



GRID SOFTWARE FOR HUAWEI UVP VERSION 367.122/370.12

RN-07939-001 _v4.4 (GRID) Revision 02 | October 2017

Release Notes



TABLE OF CONTENTS

Chapter 1. Release Notes	1
Chapter 2. Validated Platforms	2
2.1. Supported NVIDIA GPUs and Validated Server Platforms.....	2
2.2. Hypervisor Software Versions.....	2
2.3. Guest OS Support.....	3
2.3.1. Windows Guest OS Support.....	3
2.3.2. Linux Guest OS Support.....	3
Chapter 3. Known Product Limitations	4
3.1. VM running older NVIDIA vGPU drivers fails to initialize vGPU when booted.....	4
3.2. Virtual GPU fails to start if ECC is enabled.....	5
3.3. Single vGPU benchmark scores are lower than passthrough GPU.....	5
3.4. nvidia-smi fails to operate when all GPUs are assigned to GPU passthrough mode.....	6
3.5. GRID K1 and GRID K2 cards do not support monitoring of vGPU engine usage.....	6
3.6. VMs configured with large memory fail to initialize vGPU when booted.....	7
3.7. vGPU host driver RPM upgrade fails.....	8
Chapter 4. Resolved Issues	9
Chapter 5. Known Issues	10
5.1. Multiple WebGL tabs in Microsoft Internet Explorer may trigger TDR on Windows VMs.....	10

Chapter 1.

RELEASE NOTES

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

This release includes the following software:

- ▶ NVIDIA GRID Virtual GPU Manager version 367.122 for the Huawei UVP releases listed in [Hypervisor Software Versions](#)
- ▶ NVIDIA Windows drivers for vGPU version 370.12
- ▶ NVIDIA Linux drivers for vGPU version 367.122



Caution

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

The GRID vGPU Manager and Windows guest VM drivers must be installed together. Older VM drivers will not function correctly with this release of GRID vGPU Manager. Similarly, older GRID vGPU Managers will not 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:

- ▶ Miscellaneous bug fixes

Chapter 2.

VALIDATED PLATFORMS

This release of NVIDIA GRID software provides support for several NVIDIA GPUs on validated server hardware platforms, Huawei UVP hypervisor software versions, and guest operating systems.

2.1. Supported NVIDIA GPUs and Validated Server Platforms

This release of NVIDIA GRID software provides support for the following NVIDIA GPUs on Huawei UVP, running on validated server hardware platforms:

- ▶ GRID K1
- ▶ GRID K2
- ▶ Tesla M60

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



Tesla M60 and M6 GPUs support compute mode and graphics mode. GRID 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 GRID software releases.

2.2. Hypervisor Software Versions

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

Software	Version Tested
Huawei UVP	Version RC520

2.3. Guest OS Support

NVIDIA GRID software supports several Windows releases and Linux distributions as a guest OS.



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

2.3.1. Windows Guest OS Support

NVIDIA GRID software supports the following Windows releases as a guest OS on Huawei UVP:

- ▶ Windows Server 2016
- ▶ Windows Server 2012 R2
- ▶ Windows Server 2008 R2
- ▶ Windows 10 (32/64-bit)
- ▶ Windows 8.1 (32/64-bit)
- ▶ Windows 8 (32/64-bit)
- ▶ Windows 7 (32/64-bit)

2.3.2. Linux Guest OS Support

NVIDIA GRID software supports the following Linux distributions as a guest OS **only** on supported Tesla GPUs on Huawei UVP:

- ▶ Red Hat Enterprise Linux 6.6
- ▶ CentOS 6.6
- ▶ Ubuntu 14.04 LTS



GRID K1 and GRID K2 do not support vGPU on a Linux guest OS.

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 Huawei UVP 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 Huawei UVP 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

Tesla M60 GPUs support error correcting code (ECC) memory for improved data integrity. Tesla M60 GPUs in graphics mode are supplied with ECC memory disabled by default, but it may subsequently be enabled using `nvidia-smi`.

However, NVIDIA GRID vGPU does not support ECC memory. If ECC memory is enabled, NVIDIA GRID vGPU fails to start. The following error is logged in the Huawei UVP VM's `/var/log/messages` log file:

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

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. Reboot the host.

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 setting `plugin0.frame_rate_limiter=0` in the vGPU configuration file. vGPU configuration files are stored in `/usr/share/nvidia/vgx` and are named for the vGPU types they define, for example, `grid_k100.conf`.

The setting takes effect the next time any VM using the given vGPU type 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 by setting `plugin0.frame_rate_limiter=1` in the vGPU configuration file.

3.4. `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@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 Huawei UVP dom0.

Resolution

N/A

3.5. 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

3.6. 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 Huawei UVP 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 `plugin0.enable_large_sys_mem=1` in the vGPU configuration file.

vGPU configuration files are stored in `/usr/share/nvidia/vgx` and are named for the vGPU types they define, for example, `grid_k100.conf`.

The setting takes effect the next time any VM using the given vGPU type 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 `plugin0.extra_fb_reservation` in the vGPU configuration file, 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:

```
plugin0.extra_fb_reservation=128
```

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

- ▶ Removing `enable_large_sys_mem` from the vGPU configuration file
- ▶ Setting `enable_large_sys_mem=0`

3.7. vGPU host driver RPM upgrade fails

Description

Upgrading vGPU host driver RPM fails with the following message on the console:

```
[root@uvp ~]# rpm -U NVIDIA-vGPU-kepler-uvp-210.0-352.70.x86_64
error: Failed dependencies: NVIDIA-vgx-uvp conflicts with NVIDIA-vGPU-kepler-
uvp-210.0-352.70.x86_64
[root@uvp ~]#
```

Resolution

Uninstall the older vGPU RPM before installing the latest driver.

Use the following command to uninstall the older vGPU RPM:

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

Chapter 4. RESOLVED ISSUES

No resolved issues are reported in this release for Huawei UVP.

Chapter 5.

KNOWN ISSUES

5.1. 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.

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

Notice

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

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

HDMI

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

OpenCL

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

Trademarks

NVIDIA, the NVIDIA logo, NVIDIA GRID, vGPU, 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

© 2013-2017 NVIDIA Corporation. All rights reserved.