



NVIDIA DGX STATION A100

DISPOSITIVO DE GRUPO DE TRABALHO PARA A ERA DA AI



As equipes de ciência de dados estão na linha de frente das inovações em AI, desenvolvendo projetos que podem transformar empresas e o mundo. Porém, geralmente, elas precisam procurar ciclos de computação livres que possam ajudar a treinar modelos mais complexos. Essas equipes precisam de uma plataforma de AI dedicada que possa ser conectada em qualquer lugar e cujo hardware e software sejam totalmente otimizados para oferecer um desempenho revolucionário para vários usuários ao mesmo tempo em qualquer lugar do mundo.

Visão Geral da NVIDIA DGX Station™ A100

- > Um servidor de grupo de trabalho com AI que oferece um desempenho de 2,5 petaFLOPS que sua equipe pode aproveitar sem limites, seja para treinamentos, inferências ou análises de dados
- > Um sistema de nível de servidor e pronto para uso que não precisa da energia nem do resfriamento de um data center
- > Uma plataforma de AI de nível internacional de fácil instalação e que dispensa ajuda da equipe de IT
- > O único sistema em forma de workstation do mundo com 4 GPUs NVIDIA A100 Tensor Core totalmente interconectadas e uma memória de GPU de até 320GB
- > Acelere a transformação de AI com o conhecimento e a experiência da NVIDIA

A NVIDIA DGX Station A100 garante a supercomputação de AI às equipes de ciência de dados, oferecendo a tecnologia de um data center sem um data center ou uma infraestrutura de IT adicional. Desenvolvida para vários usuários simultâneos, a DGX Station A100 apresenta componentes de nível de servidor em um formato adaptado para escritórios. É o único sistema com 4 GPUs NVIDIA A100 Tensor Core totalmente interconectadas que podem ser transformadas em uma GPU Multi-Instância (MIG - Multi-Instance GPU) com uma memória total de até 320GB, podendo ser conectado a uma tomada elétrica padrão e funcionando como um dispositivo de AI potente que pode ser colocado em qualquer lugar.

Supercomputação de AI para Equipes de Ciência de Dados

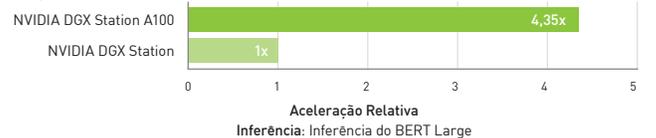
Capaz de oferecer a vários usuários um recurso centralizado de AI ao mesmo tempo com facilidade, a DGX Station A100 é o dispositivo de grupo de trabalho para a era da AI. Ela pode executar cargas de trabalho de treinamento, inferência e análise simultaneamente e, com

A DGX Station A100 Oferece um Desempenho 3,17x Mais Rápido em Inferências



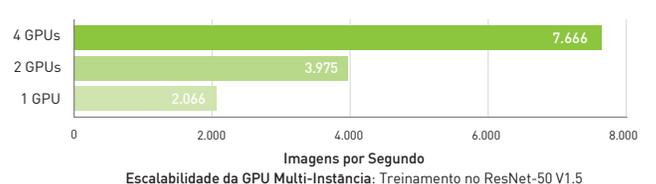
DGX Station A100 de 320GB; Tamanho do Lote = 64; Precisão Mista; com AMP; Dados Reais; Tamanho da Sequência = 128

A DGX Station A100 Oferece um Desempenho 4,35x Mais Rápido em Inferências



DGX Station A100 de 320GB; Tamanho do Lote = 256; Precisão INT8; Dados Sintéticos; Tamanho da Sequência = 128, cuDNN 8.0.4

A DGX Station A100 Oferece Escalabilidade Linear



DGX Station A100 de 320GB; Tamanho do Lote = 192; Precisão Mista; Dados Reais; Versão da cuDNN = 8.0.4; Versão da NCCL = 2.7.8; Container MXNet 20.10 do NGC

o MIG, pode oferecer até 28 GPUs distintas para usuários e tarefas individuais, de

Desempenho de um Data Center sem o Data Center

A NVIDIA DGX Station A100 é um servidor de AI de nível de data center em um formato de workstation, adaptado para uso em ambientes de escritório padrão sem energia e resfriamento especializados. Seu design conta com 4 GPUs NVIDIA A100 Tensor Core ultrapotentes, uma CPU avançada de nível de servidor, um armazenamento NVMe extremamente rápido e barramentos PCIe de 4ª geração de ponta. A DGX Station A100 também vem com o mesmo Baseboard Management Controller (BMC) que a NVIDIA DGX A100, que oferece aos administradores de sistemas a possibilidade de realizar qualquer tarefa necessária por uma conexão remota. A DGX Station A100 é o sistema de AI mais potente para ambientes de escritório, oferecendo a tecnologia de um data center sem o data center.

Um Dispositivo de AI que Você Pode Instalar em Qualquer Lugar

A NVIDIA DGX Station A100 foi desenvolvida para as equipes de ciência de dados ágeis de hoje que trabalham em escritórios corporativos, laboratórios, institutos de pesquisa ou até mesmo em casa. Enquanto para instalar uma infraestrutura de AI de grande escala são necessários investimentos consideráveis em IT e data centers de grande porte com potência e resfriamento de nível industrial, para instalar a DGX Station A100, basta conectá-la a uma tomada elétrica padrão, seja qual for o espaço de trabalho da equipe. Além disso, seu design inovador e baseado em refrigeração garante uma temperatura segura. Basta uma pessoa para instalar o dispositivo. Você mesmo pode fazer isso e começar a usar, em poucos minutos, uma plataforma de AI de nível internacional que só precisa de dois cabos para funcionar.

Modelos Maiores, Respostas Mais Rápidas

A NVIDIA DGX Station A100 não é uma workstation. É um servidor de grupo de trabalho com AI que pode ser instalado em sua mesa. Além da CPU de nível de data center com 64 núcleos, ele tem as mesmas GPUs NVIDIA A100 Tensor Core que as do servidor NVIDIA DGX A100, cada uma com 40 ou 80GB de memória, conectadas por um módulo SXM4 de alta velocidade. A NVIDIA DGX Station A100 é o único sistema adaptado para escritórios que tem 4 GPUs totalmente interconectadas com o NVIDIA® NVLink® e que podem ser transformadas em uma MIG, oferecendo até 28 GPUs distintas para tarefas paralelas e vários usuários sem afetar o desempenho do sistema.

Acesso Integrado à Experiência Única de AI

A NVIDIA DGX Station A100 é uma plataforma completa de hardware e software apoiada por milhares de especialistas em AI da NVIDIA e baseada no conhecimento adquirido no maior campo de provas DGX do mundo, a NVIDIA DGX SATURNV. O proprietário de um DGX Station A100 tem acesso direto aos DGXperts da NVIDIA, uma equipe global de profissionais especialistas em AI que oferecem instruções e expertise em design para ajudar a acelerar a transformação de AI com o conhecimento e a experiência de mais de uma década de liderança em AI da NVIDIA. Assim, as aplicações mais importantes podem começar a ser usadas com rapidez e continuar em execução com fluidez, diminuindo consideravelmente o tempo necessário para visualizar informações.

Acesse www.nvidia.com/DGXStationA100 e saiba mais sobre a NVIDIA DGX A100.

Todos os direitos reservados. NVIDIA, o logotipo da NVIDIA, DGX, DGX A100, DGX POD, DGX Station, DGX SuperPOD e NVLink são marcas comerciais ou marcas registradas da NVIDIA Corporation. Todos os nomes de produtos e empresas são marcas comerciais ou marcas registradas de seus respectivos proprietários aos quais estão associados. Todos os recursos, os preços, a disponibilidade e as especificações estão sujeitos a alterações sem aviso prévio. NOV20

ESPECIFICAÇÕES DO SISTEMA

	NVIDIA DGX Station A100 de 320GB	NVIDIA DGX Station A100 de 160GB
GPUs	4 GPUs NVIDIA A100 de 80GB	4 GPUs NVIDIA A100 de 40GB
Memória da GPU	320GB total	Total de 160GB
Performance	AI de 2,5 petaFLOPS INT8 de 5 petaOPS	
Uso de Energia do Sistema	1,5 kW a 100–120 VCA	
CPU	1 AMD 7742, 64 núcleos, 2,25 GHz (base)–3,4 GHz (desempenho máximo)	
Memória do sistema	DDR4 de 512 GB	
Rede	Duas portas LAN Ethernet 10 G Base-T 1 porta de gerenciamento Ethernet 1 G Base-T para o BMC	
Armazenamento	SO: 1 unidade NVMe de 1,92 TB Armazenamento interno: unidade NVMe U.2 de 7,68 TB	
Adaptador de Vídeo DGX	Memória de GPU de 4GB, 4 miniDisplayPorts	
Acústica do Sistema	<37 dB	
Software	SO Ubuntu Linux	
Peso do Sistema	43,1 kg (91,0 lb)	
Peso do Sistema Empacotado	57,93 kg (127,7 lb)	
Dimensões do sistema	Altura: 639 mm (25,1 pol.) Largura: 256 mm (10,1 pol.) Comprimento: 518 mm (20,4 pol.)	
Faixa de temperatura operacional	5 a 35 °C (41 a 95 °F)	

