



NVIDIA DGX H100

AI 인프라의 전형이 되는 기준

인공 지능은 까다로운 비즈니스 과제 해결에 있어 최선의 선택이 되었습니다. 거의 모든 산업군에서 고객 서비스 개선, 공급망 최적화, 비즈니스 인텔리전스의 추출 또는 최첨단 상품과 서비스의 설계에 이르기까지 AI는 조직이 혁신을 이루어 낼 수 있는 메커니즘을 제공합니다. AI 인프라의 개척자로서, NVIDIA DGX™ 시스템은 이러한 핵심 아이디어들을 실현할 수 있는 가장 강력하고도 완전한 AI 플랫폼을 제공합니다. NVIDIA DGX H100은 비즈니스 혁신과 최적화의 원동력이 됩니다. NVIDIA의 전설적인 DGX 시스템의 최신작이자 DGX SuperPOD™의 근간인 DGX H100은 획기적인 NVIDIA H100 Tensor Core GPU를 갖춘 AI의 최강자입니다. 이 시스템은 AI 처리량을 극대화하도록 설계되어, 기업이 고도로 정제되고 체계화된 확장된 플랫폼을 활용해 자연어 처리, 추천 시스템, 데이터 분석 및 그 이상의 영역에서 놀라운 성과를 얻을 수 있도록 해줍니다. 온 프렘을 비롯해 다양한 접근 및 구현 옵션을 제공하는 DGX H100은 기업의 최대 과제를 해결하기 위해 필요한 성능을 AI를 통해 제공합니다.

당신의 AI Center of Excellence의 초석

AI는 과학과 비즈니스 사이의 간극을 메꾸어 왔습니다. 더 이상 실험의 영역에 머무르지 않고, AI는 중소기업과 대기업 모두에게 혁신을 이루고 비즈니스 최적화를 달성하기 위한 용도로 활용되고 있습니다. 세계 최초로 개발된 특수 목적 AI 인프라의 4세대로서 DGX H100은 기업 AI center of excellence의 중심이 되기 위해 설계되었습니다. 이는 완전한 최적화를 마친 하드웨어와 소프트웨어 플랫폼으로, 새롭게 출시된 NVIDIA AI 소프트웨어 솔루션에 대한 전적인 지원, 제 3자 지원의 풍부한 생태계, NVIDIA 프로페셔널 서비스의 전문가 조언을 포함합니다. 전세계 거의 모든 산업군에서 수 많은 고객이 이미 사용하고 있는 DGX 시스템과 함께, DGX H100은 검증된 신뢰성을 제공합니다.

대규모 AI 구현의 장애물을 넘어서

세계 최초 NVIDIA H100 Tensor Core GPU 기반 시스템인 NVIDIA DGX H100은 AI 규모와 성능의 한계를 넘어섰습니다. 6배 높은 성능, NVIDIA ConnectX®-7 smart network interface cards (SmartNICs)와 NVIDIA BlueField®-3 데이터 처리 장치를 활용한 2배 빠른 네트워킹, NVIDIA DGX SuperPOD를 위한 고속 확장성을 자랑합니다. 차세대 아키텍처는 자연어 처리와 딥 러닝 추천 모델과 같은 가장 크고 가장 복잡한 AI 업무를 위해 더욱 더 강력해 졌습니다.

사양

GPU	8x NVIDIA H100 Tensor Core GPUs
GPU 메모리	총 640GB
성능	32 petaFLOPS FP8
NVIDIA® NVSwitch™	4x
시스템 전력 사용량	~10.2kW max
CPU	Dual x86
시스템 메모리	2TB
네트워킹	4x OSFP 포트로 8x single-port NVIDIA ConnectX-7 지원 400Gb/s InfiniBand/Ethernet 2x dual-port NVIDIA BlueField-3 DPUs VPI 1x 400Gb/s InfiniBand/Ethernet 1x 200Gb/s InfiniBand/Ethernet
관리 네트워크	10Gb/s onboard NIC with RJ45 50Gb/s Ethernet optional NIC Host baseboard management controller (BMC) with RJ45 2x NVIDIA BlueField-3 DPU BMC (with RJ45 each)
스토리지	OS: 2x 1.9TB NVMe M.2 내부 스토리지: 8x 3.84TB NVMe U.2
시스템 소프트웨어	DGX H100 시스템은 Ubuntu Linux 기반의 DGX OS가 사전 설치되며, DGX 소프트웨어 스택을 포함(모든 필요한 패키지와 드라이버는 DGX에 최적화). 고객이 Ubuntu Linux 나 Red Hat Enterprise Linux를 설치하고, 필요한 DGX 소프트웨어 스택을 별도로 설치하는 것도 가능.
운전온도범위	5-30°C (41-86°F)

리더십 수준의 인프라를 원하는 조건으로

비즈니스에서 AI는 성능과 기능만 의미하지 않습니다. 조직의 IT 환경과 활동에 잘 맞아 들어가야 합니다. DGX H100은 직접 관리를 위한 온 프렘 설치, NVIDIA DGX 실행이 가능한 데이터 센터에 공동설치, NVIDIA DGX Foundry에서 대여 및 NVIDIA 인증 관리형 사업자를 통해서도 사용이 가능합니다.

DGX를 반영한 생애주기 관리 프로그램을 통해, 조직은 업계 최신의 구현수준을 유지할 수 있는 예측 가능한 재무 모델을 가지게 됩니다. 이를 통해 DGX H100을 사용하고 획득하는 것이 전통적 IT 인프라만큼 쉬워집니다. 바쁜 IT 직원들에게 부담을 줄 필요가 없으므로 나중까지 기다리는 대신 오늘의 비즈니스에 AI를 활용할 수 있게 됩니다.

주문 정보

NVIDIA DGX H100은 2022년 하반기에 출시됩니다.
더 궁금하신 내용은 BNI&C 영업대표에게 문의하세요.

자세히 알아보세요

NVIDIA DGX H100에 대해 궁금하신 내용은 www.bninc.co.kr/nvidia/dgx.php를 방문하세요.

비엔아이엔씨 www.bninc.co.kr

서울시 영등포구 당산로 41길 11 당산 SK V1센터 W동 1208호 tel : (02)6121-8308 e-mail : sales@bninc.co.kr