

# “GPU分身の術”でより多くの方がGPUを利用可能に NVIDIA 仮想GPU (vGPU) で実現する DX推進仮想基盤



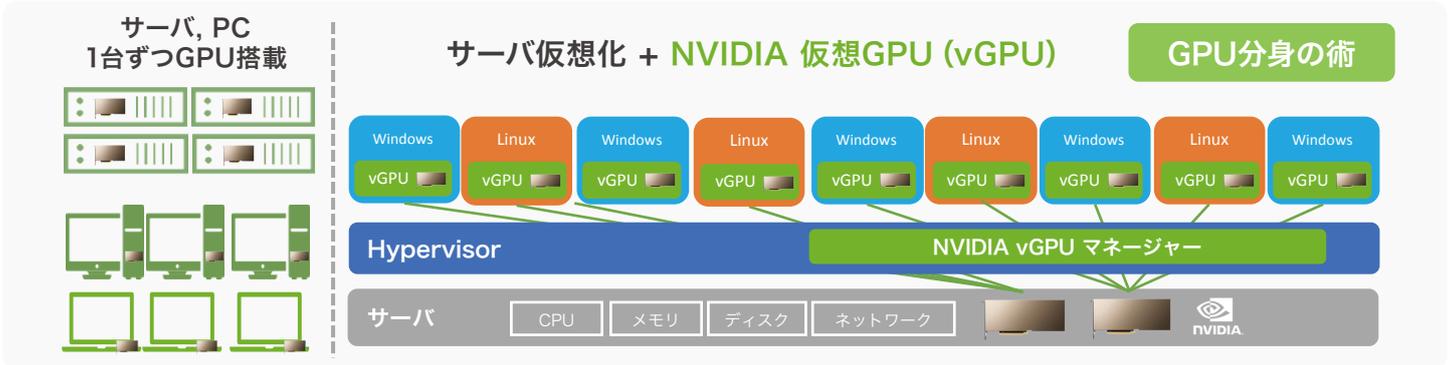
グラフィックスと計算処理をつかさどる GPUの利用ニーズが高まっています。民間企業に限らず、研究機関、中央官庁や自治体、高等教育機関など、業界を問わず需要が高まっています。

GPUはAIやグラフィックスのシステムも楽々操作することができます。一般的なオフィス業務だけでなく、研究者、設計者やエンジニア、クリエイターやデザイナー、データサイエンティストの業務まで快適に行えるようになります。



## GPUを分身させて、複数の仮想マシンで有効活用

NVIDIA vGPUは、サーバに搭載したGPUのメモリを仮想 GPU (vGPU)として分割、複数台の仮想マシンで高いコア性能を効率的に最大限共有しながら利用できます。仮想化の管理や設定の柔軟性、冗長化やセキュリティなどのメリットをそのままにベアメタルと遜色ないパフォーマンス環境を実現。サーバやPC等のすべての環境を仮想化できます。



## 仮想マシンのCPU負荷を vGPUで大幅軽減して、VDI問題を解決

- ・Web会議や動画、Webブラウザや地図、オフィスソフトやPDFなど、VDIのあらゆるパフォーマンスを改善
- ・高性能PC環境が必要な作業もVDIで利用可能に  
(生成 AI/CAD/CAE/BIM/CIM/動画/デザイン/レンダリング/デジタルツイン)

<p><b>CPU</b></p> <p><b>vGPU 無</b></p> <p>OS: Windows 10 vCPU: 4 Core vMemory: 8 GB vDisk: 50 GB vGPU: なし</p> <p>全体の処理 制御と演算</p> <p>グラフィックス処理 画面描画 画像データ変換と圧縮</p>	<p><b>CPU</b></p> <p><b>GPU</b></p> <p>OS: Windows 10 vCPU: 2 Core vMemory: 8 GB vDisk: 50 GB vGPU: A16-1B(1GB)</p> <p>全体の処理 制御と演算</p> <p>グラフィックス処理 画面描画 画像データ変換と圧縮</p>	<p><b>vGPU 有</b></p> <p><b>CPU 負荷</b> 約10~60% 削減</p> <p><b>フレームレート</b> 5~25 FPS アップ*</p> <p>* 1秒間に送られる画像数 多い方がスムーズ</p>
<p>グラフィックス処理でCPU負荷が高い 操作感が悪く処理が遅い</p>	<p>グラフィックス処理はGPUが担当 CPU負荷を軽減、操作を改善</p>	

# NVIDIA vGPU対応 データセンターGPU

ワークロードの用途、推奨利用シーンを参考にGPUを選択、サーバの型番や搭載可能枚数をご確認ください。

	H100	A100	A30	L40S	L40	A40	L4	A10	A2	T4	A16	
1ボード搭載GPU数 (GPU世代)	1 (Hopper)	1 (Ampere)	1 (Ampere)	1 (Ada Lovelace)	1 (Ada Lovelace)	1 (Ampere)	1 (Ada Lovelace)	1 (Ampere)	1 (Ampere)	1 (Turing)	4 (Ampere)	
CUDAコア数	PCIe:14,592* SXM5:16,896*	PCIe:6,912* SXM4:8,192*	3,584*	18,176	18,176	10,752	7,424	9,216	1,280	2,560	5,120 (1,280/GPU)	
Tensorコア数	PCIe:456 SXM5:528	PCIe:432 SXM4:512	224	568	568	336	232	288	40	320	160 (40/GPU)	
RTコア数	-	-	-	142	142	84	58	72	10	40	40 (10/GPU)	
メモリサイズ	PCIe: 80GB HBM2e SXM5: 80GB HBM3	80GB HBM2	24GB HBM2	48GB GDDR6	48GB GDDR6	48GB GDDR6	24GB GDDR6	24GB GDDR6	16GB GDDR6	16GB GDDR6	64GB GDDR6 (16GB/GPU)	
vGPU Profiles (GB) 分割パターン	4,5,8,10,16 20,40,80	4,5,8,10,16 20,40,80	4,6,8,12,24	1,2,3,4,6,8, 12,16,24,48 ※Max32VM	1,2,3,4,6,8, 12,16,24,48 ※Max32VM	1,2,3,4,6,8, 12,16,24,48 ※Max32VM	1,2,3,4,6, 8,12,24	1,2,3,4,6, 8,12,24	1,2,4,8,16	1,2,4,8,16	1,2,4,8,16	
NVLink Support	Yes	Yes	Yes	No	No	Yes	No	No	No	No	No	
MIG Support	Yes	Yes	Yes	No	No	No	No	No	No	No	No	
フォームファクター 物理寸法	PCIe5.0 dual slot /SXM5	PCIe4.0 dual slot /SXM4	PCIe4.0 dual slot	PCIe4.0 dual slot	PCIe4.0 dual slot	PCIe4.0 dual slot	PCIe4.0 single slot	PCIe4.0 single slot	PCIe4.0 single slot	PCIe3.0 single slot	PCIe4.0 dual slot	
消費電力	PCIe:350W SXM5:700W	PCIe:300W SXM4:400W	165W	350W	300W	300W	72W	150W	40-60W	70W	250W	
冷却	passive	Passive, liquid cool	passive	passive	passive	passive	passive	passive	passive	passive	passive	
推奨 利用 シーン	オフィス	-	-	-	-	-	Office	-	-	Office	Office	
	グラフィックス	-	-	-	CAD/Rendering/Omniverse/Digital twin			CAD/Omniverse/ Edge/IVA	CAD/ Omniverse	Edge/IVA/ Mobile Gaming	CAD	2D CAD
	コンピューティング	INT8/FP8/FP16/BF16 FP32/TF32/FP64 AI(推論/学習)/CAE/HPC	INT4/INT8/FP16/BF16 FP32/TF32/FP64 AI(推論/学習)/CAE/HPC	INT8/FP8/FP16/ BF16/FP32/TF32 AI(推論/学習)			INT8/FP8/FP16/ BF16/FP32/TF32 AI(推論)	INT4/INT8 FP16/FP32 AI(推論)	INT4/INT8 FP16/FP32 AI(推論)	INT4/INT8 FP16/FP32 AI(推論)	INT4/INT8 FP16/FP32 AI(推論)	

※最新サーバ構成、サポート状況は各サーバメーカー様にご確認ください。 \* FP32 CUDA コア数

NVIDIA  
認定システム  
カタログ



vGPU  
バージョン対応  
GPU



vGPU対応  
GPU



vGPU  
ソフトウェア  
ドキュメント



## vGPUソフトウェア エディションと購入方法

NVIDIA vGPUソフトウェアは3種類をご用意。仮想マシン上でのOSやアプリケーション、利用シーンにより選択するエディションとカウント数があります。「年間サブスクリプション」と「永久ライセンス+保守」の2種類の購入方法があります。

NVIDIA vGPU ソフトウェア エディション	Virtual Applications 仮想アプリケーション(vApps)	Virtual PC 仮想PC(vPC)	RTX Virtual Workstation 仮想ワークステーション(vWS)
カウント数	共有Windowsサーバ 総ユーザーセッション数	同時起動 仮想マシン数	同時起動 仮想マシン数
契約期間	1/3/4/5 年間サブスクリプション(サポート/アップグレード権利) or 永久ライセンス+5年保守(サポート/アップグレード権利)		
利用シーン	オフィス(SBC)	オフィス(VDI)	グラフィックス / コンピューティング
仮想デスクトップ	-	✓	✓
RDSH デスクトップ/アプリホスティング	✓	✓	✓
Compute Virtualization	-	-	✓
Windows ゲストOS	✓	✓	✓
Linux ゲストOS	✓	✓	✓
NVIDIA Graphics Driver	✓	✓	✓
NVIDIA RTX Enterprise(Quadro) Driver	-	-	✓
Multi-vGPU/NVLink	-	-	✓
ECC & Page Retirement	-	-	✓
最大画面数	1	4	4
画面あたりの最大解像度	1280×1024 *1	5K (5120×2880)	8K (7680×4320)
DirectX	✓	✓	✓
OpenGL Extensions (WebGL)	✓	✓	✓
Deep Learning Super Sampling (DLSS)	-	-	✓
Vulkan Support	-	-	✓
ISV 認証	-	-	✓
NVIDIA CUDA & OpenCL サポート	-	-	✓
GPU Pass-through サポート	✓	-	✓
Bare Metal サポート	✓	-	✓

お問い合わせ先

アセンテック 株式会社 E-mail:eigyo@ascentech.co.jp

〒101-0022 東京都千代田区神田練馬町3 大東ビル9F

TEL:03-5296-9331 FAX:03-5296-9332

Web:https://www.ascentech.co.jp