

NVIDIA vGPU Licensing



Table of Contents



[1. NVIDIA vGPU Architecture](#)

[1.1. vGPU Profiles](#)

[2. Licensing of a NVIDIA vGPU](#)

[2.1. NVIDIA License Terminology](#)

[2.1.1. SUMs](#)

[2.1.2. Perpetual License](#)

[2.1.3. Annual Subscription](#)

[2.2. NVIDIA vGPU Software Edition](#)

[2.2.1. NVIDIA grid virtual Applications](#)

[2.2.2. NVIDIA GRID Virtual PC](#)

[2.2.3. NVIDIA Quadro Virtual Data Center Workstation](#)

[2.3. Example License Calculation](#)

[2.3.1. Perpetual License](#)

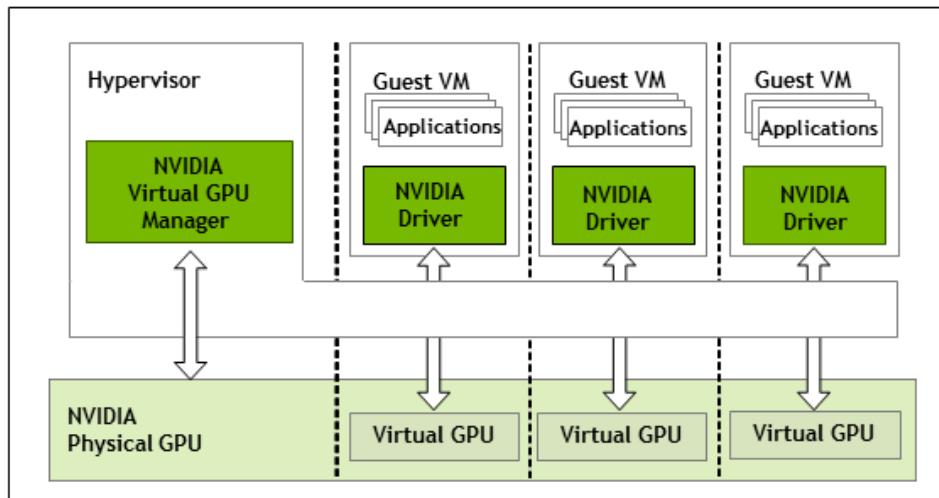
[2.3.2. Annual Subscription](#)

Through various recent projects, I had to work through the clutter of information regarding NVIDIA vGPU licensing.

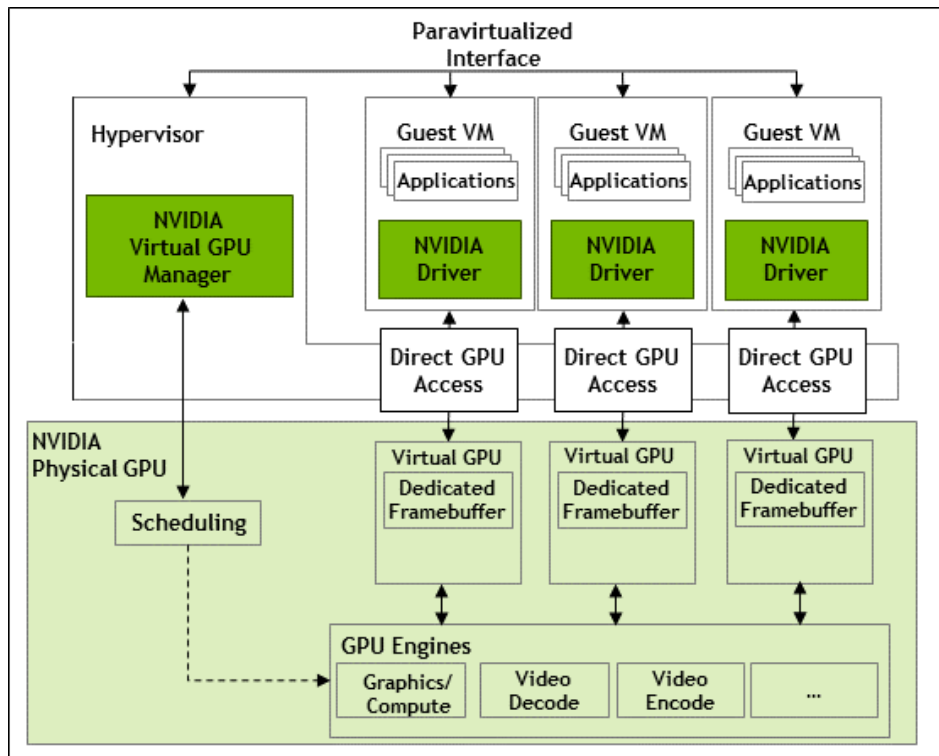
Here is a small summary of this information.

NVIDIA vGPU Architecture

Under the control of the NVIDIA GPU Virtual Manager, running in the hypervisor, the NVIDIA Physical GPU can operate multiple virtual GPU devices (vGPUs), that can be assigned directly to the Guest VM.



The Guest VMs use the NVIDIA virtual vGPUs in the same way as a physical GPU would come from the hypervisor by direct passed through. The NVIDIA Driver loaded into the guest VM provides Direct GPU Access for high-performance operations. The NVIDIA Virtual GPU Manager paravirtualized interface performs the non-performance management operations for the NVIDIA Driver.



Each NVIDIA vGPU is analogous to a conventional GPU, equipped with a fixed amount of GPU frame buffer and one or more virtual display outputs or “heads”. The vGPU frame buffer is allocated from the physical frame buffer of the physical GPU at the time of creation, and this vGPU retains exclusive access to this frame buffer part until it is destroyed.

All vGPUs residing on a physical GPU can share access to the GPU engines, including graphics (3D), video decoding, and video encoding modules.

vGPU Profiles

Each physical GPU can support several different types of vGPU Profiles (Virtual GPU Type) simultaneously. Each vGPU profile has fixed hardware key data, such as the frame buffer size, number of supported displays (Virtual Display Heads) and the maximum resolution (per display head). They are divided into different series, each corresponding to different load classes. Each series is identified by the last letter of the vGPU profile name.

- The Q-Series is designed for Designers and Power Users.
- The B-Series is designed for Power Users.
- The A-Series is designed for Users of virtual applications.

The number after the card type (P4 for Pascal Microarchitecture) in the name of the vGPU profile indicates the size of the frame buffer. For example, a P4-8Q vGPU on a Tesla P4 board is assigned 8192 MB of Frame Buffer.

1.4.1.9. Tesla P4 Virtual GPU Types

Physical GPUs per board: 1

Virtual GPU Type	Intended Use Case	Frame Buffer (Mbytes)	Virtual Display Heads	Maximum Resolution per Display Head	Maximum vGPUs per GPU	Maximum vGPUs per Board	Required License Edition
P4-8Q	Designer	8192	4	4096x2160	1	1	Quadro vDWS
P4-4Q	Designer	4096	4	4096x2160	2	2	Quadro vDWS
P4-2Q	Designer	2048	4	4096x2160	4	4	Quadro vDWS
P4-1Q	Power User, Designer	1024	2	4096x2160	8	8	Quadro vDWS
P4-2B	Power User	2048	2	4096x2160	4	4	GRID Virtual PC or Quadro vDWS
P4-2B4	Power User	2048	4	2560x1600	4	4	GRID Virtual PC or Quadro vDWS
P4-1B	Power User	1024	4	2560x1600	8	8	GRID Virtual PC or Quadro vDWS
P4-8A	Virtual Application User	8192	1	1280x1024 ²	1	1	GRID Virtual Application
P4-4A	Virtual Application User	4096	1	1280x1024 ²	2	2	GRID Virtual Application
P4-2A	Virtual Application User	2048	1	1280x1024 ²	4	4	GRID Virtual Application
P4-1A	Virtual Application User	1024	1	1280x1024 ²	8	8	GRID Virtual Application

Because of the different resource requirements of each vGPU profile, the maximum number of vGPUs that can be created simultaneously on a single physical GPU varies. For example, a Tesla P4 can board on its physical GPU up to 4 P4-2Q vGPUs, but only 2 P4-4Q vGPU profiles.

Licensing of a NVIDIA vGPU

When booting with a supported GPU, a vGPU with reduced capacity will be executed until a license is purchased.

The performance of an unlicensed vGPU is limited as follows:

- The frame rate is limited to 3 frames per second.
- The allocation of GPU resources is limited, so some applications can not run properly.

- For vGPUs that support CUDA, CUDA is disabled.

These restrictions are removed when importing a license.

NVIDIA License Terminology

It is important to note that the licenses are not to pay per VM, but in the Concurrent Use Model, per user who works with the vGPUs VM.

You must watch the peak times regarding access (for example 10 users) and calculates the license consumption.

SUMS

Behind SUMS, NVIDIA hides the Support, Upgrade and Maintenance program that is automatically purchased for each type of license.

Term	Meaning
SUMs	Support, Upgrade and Maintenance program
Perpetual License	A non-expiring, permanent software license that can be used on a perpetual basis without a need to renew. SUMS is required and is available in three, four, or five year increments. One year SUMS available only for renewals.
Annual Subscription	A software license that is active for a fixed period as defined by the terms of the subscription license, typically yearly. This includes SUMS for the duration of the license term.

PERPETUAL LICENSE

The Perpetual License never expire and you just have to renew the SUMS. Initially the Perpetual License are sold with SUMS lengths of 3, 4 or 5 years and can be renewed annually thereafter.

Enterprise Perpetual Licensing + SUMS Pricing	
GRID Virtual Applications	\$20 per CCU perpetual license
	\$5 SUMS per year
GRID Virtual PC	\$100 per CCU perpetual license
	\$25 SUMS per year
Quadro Virtual Data Center Workstation	\$450 per CCU perpetual license
	\$100 SUMS per year

ANNUAL SUBSCRIPTION

These licenses are valid for a period of time (1, 3 or 5 years) and directly include the SUMs for that period.

At the end of this period, you can not extend the Annual Subscription License, you must purchase it again.

Enterprise Annual Subscription Pricing	
GRID Virtual Applications	\$10 per CCU subscription
GRID Virtual PC	\$50 per CCU subscription
Quadro Virtual Data Center Workstation	\$250 per CCU subscription

NVIDIA vGPU Software Edition

NVIDIA distinguishes between 3 software editions, with more or less features. A mixed form of the different license models can be provided.

Example of a license mixed form

For the NX 3D application, there are 50 users with read access to the files via Citrix Published Apps and 10 users who develop the models in a Citrix Virtual Desktop Windows 10 VM in this program.

50 x vApps License

10 x Quadro vDWS License

NVIDIA virtual GPU Software Editions	
NVIDIA GRID® Virtual Applications	For organizations deploying XenApp or other RDSH solution. Designed for PC level applications and server based desktops.
NVIDIA GRID® Virtual PC	For users who want a virtual desktop but need great user experience leveraging PC Windows applications, browsers, and high definition video.
NVIDIA Quadro® vDWS	For users who want to be able to use remote professional graphics applications with full performance on any device, anywhere.

NVIDIA GRID VIRTUAL APPLICATIONS

This is useful for organizations that have deployed Citrix Virtual Apps solutions in their portfolio, for example, to allow users to view files created in 3D applications.

Limitation:

- No Desktop Virtualization (e.g. Citrix Virtual Desktop Windows 10)
- No Linux OS
- Maximum 1 Display Head
- Maximum Resolution of 1280 x 1024
- No CUDA & OpenCL

NVIDIA GRID VIRTUAL PC

NVIDIA Grid vPC is ideal for users who need a virtual desktop, to run client applications, browsers and HD video with increased graphics performance.

Limitation:

- Maximum 4 Display Heads
- Maximum Resolution per Display Head of 4096 x 2160 (by 2 Heads) or 2560 x 1600 (by 4 Heads)
- No CUDA & OpenCL
- No GPU Pass-through
- No Bare Metal Support
- Maximum Frame Buffer of 2 GB

NVIDIA QUADRO VIRTUAL DATA CENTER WORKSTATION

This edition is ideal for mainstream and high-end designers using powerful 3D applications, such as SOLIDWORKS, 3DExcite, Siemens NX or Autodesk Maya.

Limitation:

- Maximum 4 Display Heads
- Maximum Resolution per Display Head of 4096 x 2160

NVIDIA vGPU software is licensed per concurrent user. Each product includes the following feature entitlement:

Feature	GRID vApps	GRID vPC	Quadro vDWS
License Entitlement			
Concurrent User (CCU)	Yes	Yes	Yes
Capability Entitlement			
Desktop Virtualization		Yes	Yes
RDSH App Hosting	Yes	Yes	Yes ¹
RDSH Desktop Hosting	Yes	Yes	Yes ¹
Windows Guest OS	Yes	Yes	Yes
Linux Guest OS		Yes	Yes
Maximum Displays	1 ²	4	4
Maximum Resolution	1280*1024	4096*2160 (4K) ³	4096*2160 (4K)
NVIDIA Quadro Software Features			Yes
CUDA & OpenCL Supported	No		Yes ⁴
GPU Pass-through Supported ⁵	Yes		Yes
Bare Metal Supported ⁶	Yes		Yes
vGPU Profiles Supported⁷			
512 MB		Yes	Yes
1 GB	Yes	Yes	Yes
2 GB	Yes	Yes	Yes
3 GB	Yes		Yes
4 GB	Yes		Yes
6 GB	Yes		Yes
8 GB	Yes		Yes
12 GB	Yes		Yes
16 GB	Yes		Yes
24 GB	Yes		Yes

¹With packaged GRID vApps license

²Applies only to the console display in remote application environments. For details, see [Supported GPUs](#)

³Supports up to two 4K displays or four 2560x1600 displays on 2B profile. Supports up to four 2560x1600 displays on 1B profile.

⁴Supported on 8GB 1:1 profile on Maxwell and all profiles on Pascal

⁵Only supported on 1:1 profiles

Example License Calculation

PERPETUAL LICENSE

Example for 10 simultaneous user

Quadro vDWS Perpetual License for 1 CCU over 3 years

License cost for 1 user -> 450\$

SUMS for 3 years -> 300\$

Total cost for 1 user -> 750\$

For 10 user -> 7500\$

Renew the SUMS after the 3 years for another year

For 10 user -> 1000\$

Total cost for 4 years -> 8500\$

Total cost for 5 years -> 9500\$

ANNUAL SUBSCRIPTION

Example for 10 simultaneous user

Quadro vDWS Annual Subscription for 1 CCU over 3 years

License cost for 1 user for 3 years (incl. SUMS) -> 750\$

For 10 user -> 7500\$

Renew the license after the 3 years for another year

For 10 user -> 2500\$

Total cost for 4 years -> 10000\$

Total cost for 5 years -> 12500\$

Related Posts:

1. [Optimization of VMs by NUMA](#)
2. [Create a SQL Maintenance Plan for Citrix DB](#)
3. [What's new in Citrix Virtual Apps and Desktops 7 1811](#)
4. [Install Teams & OneDrive in Citrix \(Machine-Based\)](#)



Citrix, NVIDIA, Virtual
Apps and Desktops
/ Annual
Subscription, Frame
Buffer, GRID vApps,
GRID vPC,
Hypervisor, NVIDIA,
Perpetual License,
Quadro vDWS, SUMs,
vGPU, Virtual Apps,
Virtual Desktop,
Virtual GPU, XenApp,
XenDesktop

[Deyda.net](#) / [Data Protection Declaration](#) / Proudly powered by [WordPress](#)