

NVIDIA A800 40GB Active

高度なコンピューティング ワークロード向けの 画期的なパフォーマンス

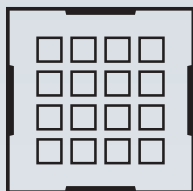
スーパーコンピューターのパワーをワークステーションにもたす。NVIDIA A800 40GB Active GPUはエンドツーエンドのデータサイエンスワークフローを加速します。NVIDIA Ampereアーキテクチャを搭載したA800 40GB Activeの提供する強力なコンピューティング、高速メモリ、拡張性により、プロフェッショナルは最も困難なデータサイエンス、AI、HPCワークロードに取り組むことができます。

エンタープライズセキュリティ、安定性、管理性、サポートを備えたエンドツーエンド ソフトウェア プラットフォームであるNVIDIA AI Enterpriseの3年間のサブスクリプションが付属しており、実稼働対応のAIおよびデータサイエンスを迅速に開発および展開するための100以上のAIフレームワーク、ライブラリ、事前トレーニング済みモデル、ツールが含まれています。

NVIDIA A800 40GB Activeと組み合わせることで、NVIDIA AI EnterpriseはAIの導入を簡素化し、最高のパフォーマンスでより迅速にビジネスの洞察を実現します。

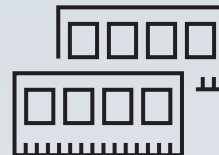
第3世代Tensorコア

倍精度 (FP64) および Tensor Float 32 (TF32) 精度をサポートする幅広い AI および HPC アプリケーション向けのパフォーマンスと多用途性により、前世代と比較して最大2倍のパフォーマンスと効率を実現します。構造的スパース性に対するハードウェアのサポートにより、推論のスループットが2倍になります。



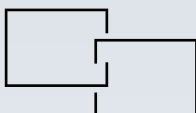
超高速 HBM2メモリ

クラス最高の1.5TB/sのメモリ帯域幅 (前世代と比較して70%以上増加)を備えた40GBの高速HBM2メモリと、40MBのレベル2キャッシュを含む大幅に増加したオンチップメモリにより、大規模な計算スループットを実現します。最も計算負荷の高いAIおよびHPCワークロードを高速化します。



第3世代 NVIDIA NVLink

GPU間の相互接続帯域幅の増加により、単一のスケラブルなメモリが提供され、コンピューティングワークロードが高速化され、より大規模なデータセットに取り組むことができます。ペアのNVIDIA A800 40GB Active GPUをNVIDIA NVLink®に接続すると、実効メモリフットプリントが80GBに増加し、最大400GB/秒(双方向)の速度でGPU間のデータ転送が可能になるため、アプリケーションのパフォーマンスが拡張されます。



マルチインスタンスGPU

専用の高帯域幅メモリ、キャッシュ、およびコンピューティングコアを備えたハードウェアレベルで完全に分離された安全なマルチテナンシー。マルチインスタンスGPU(MIG)は、GPUアクセラレーションのインフラストラクチャの利用を最大限に高め、A800 40GB Active GPUを最大7つの独立したインスタンスに分割できるようにし、複数のユーザーがGPUアクセラレーションにアクセスできるようにします。



NVIDIA A800 40GB Active

| | |
|--------------------------|---|
| 製品名 | NVIDIA A800 40GB Active |
| 製品型番 | EQRA800-40GER |
| JANコード | 4524076071611 |
| 技術仕様 | |
| GPUアーキテクチャ | NVIDIA Ampereアーキテクチャ |
| GPUメモリ | 40GB HBM2、ECC対応 |
| メモリ帯域幅 | 1,555.2 GB/s |
| メモリI/F | 5,120 bit |
| 接続バスコネクタ | PCI Express 4.0 x16 |
| CUDAコア数 | 6,912 |
| Tensorコア数 | 432 (第3世代) |
| RTコア数 | N/A |
| ピーク FP64 TFLOPS | 9.7 |
| FP64 Tensorコア TFLOPS | 19.5 |
| ピーク FP32 TFLOPS | 19.5 |
| TF32 Tensorコア TFLOPS | 311.8 |
| ピーク FP16 TFLOPS | 78.0 |
| BFLOAT16 Tensorコア TFLOPS | 362.05 |
| ピーク INT8 TOPS | 1247.4 |
| ピーク Tensor TFLOPS | 623.8 |
| 対応Compute API | CUDA 12.0 / DirectCompute / OpenCL 3.0 |
| 対応Graphics API | DirectX [®] 12 Ultimate / Shader Model 6.6 / OpenGL 4.6 / Vulkan 1.3 |
| 認証規格 | WHQL / ISO9241 / EU RoHS / JIG / REACH / HF / WEEE / RCM / BSMI / CE / FCC / ICES / KC / cUL, UL / VCCI |
| ボード外形寸法 | 266.70mm (長さ) × 111.15mm (高さ)、2スロット厚、ブラケットを含まず |
| 最大消費電力 | 240W |
| 補助電源コネクタ | PCI Express CEM5 16pin電源コネクタ×1 |
| 搭載ディスプレイコネクタ | なし ^{*1} ※NVIDIA RTXシリーズとの組み合わせを推奨 |
| 冷却方式 | アクティブ プロワーファン |
| vGPUソフトウェア対応 | NVIDIA AI Enterprise ※3年サブスクリプションライセンスが付属 (要アクティベーション ^{*2}) |
| vGPUプロファイル対応 | 仮想GPUライセンスガイド を参照 |
| NVENC NVDEC | - 5x |
| MIGサポート | 7 MIG@5GBまで対応 |
| NVIDIA NVLinkサポート | あり (NVLinkアダプター2つ接続で400 GB/sをサポート) |
| 対応OS ^{*3} | Windows [®] 11 / Windows [®] 10 / Windows [®] Server 2022 / Windows [®] Server 2019 / Linux |
| 付属品 | 製品保証書 / 2x PCIe 8pin to CEM5 16pin変換アダプタ×1 |
| 保証期間 | 3年 |
| パッケージ外形寸法 / 重量 | 328(W)×230(H)×87(D) mm / TBA |
| 推奨動作環境 | 1~8GPUを搭載可能なシステム。NVIDIA認定システムについては NVIDIA認定システムカタログ を参照。 |

*1 本製品はディスプレイ出力がありません、NVIDIA RTXシリーズなどとの組み合わせを推奨いたします。 *2 NVIDIA AI Enterpriseのライセンスを有効化するためには専用ページでアクティベーションが必要です。 *3 全て64bit OSのみ対応。ドライバーバージョンにより対応するOSバージョンが異なる場合がございます。詳細は[NVIDIA CUDA Toolkit Release Notes](#)や、該当ドライバーのリリースノート内のSupported Operating Systemsをご参照ください。

お問い合わせ先
株式会社 エルザ ジャパン

<https://www.elsa-jp.co.jp>
〒105-0014 東京都港区芝3-42-10 三田UTビル
TEL:03-5763-7391

コーポレートサイト



オンライン見積



法人様向
貸出しサービス



最新事例一覧



© 2023 Varjo All Rights Reserved. © 2023 NVIDIA Corporation All Rights Reserved.
NVIDIA, GeForce RTX, NVIDIA RTXは米国および/または他国のNVIDIA Corporation商標および/または登録商標です。
他の企業および製品名は、それらと関連性のある各企業の商標である可能性があります。
ELSA(エルザ)はテクノロジー・ジョイント株式会社の登録商標です。仕様などは改良のため予告なく変更する場合があります。