

NVIDIA DGX SUPERPOD 및 NETAPP ONTAP 스토리지



규모에 따라 AI 혁신 가속



조직에서 인공 지능(AI) 및 머신 러닝(ML) 이니셔티브를 수용함에 따라 강력하고 확장 가능하며 효율적인 인프라 솔루션에 대한 수요가 그 어느 때보다 높아졌습니다. AI CoE(AI Center of Excellence), 즉 AI Factory는 엔터프라이즈 AI 전략의 중요한 구성 요소로 부상하면서 AI 솔루션 개발 및 배포를 위한 중앙 집중식 허브 역할을 하고 있습니다. 이러한 이니셔티브의 핵심은 데이터 보안, 접근성 및 리소스 최적화를 보장하면서 점점 복잡해지는 AI 모델을 관리 및 훈련하는 데 있습니다.

사람의 입력 없이도 특정 목표를 달성할 수 있도록 의사 결정, 계획, 적응할 수 있는 AI의 진화에 따라 컴퓨팅 및 스토리지 인프라에 대한 전례 없는 수요가 생겼습니다. 조직에서는 이제 대규모 데이터 세트를 처리하고, 여러 동시 훈련 워크로드를 지원하고, 고성능 컴퓨팅 환경을 유지하면서 데이터를 보호하고 규정을 준수해야 합니다.

기존 인프라 솔루션은 이러한 요구사항을 충족하는 데 어려움을 겪고 있으므로, 운영 효율성이 떨어지며 AI 프로젝트의 가치 실현 시간이 지연됩니다.

해결책

NVIDIA DGX SuperPOD™ 및 NetApp® ONTAP® 스토리지는 NVIDIA DGX™ 시스템의 컴퓨팅 성능과 클라우드 연결 NetApp AFF A90 스토리지 시스템의 세계적인 엔터프라이즈 AI 기능을 결합하여 ML, AI 및 고성능 기술 컴퓨팅(HPC)을 위한 데이터 중심 워크플로를 지원합니다. 이 통합 솔루션은 AI 개발 및 배포를 위한 확장 가능하고 안전하며 효율적인 플랫폼을 조직에 제공합니다. 이 아키텍처는 기존의 인프라 사일로를 제거하여 데이터 과학자, 엔지니어 및 기타 이해 관계자가 원활하게 협업할 수 있습니다.

데이터 관리 및 액세스

NVIDIA DGX SuperPOD™ 및 NetApp ONTAP 스토리지는 고급 데이터 축소 기술을 구현하여 스토리지 활용률을 최적화하고 성능을 유지합니다. ONTAP의 NetApp Snapshot™ 및 FlexClone® 기능을 통해 팀은 병렬 개발 및 테스트를 위한 공간 효율적인 데이터 세트 복사본을 즉시 생성할 수 있습니다. 추가 스토리지 공간을 소비하거나 데이터 무결성을 훼손하지 않고 여러 AI 팀이 동일한 데이터 세트의 다양한 버전에서 동시에 작업할 수 있습니다.

또한 IT 팀은 솔루션의 통합 데이터 액세스 프레임워크를 활용하여 기존 워크플로와 툴을 쉽게 통합할 수 있습니다. 데이터 과학자와 엔지니어는 친숙한 프로토콜과 인터페이스를 활용하면서도 엔터프라이즈급 데이터 보호 및 관리 기능의 이점을 활용할 수 있습니다. 플랫폼의 지능형 데이터 계층화를 통해 스토리지 계층 전반에 걸쳐 데이터 배치를 자동으로 최적화하므로, 자주 액세스하는 데이터 세트의 가용성을 즉시 유지하고 사용 빈도가 낮은 데이터를 더욱 비용 효율적인 스토리지로 이동할 수 있습니다. ONTAP의 글로벌 네임스페이스 기능과 결합된 이 지능적인 데이터 관리 접근 방식은 AI 팀에 진정한 통합 효율적인 데이터 액세스 경험을 제공합니다.

확장성

NVIDIA DGX SuperPOD 및 NetApp ONTAP 스토리지는 리소스를 자동으로 조정하는 지능형 워크로드 오키스트레이션을 제공하여 동시 AI 훈련 작업 및 추론 워크로드에 대한 성능을 최적화합니다. 중복제거 및 압축과 같은 ONTAP 스토리지 효율성 기술은 데이터가 커짐에 따라 스토리지 용량을 극대화합니다.

조직은 인프라 사일로를 제거하고 AI Centers of Excellence 내에서 리소스 관리를 최적화하여 중복 시스템을 제거하고 운영 오버헤드를 줄일 수 있습니다. 이 통합 인프라는 AI 프로젝트 일정을 앞당기고 기술 투자 수익률을 개선합니다.

주요 이점

통합 데이터 관리

AI 데이터 라이프사이클에 대한 포괄적인 제어를 제공하는 중앙 집중식 플랫폼

데이터 액세스 가속화

조직 전체의 데이터에 빠르고 안전하게 액세스

원활한 인프라 확장

개념 증명부터 전사적 배포까지 필요에 따라 확장할 수 있는 통합 플랫폼

포괄적인 데이터 보호

전체 AI 라이프사이클에서 중요 정보를 보호하는 엔터프라이즈급 데이터 보호 기능

제로 트러스트 보안

모든 계층에서 AI 인프라를 보호하는 제로 트러스트 아키텍처를 사용한 보안

조직은 DGX 컴퓨팅 노드와 ONTAP 스토리지 시스템을 운영 중단 없이 추가하여 원활하게 확장할 수 있습니다. 이 아키텍처는 컴퓨팅 성능, 메모리 및 스토리지에 대한 스케일업 및 스케일아웃 방식을 모두 지원합니다. 유니파이드 패브릭은 모든 규모에서 고성능 및 낮은 지연 시간을 유지하여 AI 워크로드의 효율성을 최적화합니다.

보안

조직에서는 RBAC(역할 기반 액세스 제어), 다단계 인증, 감사 로깅 등 NetApp의 강력한 보안 수단을 사용하여 지적 재산을 보호하고 규정 준수를 유지하는 동시에 AI 이니셔티브를 안심하고 추진할 수 있습니다. 승인받은 직원과 AI 에이전트만이 중요한 모델, 훈련 데이터 및 엔터프라이즈 데이터에 액세스할 수 있습니다.



이 솔루션은 사용되지 않는 데이터와 이동 중인 데이터를 포괄적으로 암호화하여 보호하는 동시에 실시간 위협 감지 및 자동 대응을 위한 통합 보안 모니터링을 제공합니다. 내장 스냅샷 및 백업 기능을 사용하여 보안 사고 발생 시 신속하게 복구할 수 있어 다운타임 및 데이터 손실을 최소화할 수 있습니다.

이 강력한 보안 기반을 통해 기업은 뛰어난 데이터 보호를 유지하면서 더 빠르게 혁신하고, 파트너와 더 효과적으로 협업하고, AI 운영을 확장할 수 있습니다.

결론

NetApp ONTAP 스토리지 솔루션이 포함된 NVIDIA DGX SuperPOD는 AI 인프라 솔루션의 상당한 발전을 나타냅니다. IT 조직은 보안, 데이터 관리, 리소스 사용, 확장성과 관련된 주요 과제를 해결함으로써 운영 효율성, 데이터 보호 및 협업을 유지하면서 AI 이니셔티브를 가속할 수 있습니다. 이 솔루션의 통합 접근 방식은 AI 개발 파이프라인에서 일반적인 병목 현상을 제거하여 데이터 과학자와 엔지니어가 인프라 관리가 아닌 혁신에 집중할 수 있도록 지원합니다. AI가 산업 전반에 걸쳐 계속 발전하고 디지털 전환을 촉진함에 따라 강력하고 확장 가능하며 안전한 인프라 솔루션의 필요성은 갈수록 커질 것입니다. NetApp ONTAP 스토리지가 포함된 NVIDIA DGX SuperPOD는 현재 및 미래의 AI 이니셔티브를 위한 기반을 제공하여 조직이 점점 더 증가하는 AI 기반 환경에서 경쟁 우위를 유지할 수 있도록 지원합니다.



문의하기

NetApp 정보

NetApp은 유니파이드 데이터 스토리지, 통합된 데이터 서비스, CloudOps 솔루션을 결합하여 격변하는 세상에서 모든 고객에게 기회를 제공하는 지능형 데이터 인프라 회사입니다. NetApp은 사일로가 없는 인프라를 만들고, 관찰 가능성과 AI를 활용하여 업계 최고 수준의 데이터 관리를 지원합니다. 세계 최대 규모의 클라우드에 기본적으로 내장된 유일한 엔터프라이즈급 스토리지 서비스인 NetApp의 데이터 스토리지는 원활한 유연성을 제공합니다. 또한, NetApp의 데이터 서비스는 우수한 사이버 복원력, 거버넌스, 애플리케이션 민첩성을 통해 데이터 우위를 만듭니다. NetApp의 CloudOps 솔루션은 관찰 가능성과 AI를 통해 성능과 효율성의 지속적인 최적화를 제공합니다. NetApp과 함께라면 데이터 유형, 워크로드, 환경과 관계없이 데이터 인프라를 혁신하여 비즈니스의 가능성을 실현할 수 있습니다.

www.netapp.com/ko