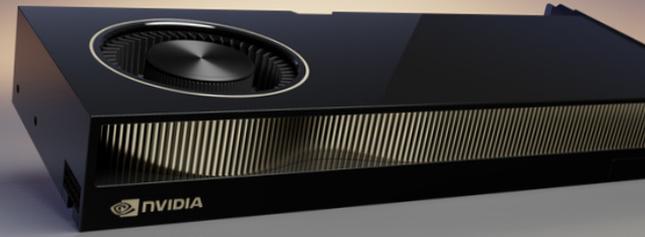


超ハイエンド プロフェッショナル向け グラフィックスカード： NVIDIA RTX 6000 Ada



アドバンテックは、NVIDIAのフルレンジのエンタープライズ向けプロフェッショナルグラフィックスカードを搭載した幅広い産業用GPUソリューションを提供しています。アドバンテックは、NVIDIA社のElite Partnerとして、最新世代GPUカードへの優先的なアクセスと、同社のテクニカルチームからの独占的なサポートを通じたプロフェッショナルなコンサルティングサービスをお客様へ提供いたします。

アドバンテックのエッジAIソリューションは、産業およびエンタープライズユーザーとの間で蓄積した豊富な経験から様々な業界の現場で広く使用されております。この記事では、人気の高いプロフェッショナルグレードのNVIDIA RTX GPUカードシリーズの比較評価を通して、お客様が産業用IoTや製造、輸送、医療、ゲーム業界などの様々なバーティカル市場に適した採用を考慮する際にお役に立てればと思います。

ハイライト

- NVIDIA RTX 6000 AdaとRTX A6000のスペック比較
- ウルトラハイエンド・プロフェッショナルグラフィックスカード：
『NVIDIA RTX 6000 Ada』の特徴
- プロフェッショナルとしての性能評価4件：
 1. プロフェッショナルモデリング・グラフィックスソフトウェア：
『SPECviewperf 2020』パフォーマンステスト
 2. AI画像生成モデル：『Stable Diffusion』計算速度テスト
 3. ディープラーニング AIモデルトレーニング：
Lambda Labs PyTorch GPU Training Benchmark
 4. 3Dパフォーマンス3DMark Time Spyテスト

NVIDIA RTX 6000 Ada とRTX A6000の仕様比較：

最新のNVIDIA® RTX™ 6000 GPUカードは、新しいAda Lovelaceアーキテクチャを使用し、より高性能なCUDAコアと第3世代RTコア、および第4世代Tensorコアを提供します。前世代のRTX A6000と比較してコア数と性能が約70%向上。単精度浮動小数点演算（FP32）の優れた能力を提供します。

NVIDIA RTX 6000 Adaは48GBのECCビデオメモリを搭載していますが、最大消費電力は300ワットに抑えられ、ゲーミング用の平均的なGPUカードの約3分の2です。

これは、エッジAIアプリケーションやハイエンドの画像処理への導入に極めて重要であり、お客様がより効率的で安定したAIアクセラレーションやグラフィックワークステーションを構築するのに役立ちます。



・スペック比較

	RTX 6000 Ada	RTX A6000
GPU アーキテクチャ	Ada Lovelace	アンペア
CUDA コア	18,176 (69%↑)	10,752
Tensor コア	568 (69%↑)	336
RT コア	142 (69%↑)	84
単精度性能	91.1 TFLOPS (135%↑)	38.7 TFLOPS
RT コア性能	210.6 TFLOPS (179%↑)	75.6 TFLOPS
Tensor コア性能	1,457 TFLOPS (370%↑)	309.7 TFLOPS
NVIDIA NVLink®	いいえ	NVIDIA RTX A6000 GPU 2基接続
GPU メモリ	48 GB GDDR6 (ECC付)	48 GB GDDR6 (ECC付)
メモリ・インターフェース	384 ビット	384 ビット
メモリ帯域幅	960 GB/秒	768 GB/秒

	RTX 6000 Ada	RTX A6000
最大消費電力	300 W	300 W
システム・インターフェース	PCI Express 4.0 x16	PCI Express 4.0 x16
ディスプレイ・レイコネクタ	4 x DisplayPort 1.4a	4 x DisplayPort 1.4a
フォームファクタ	DSFH	DSFH
スロット幅	デュアルスロット	デュアルスロット
寸法	長さ : 267 mm / 10.5 インチ 幅 : 112 mm / 4.4 インチ	長さ : 267 mm / 10.5 インチ 幅 : 112 mm / 4.4 インチ
重量	1.180 kg	1.179 kg
電源コネクタ	1 x 16 ピン PCIe	8 ピン CPU
シンク出力	NVIDIA Quadro Sync IIとの互換性	NVIDIA Quadro Sync IIとの互換性
エンコード/デコードエンジン	3 x エンコード, 3 x デコード (+AV1エンコード/デコード)	エンコードx, デコードx2 (+AV1デコード)

電源コネクタに関しては、NVIDIA RTX 6000 Adaは、新しい12VHPWR 16ピンインターフェースを導入しています。



移行期間中、NVIDIAはお客様にPCIe8ピンx2から16ピンへのアダプタケーブルを提供し、既存の電源を引き続き使用できるように配慮しています。

外観的には、RTX 6000 AdaはRTX A6000と同じピアノ塗装の鏡面デザインを維持しており、この「鏡面デザイン」RTX GPUカードは最上位ステータスの象徴です。



NVIDIA RTX 6000 Adaの上部側接続ポートは主にQuadro Syncマルチスクリーン同期ディスプレイ機能で使用され、NVLinkリンクは提供されていないことにご注意ください。

NVLink機能が必要な場合は、前世代のNVIDIA RTX A4500 / A5000 / A5500 / A6000 GPUカードを使用することをお勧めします。

最新のフラッグシップ NVIDIA RTX 6000 Ada は、前世代の RTX A6000 と比較して仕様が大幅にアップグレードされており、あらゆる面で論理的なピーク性能も大幅に向上しています。次のページではエンタープライズユーザー向けに、これらの特徴を詳しく解説します。

ウルトラハイエンドプロフェッショナル グラフィックスカード

NVIDIA RTX 6000 Ada 特徴

● 大容量・高速48GB GDDR6メモリ

48GB GDDR6 メモリを搭載することで、RTX 6000 は前世代と比較してレイトレーシング、レンダリング、AI ワークロードのスループットを25%向上させました。

RTX 6000 Ada は業界最大のグラフィック・メモリを搭載し、レイテンシに敏感なプロフェッショナル・アプリケーション向けの最大のデータセットとモデルを処理することが可能です。

● ECC（誤り訂正符号）付きメモリ

自動エラー検出・修復機能により、データの整合性を高め、モデルの精度と安定性を維持することができるため、信頼性の高いミッションクリティカルな運用が可能です。データアクセスや伝送エラーが発生するなど、長時間の計算中にワークステーションが突然処理を中断することを回避することで、ワークステーションは高い精度と信頼性での計算処理が可能です。

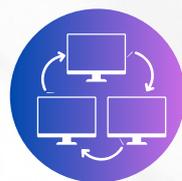
● NVIDIA Virtual GPU ソフトウェア

NVIDIA Virtual GPU (vGPU) ソフトウェアは、パーソナルワークステーションを複数の高性能仮想ワークステーションへと再構築が可能。リモートユーザーがリソースを共有してハイエンドの設計、AI、および計算ワークロードの処理が可能です。



ベアメタルの パフォーマンス

実質的にベアメタル環境と見分けがつかないパフォーマンスを提供します。



最適資源利用

フラクショナルまたはマルチGPU 仮想マシン (VM) インスタンスでGPUリソースをプロビジョニングします。



管理&監視

ライブマイグレーションなどの一般的なデータセンター管理ツールを活用。



ビジネスの改善

継続性

変化するビジネス要件やリモートチームへの対応。

● NVIDIA® RTX™ Experience

NVIDIA RTX Experienceは、最大8Kのデスクトップ録画、最新のNVIDIA RTX Enterprise ドライバアップデートのアクティブ通知、およびゲーム機能へのアクセスを含むデスクトップワークステーション用の生産性ツールスイートを提供します。

4つの主要なプロ フェッショナル向け パフォーマンス指標価

リリースに続いて、この記事のハイライトである性能評価に入ります。

このテストでは、プロフェッショナルとエンタープライズ向けの NVIDIA RTX 6000 Ada, RTX A6000, RTX A4500, RTX A4000 の合計 4 台を使用しました。



テスト用ツールとして、プロフェッショナルなモデリングと描画用の『SPECviewperf 2020』ソフトウェア, 人気のAI画像生成モデルの『Stable Diffusion』, ディープラーニングと AIモデルのトレーニング用の『Lambda Labs PyTorch GPU Training Benchmark』, そして3Dパフォーマンステスト用の『3DMark Time Spy』が含まれます。

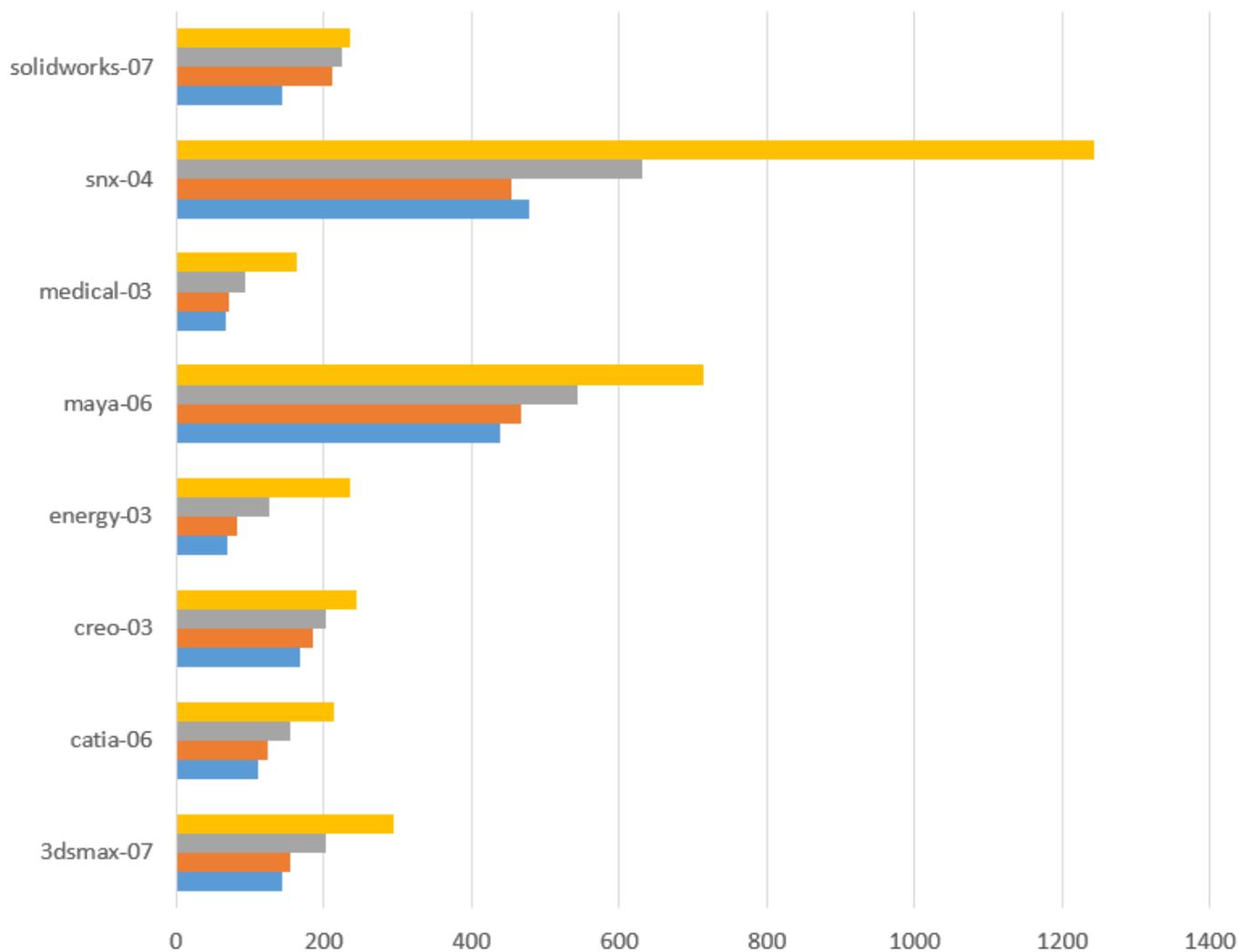
全体として、NVIDIA RTX 6000 Ada は前世代の RTX A6000 GPU と比較して、すべてのテストで大幅な性能向上を示しています。これは主に新しい Ada Lovelace アーキテクチャ、アップグレードされたCUDAコア、第3世代のRTコア、そして第4世代の Tensor コアによるものです。コア数は前世代と比較して69%増加し、レンダリング性能は3倍に、AI推論性能は2倍へと向上しました。

この記事では、NVIDIA RTX 6000 Ada のレビューに焦点を当て、性能、消費電力、システム統合における優位性を見ていきます。

1. プロフェッショナル・モデリング兼ドローイング・ ソフトウェア：『SPECviewperf 2020』 パフォーマンス・テスト

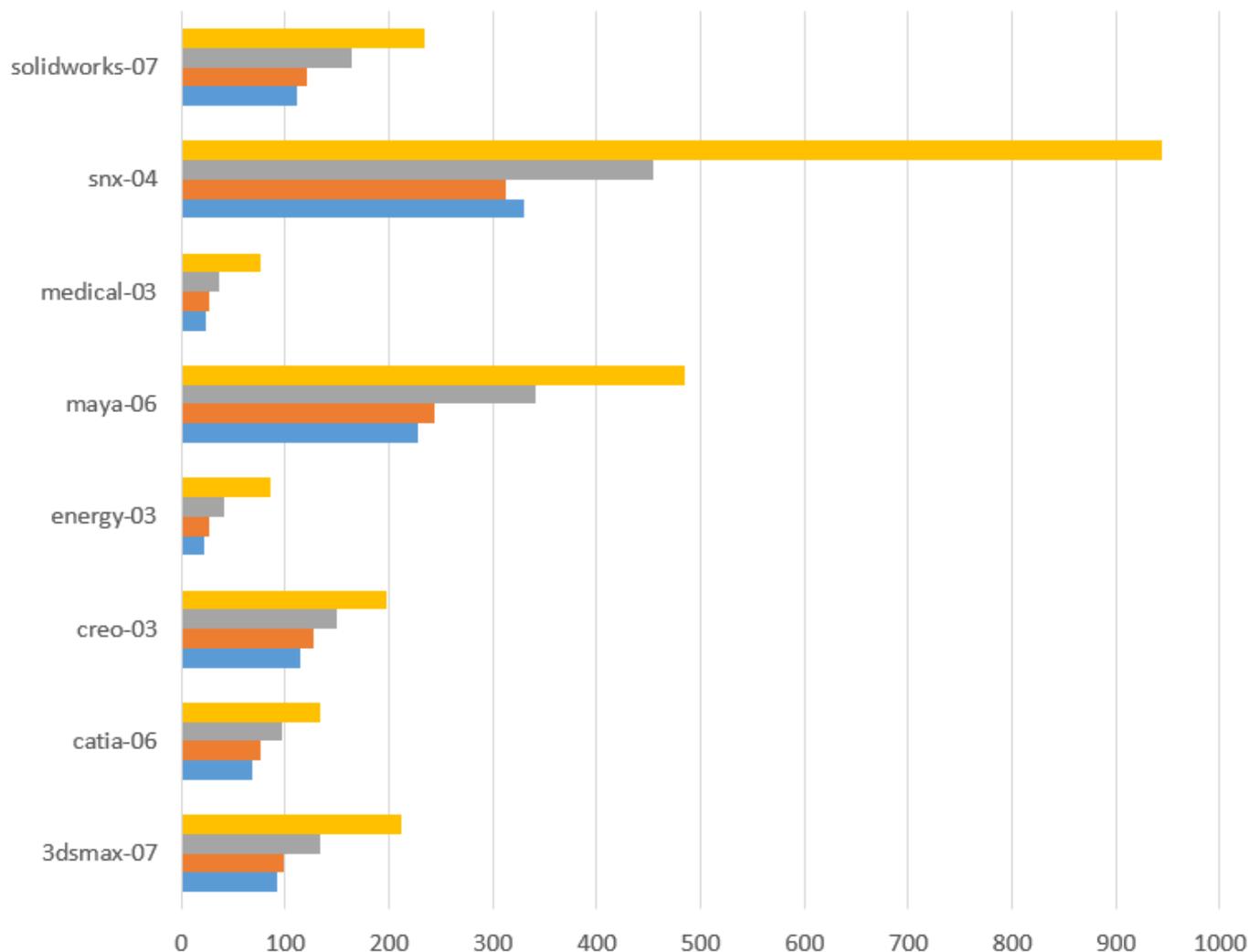
SPECviewperf 2020は3ds Max, Catia, Creo, Energy, Maya, Medical, SNX, SolidWorksなど、あらゆる種類の描画や製図、そしてエンジニアリングのシミュレーションにも対応した、各業界の専門的なソフトウェアの性能比較を目的としたベンチマークプログラムです。

SPECviewperf 2020 1080p



	3dsmax-07	catia-06	creo-03	energy-03	maya-06	medical-03	snx-04	solidworks-07
■ RTX 6000 Ada	294.4	214.25	245.27	236.45	714.45	163.51	1242.47	236
■ RTX A6000	203.81	153.84	202.79	126.63	544.47	92.75	630.48	225.24
■ RTX A4500	155.73	124.28	185.14	83.33	466.47	72.4	453.4	211.21
■ RTX A4000	143.19	111.81	167.59	69.72	438.38	68.29	479.19	144.51

SPECviewperf 2020 2160p



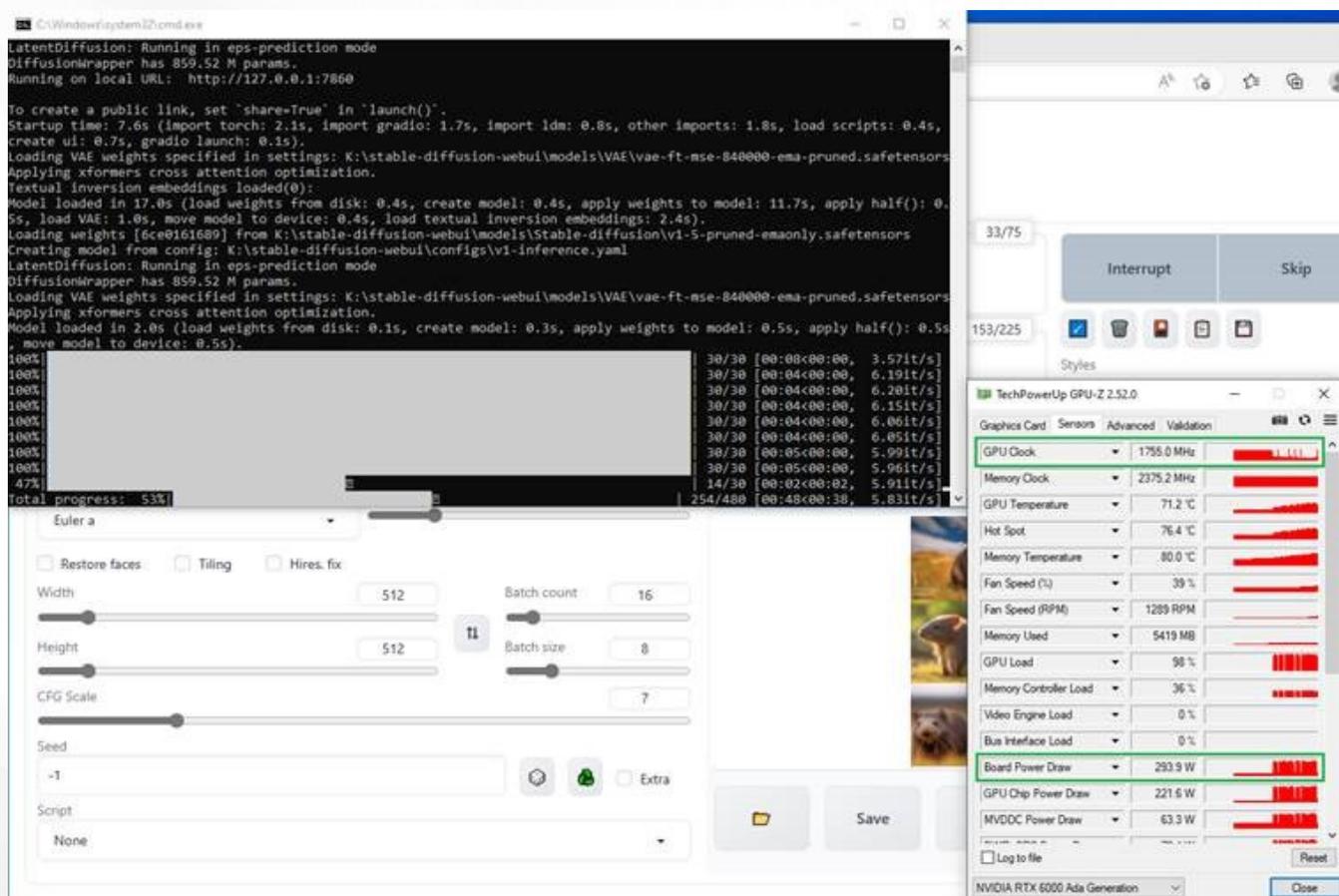
	3dsmax-07	catia-06	creo-03	energy-03	maya-06	medical-03	snx-04	solidworks-07
■ RTX 6000 Ada	211.54	133.03	198.11	86.49	484.1	76.25	944.57	234.19
■ RTX A6000	133.4	96.47	149.28	41.54	340.46	36.77	454.23	163.76
■ RTX A4500	98.03	76.49	126.6	26.62	243.74	26.2	312.46	120.86
■ RTX A4000	91.78	68.77	113.95	22.3	228.07	23.27	330.34	110.65

テストの結果、RTX 6000 Ada は他のRTX GPUカードと比較しても、すべてのソフトウェアで圧倒的なパフォーマンスを発揮しています。

製造業で一般的に使用されているモデリングソフトウェアSNXを例にとると、一般的な解像度である1080pでRTX A6000の約2倍のパフォーマンスを発揮しました。医療やエネルギー分野のアプリケーションでは、RTX 6000 Adaはテストした他のGPUカードよりも2倍以上の性能を発揮します。これは、NVIDIA RTX 6000 Adaがあらゆる業界における専門的な描画や製図、そしてモデリングのソフトウェアに非常に適していることを示しています。

2. AI画像生成モデル『Stable Diffusion』テスト

生成AIは、現在最もホットなAIアプリケーションのひとつです。Stable Diffusionは、2022年に公開されたディープラーニングの text-to-image モデルを可能にした画像生成AIです。ますます多くの企業が機密データの漏洩リスクを防ぐため、ローカル環境でAI画像を生成していく必要がでてきました。それに応じて、NVIDIA GPU カードを搭載したアドバンテックのGPUエッジコンピュータで Stable Diffusion 専用のグラフィックスワークステーションを構築する必要もでてきました。

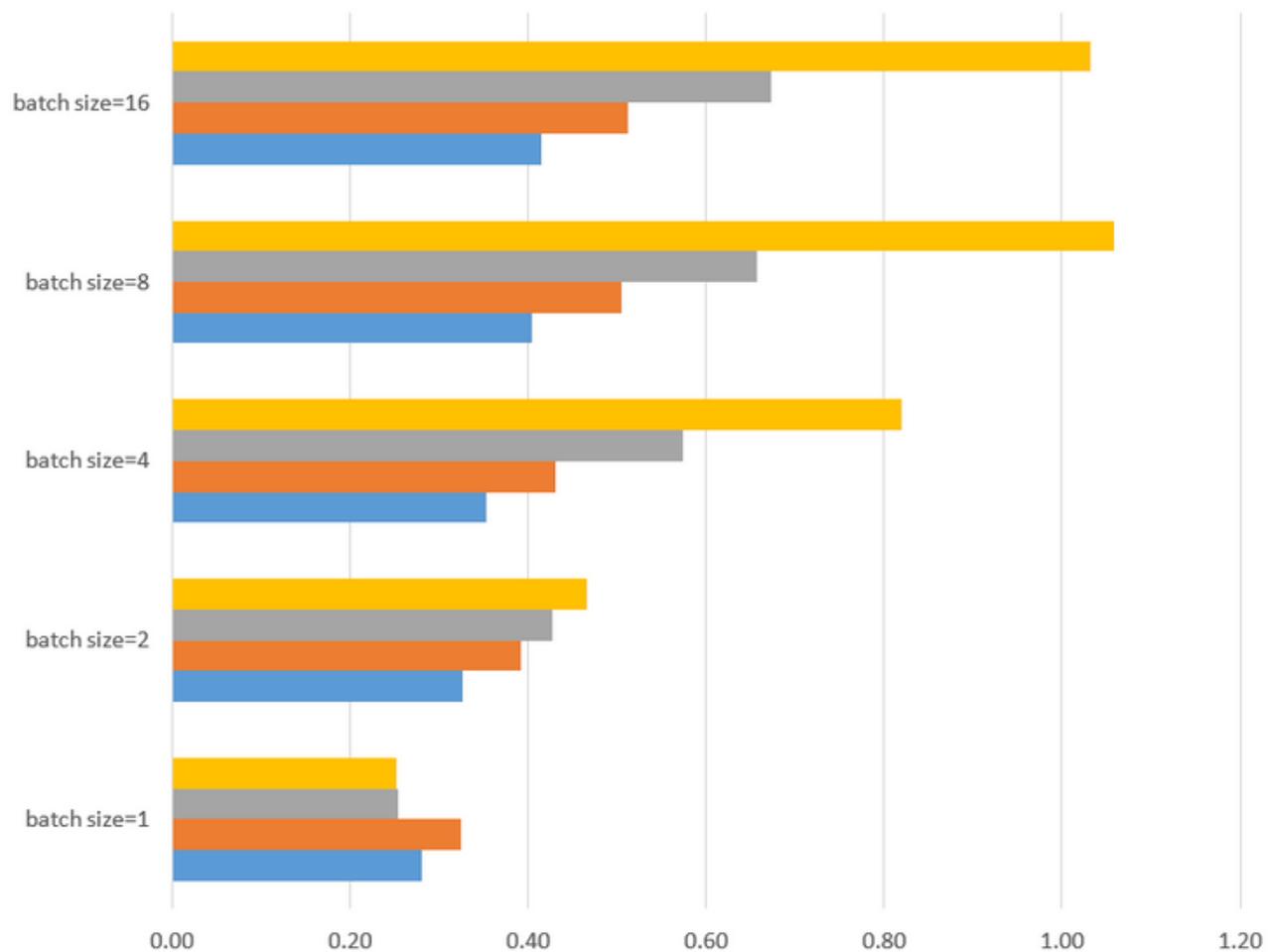


テストの結果、RTX 6000 Ada はテストしたすべてのカードの中で最も速く画像をレンダリングできることがわかりました。一度に8枚の画像を生成し（バッチサイズ=8）、1枚の画像あたり生成に必要な時間が、わずか1秒未満でした。RTX A6000 と比較して、同じ消費電力を維持しながら1秒あたりにより多くの画像を生成し（0.66画像→1.06画像）、計算時間を短縮し、画像生成プロセス効率を最適化しています。

	RTX 6000 Ada	RTX A6000
消費電力	300W	300W
img/秒	1.06	0.66
説明	低消費電力とクロックスピードの下で、1秒間に生成できる画像はわずか0.66枚	より低い消費電力とクロック速度で、1秒間に1.06枚の画像を生成可能
CUDA コア	18,176	10,752

Stable Diffusion スループット (img/s)

128画像の平均、xformers有効、サンプラー：Euler a、ステップ数：30、解像度：512*512



	batch size=1	batch size=2	batch size=4	batch size=8	batch size=16
RTX 6000 Ada	0.25	0.47	0.82	1.06	1.03
RTX A6000	0.25	0.43	0.57	0.66	0.67
RTX A4500	0.32	0.39	0.43	0.51	0.51
RTX A4000	0.28	0.33	0.35	0.40	0.42

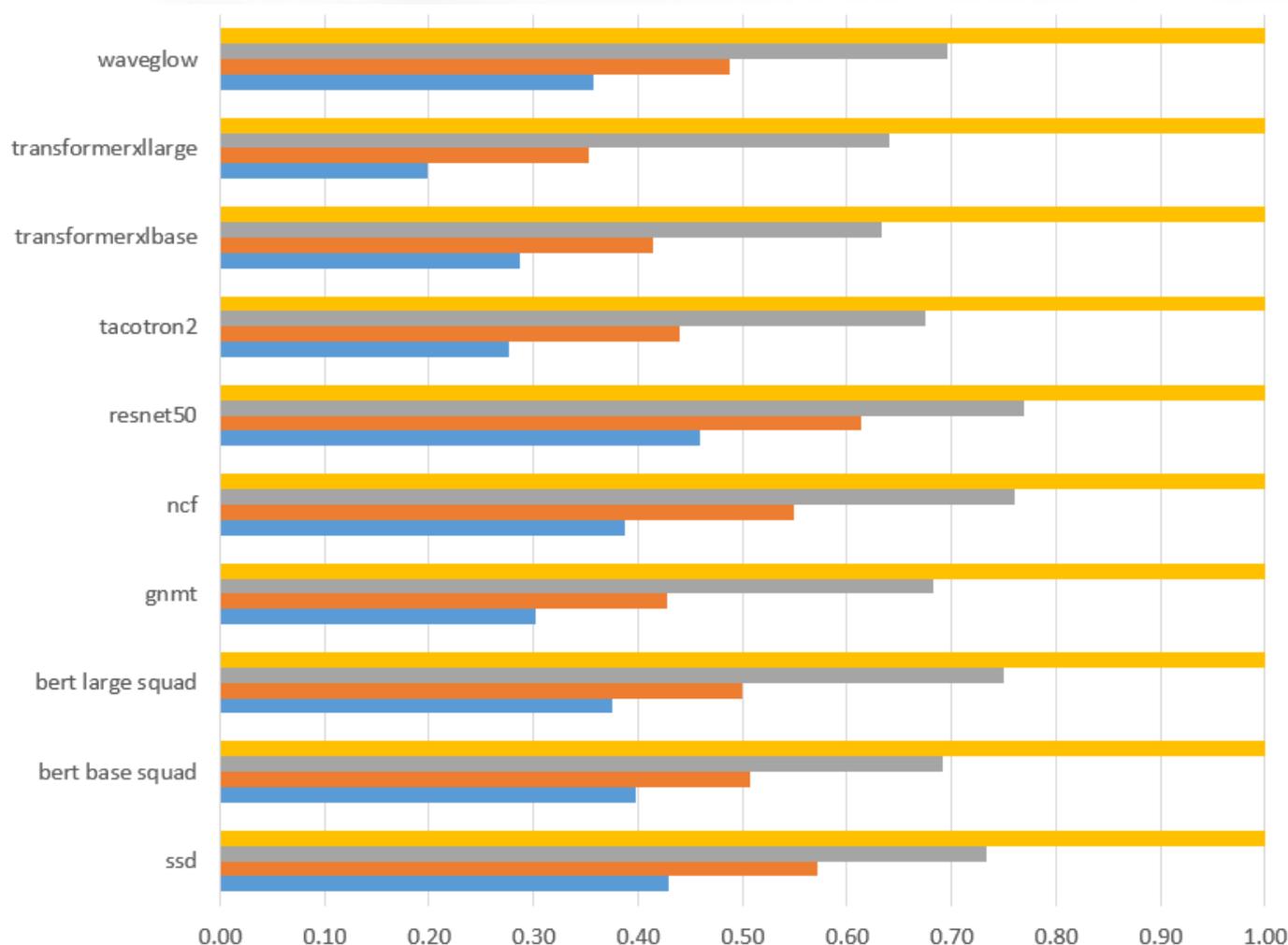
3. ディープラーニング AIモデルトレーニング『Lambda Labs PyTorch GPU トレーニング』ベンチマーク

PyTorchは深層学習モデルを構築するためのフレームワークで、コンピュータ・ビジョンや自然言語処理、さらにはTeslaの自律走行用AIモデル・トレーニングでも一般的に使用されています。このテストでも、RTX 6000 Adaは、さまざまなAIモデルのトレーニングデータの処理において、すべてのRTX GPUカードを圧倒しています。

ただし、高度に分割可能なAIモデルのトレーニングでは、システムに搭載するGPUカードの枚数が増えれば増えるほど、トレーニングの速度もそれに比例して速くなるため、GPUカードとシステムの相性は非常に重要であることは注目に値します。

FP16 トレーニング性能

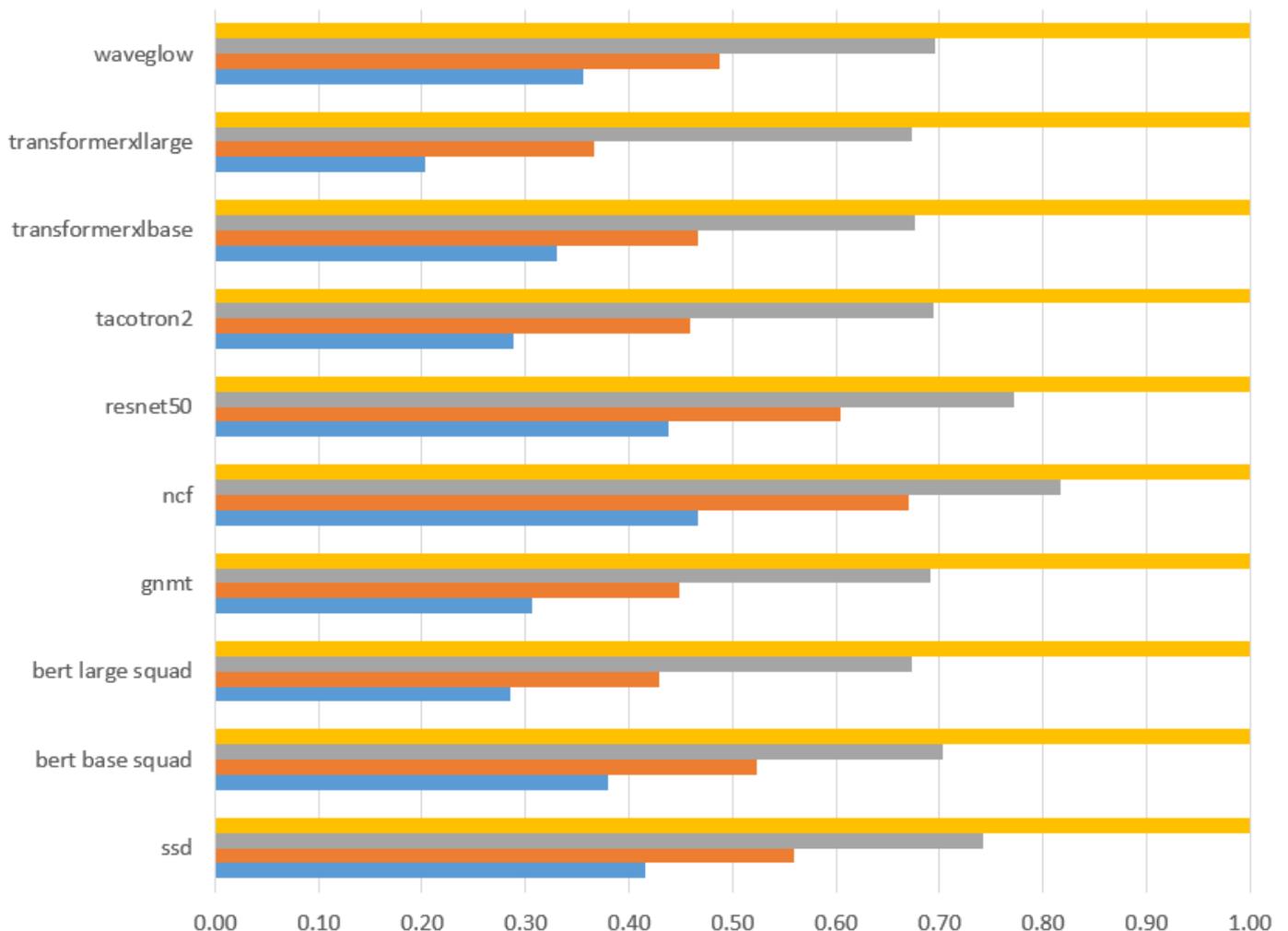
RTX 6000Adaとの比較



	ssd	bert base squad	bert large squad	gnmt	ncf	resnet50	tacotron2	transformerxlarge	transformerxlarge	waveglow
RTX 6000 Ada	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
RTX A6000	0.73	0.69	0.75	0.68	0.76	0.77	0.67	0.63	0.64	0.70

FP32 トレーニング性能

RTX 6000Adaとの比較



	ssd	bert base squad	bert large squad	gnmt	ncf	resnet50	tacotron2	transformerxlarge	transformerxlarge	waveglow
RTX 6000 Ada	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
RTX A6000	0.74	0.70	0.67	0.69	0.82	0.77	0.69	0.68	0.67	0.70
RTX A4500	0.56	0.52	0.43	0.45	0.67	0.60	0.46	0.47	0.37	0.49
RTX A4000	0.42	0.38	0.29	0.31	0.47	0.44	0.29	0.33	0.20	0.36

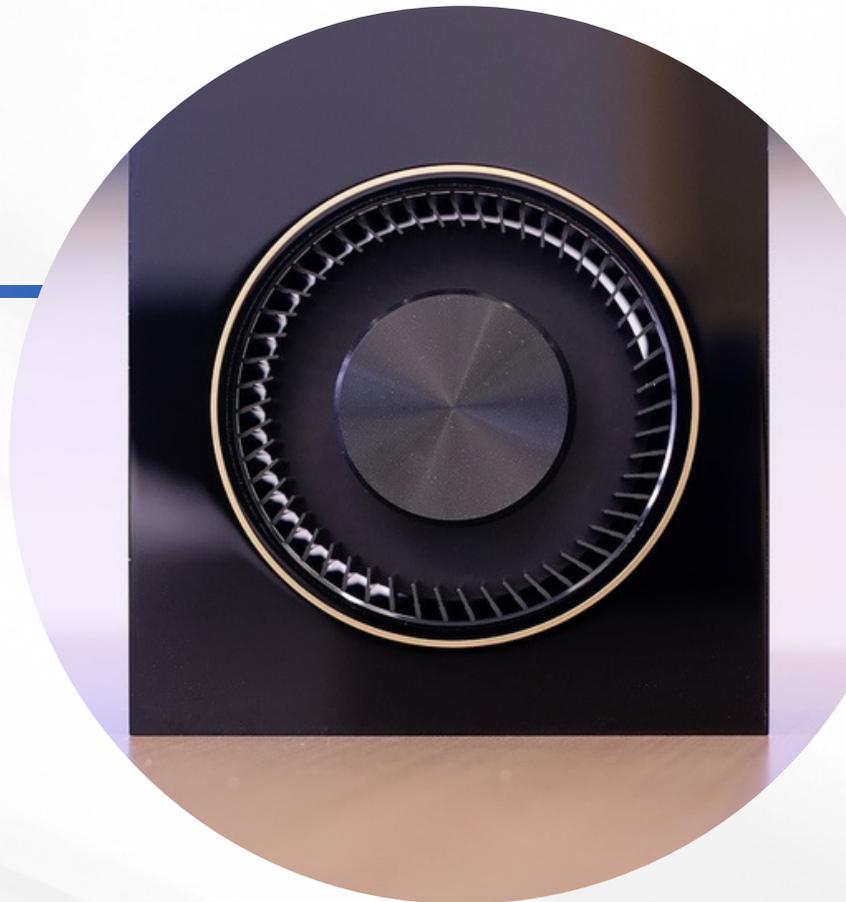
RTX 6000 Adaの利点は、デュアルスロット幅のフォームファクタと低消費電力設計にあり、お客様はより多くのGPUカードをサーバへと搭載することが可能で、それによりAIトレーニング性能を向上させることができます。例えば、2UサーバSKY-6200では、最大4枚のRTX 6000 Adaカードを挿入できます。しかし、3.5スロット幅のフォームファクタのGPUカードを使用する場合、寸法の制限から2Uサーバに搭載できないリスクがあります。

例えば、一般的に使用されている4UサーバのSKY-640 V3は、3.5スロット幅のGPUカードを搭載した場合、2枚のGPUカードしか搭載できません。しかし、RTX 6000 Adaを搭載した場合、4枚のGPUカードを搭載することができ、2枚搭載のシステムと比較して2倍のモデルトレーニング性能を提供することができます。さらに、必要な消費電力は1200Wで、一般的な2000Wサーバでの電源不足の懸念を完全に払拭できます。

カード収容枚数	消費電力	SKY-6 00 000		SKY-640 V3 (000)	
		枚数	消費電力	枚数	消費電力
RTX 6000 Ada		4	1200W	4	1200W
3.5スロットGPUカード		0	-	2	900W

特に冷却設計の面では、RTX 6000 Adaはデュアル吸気ブロー設計を採用しています。

したがって、複数の RTX 6000 Ada GPUカードを並列かつ密接にフィットする構成でシステムに搭載することで、効率的な吸気量と動作の安定性を確保することができます。

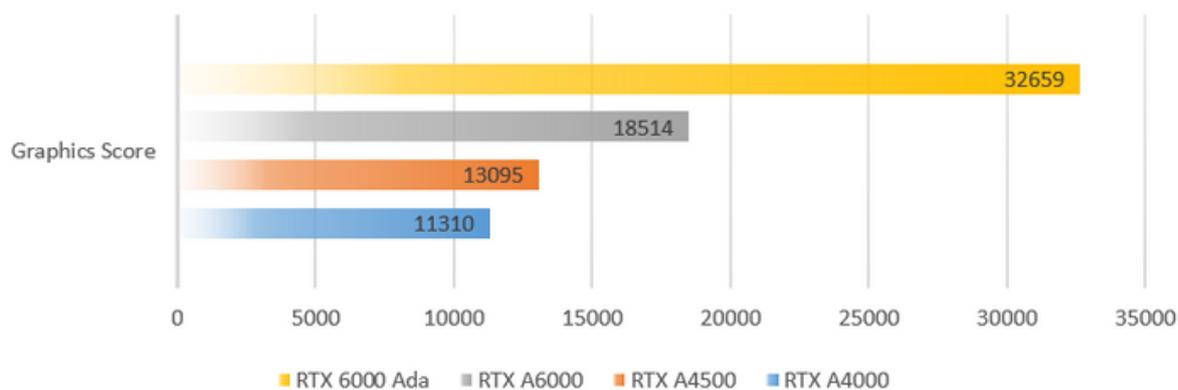


この構成が、スケーラビリティの実現と長期安定稼働の鍵を握っています。

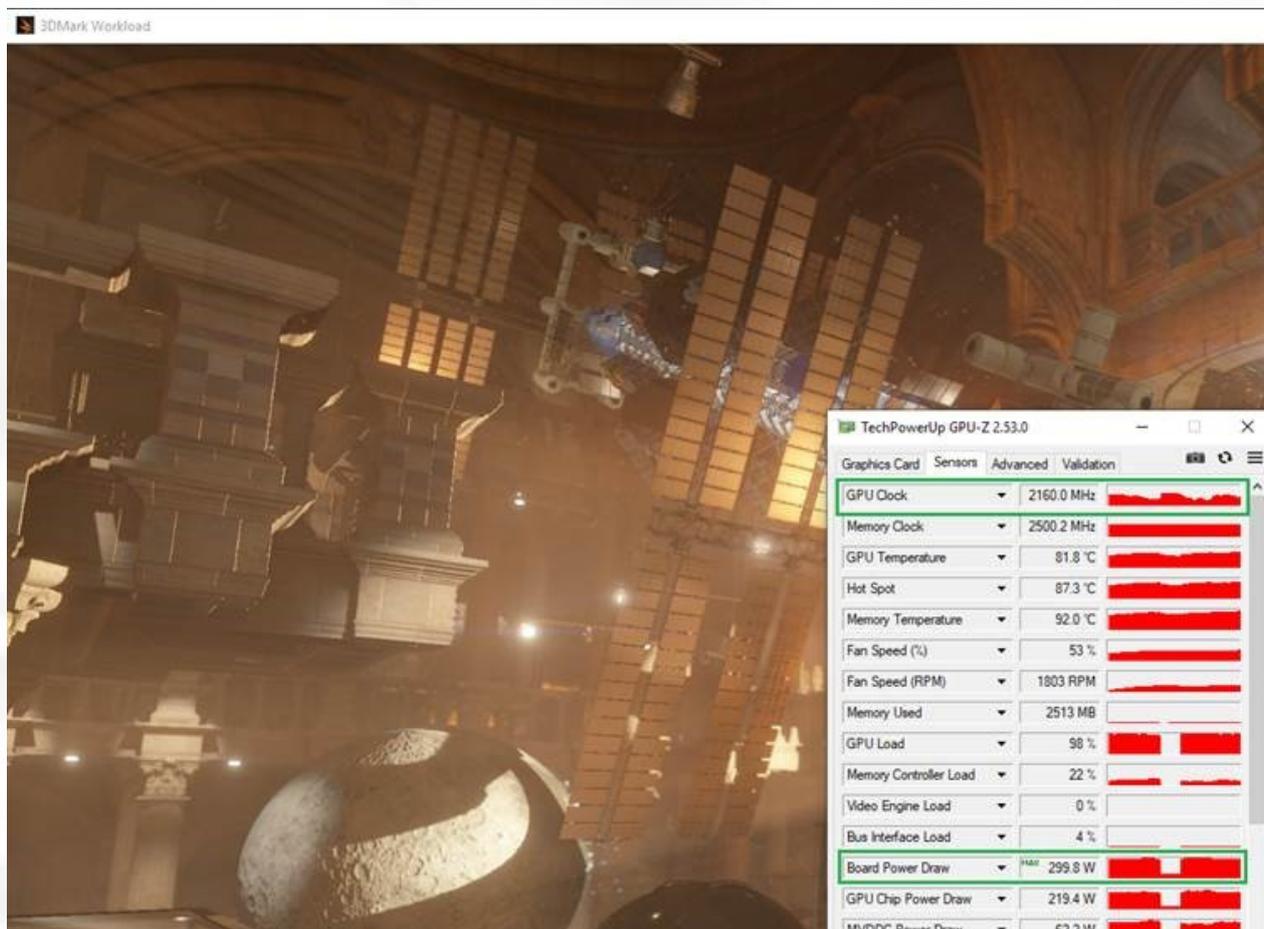
4. 3Dパフォーマンス 『3DMark Time Spy』 テスト

最後に、3DMark Time Spy テストでは、RTX 6000 Ada の性能は前世代の A6000 と比較して 76% 向上しており（18,514 → 32,659）主にコア数のアップグレードの恩恵を受けています。

3DMark Time Spy



次の図は、RTX 6000 Ada の消費電力が演算処理中にも関わらずほぼ 300W へと保たれ、GPUクロックが 1,900MHz から2,500MHz の間で変動していることを示しています。



	消費電力	GPU クロック	3DMark Time Spy グラフィックススコア
RTX 6000 Ada	300	1,900MHz~ 2,500MHz	32,659

RTX 6000 Adaは、より多くのCUDAコア、Tensorコア、RTコアを使用しています。そのため、GPUクロックが低く、変動していて低消費電力という利点をもたらしながら高い効率を維持することにつながります。これはワークステーションの長期安定稼働に大きく寄与します。

結論：世界をリードする産業用コンピューターと トップクラスのGPUを組み合わせ、最も効率的で信頼性 の高いAIワークステーションを実現

最上位のNVIDIA RTX 6000 Adaは、テストされた4つのGPUカードの中で、最高の性能、優れた双方向冷却設計、低消費電力の利点を誇り、最適なシステム統合、拡張性、ソフトウェア互換性をもたらします。プロフェッショナルのニーズを満たす幅広いアプリケーションに対応し、企業ユーザに従来の2倍以上のコンピューティング性能を提供すると同時に、24時間体制でのワークステーションの安定稼働と高い動作効率を提供します。

エッジコンピューティング、AIアクセラレーション、GPUコンピューティング、および5Gの広範囲な普及に伴い、リアルタイムAI推論コンピューティングの提供、およびレイテンシを低減するため、広範囲なAIアプリケーションをエッジ側へと展開させる必要性が高まっています。アドバンテックは、NVIDIA Elite Partnerであり、Professional RTX GPUカードを含むフルレンジのNVIDIA Enterprise PCIe GPUカードを提供できるとともに、データセンターやAIトレーニングアプリケーション向けのGPUカードも提供しています。また、最新の製品ライセンスへの優先アクセスや、ハードウェア、ソフトウェア、ドライバに関するNVIDIA社のテクニカルチームによる独占サポートも受けることができます。

さらに、アドバンテックは、エッジコンピューティングとAIソリューションにおける長年の経験を持つ産業用コンピュータの世界No.1ブランドであり、統合された技術とサービスだけでなく、お客様が選択できる幅広いシステムも提供しております。アドバンテック製品は、NVIDIAの厳格なシステム認証（NVIDIA-Certified and Qualified）に合格し、安定したシステム運用とソフトウェアの互換性を提供し、お客様の幅広いアプリケーションプロジェクトを最短時間で展開、かつ立ち上げることで、このAIとアクセラレーテッド・コンピューティングの時代に最高の価値を創造して参ります。