

PNY NVIDIA Quadro Sync

Artikelnummer 900040268

Gewicht 1kg

Länge 1mm

Breite 1mm

Höhe 1mm



Produktbeschreibung

Die rasante Zunahme der Komplexität von Workloads, der Datenmenge und die Verbreitung neuer Workloads wie generative KI läuten eine neue Ära der Datenverarbeitung ein, die wissenschaftliche Entdeckungen beschleunigt, die Produktivität verbessert und die Erstellung von Inhalten revolutioniert. Da die Modelle immer größer und komplexer werden, um die Herausforderungen der nächsten Generation zu bewältigen, müssen immer mehr Arbeitslasten auf lokalen Geräten ausgeführt werden. Workstation-Plattformen der nächsten Generation müssen High-Performance-Computing-Funktionen bieten, um diese komplexen Workloads zu unterstützen. Die NVIDIA A800 40GB Active GPU beschleunigt Data-Science-, KI- und HPC-Workflows mit 432 Tensor-Cores der dritten Generation, um die KI-Leistung zu maximieren und ultraschnelle und effiziente Inferenzen zu ermöglichen. Mit der NVIDIA NVLink Technologie der dritten Generation bietet der A800 40GB Active eine skalierbare Leistung für schwere KI-Workloads, verdoppelt den effektiven Speicherbedarf und ermöglicht GPU-zu-GPU-Datenübertragungen mit einer bidirektionalen Bandbreite von bis zu 400 Gigabyte pro Sekunde (GB/s). Dieses Board ist eine KI-fähige Entwicklungsplattform mit NVIDIA AI Enterprise und bietet Workstations, die sich ideal für die Anforderungen von erfahrenen KI-Entwicklern und Datenwissenschaftlern eignen.

- Verstärkte Grafikleistung in Verbindung mit der Leistung von NVIDIA RTX**

Zur Unterstützung von Display-Funktionen und zur Bereitstellung von Hochleistungsgrafiken für visuelle Anwendungen ist die Rechenleistung des NVIDIA A800 40GB Active für die Kombination mit NVIDIA RTX-beschleunigten Grafikprozessoren ausgelegt. Die NVIDIA RTX 4000 Ada Generation, RTX A4000 und T1000 8GB GPUs sind für den Betrieb mit dem A800 40GB Active zertifiziert und bieten leistungsstarkes Echtzeit-Raytracing und KI-beschleunigte Grafikleistung in einem Single-Slot-Formfaktor.

- Datenwissenschaft und Datenanalyse**

Beschleunigen Sie durchgängige Data-Science- und Analyse-Workflows mit hoher Leistung, um schnell aussagekräftige Erkenntnisse aus großen Datensätzen zu gewinnen. Durch die Kombination der High-Performance-Rechenleistung der A800 40GB Active mit NVIDIA AI Enterprise können Datenexperten eine große Sammlung von Bibliotheken, Tools und Technologien nutzen, um Data-Science-Workflows zu beschleunigen - von der Datenvorbereitung über das Training bis hin zur Inferenz.

- Training und Entwicklung**

Mit 40 GB HBM2-Speicher und leistungsstarken Tensor-Cores der dritten Generation, die bis zu zweimal mehr Leistung als die Vorgängergeneration bieten, liefert die A800 40 GB Active GPU eine unglaubliche Leistung, um anspruchsvolle KI-Entwicklungs- und Trainings-Workflows auf Workstation-Plattformen zu bewältigen, einschließlich Datenvorbereitung und -verarbeitung, Modelloptimierung und -abstimmung sowie Training im Frühstadium. Die NVIDIA AI Enterprise Software-Plattform beschleunigt und vereinfacht den Einsatz von KI im großen Maßstab, so dass Unternehmen einmal entwickeln und überall einsetzen können. Die Kopplung dieser leistungsstarken Softwareplattform mit der A800 40GB Active GPU ermöglicht es KI-Entwicklern, KI-Modelle auf Workstations zu erstellen, zu iterieren und zu verfeinern und dabei die Skalierung zu vereinfachen und kostspielige Rechenzentrumsressourcen für teure, umfangreiche Berechnungen zu reservieren.

- Inferenz**

Inferenz ist der Bereich, in dem KI Ergebnisse liefert und durch die Operationalisierung trainierter Modelle umsetzbare Erkenntnisse liefert. Mit 432 Tensor Cores der dritten Generation und 6.912 CUDA Cores bietet die A800 40GB Active eine doppelt so hohe Leistung bei Inferenzoperationen wie die Vorgängergeneration und unterstützt strukturelle Sparsamkeit und eine breite Palette von Genauigkeiten, einschließlich TF32, INT8 und FP64. KI-Entwickler können NVIDIA Inferenzsoftware wie NVIDIA TensorRT und NVIDIA Triton Inference Server nutzen, die Teil von NVIDIA AI Enterprise sind, um die Bereitstellung von KI-Modellen in großem Maßstab zu vereinfachen und zu optimieren.

- Generative KI**

Generative KI-Anwendungen nutzen neuronale Netze, um Muster und Strukturen in vorhandenen Daten zu erkennen, und ermöglichen es Anwendern, aus einer Vielzahl von Inputs und Outputs, darunter Bilder, Sounds, Animationen und 3D-Modelle, originelle Inhalte zu generieren. Nutzen Sie die generative KI-Lösung NeMo Framework von NVIDIA, die zusammen mit dem NVIDIA A800 40GB Active GPU in NVIDIA AI Enterprise enthalten ist, für die einfache, schnelle und anpassbare Entwicklung generativer KI-Modelle.

- Leistungsstarkes Computing**

Die A800 40GB Active GPU bietet eine unglaubliche Leistung für GPU-beschleunigte CAE-Anwendungen (Computer-Aided Engineering). Ingenieure und Produktentwickler können umfangreiche Simulationen für Finite-Elemente-Analysen (FEA), numerische Strömungsmechanik (CFD), Konstruktionsmanagement (CEM) und andere technische Analysecodes in voller FP64-Präzision mit unglaublicher Geschwindigkeit ausführen und so die Entwicklungszeiten verkürzen und die Wertschöpfung beschleunigen. Mit den zusätzlichen RTX-beschleunigten GPUs, die Anzeigefunktionen bieten, können Wissenschaftler und Ingenieure umfangreiche Simulationen und Modelle in voller Designtreue visualisieren.

- Energie und Geowissenschaften**

Mit einer FP64-Rechenleistung von 9,7 TFLOPS ermöglicht die A800 40GB Active GPU Geowissenschaftlern, die neuesten KI-unterstützten Software-Workflows für Exploration und Produktion zu nutzen und Simulationsprozesse zu beschleunigen, um einen schnelleren Einblick in unterirdische Daten zu erhalten. Für große Datensätze können zwei A800 40GB Active GPUs mit NVLink verbunden werden, um 80GB Speicher und die doppelte Rechenleistung bereitzustellen.

- Biowissenschaften**

Mit der A800 40GB Active können Fachleute in allen Bereichen der Biowissenschaften komplexe Datenverarbeitungsaufgaben beschleunigen, schnellere Entdeckungen ermöglichen und die Entscheidungsfindung verbessern. KI-beschleunigte Life-Science-Anwendungen wie Genomics-Sequenzierung, medizinische Bildgebung und personalisierte Medizin profitieren von einer schnelleren Trainings- und Inferenzleistung, um die Analyse großer Datensätze zu beschleunigen. Bei komplexen Simulationen und Datenverarbeitungsaufgaben, die eine hohe Genauigkeit erfordern, können wissenschaftliche Anwendungen wie Molekulardynamik, Arzneimittelerkundung und Genomanalyse dank der FP64-Funktionen mit höherer Genauigkeit und Präzision ausgeführt werden, was zu zuverlässigeren Ergebnissen führt.

Weitere Bilder

