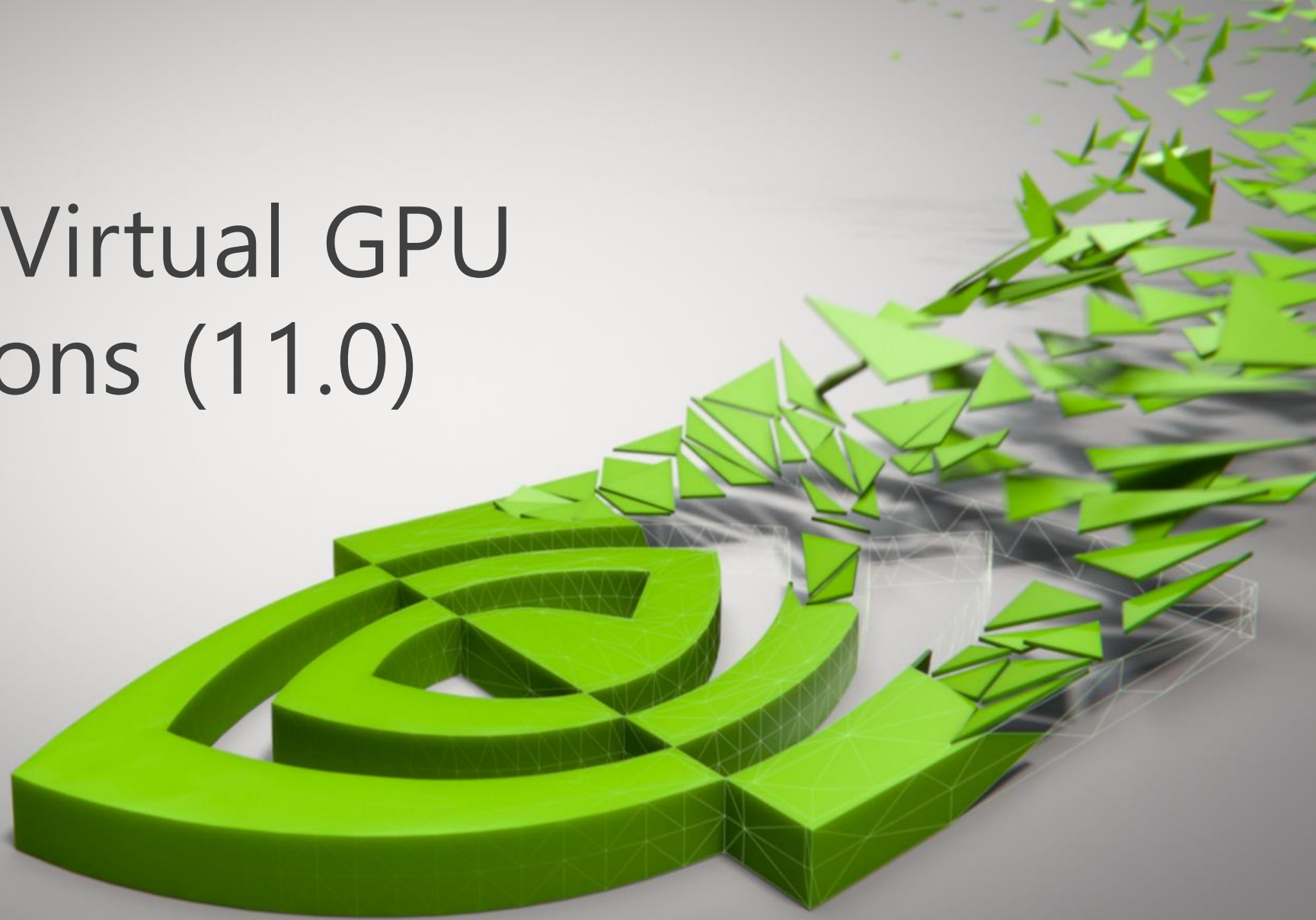


# NVIDIA Virtual GPU Solutions (11.0)

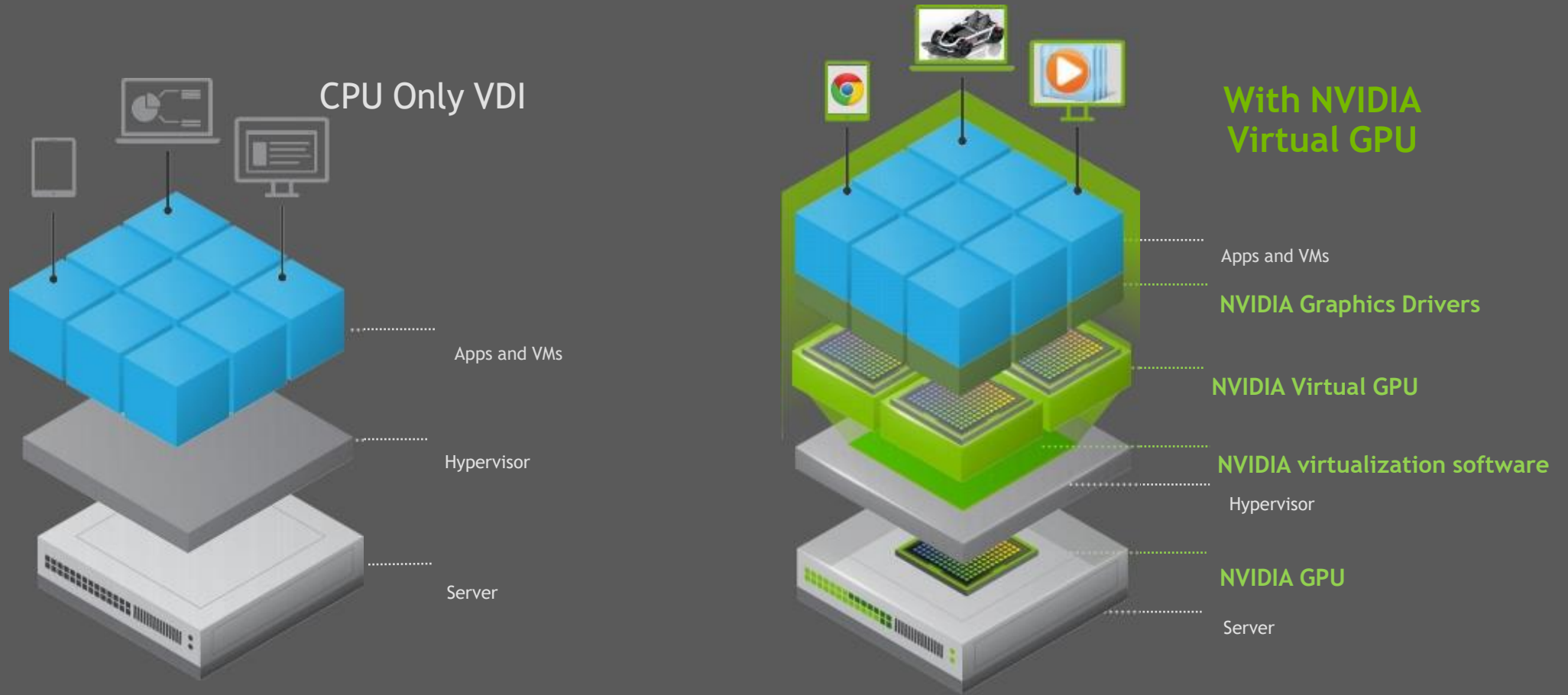


01

# NVIDIA vGPU Overview

# NVIDIA Virtual GPU Solutions 개요

NVIDIA Virtual GPU 제품은 모든 가상 데스크톱에 GPU 성능을 제공합니다.



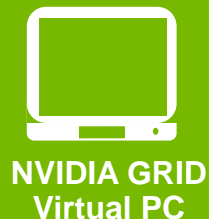
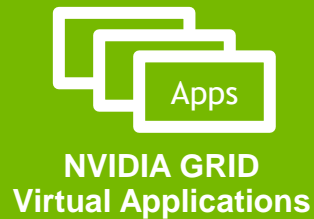
©2018 VMware, Inc.

12

# NVIDIA Virtual GPU Solutions 구성 요소

NVIDIA Virtual GPU Solutions은 기업, 연구, 교육, 의료분야등의 업무 환경에 최적화된 중앙 집중형 가상 그래픽 솔루션입니다. Platform은 크게 하드웨어 그래픽 카드, vGPU Software, SUMS 로 구성되어 있습니다.

## Support, Updates & Maintenance



## NVIDIA vGPU Software



## NVIDIA Tesla (Datacenter GPU)



## Hypervisor ( ESXi , Citrix, Hyper-V, Nutanix )



\*\* Nutanix AHV 의 경우 5.18 버전부터 지원 방향이 변경됨 - vGPU 11.0 버전 이상의 경우 vGPU 드라이버 SW 다운로드는 Nutanix사이트에서 지원

상세 문의처 정보: goodmkt@goodmit.co.kr

# 가상 그래픽의 한계

기존 VDI 환경은 그래픽 가속이 필요한 업무 관리에 한계가 있습니다.



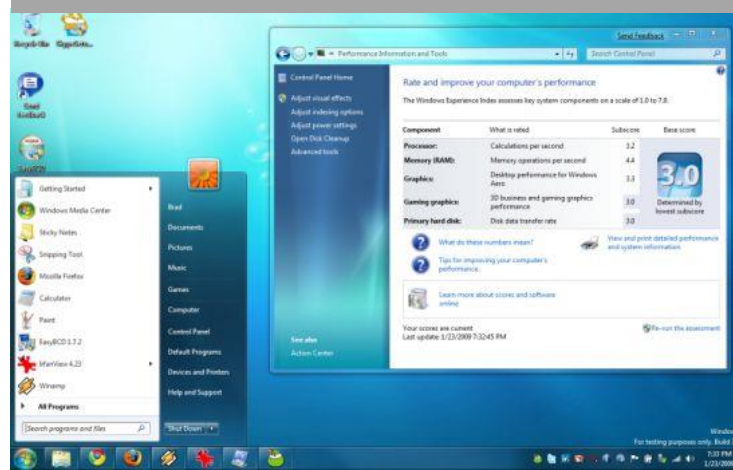
# WINDOWS 10 그래픽 사용량

현존하는 OS 중 가장 높은 그래픽 성능 요구

## WINDOWS XP



## WINDOWS 7



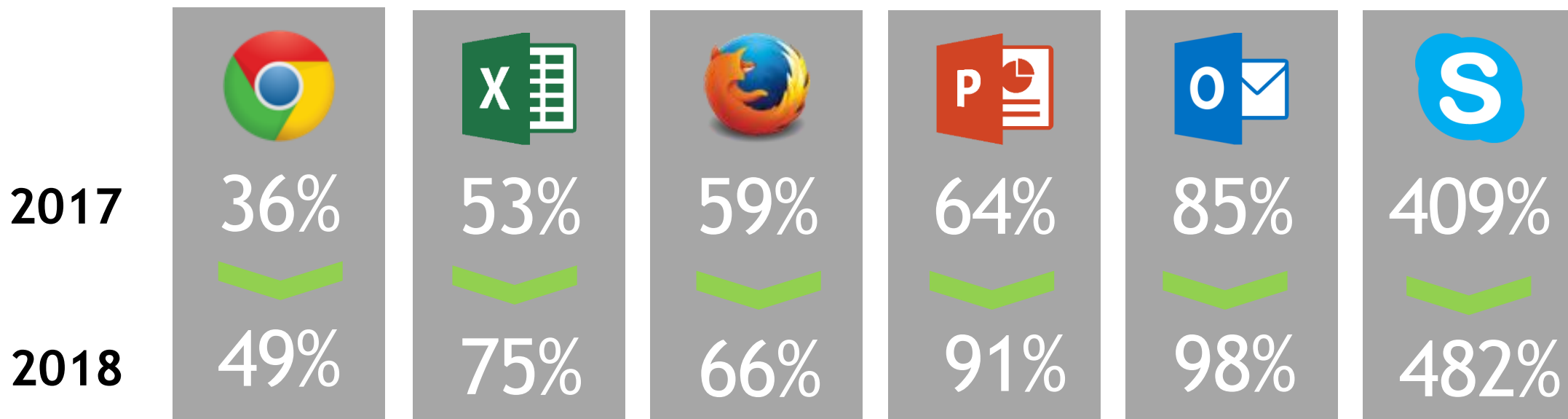
## WINDOWS 10



Windows 7\* 대비 GPU 사용량 **32%** 증가

# 디지털 워크스페이스의 최근 앱들

## Demands More GPU



2012년 이후 그래픽 성능을 요구하는 애플리케이션은 두 배로 늘었습니다

전체 대비 절반이 넘는 기업 사용자들이 최소 하나의 그래픽 가속 앱을 사용합니다

# 그래픽 가속은 선택이 아닌 필수

오늘날의 애플리케이션 및 OS는 모두 GPU 가속을 필요로 합니다



2x

2011년 이후 그래픽 가속이  
필요한 애플리케이션의 수

56%

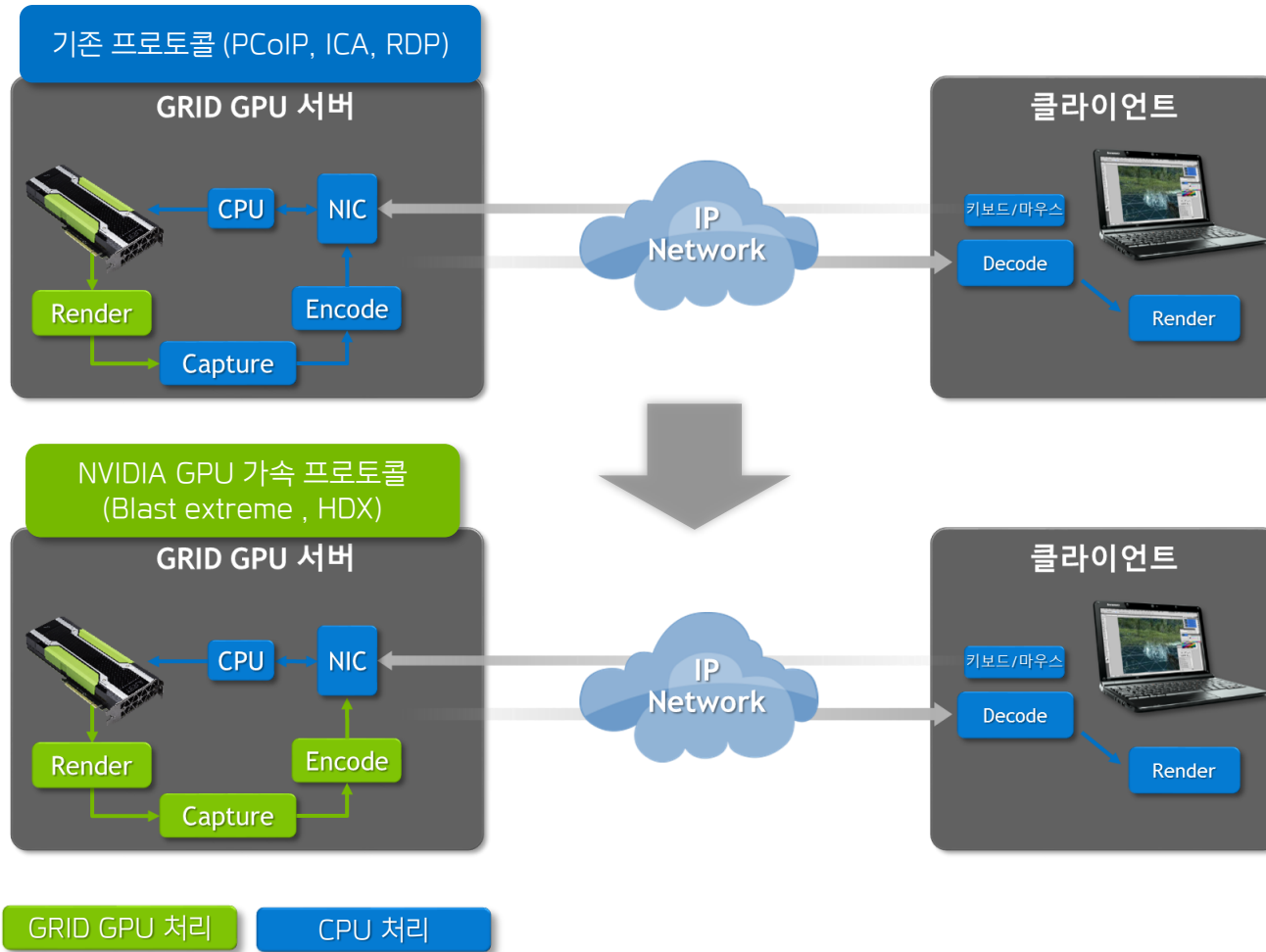
절반이 넘는 기업 사용자가 그래픽  
가속 애플리케이션 사용

50%

대부분의 기업이 2017년 1월을  
기점으로 Win10 사용 중



# NVIDIA GPU 가속 프로토콜



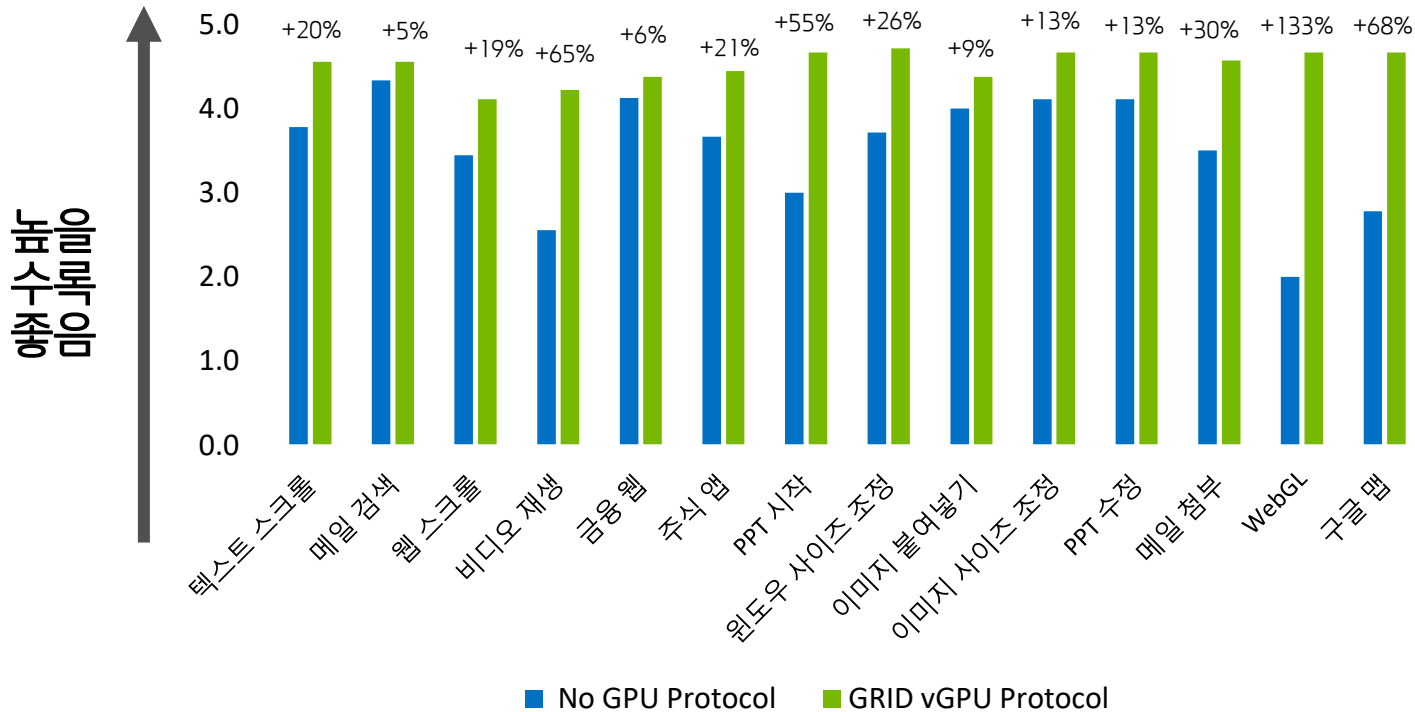
Geometria de Simulação | Inspeccionar | Ferramentas | Suplementos | Visualizar

0 itens são selecionados | Localiza

- 평균 34%** 사용자 경험 향상
- 140ms 감소** 전체 시스템 지연
- 최대 42%** 호스트 CPU 사용률 절감

# WINDOWS 10 WITH CPU vs. NVIDIA vGPU

GPU 가속 프로토콜은 사용자 경험을 평균 34% 향상시킵니다



- 사용자 경험 단위
- 5 최상급 - 거의 기존 PC와 동일
  - 4 가상 데스크톱 기준으로 좋음
  - 3 참고 사용할 만 함
  - 2 간신히 사용 가능한 수준
  - 1 도저히 사용 불가능한 성능 및 안정성

두 대의 동일한 시스템에서 테스트됨, CPU 및 GPU 시스템은 모두 최대 60~80% 수준의 부하 수준에서 테스트 진행

02

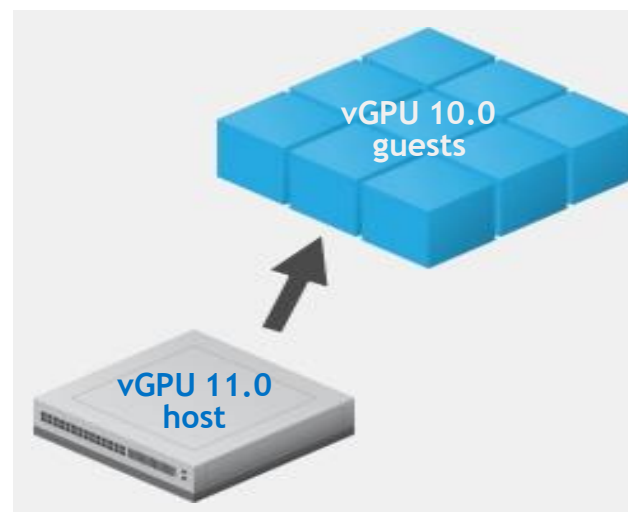
# NVIDIA vGPU 11.0 업데이트

# NVIDIA vGPU 업데이트(vGPU 11.0)



License grace period  
라이선스 유예기간

라이선스가 없는 가상 및 물리 GPU 환경에서도 일정기간 (유예기간)동안은 가속가능. 그러나 시간이 지남에 따라 성능이 저하됨.



Cross Branch Compatibility  
vGPU 10.0 and later  
(크로스 브랜치 드라이버 지원)

게스트OS 업그레이드 없이 최신 버전 SW드라이버와 함께 사용 가능.  
(단 , 호스트 드라이버는 업데이트는 필요 )

- <https://docs.nvidia.com/grid/> 참고

# NVIDIA vGPU 업데이트(vGPU 11.0)

## Enabling the Latest Hypervisor Releases

ENTERPRISE	HYPERVISOR
<ul style="list-style-type: none"><li>• Windows10 May 2020(2004) 지원</li><li>• Ubuntu 20.04 LTS Guest OS 지원 (Citrix and VMware)</li><li>• Citrix Virtual Apps and Desktop 7 2006 지원</li></ul>	<ul style="list-style-type: none"><li>• Citrix Hypervisor 8.2(Multi-vGPU)</li></ul>

# NVIDIA vGPU 업데이트(vGPU 11.0)

## Feature Support Withdrawn in Release 11.0

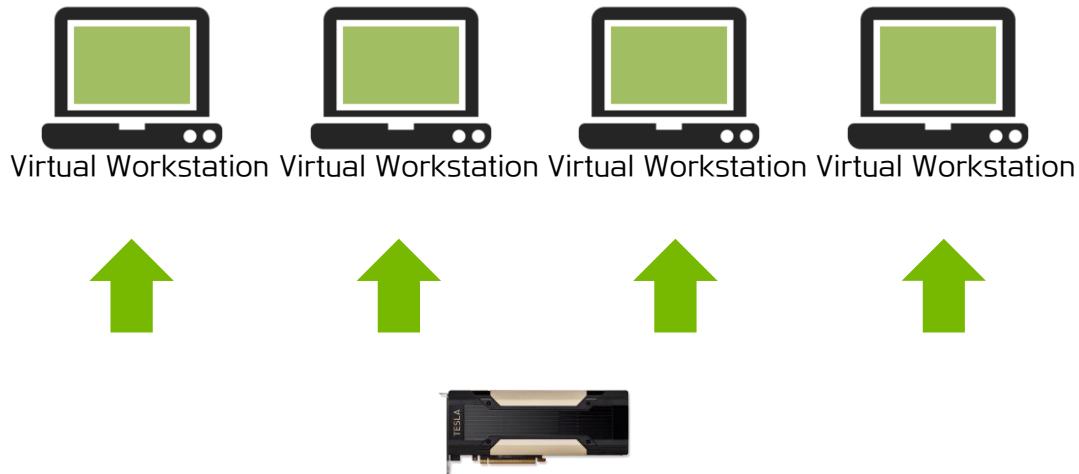
### ENTERPRISE

- Citrix Hypervisor 8.0 지원 종료
- Guest OS 릴리스 버전 지원 종료 리스트
  - Windows Server 2012
  - Windows Server 2008 R2
  - Windows 8.1
  - Windows 8
  - Windows 7

# vGPU What is multi-vGPU?

업계 최고의 성능을 자랑하는 가상 워크스테이션

여러 VM's 걸쳐 GPU 공유  
NVIDIA Quadro vDWS & vCS



NVIDIA Quadro vDWS & vCS 를 통해 하나에  
VM가 여러 Tesla GPU를 액세스 가능



# Multiply the power of vDWS & vCS

NVIDIA Quadro vDWS & vCS를 위한 탁월한 성능

- 가상 머신 당 최대 16개의 NVIDIA GPU의 성능 활용
  - ✓ RHEL 7.5 KVM 하이퍼바이저 및 NVIDIA vGPU 소프트웨어 7.0 이상 에서 지원
  - ✓ Quadro vDWS로 지원 (GPU 수에 관계 없이 VM당 라이선스 필요)
  - ✓ 동일한 GPU(no mixing) 및 최대 GPU로 지원 가능
  - ✓ Supported with Turing\*, Volta, Pascal, and Maxwell GPUs
- Quadro Virtual Data Center Workstation(Quadro vDWS) & NVIDIA vCompute Server(vCS)을 통해 멀티 vGPU의 성능을 필요로 하는 advanced design 및 visualization workflows 지원:
  - ✓ GPU Rendering
  - ✓ Simulation & Design



# 지원 가능한 vGPU 프로파일 (Multi-GPU지원)

물리적 GPU의 모든 프레임 버퍼가 할당된 Q-C시리즈 vGPU만 지원됨.

GPU Architecture	Board	vGPU
Turing	Quadro RTX 6000	Quadro RTX 6000-24Q
		Quadro RTX 6000-24C
	Quadro RTX 8000	Quadro RTX 8000-48Q
		Quadro RTX 8000-48C
	T4	T4-16Q
		T4-16C
Volta	V100 SXM2 32GB	V100DX-32Q
		V100DX-32C
	V100 PCIe 32GB	V100D-32Q
		V100D-32C
	V100S PCIe 32GB	V100S-32Q
		V100S-32c
	V100 SXM2	V100X-16Q
		V100X-16C
	V100 PCIe	V100-16Q
		V100-16C
	V100 FHHL	V100L-16Q
		V100L-16C

# 지원 가능한 vGPU 프로파일 (Multi-GPU지원)

물리적 GPU의 모든 프레임 버퍼가 할당된 Q-C시리즈 vGPU만 지원됨.

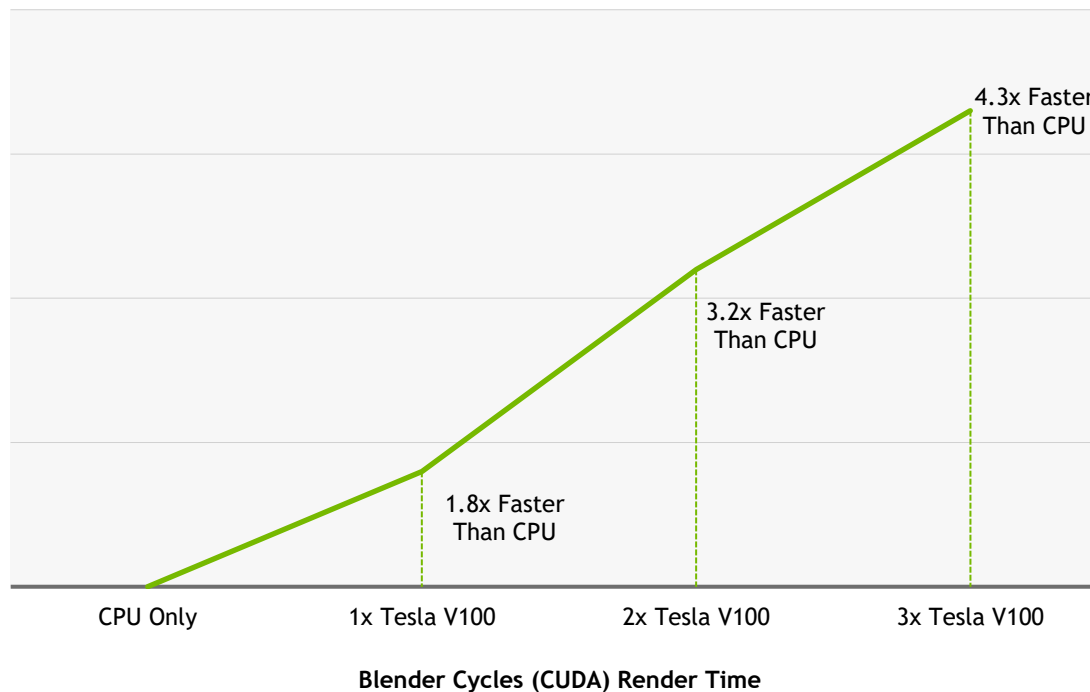
GPU Architecture	Board	vGPU
Pascal	P100 SXM2	P100X-16Q
		P100X-16C
	P100 PCIe 16GB	P100-16Q
		P100-16C
	P100 PCIe 12GB	P100C-12Q
		P100C-12C
	P40	P40-24Q
		P40-24C
	P6	P6-16Q
		P6-16C
P4	P4-8Q	
	P4-8C	
Maxwell	M60	M60-8Q
	M10	M10-8Q
	M6	M6-8Q

Maxwell 버전은 NVIDIA vCompute Server(vCS) 지원불가

# 총 렌더링 시간이 4.3x 가속화

Blender Cycles (CUDA) Render Time

3개의 Tesla V100-32Q GPU 와  
NVIDIA Quadro vDWS CPU 로  
**4.3X**나 빠른 시각화 작업 렌더링을  
제공한다.



테스트는 2x 인텔 제온 골드 (6154 3.0 GHz) CPU, 512GB RAM, RHEL 7.5, NVIDIA Quadro vDWS 소프트웨어, Tesla V100-32Q, 드라이버 - 410.39, 256GB RAM, Windows 10 x64 RS3

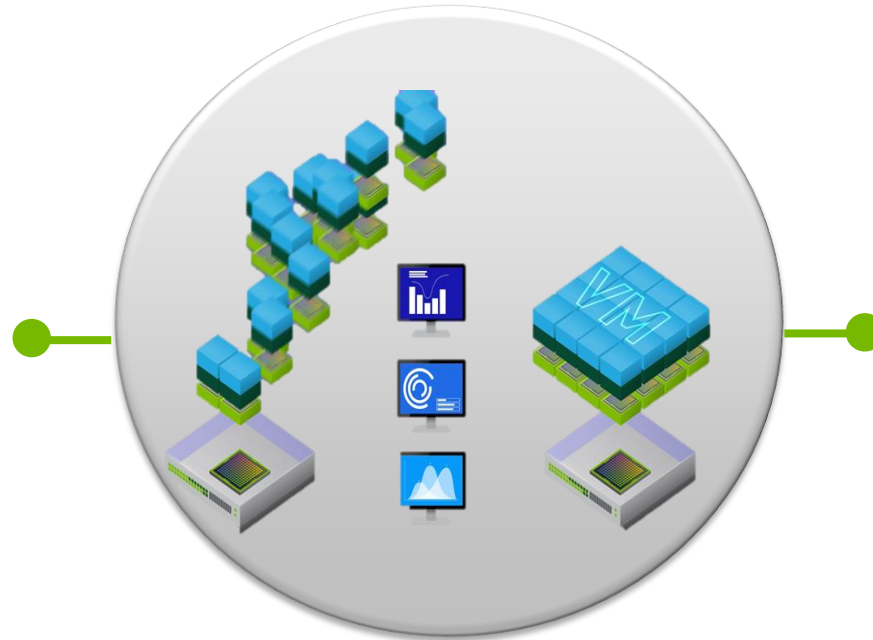
# Live Migration support for vGPU

현대적인 민첩한 데이터센터

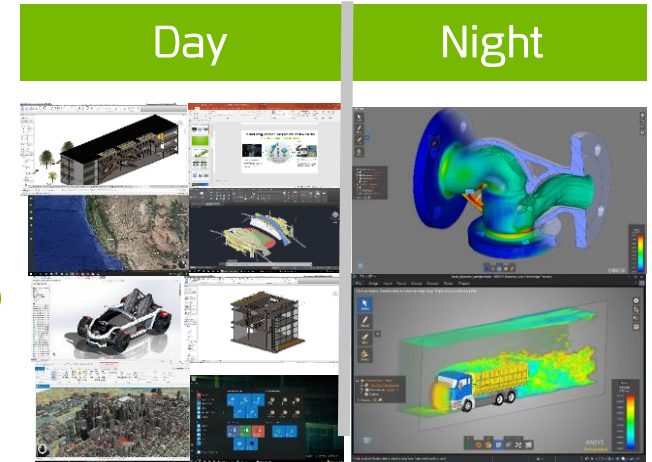


## 서버를 최상의 상태로 유지

원하는 시간에 패치나 업그레이드를 실행하고, 최종 작업자의 작업 중단이나 데이터 손실없이 워크로드 밸런싱을 손쉽게 수행할 수 있다.



VMware vMotion  
Citrix XenMotion



## 인프라 ROI 최대화

낮에 VDI를 실행하는 호스트를 변경하여 동일한 인프라를 사용하여 혼합 워크로드를 실행한다. (HPC 및 기타 컴퓨팅 워크로드 실행)

# 온라인 마이그레이션 지원 범위

<https://docs.nvidia.com/grid/latest/grid-vgpu-release-notes-citrix-xenserver/index.html>

- vGPU를 사용하는 XenMotion은 Citrix **XenServer 8.1, 8.2** 에서 지원 (P100 카드 XenMotion No Support)
- vGPU를 사용하는 vMotion은 VMware **vSphere 6.7 U1 이상**에서 지원됨 (P100 카드 vMotion No Support)

Supported VMware vSphere Hypervisor (ESXi) releases:

- Release 6.7 U1 및 호환 가능한 업데이트는 vGPU를 통한 vMotion 그리고 vGPU and suspend-resume 지원.
- Release 6.7 은 오직 suspend-resume만 지원.
- **6.7 이전의 Release 버전은 어떠한 형태의 vGPU 마이그레이션 지원하지 않음.**

## vGPU Migration Support

- Tesla M6
- Tesla M10
- Tesla M60
- Tesla P4
- Tesla P6
- Tesla P40
- Tesla V100 SXM2
- Tesla V100 SXM2 32GB
- Tesla V100 PCIe
- Tesla V100 PCIe 32GB
- Tesla V100S PCIe 32GB
- Tesla V100 FHHL
- Tesla T4
- Quadro RTX 6000
- Quadro RTX 6000 Passive
- Quadro RTX 8000
- Quadro RTX 8000 Passive

# Zero Downtime Migration

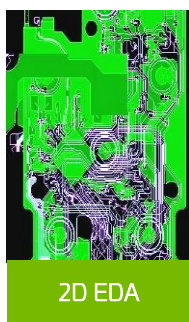


# NVIDIA vGPU로 혼합된 워크로드

생산성과 활용도 증가, 지출 감소



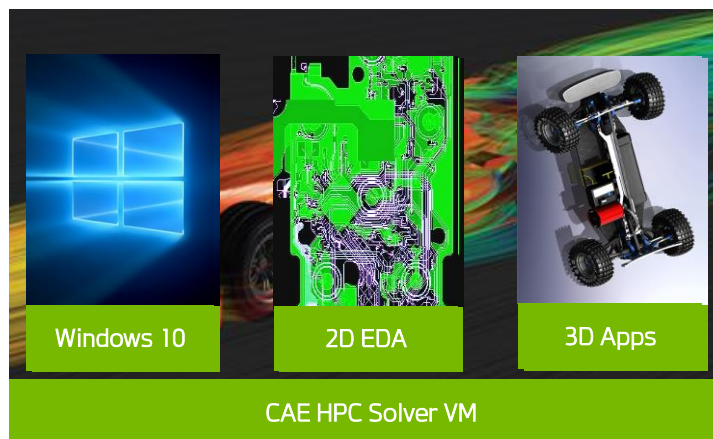
Windows 10



2D EDA



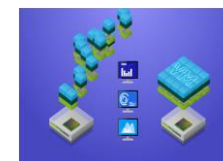
3D Apps



CAE HPC Solver VM



온라인마이그레이션  
with vGPU



End-to-end GPU  
insights with  
vROPS



03

## NVIDIA 그래픽 카드 종류



# NVIDIA 그래픽 카드 종류(TESLA & Quadro)

총 9가지 종류의 그래픽 카드가 있으며, M6과 P6은 블레이드 서버용 카드이고,  
T & V & P시리즈는 “Turing & Volta & Pascal” 아키텍처 사용으로 기존 M시리즈 보다 더 높은 성능을 제공 합니다.

	V100S / V100 SXM2	RTX 8000	RTX 6000	P40	T4	M10	P6
GPUs / Board (Architecture)	1 (Volta)	1 (Turing)	1 (Turing)	1 (Pascal)	1 (Turing)	4 (Maxwell)	1 (Pascal)
CUDA Cores	5,120	4,608	4,608	3,840	2,560	2,560 (640 per GPU)	2,048
Tensor Cores	640	576	576	--	320	--	--
RT Cores	--	72	72	--	40	--	--
Total Memory Size	32 GB HBM2	48 GB GDDR6	24 GB GDDR6	24 GB GDDR5	16 GB GDDR5	32 GB GDDR5 (8 GB per GPU)	16 GB GDDR5
vGPU Profiles	1 GB, 2 GB, 4 GB, 8 GB, 16 GB, 32 GB	1 GB, 2 GB, 3 GB, 4 GB, 6 GB, 8 GB, 12 GB, 24 GB	1 GB, 2 GB, 3 GB, 4 GB, 6 GB, 8 GB, 12 GB, 24 GB	1 GB, 2 GB, 3 GB, 4 GB, 6 GB, 8 GB, 12 GB, 24 GB	1 GB, 2 GB, 4 GB, 8 GB, 16 GB	0.5 GB, 1 GB, 2 GB, 4 GB, 8 GB	1 GB, 2 GB, 4 GB, 8 GB, 16 GB
Form Factor	PCIe, SXM2, Full Height/Length	PCIe 3.0 Dual Slot	PCIe 3.0 Dual Slot	PCIe 3.0 Dual Slot (rack servers)	PCIe 3.0 Single Slot (rack servers)	PCIe 3.0 Dual Slot (rack servers)	MXM (blade servers)
Power	250W/300W	260 W / 250 W	260 W / 250 W	250W	70W	225W	90W
Thermal	passive	Active or Passive	Active or Passive	passive	passive	passive	bare board
Recommended Use case	Quadro vDWS - Performance-Optimized; NVIDIA vComputeServer - Compute Optimized	Quadro vDWS - Performance Optimized (High end)	Quadro vDWS - Performance Optimized (High end)	Quadro vDWS - Performance Optimized (Mid to high)	Quadro vDWS - Performance Optimized (Entry to mid); GRID vPC - Density Optimized; NVIDIA vComputeServer - Compute Optimized	GRID vPC - User Density-Optimized	Quadro vDWS or GRID vPC - Blade-Optimized

# 카드 사용 요구사항

각 카드마다 호환성이 인증된 서버들이 있습니다. 반드시 인증 서버 여부를 확인하시기 바랍니다.

또한 이미 보유 중인 호환 서버를 사용하더라도 적정 파워서플라이로 업그레이드가 필요하거나 카드 장착을 위한 별도의 설치 Kit이 필요할 수 있습니다.

NVIDIA vGPU엔터프라이즈 소프트웨어는 특정한 NVIDIA Tesla GPU에서만 사용 가능하도록 디자인 되었습니다.

하기의 사항들을 확인하시기 바랍니다:

- ✓ NVIDIA vGPU 인증 서버  
: [www.nvidia.com/grid-certified-servers](http://www.nvidia.com/grid-certified-servers)
- ✓ GPU 하드웨어
- ✓ 지원되는 가상화 플랫폼

**Release note를 읽으세요!**

- ✓ 특정 NVIDIA vGPU 소프트웨어에 테스트된 정확한 가상화 소프트웨어 빌드 번호
- ✓ 기타 중요한 버전 별 정보

The screenshot displays the NVIDIA Grid Certified Servers website. At the top, it lists 'PREFERRED NVIDIA GRID CERTIFIED SERVER VENDORS' with logos for Cisco, Dell, HP, Lenovo, and Supermicro. Below this, the 'NVIDIA GRID CERTIFIED SERVERS' section includes filters for 'All Manufacturers', 'All Heights', 'All Form Factors', and 'Tesla M60', along with a search bar and a 'Share URL...' button. The main content is a table of certified servers.

MANUFACTURER	MODEL	RACK UNITS	NODE PER CHASSIS	MAXIMUM CARDS PER NODE M60	MORE INFO
ASRock Rack	3U8G	3	1	4	
Asus	Asus ESC4000 G3	2	1	4	기
	Cisco UCS C240 M4	2	1	2	기
	Cisco UCS C460 M4	2	1	2	기
	Dell PowerEdge R730	2	1	2	기
	Dell PowerEdge C4130	1	1	4	기

04

# NVIDIA vGPU Edition

# NVIDIA Virtual GPU Software Editions



XenApp or RDSH  
기반 솔루션을  
사용하고자 하는 고객.  
윈도우 기반 애플리케이션을  
최고의 성능으로 제공

NVIDIA GRID Virtual  
Applications  
(vAPP)



가상 데스크톱 환경에서  
윈도우 애플리케이션,  
인터넷 브라우저 및 고해상도  
비디오를 사용하며 뛰어난  
사용자 경험을 요구하는 경우

NVIDIA GRID Virtual PC  
(vPC)



어디서나 장치 상관 없이  
원격으로 완벽한 성능의  
전문가용 그래픽  
애플리케이션을  
사용하고자 하는 사용자

NVIDIA Quadro Virtual Data  
Center WorkStation  
(vDWS)



데이터 센터 HPC 활성화  
AI/DL 워크로드 및  
NVIDIA vGPU 보안,  
관리 및 확장성

NVIDIA Virtual Compute Server  
(vCS)

Edition	vAPP	vPC	vDWS	vCS
Target User	Knowledge Worker	Knowledge Worker	Creative & Technical Professional	Data Scientist
Compute Type	Client Computing	Client Computing	Server Workloads & Client Computing	Server Workloads (Client Computing -only Linux)
Type of deployment	Hosted Shared Desktops	VDI Desktops	VDI Desktops (Professional Graphics)	Compute & Deep Learning
라이선스 부여 방식	Concurrent User (CCU)	Concurrent User (CCU)	Concurrent User (CCU)	Per GPU (MAX 8VM)
구매방식	Annual Subscription or Perpetual License + SUMS	Annual Subscription or Perpetual License + SUMS	Annual Subscription or Perpetual License + SUMS	Annual Subscription (1,3,5 Yr)

# NVIDIA Virtual GPU Software Editions 기능비교표

NVIDIA vGPU software product 에디션 별 기능 비교표입니다.

FEATURE	NVIDIA GRID Virtual Apps	NVIDIA GRID VIRTUAL PC	NVIDIA GRID VIRTUAL WORKSTATION	NVIDIA vCompute Server
<b>License Entitlement</b>				
Concurrent User (CCU)	✓	✓	✓	
Per GPU				✓ <sup>1</sup>
<b>Capability Entitlement</b>				
Desktop Virtualization		✓	✓	
RDSH App Hosting	✓	✓	✓ <sup>2</sup>	
RDSH Desktop Hosting	✓	✓	✓ <sup>2</sup>	
Compute Virtualization			✓	✓
Windows Guest OS	✓	✓	✓	
Linux Guest OS		✓	✓	✓
Maximum Displays	1 <sup>3</sup>	4	4	1
Maximum Resolution <sup>4</sup>	1280*1024	5120 * 2880 (5K)	7680 * 4320 (8K)	4096*2160 (4K)
NVIDIA Quadro Software Feature			✓	
CUDA & OpenCL Supported			✓ <sup>5</sup>	✓
ECC & Page Retirement			✓	✓ <sup>6</sup>
Multi-GPU			✓	✓
NVLINK			✓	✓

1. GPU 최대 8개의 동시 VM
2. Package 된 GRID vAPP 라이선스
3. 원격 응용 프로그램 환경의 디스플레이에만 적용 (자세한 내용은 support GPU 참조)
4. 각프로파일에 지원되는 디스플레이 구성에 대해서는 Virtual GPU Software User Guide를 참조하십시오
5. Tesla Maxwell 카드의 8GB 1 : 1 프로파일에에서만 지원 pascal 의 모든 프로파일에에서 지원
6. ECC 지원은 파스칼 이상 지원

# NVIDIA Virtual GPU Software Editions 기능비교표

NVIDIA vGPU software product 에디션 별 기능 비교표입니다.

FEATURE	NVIDIA GRID Virtual Apps	NVIDIA GRID VIRTUAL PC	NVIDIA GRID VIRTUAL WORKSTATION	NVIDIA vCompute Server
NVLINK			✓	✓
GPU Pass-through Support <sup>7</sup>	✓		✓	✓
Bare Metal Supported <sup>8</sup>	✓		✓	
vGPU Profile Sizes supported <sup>9</sup>				
512 MB		✓	✓	
1 GB	✓	✓	✓	
2 GB	✓	✓	✓	
3 GB	✓		✓	
4 GB	✓		✓	✓
6 GB	✓		✓	✓
8 GB	✓		✓	✓
12 GB	✓		✓	✓
16 GB	✓		✓	✓
24 GB	✓		✓	✓
32 GB	✓		✓	✓
48 GB			✓	✓

7. 1:1 프로파일에서만 지원

8. NVIDIA M6 하드웨어만 기본 디스플레이 장치로 지원

9. GPU에서 지원되는 vGPU 프로파일에 대한 Virtual GPU Software User Guide를 검토하십시오.

# SUMS 제공 내용

vGPU Software에 대한 기술 지원, 소프트웨어 업데이트, 장애 시 기술 지원을 받을 수 있습니다.

구분	내용
유지 보수	최대 3년간 버그 수정 업데이트, 장애 지원 및 유연한 업그레이드를 위한 보안 패치 제공
업그레이드	새로운 기능 추가 및 신규 하드웨어 지원을 포함한 모든 주요 버전 업데이트 제공
지원 기간	최대 3년 단위
직접 지원	NVIDIA vGPU E-Mail 및 NVIDIA vGPU Support Portal
응답 시간	24 * 7
Knowledgebase 접근	✓
Web 지원	✓
E-mail 지원	✓
전화 지원	✓

# NVIDIA Virtual GPU Software Editions 구매옵션

GRID vPC/vApp 및 Quadro vDWS 라이선스는 두 가지 방식으로 구매가 가능하며 vCS 라이선스는 연 단위 라이선스로만 구매가 가능합니다. (단,그래픽카드는 별도로 구매를 하셔야 합니다.)

## SUMS (Support, Updates & Maintenance)



**NVIDIA GRID  
Virtual Applications**



**NVIDIA GRID  
Virtual PC**



**NVIDIA Quadro  
Virtual Data Center  
WorkStation**



**NVIDIA Virtual  
Compute Server**



구매 방법 1 : Annual Subscription  
(1년, 3년, 4년, 5년 단위 판매)  
연 단위 라이선스는 1년 간 소프트웨어 할당, 지원, 업데이트, 유지보수를 제공합니다. 3년 단위 계약도 가능합니다. (SUMS가 포함되어 있으며, 초기 라이선스 도입 비용이 Perpetual 방식에 비해 저렴함)



구매 방법 2 : Perpetual License + SUMS  
( 4년, 5년 단위 판매)  
영구 라이선스는 최초 구매 후 소프트웨어를 기간 제한 없이 지속적으로 사용할 수 있습니다. (단 SUMS 는 별도로 구매해야 하며, SUMS 는 매년 갱신해야 합니다.)

구매 방법 : Annual Subscription  
(1년 , 3년 , 5년 단위로만 판매)

연 단위 라이선스는 1년 간 소프트웨어 할당, 지원, 업데이트, 유지보수를 제공합니다. (SUMS가 포함되어 있으며, 초기 라이선스 도입 비용이 Perpetual 방식에 비해 저렴함)



# vGPU Software 배포 방식 라이선스 선정 방법

GRID Software 라이선스를 적용하는 원칙은 사용자가 그래픽 가속 처리를 위한 업무로 NVIDIA Tesla CARD 및 Quadro RTX 6000,8000 도입 시 반드시 NVIDIA vGPU Software 라이선스를 구매 해야함

NVIDIA vGPU Software Licensed Product	Target Users	Supported NVIDIA vGPU Software Deployments
GRID Virtual Applications (vAPP)	Citrix XenApp, VMware Horizon 또는 기타 RDSH 솔루션을 사용하는 PC 수준 응용 프로그램 및 서버 기반 데스크톱 사용자	<ul style="list-style-type: none"> <li>A-series NVIDIA vGPUs (GRID Virtual Applications)</li> <li>GPU pass through</li> <li>Microsoft DDA</li> <li>VMware vDGA</li> <li>Bare metal</li> </ul>
GRID Virtual PC (vPC)	Windows, 웹 브라우저 및 고화질 비디오 용 PC 응용 프로그램에 대한 뛰어난 사용자 경험이 필요한 비즈니스 가상 데스크톱 사용자	<ul style="list-style-type: none"> <li>B-series NVIDIA vGPUs (GRID Virtual PC or Quadro vDWS)</li> <li>Microsoft RemoteFX vGPU</li> <li>VMware vSGA</li> </ul>
vCompute Server (vCS)	인공지능(AI), 심층 학습 또는 고성능 컴퓨팅(HPC) 워크로드를 위한 컴퓨팅 집약적 가상 서버 사용자	<ul style="list-style-type: none"> <li>C-series NVIDIA vGPUs</li> <li>GPU pass through</li> <li>Microsoft DDA</li> <li>VMware vDGA</li> </ul>
Quadro vDWS (vDWS)	어디서나 모든 장치에서 완벽한 성능으로 원격 전문 그래픽 응용 프로그램에 액세스 해야하는 미드 레인지 및 하이 엔드 워크 스테이션 사용자	<ul style="list-style-type: none"> <li>Q-series NVIDIA vGPUs (Quadro vDWS)</li> <li>B-series NVIDIA vGPUs (GRID Virtual PC or Quadro vDWS)</li> <li>GPU pass through</li> <li>Microsoft DDA</li> <li>VMware vDGA</li> <li>Bare metal</li> </ul>

1. NVIDIA CARD를 Pass-through 방식으로 구성 하여도 용도가 그래픽 가속 처리를 위한 목적이라면 소프트웨어 라이선스를 반드시 구매
2. Bare metal 환경에서 Horizon RDSH App 이나 XenApp (어플리케이션 가상화) or 그래픽 가속 처리를 위한 용도이면 소프트웨어 라이선스 구매
3. Microsoft DDA 와 VMware vDGA는 기능 적으로 GPU Pass-through 동일함
4. Microsoft RemoteFX vGPU 및 VMware vSGA는 그래픽 가상화 기술 (Software 라이선스 필요)
5. Tesla CARD, Quadro RTX 6000,8000 를 적용 시 NVIDIA vGPU 드라이버를 설치 하지 않으면 정상적으로 그래픽 가속 되지 않음
6. Microsoft DDA 에는 Hyper-V 역할이 있는 Windows Server 2016 이 필요함

# vGPU Software 라이선스 부여 방법

“EULA only”는 라이선스를 수량을 기술적으로 확인하기 어려운 경우 적용 되는 원칙 (예: vApps 구성 시 RDSH 서버는 1대인 반면 동시에 접속 하는 수가 많아 별도 체크가 어려운 경우)이며 “Software”는 vGPU 라이선스 서버를 통해 확인

NVIDIA vGPU Software Deployment	Required NVIDIA vGPU Software License	Enforcement
A-series NVIDIA vGPU	GRID Virtual Applications	Software See Note (1)
B-series NVIDIA vGPU	GRID Virtual PC or Quadro vDWS	Software
C-series NVIDIA vGPU	vComputeServer or Quadro vDWS	Software See Note (2)
Q-series NVIDIA vGPU	Quadro vDWS	Software See Note (3)
GPU pass through for workstation or professional 3D graphics	Quadro vDWS	Software
GPU pass through for Compute-intensive virtual Servers	vComputeServer	Software See Note (4)
GPU pass through for PC-level applications	GRID Virtual Applications	EULA only
Microsoft DDA for workstation or professional 3D graphics	Quadro vDWS	Software
Microsoft DDA for Compute-intensive virtual Servers	vComputeServer	Software See Note (4)
Microsoft DDA for PC-level applications	GRID Virtual Applications	EULA only
VMware vDGA for workstation or professional 3D graphics	Quadro vDWS	Software
VMware vDGA for compute-intensive virtual servers	vComputeServer	Software See Note (4)
VMware vDGA for PC-level applications	GRID Virtual Applications	EULA only
Bare metal for workstation or professional 3D graphics	Quadro vDWS	Software
Bare metal for PC-level applications	GRID Virtual Applications	EULA only
Microsoft RemoteFX vGPU	GRID Virtual PC	EULA only
VMware vSGA	GRID Virtual PC	EULA only

05

# NVIDIA vGPU Sizing 가이드

# 가상화 환경에서의 NVIDIA vGPU (GRID) 제안 방법

NVIDIA vGPU 플랫폼은 크게 “데스크탑가상화 와 서버가상화 환경”에 제안 가능하며, 하이퍼바이저 및 VDI 솔루션과 NVIDIA CARD 지원 범위는 다음과 같습니다.

용도: 디스플레이 속도 or GPU Compute(CUDA)가 필요한 App

데스크탑 가상화  
+  
NVIDIA vGPU

물리 서버	데스크탑 가상화	NVIDIA CARD	NVIDIA Software 라이선스
호환성 가이드 참고	Horizon View	M10	vApps (7.0 이상) vPC (7.0 이상) vDWS (7.0 이상)
	Xen Desktop	P6	
	XenApp	T4	
	RDSH	RTX 6000 RTX 8000	
	Remote FX (bare metal)	V100	

용도: Deep Learning, AI용 컴퓨팅 등 연산 처리 작업

서버 가상화  
+  
NVIDIA vGPU

물리 서버	서버 가상화	NVIDIA CARD	NVIDIA Software 라이선스
호환성 가이드 참고	vSphere	T4	vDWS (7.0 이상) vCS (9.0 이상만 지원)
	Xen Server	P6	
	RHEL KVM (PT)	RTX 6000 RTX 8000	
	Nutanix AHV	V100	

NVIDIA Software 라이선스 7.0 이전 버전 서비스(EOL) Eed of License

<https://docs.nvidia.com/grid/7.0/product-support-matrix/>

## NVIDIA CARD 선택 시 고려 사항

- P6 : 블레이드 서버 용 카드
- M10, T4 : Windows 10, 고해상도 비디오 및 어플리케이션
- T4, V100, P40, RTX 6000, RTX 8000 :
  - 1) 전문가용 그래픽 어플리케이션이나 Deep Learning 등의 연산 처리용
  - 2) CAE (Ansys CFD, Abaqus), GPU based Rendering (Vray, Iray), multimedia editing (4K/8K video editing and transcoding) 등 CUDA (GPU Compute)가 필요한 Pro-Visualization App

## Software 라이선스 선택 시 고려 사항

- vApp : RDSH, XenApp 기반의 어플리케이션 가상화 제품
- vPC : Windows 10 유저, 고해상도 비디오 및 어플리케이션
- vDWS : 전문가용 그래픽 어플리케이션이나 3D컨텐츠 응용 프로그램 및 고급 설계자에게 이상적 (AI, DL, HPC 서버워크로드 지원 /윈도우 및 리눅스 OS 지원)
- vCS : 인공지능(AI), Deep Learning 또는 고성능(HPC)와 같은 컴퓨팅 집약적인 서버 워크로드 (리눅스 OS만 지원)

# NVIDIA vGPU 사이징을 위한 유저 Level 별 GPU 스펙 정리

NVIDIA vGPU 사이징을 위해 사용자 레벨(사용 어플리케이션)에 따라 GPU의 스펙을 아래와 같이 기준하여 고객에게 제안함

Deep learning, rendering, and GPGPU compute applications

High - End Qudadro vDWS, vCS



NVIDIA P40/RTX6000/RTX8000

NVIDIA V100

Largest CAD models, CAE, Photorealistic rendering, Seismic exploration, GPGPU compute

Large/complex CAD models, Seismic exploration, complex DCC effects, 3D Medical Imaging Recon

Large/complex CAD models, Advanced DCC, Medical Imaging



Medium size/complexity CAD models, Basic DCC, Medical Imaging, PLM

Small/simple CAD models, video, Entry PLM

Entry - Mid Range Qudadro vDWS

Office, Sketchup	PACS/Diagnostics	Schlumberger, Halliburton, DeltaGen, Catia Live Rendering
AutoCAD, Revit, Inventor		Ansys, Abaqus, Simulia
	Solidworks, Siemens NX, Creo, Catia, ArcGIS Pro	
Adobe CC Photoshop, Illustrator	Adobe CC Premiere Pro, After Effects, Autodesk Maya, 3ds Max, Mari, Nuke	

# NVIDIA vGPU Software 라이선스 별 서버 제안 사양

## NVIDIA Recommend GPU and Server Spec 가이드

라이선스 에디션	가상데스크탑 스펙	제안 스펙	수량	GPU 프로파일	서버당 유저 수
vApps	2~4vCPU, Mem 4GB, Disk 60GB, 50 IOPS	Intel Xenon Gold 6248 + NVIDIA Tesla M10, T4	M10 or T4 x 1~2	vPC (1B)	16 ~ 64 (1B)
vPC	2~8vCPU, Mem 4~8GB, Disk 60GB, 150 IOPS	Intel Xenon Gold 6248 + NVIDIA Tesla M10, T4	M10 or T4 x 1~2	vPC (1B) or vPC (2B)	32 ~ 64 (1B) 16 ~ 32 (2B)
vDWS	4~16vCPU, Mem 8~16GB, Disk 80GB, 200 IOPS	Intel Xenon Gold 6254 + NVIDIA Tesla T4, RTX 6000 ~8000, V100	T4 x 2 or RTX6000,8000, V100 x 2	vDWS (2Q) ~ vDWS (16Q)	4 ~ 32
vCS	8~16vCPU, Mem 16~32GB, Disk 100GB, 400 IOPS	Intel Xenon Gold 6254 + NVIDIA Tesla V100, V100s	V100 x 2 or V100s x 2	vCS (8C) ~ vCS (32C)	2 ~ 8

- 서버 1대 기준 사이징 작성
- Thick Disk 기준
- CUDA enable 어플리케이션 제안 시 T4,P40, RTX 6000,8000, V100 선택

• [http://www.teamrge.com/teamrge/wp-content/uploads/2019/07/ErikBohnhorst\\_Choosing\\_the\\_right\\_NVIDIA\\_-GPU\\_Shared\\_with\\_TeamRGE.pdf](http://www.teamrge.com/teamrge/wp-content/uploads/2019/07/ErikBohnhorst_Choosing_the_right_NVIDIA_-GPU_Shared_with_TeamRGE.pdf) 참고

# VDI FOR KNOWLEDGE & POWER WORKERS

Best User Experience for the Remote Workforce

## COST OPTIMIZED VDI SOLUTION

2-socket, 2U rack server  
2-4 vCPUs  
6 GB – 8GB memory (per user)  
Up to 3x NVIDIA M10 GPUs(1B)  
UP to 96x CCU licenses **NVIDIA GRID vPC Software**  
subscription

## FLEXIBILITY OPTIMIZED VDI SOLUTION

2-socket, 2U rack server  
2-4 vCPUs  
6 GB – 8GB memory (per user)  
Up to 6x NVIDIA T4 GPUs(1B)  
UP to 96x CCU licenses **NVIDIA GRID vPC Software**  
subscription

# HIGH-PERFORMANCE VIRTUAL WORKSTATION

Best User Experience for the Remote Workforce from the Data Center

## LIGHT TO MEDIUM USER


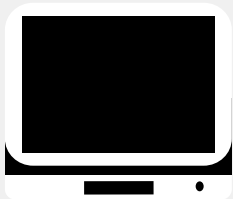


2-socket, 2U rack server  
8 vCPUs  
16 GB – 32GB memory (per user)  
Up to 2x NVIDIA RTX 6000/ 8000 (4Q or 6Q) or  
6x NVIDIA T4 GPUs (4Q)  
UP to 16 – 24 CCU licenses **NVIDIA Quadro vDWS Software**  
subscription

## HEAVY USERS

2-socket, 2U rack server  
12+ vCPUs  
96 GB memory (per user)  
Up to 2x NVIDIA RTX 6000/ 8000 (8Q or 12Q) or  
6x NVIDIA T4 GPUs (4Q)  
UP to 8 – 12 CCU licenses **NVIDIA Quadro vDWS Software**  
subscription



# NVIDIA vGPU Editions Recommend Application List

 <p>XenApp or RDSH 기반 솔루션을 사용하고자 하는 고객. 윈도우 기반 애플리케이션을 최고의 성능으로 제공</p>	 <p>가상 데스크톱 환경에서 윈도우 애플리케이션, 인터넷 브라우저 및 고해상도 비디오를 사용하며 뛰어난 사용자 경험을 요구하는 경우</p>	 <p>어디서나 장치 상관 없이 원격으로 완벽한 성능의 전문가용 그래픽 애플리케이션을 사용하고자 하는 사용자</p>	 <p>데이터 센터 HPC 활성화 AI/DL 워크로드 및 NVIDIA vGPU 보안, 관리 및 확장성</p>
<p>NVIDIA GRID Virtual Applications (vAPP)</p>	<p>NVIDIA GRID Virtual PC (vPC)</p>	<p>NVIDIA Quadro Virtual Data Center WorkStation (vDWS)</p>	<p>NVIDIA Virtual Compute Server (vCS)</p>

## NVIDIA vGPU Editions Recommend Application List

Editions	Recommend Application
NVIDIA GRID Virtual Applications (vAPP)	MS Office, Adobe Protoshop, Web Browser, Media Player
NVIDIA GRID Virtual PC (vPC)	MS Office, Adobe Protoshop, Web Browser, Media Player
NVIDIA Quadro vDWS (vDWS)	Autodesk AutoCAD, SOLIDWORKS, Petrel, CATIA, ArcGIS, SIEMENS NX, Autodesk MAYA, 3DEXCITE
NVIDIA vCompute Server (vCS)	Python, TensorFlow, MXNET, RAPIDS, PYTORCH, Chainer, ONNX,

06

# Appendix

# NVIDIA Virtual GPU Software Editions

Deciding the right license based on capability and entitlement

CUDA 또는 전문적인 그래픽 어플리케이션을 사용 중이거나 예정입니까?

예 >

Quadro vDWS License

아니요



데스크탑/워크스테이션 OS를 사용하여 데스크톱을 서비스하십니까?

예 >

GRID vPC License

아니요



RDSH Apps 또는 어플리케이션 서비스를 제공하거나, 할 예정입니까?

예 >

GRID vApps License

아니요



AI/DL/용도 또는 사용 예정입니까? (그래픽 용도 X)

예 >




NVIDIA vCS License

Use case (사용 용도 & 사례)	AI, Deep Learning & HPC	Knowledge Worker	Creative & Technical Professional
Compute Type	Server Workloads (서버 전용)	Client Computing (개인 사용자 컴퓨팅)	Client Computing (개인 사용자 컴퓨팅)
Virtual GPU Software Editions (vGPU 라이선스 Type)	vCompute Server	GRID vPC License & GRID vApps License	Quadro vDWS License
Recommended GPU	V100 or T4	T4 or M10	T4 or V100

# KEY User Groups (AEC : Architecture, Engineering and Construction)

Target User	 <p>Architects, Engineers, Designers</p>	 <p>Marketing, Creative, Design, illustrators</p>	 <p>Accounting, Finance, Human Resources</p>
User Cases	<p>매우 큰 3D 모델 및 이미지를 원격으로 보고 편집</p>	<p>가상화된 디자인 및 Adobe Creative Cloud와 같은 크리에이티브 응용프로그램을 사용하는 범용 VDI</p>	<p>가상화된 Windows 10 및 일반 사무용 생산성 응용 프로그램을 사용하는 범용 VDI</p>
Recommend	<p>Quadro vDWS with up to four T4, or V100 per VM (supports up to four 4K displays) Quadro vDWS with RTX 6000 or RTX 8000 for design workstations and rendering</p>	<p>GRID vPC/vApps on Tesla M10 for single purpose, lowest cost solution for VDI, T4 for multi-workload solution for VDI and compute, or P6 for blade servers</p>	<p>GRID vPC/vApps on Tesla M10 for single purpose, lowest cost solution for VDI, T4 for multi-workload solution for VDI and compute, or P6 for blade servers</p>

# KEY User Groups (Financial Services)

<p>Target User</p>	 <p>Power Traders &amp; Financial Analysts</p>	 <p>Traders, Financial Analysts, Super Users from other Departments</p>	 <p>Retail Bankers, Investment Managers, Financial Advisors, Insurance Agents, and Others</p>
<p>User Cases</p>	<p>다중 모니터에서 네트워크가 헤비한 응용프로그램 (Bloomberg, Reuters, Eikon)을 보안, 이중화 및 연속성으로 실행</p>	<p>생산성 응용 프로그램과 금융 응용 프로그램을 동시에 실행하고 Windows 10에서 멀티 모니터 지원을 사용하는 범용 VDI</p>	<p>맞춤형 금융 응용 프로그램 및 일반 사무용 생산성 응용 프로그램을 사용하는 범용 VDI</p>
<p>Recommend</p>	<p>Quadro vDWS on NVIDIA T4 for high end. P6 for blades (supports up to four 4K displays )</p>	<p>GRID vPC/vApps on NVIDIA M10 for rack servers or P6 for blades, T4 for high end (supports up to two 4K or four HD displays)</p>	<p>GRID vPC/vApps on NVIDIA M10 for rack servers or P6 for blades, T4 for high end (supports up to two 4K or four HD displays)</p>




# KEY User Groups (Government)

Target User	 Engineers, Analysts, Data Scientists	 Simulation and Training	 Knowledge Worker
User Cases	매우 큰 3D 모델 및 이미지를 원격으로 보고, 분석, 편집	차량, 비행 시뮬레이터, 집단, 개인 및 사이버 트레이닝	가상화된 Windows 10 및 일반 사무용 생산성 응용 프로그램을 사용하는 범용 VDI
Recommend	Quadro vDWS on Tesla P40, P4, M60, P6 or M6 (supports up to for 4K displays)	Quadro vDWS on Tesla P40, P4, M60, P6 or M6 (supports up to for 4K displays)	GRID vPC/vApps on Tesla M10, M6 or P6 (supports up to four HD or two 4K displays)

# KEY User Groups (Healthcare)

Target User	 <p>Data Scientists, Researchers</p>	 <p>Radiologists, Medical Imaging Specialists</p>	 <p>Doctors, Clinicians, Nurses, Staff</p>
User Cases	<p>AI 기반 애플리케이션을 사용하여 의료 영상 분석, DNA 시퀀스 추정, 약물 발견 또는 질병의 결과 예측</p>	<p>크고 복잡한 의료 영상 원격 보기 및 편집 (PACS)</p>	<p>가상화된 EMR 응용프로그램과 일반 사무실생산성 응용프로그램을 사용하는 범용 VDI</p>
Recommend	<p>NVIDIA vComputeServer on NVIDIA T4 for entry to mid range, V100 for high-end and P6 for blades</p>	<p>Quadro vDWS on NVIDIA T4 or P6 for blades (supports up to four 4K displays)</p>	<p>GRID vPC/vApps on NVIDIA T4 or M10 and P6 for blades (vPC supports up to two 4K or four HD displays)</p>

# KEY User Groups (Education)

<p>Target User</p>	 <p>Architecture, engineering, and design students</p>	 <p>General student body, faculty, and researchers</p>	 <p>University administration and staff</p>
<p>User Cases</p>	<p>전통적으로 온캠퍼스 연구소에서 제공되던 Autodesk AutoCAD, Dassault Systemes SOLIDWORKS, MATLAB 와 같은 CAD 또는 3D 소프트웨어에 어디서든, 어떤 장치로든 접근 및 가상 실습환경에서 학습 증가</p>	<p>Windows 10을 실행하는 범용 VDI 및 현대 생산성 응용 프로그램, 비디오와 멀티미디어 스트리밍, 쌍방향 학습 플랫폼 사용</p>	<p>학생 액세스정보 및 운영 생산성 소프트웨어에 접근하기 위한 Windows 10을 실행하는 범용 VDI</p>
<p>Recommend</p>	<p>Quadro vDWS for Education on NVIDIA® T4 or P40, V100 for high-end, Quadro RTX 6000, RTX 8000 for rendering and design, and P6 for blades (supports up to four 4K displays)</p>	<p>GRID vPC on Tesla M10, T4 and P6 (supports up to two 4K or four HD displays), GRID vApps on Tesla T4, M60, and P6 for blades (NVIDIA GRID vPC/vApps is included in the Quadro vDWS for Education license)</p>	<p>GRID vPC (included in the Quadro vDWS for Education license) on Tesla M10, T4 and P6 for blades (supports up to two 4K or four HD displays)</p>



# vGPU Software 라이선스 부여 방법

“EULA only”는 라이선스를 수량을 기술적으로 확인하기 어려운 경우 적용 되는 원칙 (예: vApps 구성 시 RDSH 서버는 1대인 반면 동시에 접속 하는 수가 많아 별도 체크가 어려운 경우)이며 “Software”는 GRID 라이선스 서버를 통해 확인

1. A -Series NVIDIA vGPU 배포의 경우 각 동시 사용자마다 하나의 GRID 가상 응용 프로그램 라이선스가 필요하지만 이러한 사용자 중 하나에 대한 라이선스는 소프트웨어를 통해 시행됩니다.  
나머지 사용자에게 대한 라이선스는 EULA를 통해 시행됩니다.
2. C-Series NVIDIA vGPU 배포의 경우 소프트웨어를 통해 VM 당 하나의 라이선스가 적용됩니다.  
이 라이선스는 단일 GPU에서 최대 8 개의 vGPU 인스턴스 또는 모든 물리적 GPU의 프레임 버퍼에 할당 된 하나의 vGPU의 VM에 할당하는 데 유효합니다.  
여러 C-Series vGPU가 단일 VM에 할당 된 경우 VM에 할당 된 각 vGPU마다 하나의 라이선스가 필요합니다.  
하나의 라이선스는 소프트웨어를 통해 시행됩니다. 나머지 라이선스는 EULA를 통해 시행됩니다.
3. 여러 Q-Series vGPU가 단일 VM에 할당 된 경우 VM에 할당 된 각 vGPU에 대한 라이선스가 아니라 VM에 대한 라이선스 하나만 필요합니다 .
4. 컴퓨팅 집약적 가상 서버를 위한 GPU pass-through, Microsoft DDA 및 VMware vDGA 배포의 경우 소프트웨어를 통해 VM 당 하나의 라이선스가 적용됩니다.  
이 라이선스는 하나의 물리적 GPU를 VM에 할당하는 데 유효합니다. VM에 할당 된 각 추가 물리적 GPU마다 하나의 추가 라이선스가 필요합니다. 추가 라이선스는 EULA를 통해 시행됩니다.

# NVIDIA vGPU FAQ (1/5)

## Q01 VM에 라이선스를 어떻게 배포하나요?

- A. 고객이 NVIDIA GRID 라이선스를 구매하면 라이선스 사용권과 함께 PAK (Product Activation Key)를 받게 됩니다. 라이선스는 NVIDIA 라이선스 포털의 PAK를 사용하여 다운로드하고 라이선스 서버에 설치합니다. VM은 네트워크 (외부 또는 고객의 내부 네트워크)를 통해 라이선스 서버에 연결하여 이러한 라이선스를 체크 아웃합니다. 라이선스 서버에 연결할 수 없거나 라이선스가 부족한 경우에도 VM은 실행 되지만 정상적인 그래픽 가속이 되지 않습니다.

## Q02 라이선스가 부족하면 어떻게 되나요?

- A. 고객이 구입 한 라이선스를 초과하더라도 기존 GRID vGPU 프로필을 적용된 VM은 계속 작동합니다. 고객은 라이선스 사용 현황을 추적해야 하며 구매 한 라이선스보다 더 많은 라이선스를 지속적으로 사용하는 경우 라이선스가 추가로 구매되어 EULA를 준수하는지 확인해야 합니다. 새 라이선스는 언제든지 라이선스 서버에 추가 할 수 있습니다.

## Q03 라이선스 서버가 중단되면 어떻게 되나요?

- A. VM은 라이선스 서버에서 라이선스를 계속 확인하지만 VM의 실행에는 라이선스가 필요하지 않습니다. 라이선스 서버가 중단되거나 액세스 할 수 없는 경우 VM은 GRID 기능이 정상적으로 동작 되지 않을 수 있습니다. (라이선스 서버 이중화 기능 제공)

## Q04 라이선스 서버에 물리적 서버가 필요하나요?

- A. GRID 라이선스 서버는 기존 서버의 VM에 설치할 수 있습니다. 이 라이선스 VM은 별도의 GRID 서버 또는 일반적으로 다른 소프트웨어 제품에 사용되는 기존 라이선스 서버에서 실행할 수 있습니다. (Windows, Linux 버전 제공)

# NVIDIA vGPU FAQ (2/5)

## Q05 vApp 라이선스는 어떻게 적용 되나요?

- A. 라이선스는 EULA를 통해 시행됩니다. 다른 GRID 제품과 달리 vAPP 라이선스는 엄격한 집행이 적용되지 않습니다. GRID의 다른 에디션과 달리 어플리케이션 가상화 제품은 체크 인 및 체크 아웃 기능이 없기 때문에 고객의 EULA 준수 여부를 확인하는 것이 중요합니다. 고객은 GRID 사용 가능 XenApp 또는 다른 RDSH 솔루션 서버에 액세스하는 모든 동시 사용자에게 대한 라이선스를 구입하는 것이 좋습니다.

## Q06 Tesla Card 가 아닌 다른 카드로도 대체가 가능한가요? (Ex : Quadro, Geforce )

- A. NVIDIA GRID 는 GRID 전용 카드인 Tesla 시리즈와 **Quadro RTX 6000 or 8000 만** 사용 가능 합니다.

## Q07 Pass-through 모드로 구성 시 호스트 드라이버 설치가 필요하나요?

- A. pass-through 모드 구성 시 해당 호스트 드라이버 설치의 필요 없습니다. 기존에 GPU 프로파일 할당하는 방식이 아니므로 VM 편집 ) Share PCI Device가 아닌 물리적인 PCI Device 형식으로 추가를 해주시면 됩니다.

## Q08 Pass-through 모드 사용시 라이선스가 필요한가요?

- A. 그래픽 가속 처리 기능을 완벽히 작동하려면 pass-through 또는 베어 메탈 형식으로 사용시에도 동일한 라이선스가 필요합니다. (vPC Edition 사용 불가) vDWS or vCS 라이선스 필요 (단 그래픽가속 아닌 연산처리 용도 이면 필요 없음)

## Q09 NVIDIA vGPU에서도 H.264 및 H265 인코딩 디코딩 지원 되나요?

- A. NVIDIA vGPU에서 인코딩 및 디코딩이 지원 합니다. 자세한 사항은 NVIDIA 인코딩 / 디코딩 매트릭스 참조  
<https://developer.nvidia.com/video-encode-decode-gpu-support-matrix>

# NVIDIA vGPU FAQ (3/5)

## Q10 NVIDIA vGPU도 Linux VDI를 지원하나요?

- A. 네 가능합니다. 사용가능 하이퍼바저 별 Linux Guest OS 해당 URL 참고 부탁드립니다.  
<https://docs.nvidia.com/grid/index.html>

## Q11 라이선스서버 재 구성 시 기존라이선스서버 삭제는 어떻게 하나요?

- A. 우선 처음 NPN 라이선스 포탈에서 등록 하신 라이선스서버의 고유의 MAC 값은 삭제가 불가능 합니다.  
기존 라이선스서버의 고유 MAC 값의 등록된 라이선스 회수를 먼저 하신 후, 신규 라이선스서버 등록을 새로 해주셔야 합니다.

## Q12 라이선스 추가 구매 시 등록 방법 문의

- A. 1. 기존에 라이선스서버에 등록 된 라이선스 회수  
2. 추가 발급 받으신 메일 계정의 Activation Key 를 재 등록  
3. 기존 라이선스서버의 발급받으신 신규 라이선스를 Map-add ON 다시 등록

## Q13 AI & Deep Learning 용도로 GRID 사용이 가능하나요?

- A. 네 가능합니다. AI & Deep Learning 용 GPU 가 Pascal 아키텍처 이후 패스스루 및 베이메탈 형식과 비슷한 성능을 발휘 합니다. (동일한 성능의 Profile 적용 시)

# NVIDIA vGPU FAQ (4/5)

## Q14 vComputeServer는 GRID vPC/vApp 및 Quadro vDWS와 어떻게 다른가요?

- A. GRID vPC/vApp 및 Quadro vDWS는 지식인 및 창의적 또는 기술 전문가를 위해 설계된 가상 그래픽을 위한 클라이언트 컴퓨팅 제품이며 vComputeServer는 AI, 딥러닝 및 데이터 과학과 같은 컴퓨팅 집약적인 서버 워크로드를 위한 제품입니다.

## Q15 vComputeServer는 GRID vPC/vApp 및 Quadro vDWS와 동일한 방식으로 라이선스가 부여되나요?

- A. 아니요. vComputeServer는 GRID vPC/vApp 및 Quadro vDWS와 다르게 라이선스가 부여됩니다. GRID vPC/vApp 및 Quadro vDWS는 CCU(동시 사용자)에 의해 영구 라이선스 또는 연간 구독으로 라이선스가 부여됩니다. . vComputeServer는 서버 컴퓨팅 워크로드를 위한 라이선스로 라이선스는 사용자 대신 GPU에 연결되며 vComputeServer는 GPU별로 연간 구독으로 라이선스가 부여됩니다. 라이선스에 대한 자세한 내용은 NVIDIA Virtual GPU 패키지, 가격 및 라이선싱 가이드를 참조하십시오. (<https://images.nvidia.com/content/grid/pdf/Virtual-GPU-Packaging-and-Licensing-Guide.pdf>)

## Q16 vComputeServer에서 지원되는 NVIDIA GPU는 무엇인가요?

- A. vComputeServer에 권장되는 NVIDIA GPU카드는 NVIDIA V100S, V100 및 NVIDIA T4입니다. 쿼드로 RTX 6000과 RTX 8000도 지원되며, NVIDIA P40, P100, P6 등이 포함된 파스칼 기반의 GPU도 지원됩니다. ( M, K 시리즈 사용불가)

NVIDIA 글로벌 홈페이지 내용 참고 : <https://www.nvidia.com/en-us/data-center/virtual-compute-server/>

# NVIDIA vGPU FAQ (5/5)

Q17 vComputeServer 라이선스 가격표상 1 GPU (Max 8 VMs) 의 의미는 무엇인가요?

- A. 단일 vCS 라이선스를 통해 최대 8개의 동시 VM을 사용할 수 있으며 GPU 모델 별로 제공됩니다.
- B. 기본 4GB 프로파일 부터 적용이 가능하며 Max 8 VMs 이란 의미는 Tesla v100 (32GB)의 경우 4GB 적용 시 최대 8VMs 지원이 가능하다는 것을 의미합니다. RTX8000 (48GB) 4GB 적용 시 수치상 12VMs 이나 최대 8VMs만 지원
- C. 라이선스 수량 안내 시 GPU 단위로 안내하면 되며 , VM 수량은 무관합니다. ( Multi-GPU의 경우에도 동일합니다. )

# 감사합니다

제품 문의(e-mail) : [goodmkt@goodmit.co.kr](mailto:goodmkt@goodmit.co.kr)

제품 문의(Phone) : 070-7164-4253