



# NVIDIA DGX SuperPOD

AI 엔터프라이즈를 위한 터키 데이터 센터 솔루션

NVIDIA DGX SuperPOD™는 선도적인 인프라와 함께 까다로운 AI 및 고성능 컴퓨팅(HPC) 워크로드를 위한 민첩하고 확장 가능한 성능을 제공합니다. DGX SuperPOD를 사용하면 몇 달이 아닌 불과 몇 주 내에 업계에서 검증된 결과를 포함한 풀 서비스를 경험할 수 있습니다. DGX SuperPOD는 단순한 하드웨어 컬렉션을 넘어 풀스택 데이터 센터 플랫폼으로서 협업 및 대규모 성능 극대화에 최적화된 업계 선도적 컴퓨팅, 스토리지, 네트워킹, 소프트웨어 및 인프라 관리를 포함합니다. 또한 원활한 배포 및 작동을 보장하는 화이트 글러브 구현 서비스를 통해 간편한 사용을 지원합니다.

## 대규모 다중 노드 AI 인프라의 어려움 해소

DGX 플랫폼의 일부인 NVIDIA DGX SuperPOD는 독보적인 수준의 다중 노드 훈련을 제공하도록 설계되었습니다. 전통적인 대규모 컴퓨팅 클러스터는 구성의 크기가 커지고 계속해서 증가하는 노드 전반에서 컴퓨팅이 병렬화되면서 발생하는 GPU 간 통신 확장의 복잡성에 제약을 받습니다. 이는 성능 출력 저하를 야기합니다. DGX SuperPOD는 이러한 확장 문제를 해결하기 위해 다중 노드 AI 인프라의 고유한 요구에 따라 시스템 내 모든 구성 요소의 최적화를 실시합니다. 세계에서 가장 빠른 슈퍼컴퓨터 중 하나인 NVIDIA의 Eos와 DGX SuperPOD 아키텍처에 기반한 기타 클러스터는 다년간 TOP500 및 Green500 목록<sup>1</sup>의 상위권에 이름을 올리는 한편 MLPerf 벤치마크 기록을 수립하고 있습니다.<sup>2</sup>

## NVIDIA Base Command의 지원

NVIDIA Base Command는 DGX 플랫폼을 지원하여 조직이 NVIDIA가 제공하는 최상의 소프트웨어 혁신을 활용할 수 있도록 합니다. 엔터프라이즈는 엔터프라이즈급 오케스트레이션 및 클러스터 관리, 컴퓨팅, 스토리지 및 네트워크 인프라를 가속화하는 라이브러리, AI 워크로드에 최적화된 운영 체제를 포함하는 입증된 플랫폼을 이용해 DGX 인프라에 잠재된 가능성을 최대한 구현할 수 있습니다. 또한 DGX 인프라는 AI 개발 및 배포 간소화를 위해 최적화된 소프트웨어 제품군인 NVIDIA AI Enterprise를 포함하며, AI 워크플로 및 MLOps 도구와 통합됩니다.

## NVIDIA DGX SuperPOD

### 하드웨어

- > NVIDIA DGX™ H100 또는 DGX A100 시스템
- > NVIDIA 네트워킹
- > 고성능 스토리지

### 소프트웨어

- > NVIDIA AI Enterprise
- > NVIDIA Base Command

### 라이프사이클 서비스\*

- > 계획/배포\*\*
  - 용량 계획
  - 데이터 센터 설계
  - 성능 예측
  - 사이트 평가/준비
  - 설치
  - 설치 후 테스트
  - 프로비저닝/관리
- > 훈련/최적화
  - 애플리케이션 성능 테스트
  - 사이트 문서 패키지
  - 사용자/DevOps 훈련
  - 워크로드에 기반한 NVIDIA 딥러닝 인스티튜트 훈련
  - 맞춤형 시스템 실행서
  - 인수인계 세션

\* NVIDIA 및 파트너 서비스 조합

\*\* 온프레미스 또는 DGX 레디 데이터 센터 배포

## 테스트 및 검증된 솔루션, NVIDIA DGX SuperPOD

DGX SuperPOD는 NVIDIA 방식으로 수행된 AI 인프라이자 엔터프라이즈의 성능 및 안정성 요구 사항을 충족하는 예측 가능한 솔루션입니다. NVIDIA에서 DGX SuperPOD를 광범위하게 테스트하고 실제 엔터프라이즈 AI 워크로드로 최대한도까지 밀어붙이는 모든 작업을 책임지기 때문에 애플리케이션 성능을 걱정할 필요가 없습니다.

## NVIDIA의 지원으로 전체 라이프사이클 동안 제공되는 전문 지식

엔터프라이즈는 아키텍처 설계뿐만 아니라 비즈니스 운영에 도움이 되는 가속 컴퓨팅 인프라를 구축할 수 있는 지름길, 즉 내부 IT 환경을 중심으로 최적화된 신속한 터키 구현 경험이 필요합니다. 이를 통해 데이터 사이언티스트는 지연 없이 바로 작업을 시작해 지속적인 개선을 구현할 수 있습니다. NVIDIA DGX SuperPOD는 설치에서 인프라 관리, 워크로드 확장, 프로덕션 AI 간소화에 이르는 전체 라이프사이클 기간 동안 엔터프라이즈를 위한 전문 서비스를 제공합니다. DGX SuperPOD는 또한 계속해서 업그레이드되고 있습니다. NVIDIA 엔지니어들은 DGX SuperPOD를 지원하는 소프트웨어의 지속적인 혁신 및 업데이트를 통해 모든 시스템이 시운전 대비 빠른 속도로 구동되도록 합니다.

### NVIDIA DGX – AI 소프트웨어 스택



## AI를 위해 최적화된 단일 솔루션 내 고성능 인프라

NVIDIA DGX SuperPOD는 AI 컴퓨팅, 네트워크 패브릭, 스토리지 및 소프트웨어로 구성된 최적화된 설계 조합을 제공합니다. NVIDIA DGX H100 또는 DGX A100 시스템에 구축된 컴퓨팅 기반을 통해 전례 없는 컴퓨팅 밀도, 성능 및 유연성을 경험할 수 있습니다. 세계 최고 수준의 가속기가 탑재된 NVIDIA DGX H100 및 DGX A100 시스템은 엔터프라이즈가 훈련, 추론 및 분석을 배포가 쉬운 통합 AI 인프라에 결합할 수 있도록 합니다.

DGX SuperPOD의 고성능 네트워크 패브릭은 극도로 낮은 레이턴시가 특징인 NVIDIA InfiniBand 네트워킹을 활용합니다. 이 강력한 기술은 낮은 운영 비용과 인프라 복잡성을 바탕으로 대규모 AI 워크로드를 위한 우수한 성능과 확장성을 제공합니다.

AI 슈퍼컴퓨터를 최대 용량으로 실행하기 위해서는 초고속 스토리지가 필요합니다. 제대로 설계된 시스템의 경우, 스토리지 솔루션은 텍스트, 표 형식 데이터, 오디오, 비디오와 같은 다양한 데이터 유형을 일관된 성능으로 병렬 처리할 수 있어야 합니다. NVIDIA DGX SuperPOD용으로 인증된 스토리지는 AI 워크로드의 고유한 요구 사항에 따른 선정 및 테스트를 거친 후 각 환경에 최적화되어 성공적인 결과를 보장합니다.

## AI 성공을 촉진하는 경험

수천 명에 이르는 NVIDIA 연구원 및 엔지니어들이 DGX SuperPOD를 활용해 시장에 새로운 혁신 제품을 선보이고 있습니다. DGX SuperPOD는 이를 바탕으로 한 NVIDIA의 독보적인 AI 슈퍼컴퓨터 설계 및 사용 경험을 통합합니다. NVIDIA DGX SuperPOD는 인프라가 아닌 혁신에 초점을 맞춰 NVIDIA의 AI 수행 방식으로 설계, 배포 및 관리되는 엔터프라이즈용 터키 데이터 센터 솔루션을 제공합니다.

## 시작할 준비가 되셨나요?

NVIDIA DGX SuperPOD에 대한 추가 정보는 [www.nvidia.com/dgx-superpod](https://www.nvidia.com/dgx-superpod)에서 확인할 수 있습니다.

1. 자세한 정보는 [top500.org](https://top500.org) 참조
2. 더 많은 내용은 [mlperf.org](https://mlperf.org) 참조

© 2024 NVIDIA Corporation. All rights reserved. NVIDIA, NVIDIA 로고, Base Command, DGX, DGX SuperPOD 및 NGC는 미국 및 기타 국가에서 NVIDIA Corporation의 상표 및/또는 등록 상표입니다. 다른 회사 및 제품 이름은 연관된 각 회사의 상표일 수 있습니다. 기타 모든 상표는 해당 소유자의 자산입니다. 3164752. FEB24

