

모든 엔터프라이즈를 위한 AI 가속화

NVIDIA DGX™ B300

견적문의 바로가기 ↗

NVIDIA DGX B300은 생성형 AI의 컴퓨팅 수요를 충족하도록 맞춤 제작된 AI 인프라 솔루션입니다. NVIDIA Blackwell Ultra GPU를 기반으로 하는 DGX B300은 최신 데이터센터에 원활하게 맞도록 설계된 새로운 폼 팩터에서 이전 세대보다 11배 향상된 추론 및 4배 향상된 훈련 성능을 제공합니다. NVIDIA MGX™와 기존 엔터프라이즈 랙과 호환되며 풀스택 소프트웨어와 함께 AI 배포를 간소화하고 효율화하여 모든 기업이 하이퍼스케일러처럼 운영할 수 있도록 지원합니다.

AI 팩토리 시대의 핵심 인프라

비즈니스의 미래는 AI 추론 시대와 AI 팩토리로의 데이터 센터 진화라는 이중 혁신에 직면하면서, 인프라에 대한 근본적인 변화가 불가피해졌습니다. 하지만 이 새로운 변화로 인해 기업들은 막대한 컴퓨팅 자원이 요구되며, 인프라 구축 및 운영의 복잡성, 전문 인력과 시스템 통합 능력 부족, 높은 에너지 비용 등 현실적인 어려움에 직면하게 됩니다.

NVIDIA DGX B300은 이러한 AI 시대의 핵심 구성 요소로서, 선구적인 기업들이 AI 추론의 요구 사항을 충족하는 자체 AI 팩토리를 구축하도록 지원합니다. NVIDIA Blackwell Ultra 아키텍처로 구동되는 DGX B300은 엔터프라이즈 규모의 설치 공간에서 하이퍼스케일러 성능을 제공하도록 특별히 제작된 AI 강력한 시스템입니다. 추론 워크로드에 대해 144 페타플롭스(petaFLOPS)의 FP4 추론 성능을 제공합니다. 새롭게 설계된 NVIDIA MGX™를 준수하는 공랭식 새시는 최신 데이터 센터에 원활하게 통합할 수 있도록 합니다. NVIDIA Mission Control 소프트웨어와 결합되어 AI 운영을 단순화하고, 기업이 생성형 AI의 복잡성을 숙달하고 투자 수익을 확보하는 데 필요한 풀 스택(full-stack) 솔루션을 제공합니다.

실시간 AI 강력한 시스템

DGX B300은 AI 팩토리의 근본적인 구성 요소가 되도록 설계되었습니다. 이는 모든 규모의 AI 혁신가들이 이전에 글로벌 규모의 AI 조직에만 국한되었던 생성형 AI 기능을 활용하여 대규모로 인텔리전스를 제조할 수 있도록 합니다. NVIDIA Blackwell Ultra GPU, NVIDIA® ConnectX®-8 네트워킹, NVIDIA Mission Control 소프트웨어로 구동되는 완벽하게 통합된 시스템인 DGX B300은 전례 없는 성능과 하이퍼스케일급 효율성을 제공합니다. 탁월한 학습 성능과 선도적인 실시간 추론을 결합함으로써, DGX B300은 모든 조직이 AI 추론 시대를 위한 확장 가능한 인프라를 구축할 수 있도록 지원합니다.

현대 데이터 센터를 위한 청사진

DGX B300은 재설계된 OCP(Open Compute Project)를 준수하는 새시로 시작하여, 모든 데이터 센터에 하이퍼스케일러 설계 원칙을 도입함으로써 최신 AI 팩토리의 청사진 역할을 합니다. 처음으로 DGX B300은 NVIDIA MGX 호환 새시 또는 보다 전통적인 AC 전원 설계를 사용하여 배포할 수 있어, DGX B300을 모든 인프라 환경에서 사용할 수 있도록 보장합니다. 이 최첨단 설계와 실용적인 서비스 용이성을 결합함으로써, DGX B300은 기업이 전례 없는 효율성과 선택권을 가지고 자신만의 방식으로 AI 팩토리를 구축할 수 있도록 합니다.

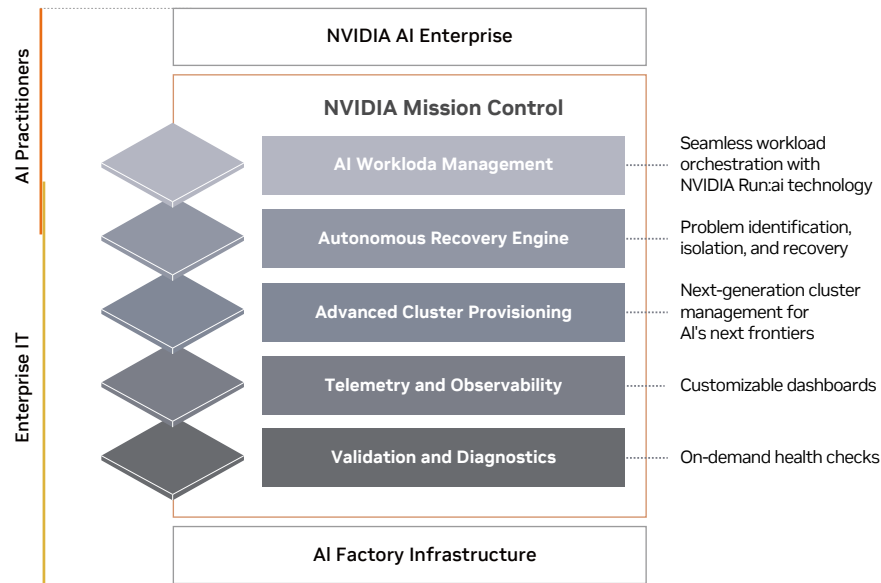


| NVIDIA DGX B300 주요 기능

- 8x NVIDIA Blackwell Ultra SXM 탑재
- 2.1TB의 GPU 메모리 공간
- 72 petaFLOPS의 훈련 성능
- 144 petaFLOPS의 추론 성능
- NVIDIA 네트워킹
- Intel Xeon 6776P 프로세서
- [NVIDIA DGX BasePOD](#) 및 [NVIDIA DGX SuperPOD](#)의 기반
- NVIDIA AI Enterprise 및 NVIDIA Mission Control™ 소프트웨어 포함

모델 실행, 필수 작업 자동화 NVIDIA Mission Control™

NVIDIA NVIDIA Mission Control은 엔터프라이즈 AI 인프라의 운영을 간소화하고, 다운타임을 줄이며, 모델 개발을 가속화하도록 설계된 통합 AI 공장 관리 플랫폼입니다. NVIDIA의 운영 모범 사례와 AI 클러스터 자동화를 단일 제어 영역으로 통합합니다.



NVIDIA DGX software stack.

DGX B300 시스템의 가치 극대화

리더시스템즈의 NVIDIA 엔터프라이즈 서비스는 DGX 인프라에 대한 지원, 교육 및 인프라 전문가를 제공합니다. AI 여정의 모든 단계에서 NVIDIA 전문가가 함께하기 때문에, 엔터프라이즈 서비스는 프로젝트를 빠르고 성공적으로 시작하도록 지원할 수 있습니다.

성공을 위한 핵심: NVIDIA 교육 프로그램

혁신적인 기술, 특히 NVIDIA DGX 시스템의 잠재력을 최대한 활용하기 위해서는 교육이 필수적입니다. NVIDIA Elite 파트너, 리더시스템즈는 NVIDIA DGX 고객을 위한 [NVIDIA Deep Learning Institute\(DLI\)](#)를 제공하여 맞춤 기술 교육을 받을 수 있도록 합니다. AI 전문가가 진행하는 실습 위주 기술 교육을 통해 개인과 조직은 AI 분야에서 첨단 혁신을 달성하는 방법을 배울 수 있습니다.

NVIDIA DGX B300 GPU Specifications

GPU	8× NVIDIA Blackwell Ultra SXM
GPU Memory	2.1 TB total, 64 TB/s HBM3e bandwidth
Performance	FP4 Tensor Core* - 144 PFLOPS 108 PFLOPS FP8/FP6 Tensor Core** - 72 PFLOPS
NVIDIA® NVSwitch™	2×
NVIDIA NVLink Bandwidth	14.4 TB/s aggregate bandwidth
System Power	14.5 kW (Busbar) / 15.1 kW (PSU)
CPU	Intel® Xeon® 6776P processors
System Memory	2 TB default (up to 4 TB)
Networking	- 8x OSFP ports serving 8x NVIDIA ConnectX-8 VPI > Up to 800 Gb/s of NVIDIA InfiniBand/Ethernet 2x dual-port QSFP112 NVIDIA BlueField®-3 DPU > Up to 400 Gb/s of NVIDIA InfiniBand/Ethernet
Management Network	1GbE onboard network interface card (NIC) with RJ45 1GbE RJ45 host baseboard management controller (BMC)
Storage	OS: 2× 1.9 TB NVMe M.2 Internal storage: 8× 3.84 TB NVMe E1.S
Software	NVIDIA AI Enterprise (optimized AI software) / NVIDIA Mission Control (AI data center operations and orchestration with NVIDIA Run:ai technology) / NVIDIA DGX OS (operating system) Supports Red Hat Enterprise Linux / Rocky / Ubuntu
Rack Units (RU)	10
Operating Temperature	10C°~35C°
Support	3년간 비즈니스 표준 하드웨어 및 소프트웨어 지원

* 표기된 사양은 최소 | 밀집 구성 기준입니다. ** 수치는 최소 구성 기준이며, 밀집 구성 시 해당 값의 절반입니다.

[주]리더시스템즈

경기도 하남시 미사강변한강로 155, 미사강변 SK V1센터 504~513호
문의전화 1544-9880 | 메일 sales@leaderssys.com www.leaderssys.com

