



NVIDIA DGX H200

NVIDIA DGX 시스템은 AI 인프라의
최고 수준을 대표합니다

인공지능은 고객 서비스 개선, 공급망 최적화, 비즈니스 인텔리전스 추출, 생성형 AI와 기타 트랜스포머 모델을 활용한 첨단 제품 및 서비스 설계 등 다양한 비즈니스 과제를 해결하는 핵심적인 접근 방식으로 자리잡았습니다. 거의 모든 산업 분야의 조직들이 혁신을 실현할 수 있는 메커니즘을 제공합니다. AI 인프라의 선두주자인 NVIDIA DGX는 이러한 필수적인 아이디어를 현실화하기 위한 가장 강력하고 완벽한 AI 플랫폼을 제공합니다.

NVIDIA DGX H200, 비즈니스 혁신과 최적화를 가속화

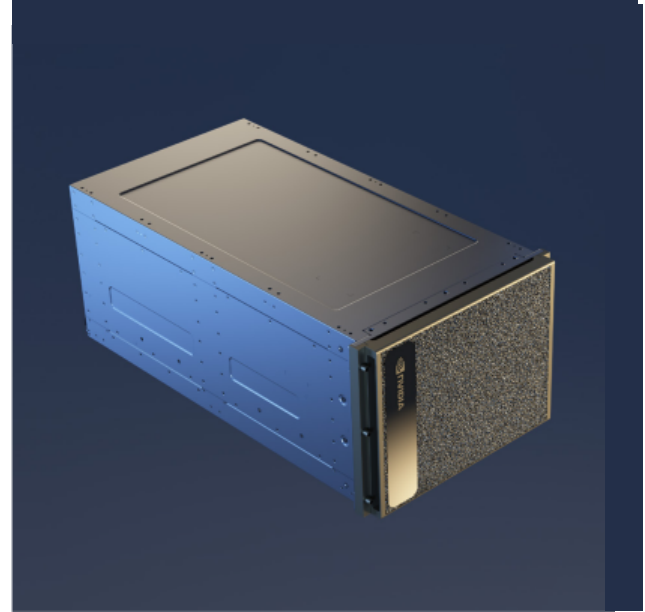
NVIDIA의 전설적인 **DGX 플랫폼**의 일부이며 **NVIDIA DGX SuperPOD** 및 **DGX BasePOD™**의 기반인 DGX H200은 획기적인 NVIDIA H200 Tensor Core GPU를 탑재한 AI 강력한 플랫폼입니다. 이 시스템은 AI 처리량을 극대화하도록 설계되어 기업이 자연어 처리, 추천 시스템, 데이터 분석 등 다양한 분야에서 혁신을 달성할 수 있도록 고도로 정제되고 체계적이며 확장 가능한 플랫폼을 제공합니다. 온-프레미스 및 다양한 액세스 및 배포 옵션을 통해 제공되는 DGX H200은 기업이 AI로 가장 큰 과제를 해결하는 데 필요한 성능을 제공합니다.

AI 센터의 핵심

AI는 과학과 비즈니스 사이의 간극을 메웠습니다. 더 이상 실험의 영역이 아닌, AI는 대기업과 소기업 모두가 혁신을 주도하고 비즈니스를 최적화하기 위해 매일 사용하고 있습니다. 세계 최초의 목적 기반 AI 인프라 포트폴리오의 일부인 DGX H200은 기업 AI 센터의 핵심으로 설계되었습니다. **NVIDIA Enterprise Support**를 포함한 완전히 최적화된 하드웨어 및 소프트웨어 플랫폼으로, NVIDIA AI 소프트웨어 솔루션의 새로운 범위, 풍부한 타사 지원 생태계, NVIDIA 전문가의 전문적인 조언에 대한 액세스를 제공하여 조직이 가장 크고 복잡한 비즈니스 문제를 AI로 해결할 수 있도록 합니다. DGX H200은 DGX 플랫폼이 전 세계 거의 모든 산업 분야의 수천 명의 고객이 사용하고 있는 검증된 신뢰성을 제공합니다.

AI의 규모와 성능의 한계를 돌파하다

NVIDIA DGX H200은 AI의 규모와 성능의 한계를 돌파합니다. 32페타플롭스의 AI 성능, NVIDIA ConnectX-7 스마트 네트워크 인터페이스 카드(SmartNIC)를 탑재한 DGX A100보다 2배 빠른 네트워킹, NVIDIA DGX SuperPOD 및 DGX BasePOD를 위한 고속 확장성을 제공합니다. DGX H200은 생성형 AI, 자연어 처리, 딥 러닝 추천 모델과 같은 가장 크고 복잡한 AI 학습 및 추론 작업을 위한 1,128GB의 GPU 메모리로 강화되었습니다.

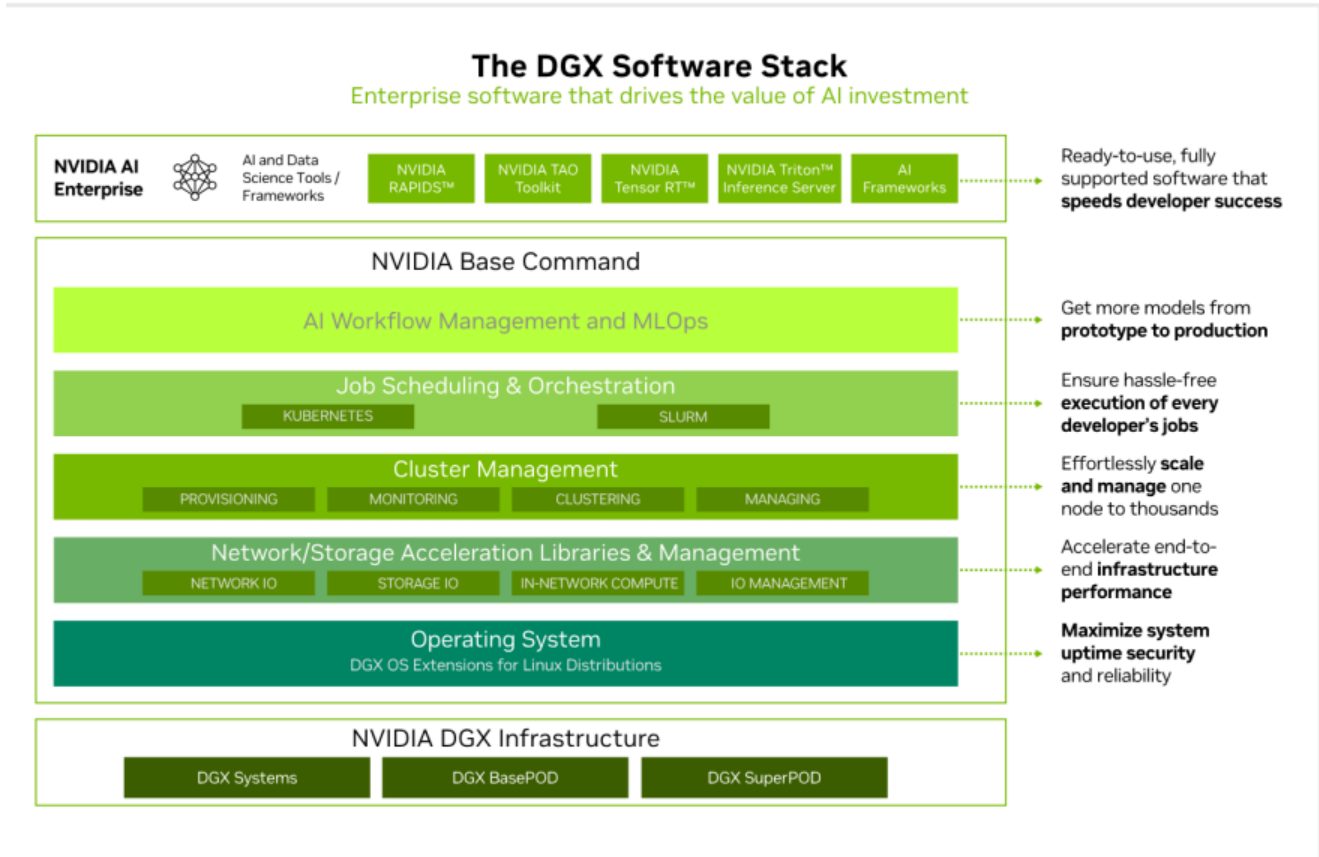


Specifications	
GPU	8x NVIDIA H200 Tensor Core GPUs, with 141GB of GPU memory each
GPU memory	1,128GB total
Performance	32 petaFLOPS FP8
NVIDIA NVSwitch™	4x
System power usage	10.2kW max*
CPU	Dual Intel® Xeon® Platinum 8480C Processors 112 Cores total, 2.00 GHz (Base), 3.80 GHz (Max Boost)
System memory	2TB
Networking	4x OSFP ports serving 8x single-port NVIDIA ConnectX-7 VPI ➢ Up to 400Gb/s InfiniBand/Ethernet 2x dual-port QSFP112 NVIDIA ConnectX-7 VPI ➢ Up to 400Gb/s InfiniBand/Ethernet
Management network	10Gb/s onboard NIC with RJ45 100Gb/s Ethernet NIC Host baseboard management controller (BMC) with RJ45
Storage	OS: 2x 1.92TB NVMe M.2
Internal storage:	8x 3.84TB NVMe U.2
Software	NVIDIA AI Enterprise – Optimized AI software NVIDIA Base Command – Orchestration, scheduling, and cluster management DGX OS / Ubuntu / Red Hat Enterprise Linux / Rocky – Operating System
Support	Comes with 3-year business-standard hardware and software support
System weight	287.6lbs (130.45kgs)
Packaged weight	376lbs (170.45kgs)
System dimensions	Height: 14.0in (356mm) Width: 19.0in (482.2mm) Length: 35.3in (897.1mm)
Operating temperature range	5–30°C (41–86°F)

*10.2kW for standard configuration. DGX H200 CTS (Custom Thermal Solution) supports up to 14.3kW.

NVIDIA Base Command의 힘

NVIDIA Base Command는 DGX 플랫폼을 구동하여 조직이 NVIDIA 소프트웨어 혁신의 최고를 활용할 수 있도록 합니다. 기업은 엔터프라이즈급 오케스트레이션 및 클러스터 관리, 컴퓨팅, 스토리지, 네트워크 인프라를 가속화하는 라이브러리, AI 워크로드에 최적화된 운영 체제를 포함하는 검증된 플랫폼으로 DGX 인프라의 잠재력을 최대한 발휘할 수 있습니다. 또한 DGX 인프라는 AI 개발 및 배포를 간소화하도록 최적화된 소프트웨어 제품군인 NVIDIA AI Enterprise를 포함합니다.



리더십급 인프라, 당신의 조건에 맞게

비즈니스를 위한 AI는 성능과 기능 이상의 의미를 지닙니다. 조직의 IT 환경과 관행에 깔끔하게 적합해야 합니다. DGX H200은 직접 관리를 위해 온-프레미스로 설치되거나, **NVIDIA DGX-Ready 데이터 센터**에 코로케이션되거나, **NVIDIA 인증 관리 서비스 제공자**를 통해 액세스할 수 있습니다. **DGX-Ready Lifecycle Management** 프로그램을 통해 조직은 예측 가능한 재무 모델을 얻어 배포를 최첨단 상태로 유지할 수 있습니다. 이를 통해 DGX H200은 기존 IT 인프라만큼 사용하고 획득하기 쉽고, 바쁜 IT 직원에게 추가적인 부담을 주지 않아 조직이 미래를 기다리지 않고 오늘날 비즈니스를 위해 AI를 활용할 수 있습니다.

Ready to Get Started?

NVIDIA DGX H200에 대해 더 알아보고 싶다면, 방문해주세요

nvidia.com/DGX-H200

(주)한국인프라

서울 강남구 삼성로 150, 극동교회 3층
문의전화 02-6204-5000 | 메일 nvidia@krinfra.co.kr www.krinfra.co.kr

 (주)한국인프라
Infra Information Technology Co., Ltd.

