. |
| Citrix XenDesktop | Version 7.6, 7.8 in HDX 3D Pro mode. |

## 2.2. Linux Support

GRID vGPU with Linux guest VMs is supported on Tesla M60, Tesla M10, and Tesla M6 on Citrix XenServer 7.0, with the following distributions:

- Red Hat Enterprise Linux 7
- CentOS 7
- Ubuntu 14.04 LTS

## 2.3. Hardware Configuration

Tesla M60 and M6 GPUs support compute and graphics modes, which can be configured by using the `gpumodeswitch` tool provided with GRID software releases. GRID vGPU requires that M60 and M6 GPUs are configured in graphics mode.

# Chapter 3.
# KNOWN PRODUCT LIMITATIONS

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

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

### Description

A VM running older NVIDIA drivers, such as those from a previous vGPU release, will fail to initialize vGPU when booted on a Citrix XenServer platform running the current release of GRID Virtual GPU Manager.

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

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

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

- An error message:

  ```
  vmiop_log: error: Unable to fetch Guest NVIDIA driver information
  ```
- A version mismatch between guest and host drivers:

  ```
  vmiop_log: error: Guest VGX version(1.1) and Host VGX version(1.2) do not
   match
  ```
- A signature mismatch:

  ```
  vmiop_log: error: VGPU message signature mismatch.
  ```

### Resolution

Install the latest NVIDIA vGPU release drivers in the VM.

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

### Description

GRID K2, Tesla M60, and Tesla M6 support error correcting code (ECC) for improved data integrity. If ECC is enabled, virtual GPU fails to start. The following error is logged in the Citrix XenServer VM's `/var/log/messages` log file:

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

Virtual GPU is not currently supported with ECC active. GRID K2 cards and Tesla M60, M6 cards in graphics mode ship with ECC disabled by default, but ECC may subsequently be enabled using `nvidia-smi`.

### Resolution

Ensure that ECC is disabled on all GPUs.

1. Use `nvidia-smi` to list the status of all GPUs, and check for ECC noted as enabled on GPUs.
2. Change the ECC status to off on each GPU for which ECC is enabled by executing the following command:

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

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

## 3.3. Single vGPU benchmark scores are lower than passthrough GPU

### Description

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

Aside from performance differences that may be attributed to a vGPU's smaller framebuffer size, vGPU incorporates a performance balancing feature known as Frame Rate Limiter (FRL), which is enabled on all vGPUs. 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 passthrough GPU.

### Resolution

FRL is controlled by an internal vGPU setting. NVIDIA does not validate vGPU with FRL disabled, but for validation of benchmark performance, FRL can be temporarily disabled by specifying `frame_rate_limiter=0` in the VM's `platform:vgpu_extra_args` parameter:

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

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

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

- Removing the `vgpu_extra_args` key from the `platform` parameter
- Removing `frame_rate_limiter=0` from the `vgpu_extra_args` key
- Setting `frame_rate_limiter=1`. For example:

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

## 3.4. Virtual GPU fails to start when GPUs are mapped above 4G

### Version

XenServer 6.2

### Status

Fixed in XenServer 6.5

### Description

GRID vGPU on Citrix XenServer 6.2 does not support operation with GPUs mapped above the 4 gigabyte (4G) boundary in the system's physical address space.

If GPUs are mapped above 4G, the GRID vGPU Manager rpm will warn at the time of installation:

```
Warning: vGPU does not support GPUs mapped in 64-bit address space. Please
 disable 64-bit MMIO from the system's BIOS. Refer to vGPU release notes for
 details.
```

Also, the NVIDIA kernel driver will fail to load in XenServer's dom0, so the `nvidia` module won't appear in the module listing produced by `lsmod`. Additionally, the following warning messages will be present in the output of `dmesg`:

```
NVRM: This PCI I/O region assigned to your NVIDIA device is invalid:
NVRM: BAR1 is 128M @ 0xf800000000000000 (PCI:03ff:00:07.0)
NVRM: This is a 64-bit BAR mapped above 4GB by the system
NVRM: BIOS or the Linux kernel. The NVIDIA Linux/x86
NVRM: graphics driver and other system software components
NVRM: do not support this configuration.
```

### Resolution

Ensure that GPUs are mapped below the 4G boundary by disabling your server's SBIOS option that controls 64-bit memory-mapped I/O support. This option may be labeled **Enable4G>Decode** or **Enable 64-bit MMIO**.

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

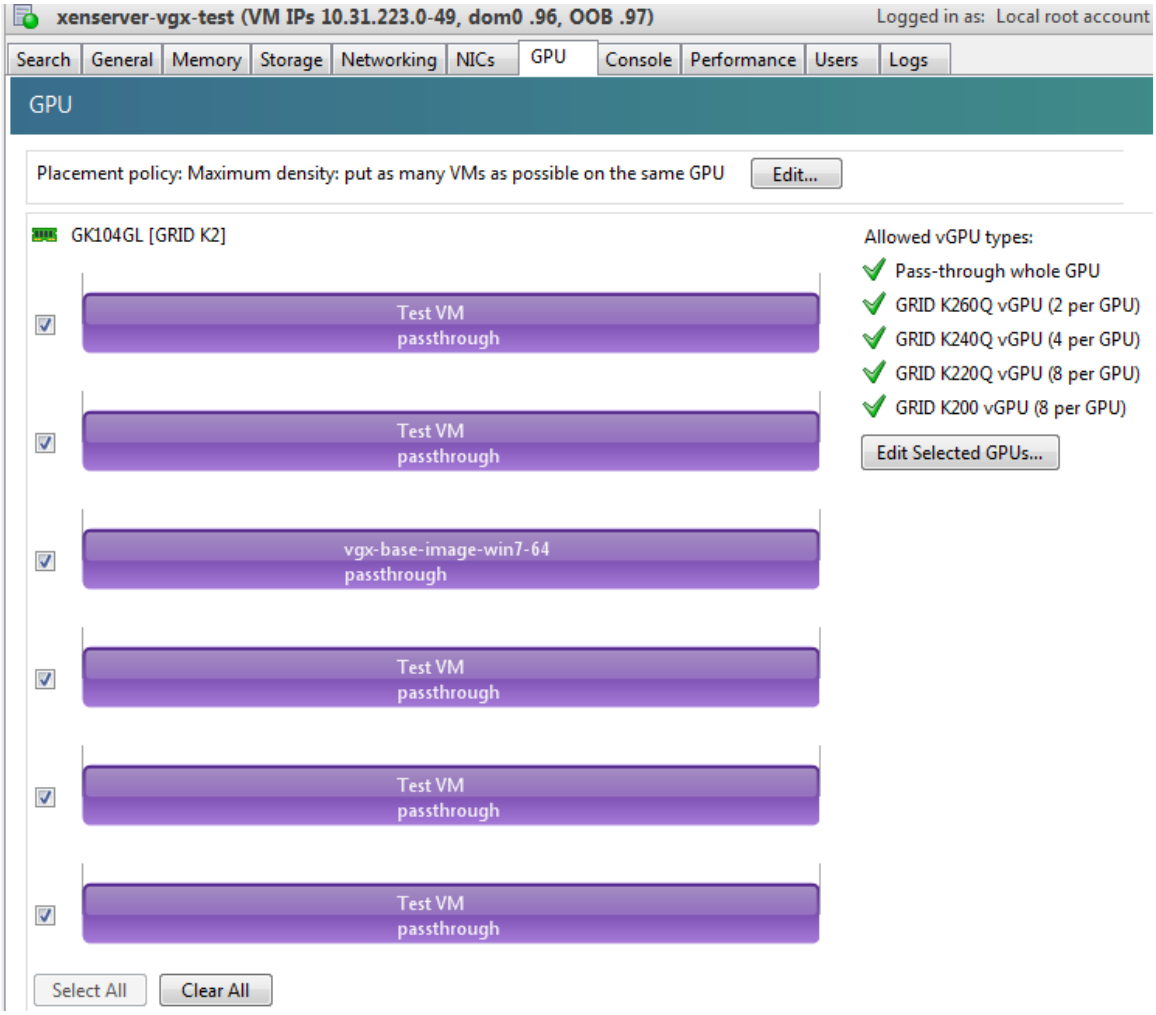
### Description

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

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

This is because GPUs operating in passthrough mode are not visible to `nvidia-smi` and the NVIDIA kernel driver operating in the Citrix XenServer dom0.

To confirm that all GPUs are operating in passthrough, use XenCenter's GPU tab to review current GPU assignment:




### Resolution

N/A

## 3.6. GRID K1 and GRID K2 cards do not support monitoring of vGPU engine usage

### Description

GRID K1 and GRID K2 cards do not support monitoring of vGPU engine usage. All tools and APIs for any vGPU running on GRID K1 or GRID K2 cards report 0 for the following usage statistics:

- 3D/Compute
- Memory controller bandwidth
- Video encoder

- Video decoder

## 3.7. Windows Aero is disabled on XenDesktop session using 3 or 4 monitors in 2560×1600 resolution

### Description

Windows Aero may be disabled when XenDesktop 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.

## 3.8. 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 GRID GPU is present in **Windows Device Manager** but displays a warning sign, and the following device status:

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

The Citrix XenServer 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 VM's `platform:vgpu_extra_args` parameter:

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

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

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

The reservation can be reverted back to its default setting in one of the following ways:

- Removing the `vgpu_extra_args` key from the `platform` parameter
- Removing `enable_large_sys_mem` from the `vgpu_extra_args` key
- Setting `enable_large_sys_mem=0`

## 3.9. vGPU host driver RPM upgrade fails

### Description

Upgrading vGPU host driver RPM fails with the following message 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

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

# Chapter 5. KNOWN ISSUES

## 5.1. 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 when multiple display heads are used with Citrix XenDesktop or VMware Horizon on a Windows 10 guest VM.

When this error occurs, the NVIDIA host driver reports Xid error 31 and Xid error 43 in XenServer's `/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
- GRID K100, K120Q
- GRID K200, K220Q

### Version

### Workaround

### Status

Open

### Ref. #

200130864

## 5.2. With no NVIDIA driver installed, XenServer misidentifies Tesla M10 cards

### Description

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

### Version

Citrix XenServer 6.5 and 7.0

### Workaround

None

### Status

Not an NVIDIA bug

### Ref. #

NVIDIA-420/1792341

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

### Description

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

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

### Version

### Workaround

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

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

   ```
   SELINUX=permissive
   ```

2. Reboot the system.

```
~]# reboot
```

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

**Status**

Not an NVIDIA bug

**Ref. #**

200167868

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

**Description**

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

**Workaround**

None

**Status**

Open

**Ref. #**

1306623

## 5.5. Local VGA console is momentarily unblanked when XenDesktop changes resolution of the VM desktop

**Description**

When XenDesktop establishes a remote connection to a VM using vGPU, the VM's local VGA console display in XenCenter is blanked (assuming the VM local console has not been disabled by setting `platform:vgpu_extra_args="disable_vnc=1"`). If the

XenDesktop session changes resolution of the VM's desktop, the local VGA console momentarily unblanks, allowing a XenCenter user to briefly view the desktop.

### Workaround

Disable the VM's local VGA console

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

### Status

Open

### Ref. #

NVIDIA-145/1375164

## 5.6. VM bugchecks on shutdown/restart when XenDesktop is installed and NVIDIA driver is uninstalled or upgraded.

### Description

If the XenDesktop 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 XenDesktop.

### Status

Open

### Ref. #

NVIDIA-295/200018125

## 5.7. Application frame rate may drop when running XenDesktop at 2560×1600 resolution.

### Description

An application's rendering frame rate may drop when running XenDesktop 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

## 5.8. Windows VM BSOD

### Description

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

XenServer's `/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:         0xc1d00001, 0xff010000, 0x1a77ba000, 0x0, 0x1,
```

```
vmiop_log: error:          0x1, 0x1000, 0x10202, 0xc1d00001, 0xff010000,
vmiop_log: error:          0xcaf00004, 0x0
vmiop_log: error: Timeout occurred, reset initiated.
```

### Version

XenServer 6.2

### Fix

Ensure that you are running the latest OEM firmware for your GRID boards.

### Status

Closed

### Ref. #

NVIDIA-327/1632120

## 5.9. Windows VM BSOD when upgrading NVIDIA drivers over a XenDesktop session

### Description

Windows VM bugchecks when NVIDIA guest drivers are upgraded over a XenDesktop session.

If the VM is restarted after the bugcheck, the upgraded driver loads correctly and full functionality is available.

### Version

### Fix

Upgrade XenDesktop to 7.6 Feature Pack 3

### Status

Closed

### Ref. #

NVIDIA-370/200130780

## 5.10. 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 XenServer 6.5, and does include XenCenter integration.

### Version

Affects the XenCenter integration with XenServer 6.5 only.

Resolved in the XenCenter integration with XenServer 7.0.

### Workaround

Refer to XenServer vGPU Management in *GRID Virtual GPU User Guide* for how to configure vGPU by using the `xe` CLI.

### Status

Closed

### Ref. #

NVIDIA-360

## 5.11. 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.

### Version

### Workaround

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

### Status

Closed

### Ref. #

NVIDIA-350/200123378

## 5.12. Multiple WebGL tabs in Microsoft Internet Explorer may trigger TDR on Windows VMs

### Description

Running intensive WebGL applications in multiple IE tabs may trigger a TDR on Windows VMs.

### Version

### Workaround

Disable hardware acceleration in IE.

To enable software rendering in IE, refer to the Microsoft knowledge base article How to enable or disable software rendering in Internet Explorer.

### Status

Open

### Ref. #

200148377

# 5.13. XenDesktop shows only a black screen when connected to a vGPU VM

### Description

XenDesktop sometimes displays only a black screen when it is connected to an NVIDIA vGPU VM. The probable cause is that the display that is connected to the NVIDIA vGPU is entering a lower power state.

### Fix

Disable all display-related power management settings.

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

### Status

Not an NVIDIA bug

### Ref. #

1719877

### Notice



### HDMI



### OpenCL



### Trademarks



### Copyright