



NVIDIA DGX 시스템을 탑재한 **NetApp AI**Pod NetApp artificial intelligence solutions

NetApp
August 18, 2025

목차

NVIDIA DGX 시스템을 탑재한 NetApp AIPOd	1
NVIDIA DGX 시스템을 탑재한 NVA-1173 NetApp AIPOd - 소개	1
요약	1
NVIDIA DGX 시스템을 탑재한 NVA-1173 NetApp AIPOd - 하드웨어 구성 요소	2
NetApp AFF 스토리지 시스템	2
NVIDIA DGX BasePOD	3
NVIDIA DGX 시스템을 탑재한 NVA-1173 NetApp AIPOd - 소프트웨어 구성 요소	5
NVIDIA 소프트웨어	6
NetApp 소프트웨어	7
NVIDIA DGX H100 시스템을 탑재한 NVA-1173 NetApp AIPOd - 솔루션 아키텍처	9
DGX 시스템을 갖춘 NetApp AIPOd	9
네트워크 디자인	10
DGX H100 시스템의 스토리지 액세스 개요	11
저장 시스템 설계	11
관리 플레인 서버	12
NVIDIA DGX 시스템을 탑재한 NVA-1173 NetApp AIPOd - 배포 세부 정보	12
스토리지 네트워크 구성	14
스토리지 시스템 구성	15
NVIDIA DGX 시스템을 탑재한 NVA-1173 NetApp AIPOd - 솔루션 검증 및 크기 조정 지침	20
솔루션 검증	20
스토리지 시스템 크기 지침	21
NVIDIA DGX 시스템을 탑재한 NVA-1173 NetApp AIPOd - 결론 및 추가 정보	21
결론	21
추가 정보	21
감사의 말	22

NVIDIA DGX 시스템을 탑재한 NetApp AI Pod

NVIDIA DGX 시스템을 탑재한 NVA-1173 NetApp AI Pod - 소개

POWERED BY



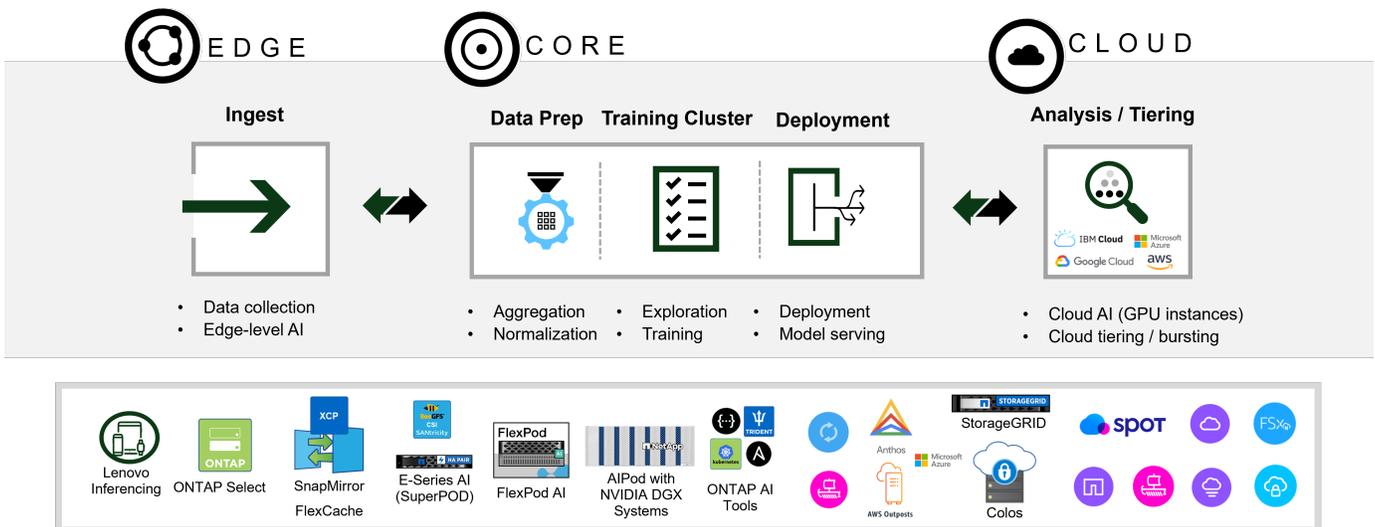
NVIDIA®

NetApp 솔루션 엔지니어링

요약

NVIDIA DGX™ 시스템과 NetApp 클라우드 연결 스토리지 시스템을 탑재한 NetApp™ AI Pod 설계 복잡성과 추적을 제거하여 머신 러닝(ML) 및 인공지능(AI) 워크로드를 위한 인프라 배포를 간소화합니다. 차세대 워크로드에 탁월한 컴퓨팅 성능을 제공하도록 설계된 NVIDIA DGX BasePOD™를 기반으로 하는 NVIDIA DGX 시스템을 탑재한 AI Pod는 NetApp AFF 스토리지 시스템을 추가하여 고객이 소규모로 시작하여 중단 없이 확장할 수 있도록 하는 동시에 엣지에서 코어, 클라우드로 데이터를 지능적으로 관리할 수 있도록 합니다. NetApp AI Pod 아래 그림에 표시된 대로 NetApp AI 솔루션의 더 큰 포트폴리오의 일부입니다.

NetApp AI 솔루션 포트폴리오



이 문서에서는 AI Pod 참조 아키텍처의 주요 구성 요소, 시스템 연결 및 구성 정보, 검증 테스트 결과, 솔루션 크기 조정 지침을 설명합니다. 이 문서는 ML/DL 및 분석 워크로드를 위한 고성능 인프라를 구축하는 데 관심이 있는 NetApp 및 파트너 솔루션 엔지니어와 고객 전략적 의사 결정자를 대상으로 합니다.

NVIDIA DGX 시스템을 탑재한 NVA-1173 NetApp AI Pod - 하드웨어 구성 요소

이 섹션에서는 NVIDIA DGX 시스템이 탑재된 NetApp AI Pod 의 하드웨어 구성 요소에 대해 중점적으로 설명합니다.

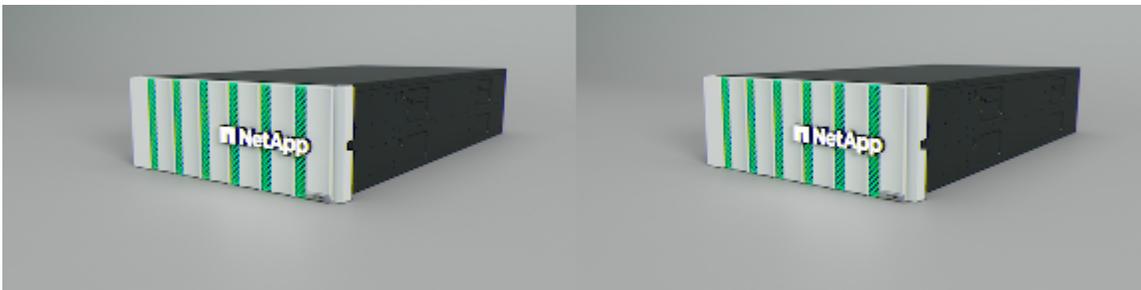
NetApp AFF 스토리지 시스템

NetApp AFF 최첨단 스토리지 시스템을 사용하면 IT 부서가 업계 최고의 성능, 뛰어난 유연성, 클라우드 통합, 동급 최고의 데이터 관리 기능을 통해 엔터프라이즈 스토리지 요구 사항을 충족할 수 있습니다. 플래시에 맞춰 특별히 설계된 AFF 시스템은 비즈니스에 중요한 데이터를 가속화하고 관리하며 보호하는 데 도움이 됩니다.

AFF A90 저장 시스템

NetApp ONTAP 데이터 관리 소프트웨어 기반의 NetApp AFF A90 내장형 데이터 보호 기능, 옵션으로 제공되는 랜섬웨어 방지 기능, 가장 중요한 비즈니스 워크로드를 지원하는 데 필요한 높은 성능과 복원력을 제공합니다. 이는 임무 수행에 중요한 운영의 중단을 제거하고, 성능 조정을 최소화하며, 랜섬웨어 공격으로부터 데이터를 보호합니다. 다음을 제공합니다. • 업계 최고의 성능 • 타협 없는 데이터 보안 • 간소화된 중단 없는 업그레이드

NetApp AFF A90 스토리지 시스템



업계 최고의 성능

AFF A90 딥 러닝, AI, 고속 분석과 같은 차세대 워크로드는 물론 Oracle, SAP HANA, Microsoft SQL Server, 가상화된 애플리케이션과 같은 기존 엔터프라이즈 데이터베이스도 쉽게 관리합니다. 이 솔루션은 HA 쌍당 최대 240만 IOPS와 최저 100µs의 대기 시간을 제공하여 비즈니스에 중요한 애플리케이션을 최고 속도로 실행하고 이전 NetApp 모델보다 성능을 최대 50%까지 향상시킵니다. RDMA를 통한 NFS, pNFS 및 세션 트렁킹을 통해 고객은 기존 데이터 센터 네트워킹 인프라를 사용하여 차세대 애플리케이션에 필요한 높은 수준의 네트워크 성능을 달성할 수 있습니다. 고객은 SAN, NAS 및 개체 스토리지에 대한 통합된 다중 프로토콜 지원을 통해 확장하고 성장할 수 있으며, 온프레미스 또는 클라우드의 데이터에 대한 통합된 단일 ONTAP 데이터 관리 소프트웨어를 통해 최대한의 유연성을 제공할 수 있습니다. 또한 Active IQ 와 Cloud Insights 가 제공하는 AI 기반 예측 분석을 통해 시스템 상태를 최적화할 수 있습니다.

타협 없는 데이터 보안

AFF A90 시스템에는 NetApp 통합 및 애플리케이션 일관성을 갖춘 전체 데이터 보호 소프트웨어가 포함되어 있습니다. 이 솔루션은 사전 예방 및 공격 후 복구를 위한 내장형 데이터 보호 기능과 최첨단 랜섬웨어 방지 솔루션을 제공합니다. 악성 파일이 디스크에 기록되는 것을 차단할 수 있으며, 저장소 이상을 쉽게 모니터링하여 통찰력을 얻을 수 있습니다.

간소화된 무중단 업그레이드

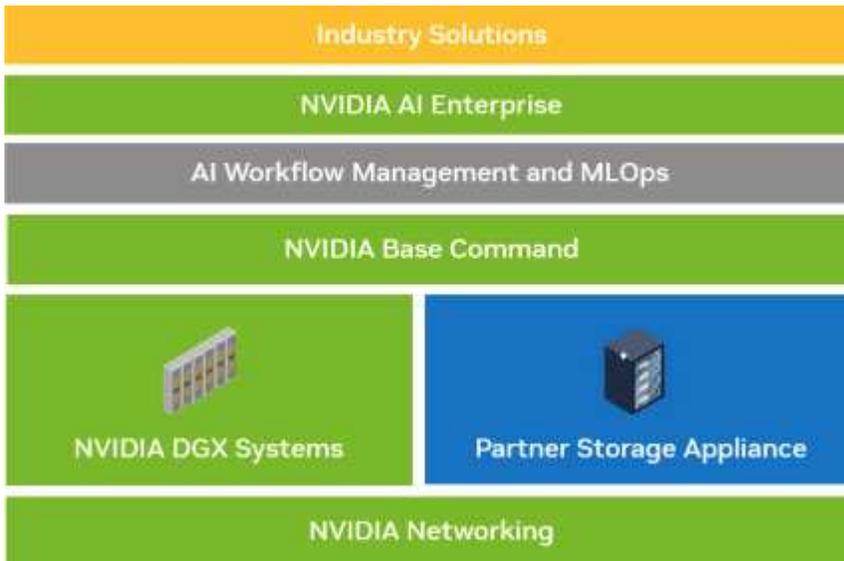
AFF A90 기존 A800 고객에게 중단 없는 새시 내부 업그레이드로 제공됩니다. NetApp 고급 안정성, 가용성, 서비스 용이성 및 관리 용이성(RASM) 기능을 통해 임무 수행에 중요한 운영을 간편하게 갱신하고 중단을 제거할 수 있도록 지원합니다. 또한 NetApp ONTAP 소프트웨어가 모든 시스템 구성 요소에 대한 펌웨어 업데이트를 자동으로 적용하므로 운영 효율성을 더욱 높이고 IT 팀의 일상 업무를 간소화합니다.

가장 큰 규모의 배포의 경우, AFF A1K 시스템은 가장 높은 성능과 용량 옵션을 제공하는 반면, AFF A70, AFF C800 과 같은 다른 NetApp 스토리지 시스템은 더 낮은 비용으로 더 작은 규모의 배포를 위한 옵션을 제공합니다.

NVIDIA DGX BasePOD

NVIDIA DGX BasePOD NVIDIA 하드웨어 및 소프트웨어 구성 요소, MLOps 솔루션, 타사 스토리지로 구성된 통합 솔루션입니다. NVIDIA 제품과 검증된 파트너 솔루션을 통해 확장형 시스템 설계의 모범 사례를 활용함으로써 고객은 AI 개발을 위한 효율적이고 관리하기 쉬운 플랫폼을 구현할 수 있습니다. 그림 1은 NVIDIA DGX BasePOD 의 다양한 구성 요소를 강조하여 보여줍니다.

NVIDIA DGX BasePOD 솔루션



NVIDIA DGX H100 시스템

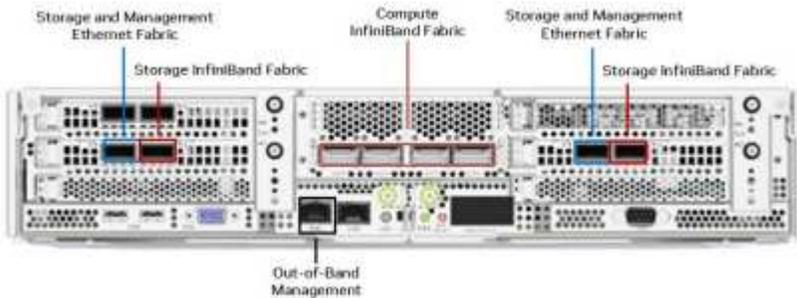
NVIDIA DGX H100™ 시스템은 NVIDIA H100 Tensor Core GPU의 획기적인 성능으로 가속화된 AI 강자입니다.

NVIDIA DGX H100 시스템



DGX H100 시스템의 주요 사양은 다음과 같습니다. • 8개의 NVIDIA H100 GPU. • GPU당 80GB GPU 메모리, 총 640GB. • NVIDIA NVSwitch 칩 4개. • PCIe 5.0을 지원하는 듀얼 56코어 Intel Xeon Platinum 8480 프로세서. • 2TB DDR5 시스템 메모리. • 8개의 싱글 포트 NVIDIA ConnectX@-7(InfiniBand/Ethernet) 어댑터와 2개의 듀얼 포트 NVIDIA ConnectX-7(InfiniBand/Ethernet) 어댑터를 제공하는 4개의 OSFP 포트. • DGX OS용 1.92TB M.2 NVMe 드라이브 2개, 스토리지/캐시용 3.84TB U.2 NVMe 드라이브 8개. • 최대 전력 10.2kW. DGX H100 CPU 트레이의 후면 포트는 아래와 같습니다. OSFP 포트 4개는 InfiniBand 컴퓨팅 패브릭을 위한 8개의 ConnectX-7 어댑터를 지원합니다. 각 듀얼 포트 ConnectX-7 어댑터 쌍은 스토리지 및 관리 패브릭에 대한 병렬 경로를 제공합니다. 대역 외 포트는 BMC 액세스에 사용됩니다.

NVIDIA DGX H100 후면 패널



NVIDIA 네트워킹

NVIDIA Quantum-2 QM9700 스위치

NVIDIA Quantum-2 QM9700 InfiniBand 스위치



400Gb/s InfiniBand 연결 기능을 갖춘 NVIDIA Quantum-2 QM9700 스위치는 NVIDIA Quantum-2 InfiniBand BasePOD 구성의 컴퓨팅 패브릭에 전원을 공급합니다. ConnectX-7 단일 포트 어댑터는 InfiniBand 컴퓨팅 패브릭에 사용됩니다. 각 NVIDIA DGX 시스템은 각 QM9700 스위치에 이중으로 연결되어 있어 시스템 간에 여러 개의 고대역폭, 저지연 경로를 제공합니다.

NVIDIA Spectrum-3 SN4600 스위치

NVIDIA Spectrum-3 SN4600 스위치



NVIDIA Spectrum™-3 SN4600 스위치는 총 128개의 포트(스위치당 64개)를 제공하여 DGX BasePOD의 대역 내 관리를 위한 중복 연결을 제공합니다. NVIDIA SN4600 스위치는 1GbE~200GbE 사이의 속도를 제공할 수 있습니다. 이더넷을 통해 연결된 스토리지 어플라이언스의 경우 NVIDIA SN4600 스위치도 사용됩니다. NVIDIA DGX 듀얼 포트 ConnectX-7 어댑터의 포트는 대역 내 관리와 스토리지 연결에 모두 사용됩니다.

NVIDIA Spectrum SN2201 스위치

NVIDIA Spectrum SN2201 스위치



NVIDIA Spectrum SN2201 스위치는 대역 외 관리를 위한 연결을 제공하기 위해 48개의 포트를 제공합니다. 대역 외 관리 기능은 DGX BasePOD의 모든 구성 요소에 대한 통합 관리 연결을 제공합니다.

NVIDIA ConnectX-7 어댑터

NVIDIA ConnectX-7 어댑터



NVIDIA ConnectX-7 어댑터는 25/50/100/200/400G의 처리량을 제공할 수 있습니다. NVIDIA DGX 시스템은 단일 및 이중 포트 ConnectX-7 어댑터를 모두 사용하여 400Gb/s InfiniBand 및 이더넷을 통한 DGX BasePOD 배포에 유연성을 제공합니다.

NVIDIA DGX 시스템을 탑재한 NVA-1173 NetApp AIPOD - 소프트웨어 구성 요소

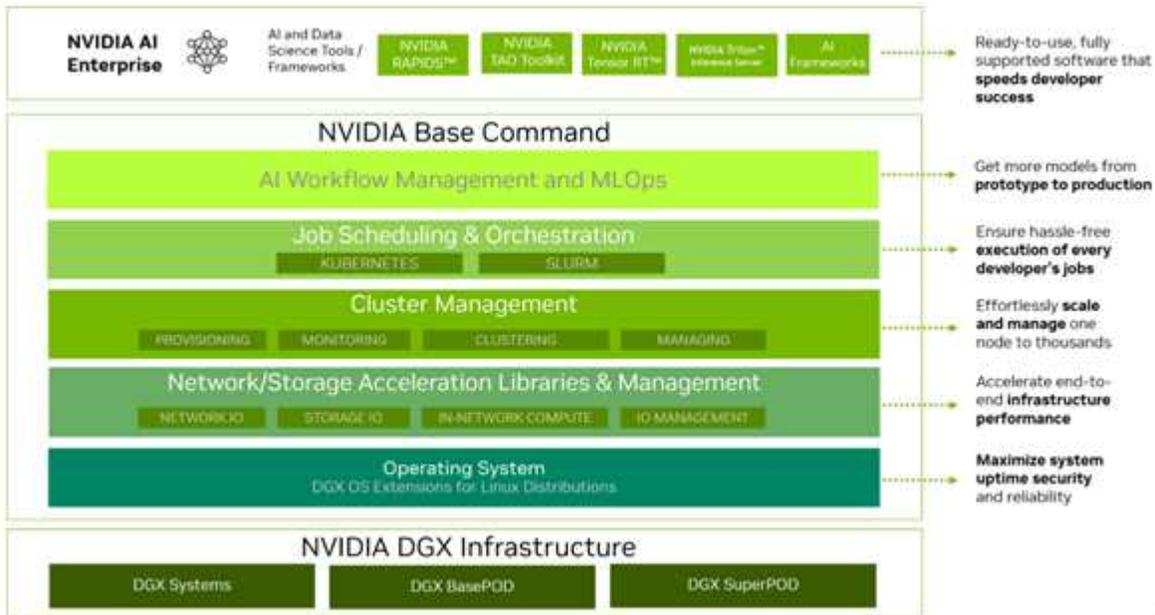
이 섹션에서는 NVIDIA DGX 시스템이 탑재된 NetApp AIPOD의 소프트웨어 구성 요소에 대해 중점적으로 설명합니다.

NVIDIA 소프트웨어

NVIDIA 베이스 명령

NVIDIA Base Command™는 모든 DGX BasePOD를 구동하여 조직이 NVIDIA 소프트웨어 혁신의 장점을 최대한 활용할 수 있도록 지원합니다. 기업은 엔터프라이즈급 오케스트레이션 및 클러스터 관리, 컴퓨팅, 스토리지 및 네트워크 인프라를 가속화하는 라이브러리, AI 워크로드에 최적화된 운영 체제(OS)를 포함하는 검증된 플랫폼을 통해 투자의 잠재력을 최대한 발휘할 수 있습니다.

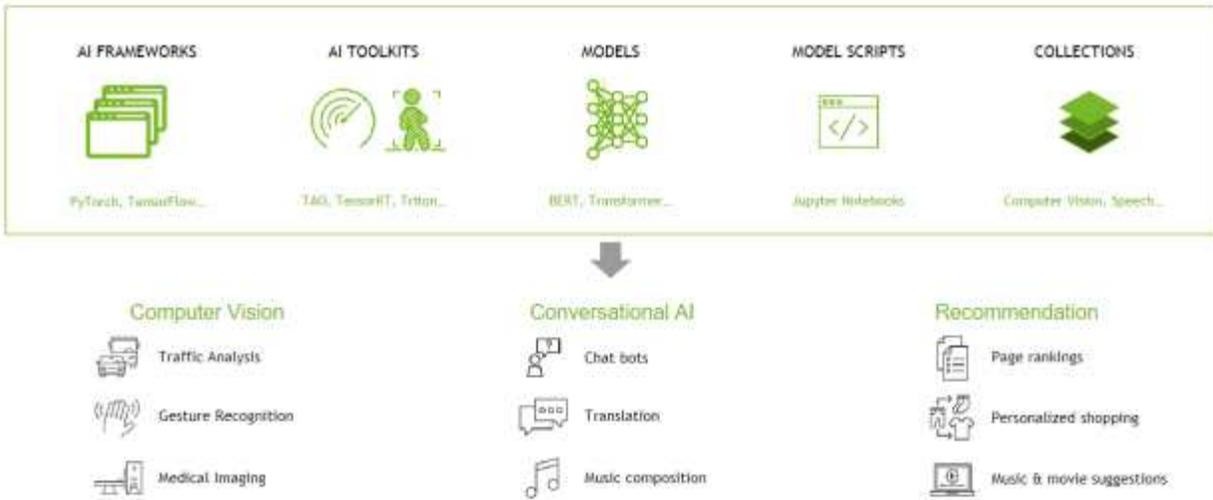
NVIDIA BaseCommand 솔루션



NVIDIA GPU 클라우드(NGC)

NVIDIA NGC는 다양한 수준의 AI 전문 지식을 갖춘 데이터 과학자, 개발자, 연구자의 요구를 충족하는 소프트웨어를 제공합니다. NGC에 호스팅된 소프트웨어는 공통적인 취약성 및 노출(CVE), 암호 및 개인 키의 집계된 집합에 대한 검사를 거칩니다. DGX 시스템은 여러 GPU에 확장 가능하고, 많은 경우 다중 노드에 확장 가능하도록 테스트 및 설계되어 사용자가 DGX 시스템에 대한 투자를 극대화할 수 있도록 보장합니다.

NVIDIA GPU 클라우드



NVIDIA AI 엔터프라이즈

NVIDIA AI Enterprise는 모든 기업이 생성적 AI를 활용할 수 있도록 하는 엔드 투 엔드 소프트웨어 플랫폼으로, NVIDIA DGX 플랫폼에서 실행되도록 최적화된 생성적 AI 기반 모델을 위한 가장 빠르고 효율적인 런타임을 제공합니다. 프로덕션 수준의 보안, 안정성, 관리 용이성을 갖춰 생성적 AI 솔루션 개발을 간소화합니다. NVIDIA AI Enterprise는 DGX BasePOD에 포함되어 기업 개발자가 사전 학습된 모델, 최적화된 프레임워크, 마이크로서비스, 가속 라이브러리 및 기업 지원에 액세스할 수 있도록 합니다.

NetApp 소프트웨어

NetApp ONTAP

NetApp의 최신 스토리지 관리 소프트웨어인 ONTAP 9를 사용하면 기업이 인프라를 현대화하고 클라우드 지원 데이터 센터로 전환할 수 있습니다. ONTAP 업계 최고의 데이터 관리 역량을 활용하여 데이터가 어디에 있던 단일 도구 세트를 사용하여 데이터를 관리하고 보호할 수 있도록 지원합니다. 또한 필요한 곳, 즉 엣지, 코어, 클라우드로 데이터를 자유롭게 이동할 수 있습니다. ONTAP 9에는 데이터 관리를 간소화하고, 중요 데이터를 가속화하고 보호하며, 하이브리드 클라우드 아키텍처 전반에서 차세대 인프라 기능을 구현하는 다양한 기능이 포함되어 있습니다.

데이터 가속화 및 보호

ONTAP 뛰어난 수준의 성능과 데이터 보호 기능을 제공하며 다음과 같은 방식으로 이러한 기능을 확장합니다.

- 성능과 낮은 지연 시간. ONTAP RDMA를 통한 NFS, 병렬 NFS(pNFS), NFS 세션 트렁킹을 사용하는 NVIDIA GPUDirect Storage(GDS)에 대한 지원을 포함하여 가능한 가장 낮은 지연 시간으로 가능한 가장 높은 처리량을 제공합니다.
- 데이터 보호. ONTAP 모든 플랫폼에서 공통적으로 관리 가능하며, 내장형 데이터 보호 기능과 업계에서 가장 강력한 랜섬웨어 방지 기능을 제공합니다.
- NetApp 볼륨 암호화(NVE). ONTAP 온보드 및 외부 키 관리 지원을 통해 기본 볼륨 수준 암호화를 제공합니다.
- 스토리지 멀티테넌시 및 다중 요소 인증. ONTAP 최고 수준의 보안을 통해 인프라 리소스를 공유할 수 있도록 합니다.

데이터 관리 간소화

적절한 리소스가 AI 애플리케이션과 AI/ML 데이터 세트 교육에 사용될 수 있도록 기업 IT 운영과 데이터 과학자에게 데이터 관리가 매우 중요합니다. NetApp 기술에 대한 다음 추가 정보는 이 검증 범위를 벗어나지만 배포에 따라 관련이

있을 수 있습니다.

ONTAP 데이터 관리 소프트웨어에는 다음과 같은 기능이 포함되어 있어 운영을 간소화하고 단순화하며 총 운영 비용을 절감할 수 있습니다.

- 스냅샷과 클론을 사용하면 ML/DL 워크플로에 대한 협업, 병렬 실험 및 향상된 데이터 거버넌스가 가능합니다.
- SnapMirror 하이브리드 클라우드와 다중 사이트 환경에서 원활한 데이터 이동을 지원하여 필요한 곳에 필요한 시간에 데이터를 제공합니다.
- 인라인 데이터 압축 및 확장된 중복 제거. 데이터 압축은 저장 블록 내부의 낭비되는 공간을 줄이고, 중복 제거는 효과적인 용량을 크게 증가시킵니다. 이는 로컬에 저장된 데이터와 클라우드에 계층화된 데이터 모두에 적용됩니다.
- 최소, 최대 및 적응형 서비스 품질(AQoS). 세분화된 서비스 품질(QoS) 제어는 공유 빈도가 높은 환경에서 중요한 애플리케이션의 성능 수준을 유지하는 데 도움이 됩니다.
- NetApp FlexGroups를 사용하면 스토리지 클러스터의 모든 노드에 데이터를 분산하여 대규모 데이터 세트에 대해 막대한 용량과 더 높은 성능을 제공할 수 있습니다.
- NetApp FabricPool. Amazon Web Services(AWS), Azure, NetApp StorageGRID 스토리지 솔루션을 포함한 퍼블릭 및 프라이빗 클라우드 스토리지 옵션에 콜드 데이터의 자동 계층화를 제공합니다. FabricPool 에 대한 자세한 내용은 다음을 참조하세요. "[TR-4598: FabricPool 모범 사례](#)".
- NetApp FlexCache. 파일 배포를 간소화하고, WAN 지연 시간을 줄이고, WAN 대역폭 비용을 낮추는 원격 볼륨 캐싱 기능을 제공합니다. FlexCache 사용하면 여러 사이트에 걸쳐 분산된 제품 개발을 수행할 수 있을 뿐만 아니라 원격 위치에서도 회사 데이터 세트에 더 빨리 액세스할 수 있습니다.

미래 지향적 인프라

ONTAP 다음과 같은 기능을 통해 까다롭고 끊임없이 변화하는 비즈니스 요구 사항을 충족하는 데 도움이 됩니다.

- 원활한 확장과 중단 없는 운영. ONTAP 기존 컨트롤러와 확장형 클러스터에 용량을 온라인으로 추가하는 것을 지원합니다. 고객은 비용이 많이 드는 데이터 마이그레이션이나 중단 없이 NVMe 및 32Gb FC와 같은 최신 기술로 업그레이드할 수 있습니다.
- 클라우드 연결. ONTAP 모든 퍼블릭 클라우드에서 소프트웨어 정의 스토리지(ONTAP Select)와 클라우드 기반 인스턴스(Google Cloud NetApp Volumes)에 대한 옵션을 제공하는 가장 클라우드에 연결된 스토리지 관리 소프트웨어입니다.
- 새로운 애플리케이션과의 통합. ONTAP 기존 엔터프라이즈 앱을 지원하는 동일한 인프라를 사용하여 자율주행차, 스마트 시티, 산업 4.0과 같은 차세대 플랫폼과 애플리케이션을 위한 엔터프라이즈급 데이터 서비스를 제공합니다.

NetApp DataOps 툴킷

NetApp DataOps Toolkit은 고성능, 확장형 NetApp 스토리지에 의해 지원되는 개발/교육 작업 공간과 추론 서버의 관리를 간소화하는 Python 기반 도구입니다. DataOps Toolkit은 독립형 유틸리티로 작동할 수 있으며 NetApp Trident 활용하여 스토리지 작업을 자동화하는 Kubernetes 환경에서 더욱 효과적입니다. 주요 기능은 다음과 같습니다.

- 고성능, 확장형 NetApp 스토리지로 지원되는 새로운 대용량 JupyterLab 작업 공간을 빠르게 프로비저닝하세요.
- 엔터프라이즈급 NetApp 스토리지로 지원되는 새로운 NVIDIA Triton Inference Server 인스턴스를 빠르게 프로비저닝하세요.
- 실험이나 빠른 반복을 가능하게 하기 위해 대용량 JupyterLab 작업 공간을 거의 즉각적으로 복제합니다.
- 대용량 JupyterLab 작업 공간의 거의 즉각적인 스냅샷을 백업 및/또는 추적/기준 설정에 사용할 수 있습니다.
- 대용량, 고성능 데이터 볼륨에 대한 거의 즉각적인 프로비저닝, 복제 및 스냅샷이 가능합니다.

NetApp Trident

Trident Anthos를 포함한 컨테이너와 Kubernetes 배포판을 위한 완벽하게 지원되는 오픈 소스 스토리지 오케스트레이터입니다. Trident NetApp ONTAP 포함한 전체 NetApp 스토리지 포트폴리오와 호환되며 NFS, NVMe/TCP 및 iSCSI 연결도 지원합니다. Trident 최종 사용자가 스토리지 관리자의 개입 없이 NetApp 스토리지 시스템에서 스토리지를 프로비저닝하고 관리할 수 있도록 하여 DevOps 워크플로를 가속화합니다.

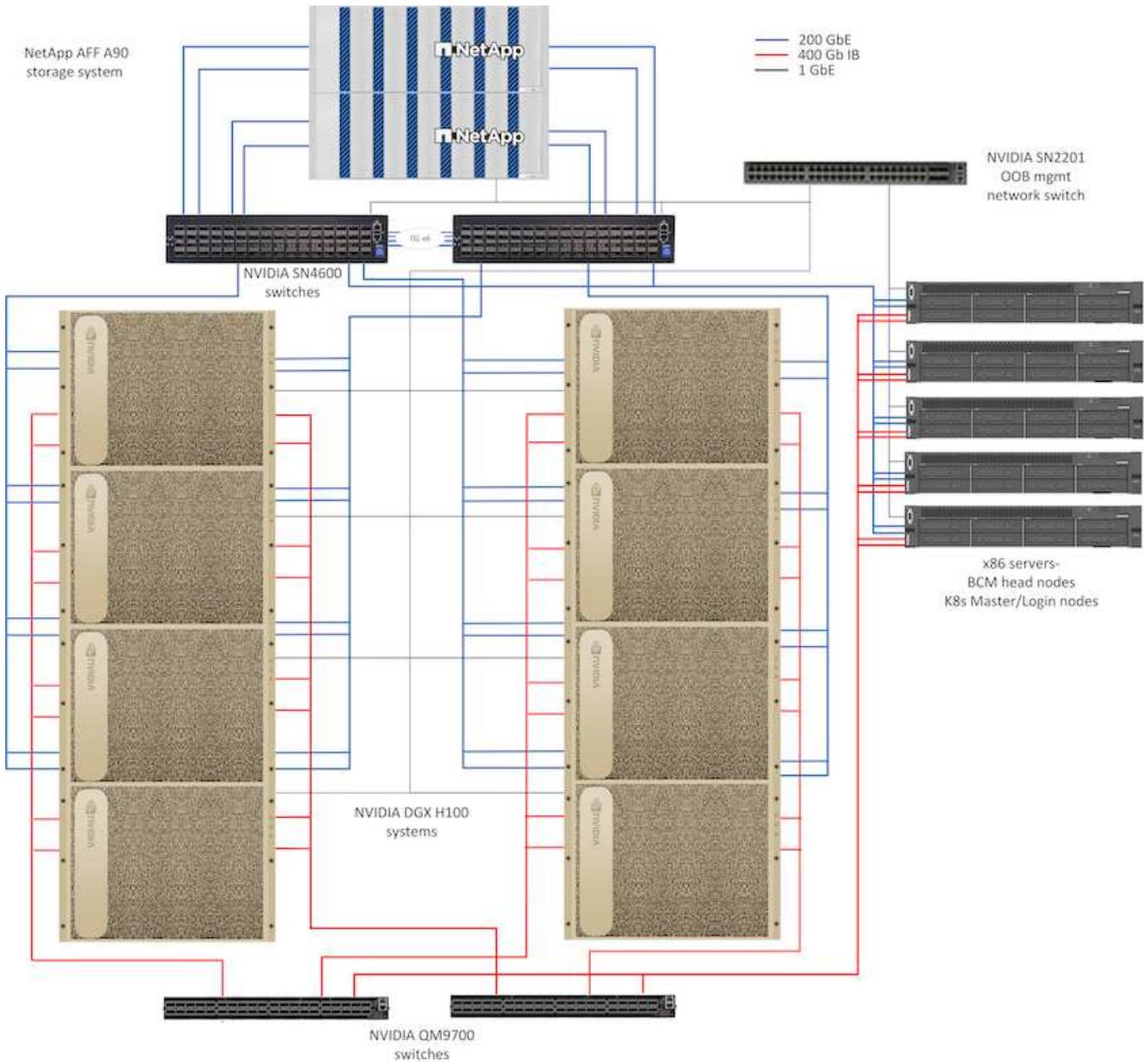
NVIDIA DGX H100 시스템을 탑재한 NVA-1173 NetApp AI Pod - 솔루션 아키텍처

이 섹션에서는 NVIDIA DGX 시스템을 탑재한 NetApp AI Pod 의 아키텍처에 중점을 둡니다.

DGX 시스템을 갖춘 NetApp AI Pod

이 참조 아키텍처는 컴퓨팅 노드 간에 400Gb/s InfiniBand (IB) 연결을 통해 컴퓨팅 클러스터 상호 연결 및 스토리지 액세스를 위해 별도의 패브릭을 활용합니다. 아래 그림은 DGX H100 시스템이 포함된 NetApp AI Pod 의 전체 솔루션 토폴로지를 보여줍니다.

NetApp AI Pod 솔루션 토폴로지



네트워크 디자인

이 구성에서 컴퓨팅 클러스터 패브릭은 높은 가용성을 위해 서로 연결된 한 쌍의 QM9700 400Gb/s IB 스위치를 사용합니다. 각 DGX H100 시스템은 8개의 연결을 사용하여 스위치에 연결되며, 짝수 포트는 한 스위치에 연결되고 홀수 포트는 다른 스위치에 연결됩니다.

스토리지 시스템 액세스, 대역 내 관리 및 클라이언트 액세스를 위해 SN4600 이더넷 스위치 한 쌍이 사용됩니다. 스위치는 스위치 간 링크로 연결되고 여러 VLAN으로 구성되어 다양한 트래픽 유형을 분리합니다. 특정 VLAN 간에 기본 L3 라우팅이 활성화되어 동일한 스위치에 있는 클라이언트와 스토리지 인터페이스 간, 그리고고가용성을 위한 스위치 간 다중 경로가 가능합니다. 대규모 배포의 경우, 필요에 따라 스파인 스위치용 추가 스위치 쌍과 추가 리프를 추가하여 이더넷 네트워크를 리프-스파인 구성으로 확장할 수 있습니다.

컴퓨팅 상호 연결 및 고속 이더넷 네트워크 외에도 모든 물리적 장치는 대역 외 관리를 위해 하나 이상의 SN2201 이더넷 스위치에 연결됩니다. 를 참조하십시오 ["배포 세부 정보"](#) 네트워크 구성에 대한 자세한 내용은 페이지를 참조하세요.

DGX H100 시스템의 스토리지 액세스 개요

각 DGX H100 시스템에는 관리 및 스토리지 트래픽을 위한 듀얼 포트 ConnectX-7 어댑터 2개가 제공되며, 이 솔루션의 경우 각 카드의 두 포트가 동일한 스위치에 연결됩니다. 각 카드의 한 포트는 LACP MLAG 본드로 구성되고, 한 포트는 각 스위치에 연결되며, 대역 내 관리, 클라이언트 액세스, 사용자 수준 스토리지 액세스를 위한 VLAN이 이 본드에 호스팅됩니다.

각 카드의 다른 포트는 AFF A90 스토리지 시스템에 연결하는 데 사용되며, 작업 부하 요구 사항에 따라 여러 구성으로 사용할 수 있습니다. NVIDIA Magnum IO GPUDirect 스토리지를 지원하기 위해 RDMA를 통한 NFS를 사용하는 구성의 경우, 포트는 별도의 VLAN에 있는 IP 주소와 함께 개별적으로 사용됩니다. RDMA가 필요하지 않은 배포의 경우 스토리지 인터페이스를 LACP 본딩으로 구성하여 높은 가용성과 추가 대역폭을 제공할 수도 있습니다. RDMA 사용 여부와 관계없이 클라이언트는 NFS v4.1 pNFS 및 세션 트렁킹을 사용하여 스토리지 시스템을 마운트하여 클러스터의 모든 스토리지 노드에 대한 병렬 액세스를 활성화할 수 있습니다. 를 참조하십시오 ["배포 세부 정보"](#) 클라이언트 구성에 대한 자세한 내용은 페이지를 참조하세요.

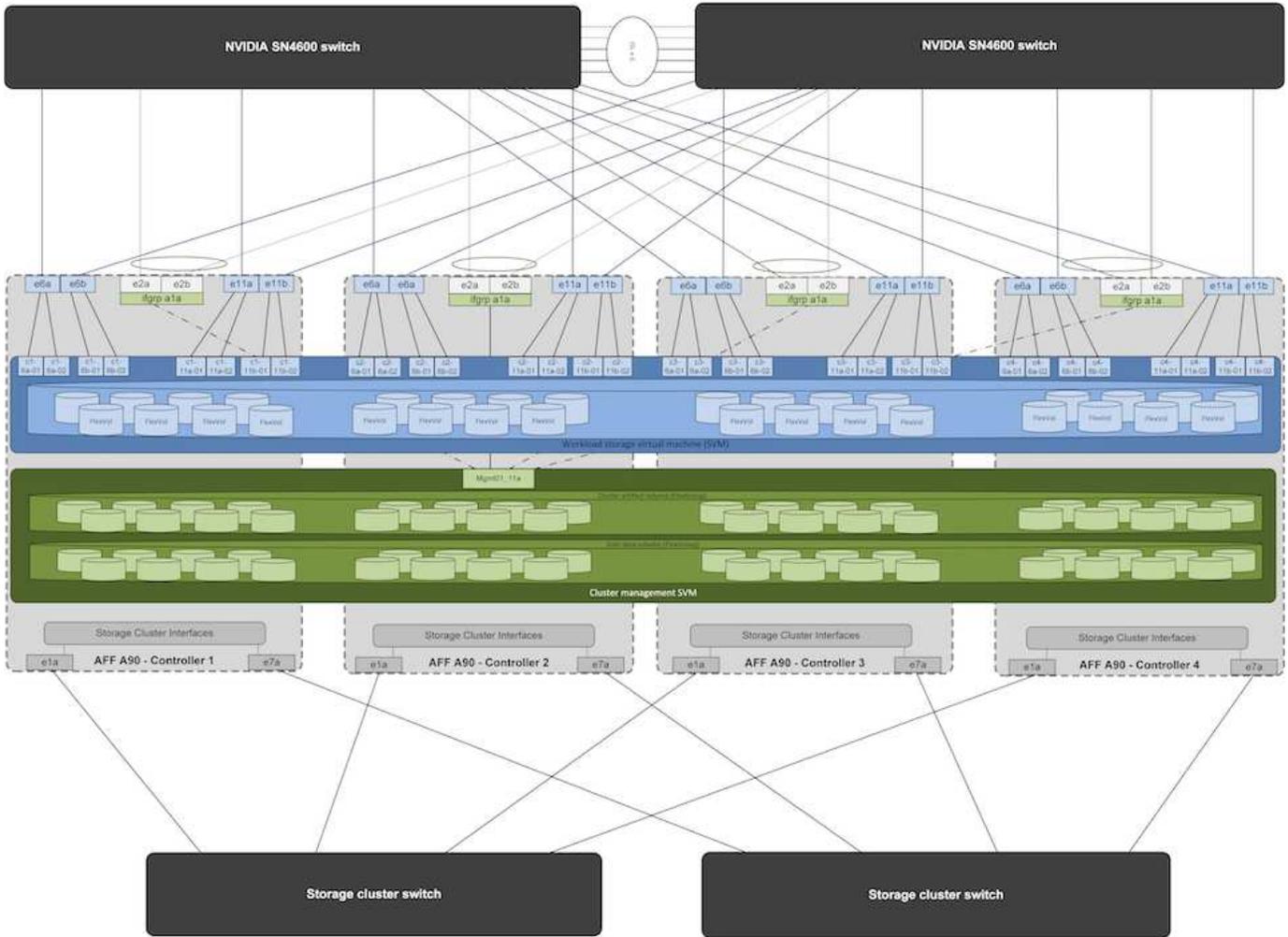
DGX H100 시스템 연결에 대한 자세한 내용은 다음을 참조하세요. ["NVIDIA BasePOD 문서"](#).

저장 시스템 설계

각 AFF A90 스토리지 시스템은 각 컨트롤러에서 6개의 200GbE 포트를 사용하여 연결됩니다. 각 컨트롤러의 4개 포트는 DGX 시스템의 워크로드 데이터 액세스에 사용되고, 각 컨트롤러의 2개 포트는 클러스터 관리 아티팩트 및 사용자 홈 디렉토리에 대한 관리 플레인 서버의 액세스를 지원하기 위해 LACP 인터페이스 그룹으로 구성됩니다. 스토리지 시스템의 모든 데이터 액세스는 NFS를 통해 제공되며, AI 워크로드 액세스에 전담된 스토리지 가상 머신(SVM)과 클러스터 관리 용도에 전담된 별도의 SVM이 있습니다.

관리 SVM에는 각 컨트롤러에 구성된 2포트 인터페이스 그룹에 호스팅되는 단일 LIF만 필요합니다. 다른 FlexGroup 볼륨은 클러스터 노드 이미지, 시스템 모니터링 기록 데이터, 최종 사용자 홈 디렉토리 및 같은 클러스터 관리 아티팩트를 보관하기 위해 관리 SVM에 프로비저닝됩니다. 아래 그림은 저장 시스템의 논리적 구성을 보여줍니다.

NetApp A90 스토리지 클러스터 논리적 구성



관리 플레인 서버

이 참조 아키텍처에는 관리 플레인용으로 사용되는 5개의 CPU 기반 서버도 포함되어 있습니다. 이 시스템 중 두 개는 클러스터 배포 및 관리를 위한 NVIDIA Base Command Manager의 헤드 노드로 사용됩니다. 나머지 세 시스템은 Slurm을 사용하여 작업 스케줄링을 수행하는 배포를 위한 Kubernetes 마스터 노드나 로그인 노드와 같은 추가 클러스터 서비스를 제공하는 데 사용됩니다. Kubernetes를 활용한 배포에서는 NetApp Trident CSI 드라이버를 활용하여 AFF A900 스토리지 시스템에서 관리 및 AI 워크로드를 위한 영구 스토리지와 함께 자동화된 프로비저닝 및 데이터 서비스를 제공할 수 있습니다.

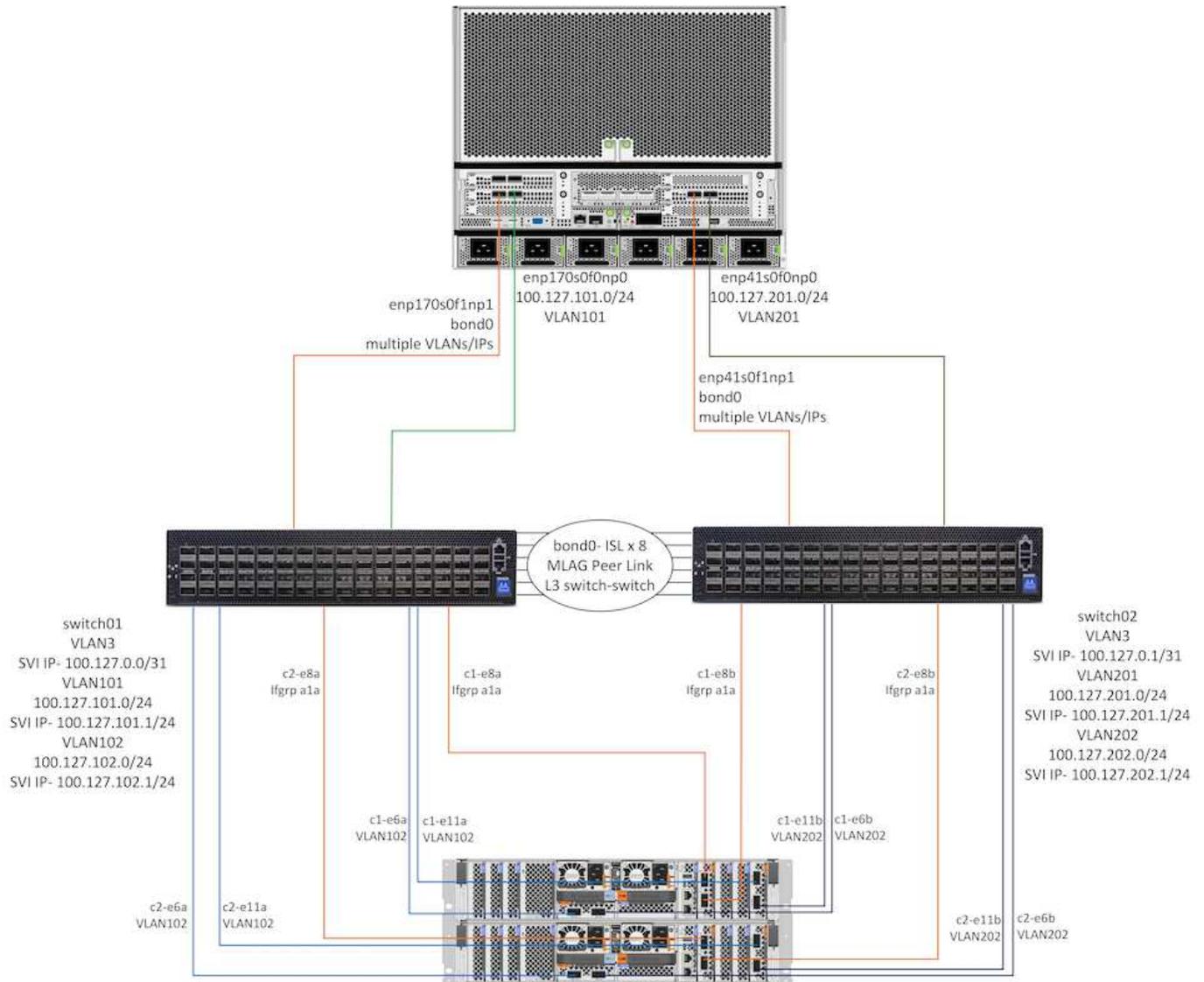
각 서버는 클러스터 배포 및 관리를 가능하게 하기 위해 IB 스위치와 이더넷 스위치에 물리적으로 연결되며, 앞서 설명한 대로 클러스터 관리 아티팩트를 저장하기 위해 관리 SVM을 통해 스토리지 시스템에 NFS 마운트로 구성됩니다.

NVIDIA DGX 시스템을 탑재한 NVA-1173 NetApp AIPod - 배포 세부 정보

이 섹션에서는 이 솔루션의 검증 과정에 사용된 배포 세부 정보를 설명합니다. 사용된 IP 주소는 예시이며 배포 환경에 따라 수정해야 합니다. 이 구성을 구현하는 데 사용된 특정 명령에 대한 자세한 내용은 해당 제품 설명서를 참조하세요.

아래 다이어그램은 1개의 DGX H100 시스템과 1개의 HA 쌍의 AFF A90 컨트롤러에 대한 자세한 네트워크 및 연결 정보를 보여줍니다. 다음 섹션의 배포 지침은 이 다이어그램의 세부 정보를 기반으로 합니다.

NetApp Alpod 네트워크 구성



다음 표는 최대 16개의 DGX 시스템과 2개의 AFF A90 HA 쌍에 대한 케이블 할당의 예를 보여줍니다.

스위치 및 포트	장치	장치 포트
switch1 포트 1-16	DGX-H100-01부터 -16까지	enp170s0f0np0, 슬롯1 포트 1
switch1 포트 17-32	DGX-H100-01부터 -16까지	enp170s0f1np1, 슬롯1 포트 2
switch1 포트 33-36	AFF-A90-01부터 -04까지	포트 e6a
switch1 포트 37-40	AFF-A90-01부터 -04까지	포트 e11a
switch1 포트 41-44	AFF-A90-01부터 -04까지	포트 e2a
switch1 포트 57-64	ISL에서 switch2로	포트 57-64
switch2 포트 1-16	DGX-H100-01부터 -16까지	enp41s0f0np0, 슬롯 2 포트 1
switch2 포트 17-32	DGX-H100-01부터 -16까지	enp41s0f1np1, 슬롯 2 포트 2
switch2 포트 33-36	AFF-A90-01부터 -04까지	포트 e6b

스위치 및 포트	장치	장치 포트
스위치2 포트 37-40	AFF-A90-01부터 -04까지	포트 e11b
스위치2 포트 41-44	AFF-A90-01부터 -04까지	포트 e2b
스위치2 포트 57-64	ISL에서 switch1로	포트 57-64

다음 표는 이 검증에 사용된 다양한 구성 요소의 소프트웨어 버전을 보여줍니다.

장치	소프트웨어 버전
NVIDIA SN4600 스위치	큐물러스 리눅스 v5.9.1
NVIDIA DGX 시스템	DGX OS v6.2.1(우분투 22.04 LTS)
멜라녹스 OFED	24.01
NetApp AFF A90	NetApp ONTAP 9.14.1

스토리지 네트워크 구성

이 섹션에서는 이더넷 저장 네트워크 구성에 대한 주요 세부 정보를 설명합니다. InfiniBand 컴퓨팅 네트워크 구성에 대한 정보는 다음을 참조하세요. "[NVIDIA BasePOD 문서](#)". 스위치 구성에 대한 자세한 내용은 다음을 참조하세요. "[NVIDIA Cumulus Linux 문서](#)".

SN4600 스위치를 구성하는 데 사용되는 기본 단계는 아래와 같습니다. 이 프로세스에서는 케이블 연결과 기본 스위치 설정(관리 IP 주소, 라이선싱 등)이 완료되었다고 가정합니다.

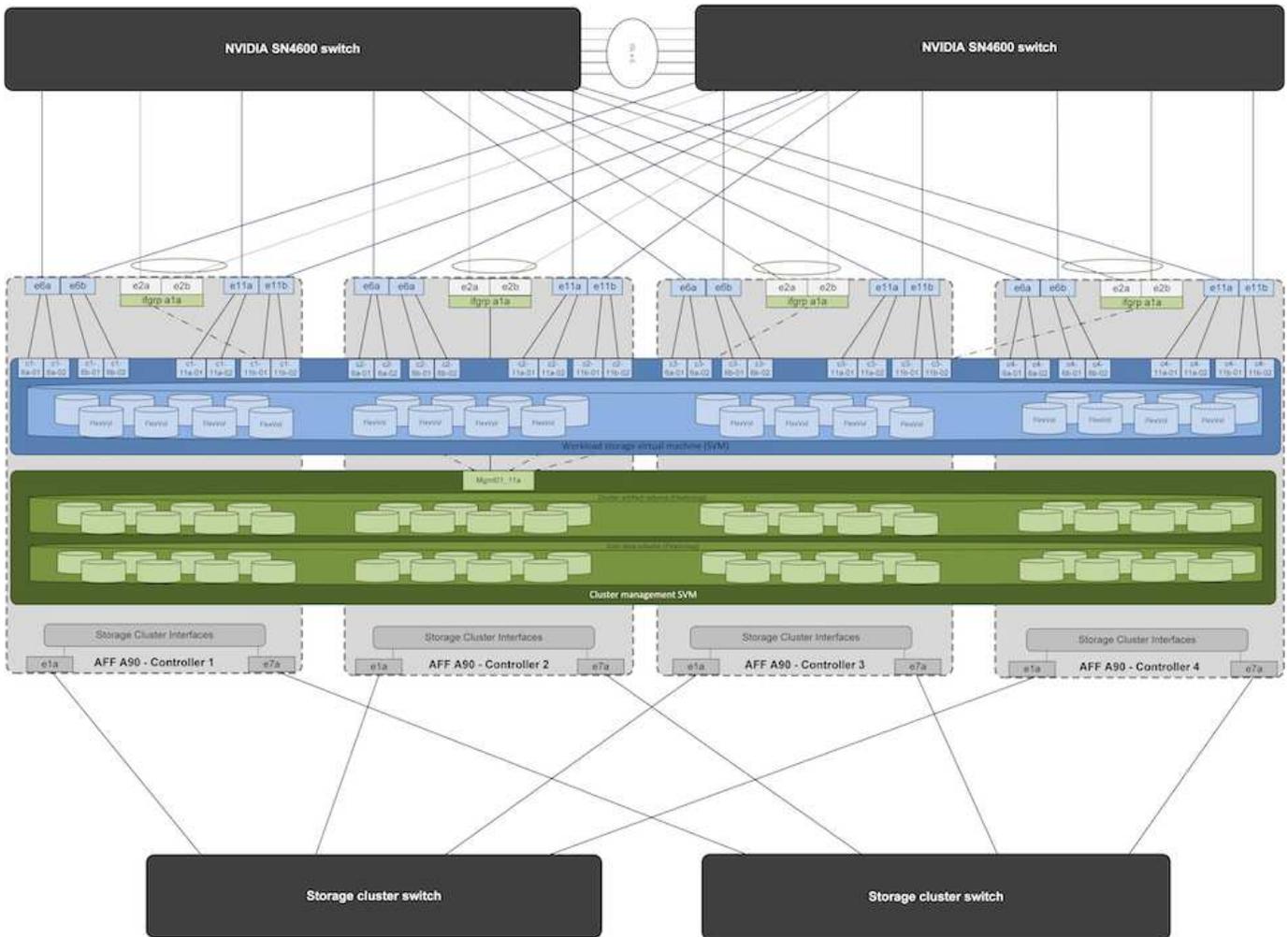
- 스위치 간 ISL 본드를 구성하여 다중 링크 집계(MLAG) 및 장애 조치 트래픽을 활성화합니다.
 - 이 검증에서는 테스트 중인 스토리지 구성에 충분한 대역폭을 제공하기 위해 8개의 링크를 사용했습니다.
 - MLAG 활성화에 대한 구체적인 지침은 Cumulus Linux 문서를 참조하세요.
- 두 스위치의 각 클라이언트 포트와 스토리지 포트 쌍에 대해 LACP MLAG를 구성합니다.
 - DGX-H100-01의 경우 각 스위치에 포트 swp17(enp170s0f1np1 및 enp41s0f1np1), DGX-H100-02의 경우 포트 swp18 등(bond1-16)
 - AFF-A90-01(e2a 및 e2b)의 경우 각 스위치에 포트 swp41, AFF-A90-02의 경우 포트 swp42 등(bond17-20)
 - `nv set interface bondX bond member swpX`
 - `nv set interface bondx bond mlag id X`
- 모든 포트와 MLAG 본드를 기본 브리지 도메인에 추가합니다.
 - `nv set int swp1-16,33-40 브리지 도메인 br_default`
 - `nv set int bond1-20 브리지 도메인 br_default`
- 각 스위치에서 RoCE 활성화
 - `nv 세트 로체 모드 무손실`
- VLAN 구성 - 클라이언트 포트용 2개, 스토리지 포트용 2개, 관리용 1개, L3 스위치 간 스위치용 1개
 - 스위치 1-
 - 클라이언트 NIC 장애 발생 시 L3 스위치 간 라우팅을 위한 VLAN 3

- 각 DGX 시스템의 스토리지 포트 1에 대한 VLAN 101(enp170s0f0np0, slot1 포트 1)
 - 각 AFF A90 스토리지 컨트롤러의 포트 e6a 및 e11a에 대한 VLAN 102
 - 각 DGX 시스템 및 스토리지 컨트롤러에 대한 MLAG 인터페이스를 사용하여 관리하기 위한 VLAN 301
- 스위치 2-
 - 클라이언트 NIC 장애 발생 시 L3 스위치 간 라우팅을 위한 VLAN 3
 - 각 DGX 시스템의 스토리지 포트 2에 대한 VLAN 201(enp41s0f0np0, slot2 포트 1)
 - 각 AFF A90 스토리지 컨트롤러의 포트 e6b 및 e11b에 대한 VLAN 202
 - 각 DGX 시스템 및 스토리지 컨트롤러에 대한 MLAG 인터페이스를 사용하여 관리하기 위한 VLAN 301
6. 클라이언트 VLAN의 클라이언트 포트, 스토리지 VLAN의 스토리지 포트 등 각 VLAN에 적절한 물리적 포트를 할당합니다.
- nv set int <swpX> 브리지 도메인 br_default 액세스 <vlan id>
 - 필요에 따라 본딩된 인터페이스를 통해 여러 VLAN을 활성화하기 위해 MLAG 포트는 트렁크 포트로 유지되어야 합니다.
7. 각 VLAN에 SVI(스위치 가상 인터페이스)를 구성하여 게이트웨이 역할을 하고 L3 라우팅을 활성화합니다.
- 스위치 1-
 - nv set int vlan3 ip 주소 100.127.0.0/31
 - nv set int vlan101 ip 주소 100.127.101.1/24
 - nv set int vlan102 ip 주소 100.127.102.1/24
 - 스위치 2-
 - nv set int vlan3 ip 주소 100.127.0.1/31
 - nv set int vlan201 ip 주소 100.127.201.1/24
 - nv set int vlan202 ip 주소 100.127.202.1/24
8. 정적 경로 생성
- 동일한 스위치의 서브넷에 대해 정적 경로가 자동으로 생성됩니다.
 - 클라이언트 링크 장애 발생 시 스위치 간 라우팅을 위해 추가 정적 경로가 필요합니다.
 - 스위치 1-
 - nv set vrf 기본 라우터 정적 100.127.128.0/17 via 100.127.0.1
 - 스위치 2-
 - nv set vrf 기본 라우터 정적 100.127.0.0/17 via 100.127.0.0

스토리지 시스템 구성

이 섹션에서는 이 솔루션을 위한 A90 스토리지 시스템 구성에 대한 주요 세부 정보를 설명합니다. ONTAP 시스템 구성에 대한 자세한 내용은 다음을 참조하세요. "[ONTAP 문서](#)". 아래 다이어그램은 저장 시스템의 논리적 구성을 보여줍니다.

NetApp A90 스토리지 클러스터 논리적 구성



저장 시스템을 구성하는 데 사용되는 기본 단계는 아래와 같습니다. 이 프로세스에서는 기본 스토리지 클러스터 설치가 완료되었다고 가정합니다.

1. 사용 가능한 모든 파티션에서 1개의 예비 파티션을 뺀 각 컨트롤러에 1개의 집계를 구성합니다.
 - `aggr create -node <노드> -aggregate <노드>_data01 -diskcount <47>`
2. 각 컨트롤러에서 ifgrps 구성
 - `net port ifgrp create -node <노드> -ifgrp a1a -mode multimode_lacp -distr-function port`
 - `net port ifgrp add-port -node <노드> -ifgrp <ifgrp> -ports <노드>:e2a,<노드>:e2b`
3. 각 컨트롤러의 ifgrp에서 mgmt vlan 포트를 구성합니다.
 - 넷 포트 VLAN 생성 -노드 aff-a90-01 -포트 a1a -vlan-id 31
 - 넷 포트 VLAN 생성 -노드 aff-a90-02 -포트 a1a -vlan-id 31
 - 넷 포트 VLAN 생성 -노드 aff-a90-03 -포트 a1a -vlan-id 31
 - 넷 포트 VLAN 생성 -노드 aff-a90-04 -포트 a1a -vlan-id 31
4. 브로드캐스트 도메인 생성
 - 브로드캐스트 도메인 생성 -브로드캐스트 도메인 vlan21 -mtu 9000 -포트 aff-a90-01:e6a, aff-a90-01:e11a, aff-a90-02:e6a, aff-a90-02:e11a, aff-a90-03:e6a, aff-a90-03:e11a, aff-a90-04:e6a, aff-a90-04:e11a
 - 브로드캐스트 도메인 생성 -브로드캐스트 도메인 vlan22 -mtu 9000 -포트 aaff-a90-01:e6b, aff-a90-01:e11b,

aff-a90-02:e6b, aff-a90-02:e11b, aff-a90-03:e6b, aff-a90-03:e11b, aff-a90-04:e6b, aff-a90-04:e11b

- 브로드캐스트 도메인 생성 -브로드캐스트 도메인 vlan31 -mtu 9000 -포트 aff-a90-01:a1a-31,aff-a90-02:a1a-31,aff-a90-03:a1a-31,aff-a90-04:a1a-31

5. 관리 SVM 생성 *

6. 관리 SVM 구성

- LIF 생성
 - net int create -vserver basepod-mgmt -lif vlan31-01 -home-node aff-a90-01 -home-port a1a-31 -address 192.168.31.X -netmask 255.255.255.0
- FlexGroup 볼륨 생성-
 - vol create -vserver basepod-mgmt -volume home -size 10T -auto-provision-as flexgroup -junction -path /home
 - vol create -vserver basepod-mgmt -volume cm -size 10T -auto-provision-as flexgroup -junction -path /cm
- 수출 정책 생성
 - export-policy 규칙 생성 -vserver basepod-mgmt -policy default -client-match 192.168.31.0/24 -rorule sys -rwrule sys -superuser sys

7. 데이터 SVM 생성 *

8. 데이터 SVM 구성

- RDMA 지원을 위한 SVM 구성
 - vserver nfs 수정 -vserver basepod-data -rdma 활성화
- LIF를 생성하다
 - net int create -vserver basepod-data -lif c1-6a-lif1 -home-node aff-a90-01 -home-port e6a -address 100.127.102.101 -netmask 255.255.255.0
 - net int create -vserver basepod-data -lif c1-6a-lif2 -home-node aff-a90-01 -home-port e6a -address 100.127.102.102 -netmask 255.255.255.0
 - net int create -vserver basepod-data -lif c1-6b-lif1 -home-node aff-a90-01 -home-port e6b -address 100.127.202.101 -netmask 255.255.255.0
 - net int create -vserver basepod-data -lif c1-6b-lif2 -home-node aff-a90-01 -home-port e6b -address 100.127.202.102 -netmask 255.255.255.0
 - net int create -vserver basepod-data -lif c1-11a-lif1 -home-node aff-a90-01 -home-port e11a -address 100.127.102.103 -netmask 255.255.255.0
 - net int create -vserver basepod-data -lif c1-11a-lif2 -home-node aff-a90-01 -home-port e11a -address 100.127.102.104 -netmask 255.255.255.0
 - net int create -vserver basepod-data -lif c1-11b-lif1 -home-node aff-a90-01 -home-port e11b -address 100.127.202.103 -netmask 255.255.255.0
 - net int create -vserver basepod-data -lif c1-11b-lif2 -home-node aff-a90-01 -home-port e11b -address 100.127.202.104 -netmask 255.255.255.0
 - net int create -vserver basepod-data -lif c2-6a-lif1 -home-node aff-a90-02 -home-port e6a -address 100.127.102.105 -netmask 255.255.255.0
 - net int create -vserver basepod-data -lif c2-6a-lif2 -home-node aff-a90-02 -home-port e6a -address 100.127.102.106 -netmask 255.255.255.0

- net int create -vserver basepod-data -lif c2-6b-lif1 -home-node aff-a90-02 -home-port e6b -address 100.127.202.105 -netmask 255.255.255.0
- net int create -vserver basepod-data -lif c2-6b-lif2 -home-node aff-a90-02 -home-port e6b -address 100.127.202.106 -netmask 255.255.255.0
- net int create -vserver basepod-data -lif c2-11a-lif1 -home-node aff-a90-02 -home-port e11a -address 100.127.102.107 -netmask 255.255.255.0
- net int create -vserver basepod-data -lif c2-11a-lif2 -home-node aff-a90-02 -home-port e11a -address 100.127.102.108 -netmask 255.255.255.0
- net int create -vserver basepod-data -lif c2-11b-lif1 -home-node aff-a90-02 -home-port e11b -address 100.127.202.107 -netmask 255.255.255.0
- net int create -vserver basepod-data -lif c2-11b-lif2 -home-node aff-a90-02 -home-port e11b -address 100.127.202.108 -netmask 255.255.255.0

9. RDMA 액세스를 위한 LIF 구성

- ONTAP 9.15.1을 배포하는 경우 물리적 정보에 대한 RoCE QoS 구성에는 ONTAP CLI에서 사용할 수 없는 OS 수준 명령이 필요합니다. RoCE 지원을 위한 포트 구성과 관련된 도움이 필요하면 NetApp 지원팀에 문의하세요. RDMA를 통한 NFS는 문제없이 작동합니다.
- ONTAP 9.16.1부터 물리적 인터페이스는 엔드투엔드 RoCE 지원을 위한 적절한 설정으로 자동으로 구성됩니다.
- net int 수정 -vserver basepod-data -lif * -rdma-protocols roce

10. 데이터 SVM에서 NFS 매개변수 구성

- nfs 수정 -vserver basepod-data -v4.1 활성화 -v4.1-pnfs 활성화 -v4.1-trunking 활성화 -tcp-max-transfer-size 262144

11. FlexGroup 볼륨 생성

- vol create -vserver basepod-data -volume data -size 100T -auto-provision-as flexgroup -junction-path /data

12. 수출 정책 생성

- export-policy 규칙 생성 -vserver basepod-data -policy default -client-match 100.127.101.0/24 -rorule sys -rwrule sys -superuser sys
- export-policy 규칙 생성 -vserver basepod-data -policy default -client-match 100.127.201.0/24 -rorule sys -rwrule sys -superuser sys

13. 경로 생성

- 경로 추가 -vserver basepod_data -대상 100.127.0.0/17 -게이트웨이 100.127.102.1 메트릭 20
- 경로 추가 -vserver basepod_data -대상 100.127.0.0/17 -게이트웨이 100.127.202.1 메트릭 30
- 경로 추가 -vserver basepod_data -대상 100.127.128.0/17 -게이트웨이 100.127.202.1 메트릭 20
- 경로 추가 -vserver basepod_data -대상 100.127.128.0/17 -게이트웨이 100.127.102.1 메트릭 30

RoCE 스토리지 액세스를 위한 DGX H100 구성

이 섹션에서는 DGX H100 시스템 구성에 대한 주요 세부 정보를 설명합니다. 이러한 구성 항목 중 다수는 DGX 시스템에 배포된 OS 이미지에 포함되거나 부팅 시 Base Command Manager에서 구현될 수 있습니다. BCM에서 노드 및 소프트웨어 이미지 구성에 대한 자세한 내용은 참조용으로 여기에 나열되어 있습니다. "[BCM 문서](#)".

1. 추가 패키지 설치
 - 아이피미틀
 - 파이썬3-pip
2. Python 패키지 설치
 - 파라미코
 - 맷플롯립
3. 패키지 설치 후 dpkg 재구성
 - dpkg --configure -a
4. MOFED 설치
5. 성능 튜닝을 위한 mst 값 설정
 - mstconfig -y -d <aa:00.0,29:00.0> ADVANCED_PCI_SETTINGS=1 NUM_OF_VFS=0
MAX_ACC_OUT_READ=44로 설정
6. 설정 수정 후 어댑터 재설정
 - mlxfwreset -d <aa:00.0,29:00.0> -y 재설정
7. PCI 장치에 MaxReadReq 설정
 - setpci -s <aa:00.0,29:00.0> 68.W=5957
8. RX 및 TX 링 버퍼 크기 설정
 - ethtool -G <enp170s0f0np0,enp41s0f0np0> rx 8192 tx 8192
9. mlx_qos를 사용하여 PFC 및 DSCP 설정
 - mlx_qos -i <enp170s0f0np0,enp41s0f0np0> --pfc 0,0,0,1,0,0,0,0 --trust=dscp --케이블_길이=3
10. 네트워크 포트의 RoCE 트래픽에 대한 ToS 설정
 - 예코 106 > /sys/class/infiniband/<mlx5_7,mlx5_1>/tc/1/트래픽_클래스
11. 적절한 서브넷의 IP 주소로 각 스토리지 NIC를 구성합니다.
 - 저장용 NIC 1의 경우 100.127.101.0/24
 - 저장용 NIC 2의 경우 100.127.201.0/24
12. LACP 본딩을 위한 인밴드 네트워크 포트 구성(enp170s0f1np1, enp41s0f1np1)
13. 각 스토리지 서브넷에 대한 기본 및 보조 경로에 대한 정적 경로를 구성합니다.
 - 경로 추가 -net 100.127.0.0/17 gw 100.127.101.1 메트릭 20
 - 경로 추가 -net 100.127.0.0/17 gw 100.127.201.1 메트릭 30
 - 경로 추가 -net 100.127.128.0/17 gw 100.127.201.1 메트릭 20
 - 경로 추가 -net 100.127.128.0/17 gw 100.127.101.1 메트릭 30
14. 마운트 /home 볼륨
 - mount -o vers=3,nconnect=16,rsz=262144,wsz=262144 192.168.31.X:/home /home
15. /데이터 볼륨 마운트
 - 데이터 볼륨을 마운트할 때 다음 마운트 옵션이 사용되었습니다.

- vers=4.1 # 여러 스토리지 노드에 대한 병렬 액세스를 위해 pNFS를 활성화합니다.
- proto=rdma # 기본 TCP 대신 RDMA로 전송 프로토콜을 설정합니다.
- max_connect=16 # NFS 세션 트렁킹을 활성화하여 스토리지 포트 대역폭을 집계합니다.
- write=eager # 버퍼링된 쓰기의 쓰기 성능을 향상시킵니다.
- rsize=262144, wsize=262144 # I/O 전송 크기를 256k로 설정합니다.

NVIDIA DGX 시스템을 탑재한 NVA-1173 NetApp AI Pod - 솔루션 검증 및 크기 조정 지침

이 섹션에서는 NVIDIA DGX 시스템을 사용하는 NetApp AI Pod 에 대한 솔루션 검증 및 크기 조정 지침에 중점을 둡니다.

솔루션 검증

이 솔루션의 스토리지 구성은 오픈 소스 도구인 FIO를 사용하여 일련의 합성 워크로드를 통해 검증되었습니다. 이러한 테스트에는 DGX 시스템이 딥 러닝 학습 작업을 수행하면서 생성되는 스토리지 작업 부하를 시뮬레이션하기 위한 읽기 및 쓰기 I/O 패턴이 포함됩니다. DGX 시스템 클러스터를 시뮬레이션하기 위해 FIO 워크로드를 동시에 실행하는 2소켓 CPU 서버 클러스터를 사용하여 스토리지 구성을 검증했습니다. 각 클라이언트는 이전에 설명한 것과 동일한 네트워크 구성으로 구성되었으며, 다음 세부 정보가 추가되었습니다.

이 검증에는 다음 마운트 옵션이 사용되었습니다.

버전=4.1	여러 스토리지 노드에 대한 병렬 액세스를 위해 pNFS를 활성화합니다.
프로토=rdma	기본 TCP 대신 RDMA로 전송 프로토콜을 설정합니다.
포트=20049	RDMA NFS 서비스에 대한 올바른 포트를 지정하세요
최대 연결=16	NFS 세션 트렁킹을 통해 스토리지 포트 대역폭을 집계할 수 있습니다.
쓰기=열망	버퍼링된 쓰기의 쓰기 성능을 향상시킵니다.
rsize=262144, wsize=262144	I/O 전송 크기를 256k로 설정합니다.

또한 클라이언트는 NFS max_session_slots 값이 1024로 구성되었습니다. NFS over RDMA를 사용하여 솔루션을 테스트했으므로 스토리지 네트워크 포트는 액티브/패시브 본드로 구성되었습니다. 이 검증에는 다음과 같은 결합 매개변수가 사용되었습니다.

모드=활성 백업	결합을 활성/수동 모드로 설정합니다
primary=<인터페이스 이름>	모든 클라이언트의 기본 인터페이스는 스위치에 분산되었습니다.
mii-모니터-간격=100	100ms의 모니터링 간격을 지정합니다.
fail-over-mac-policy=active	활성 링크의 MAC 주소가 본드의 MAC임을 지정합니다. 이는 본딩된 인터페이스를 통한 RDMA의 적절한 작동에 필요합니다.

스토리지 시스템은 각 HA 쌍에 24개의 1.9TB NVMe 디스크 드라이브가 연결된 2개의 NS224 디스크 셀프와 함께 2개의 A900 HA 쌍(컨트롤러 4개)으로 구성된 것과 같이 구성되었습니다. 아키텍처 섹션에서 설명한 대로 모든 컨트롤러의 저장 용량은 FlexGroup 볼륨을 사용하여 결합되었으며 모든 클라이언트의 데이터는 클러스터의 모든 컨트롤러에 분산되었습니다.

스토리지 시스템 크기 지침

NetApp DGX BasePOD 인증을 성공적으로 완료했으며, 테스트된 두 개의 A90 HA 쌍은 16개의 DGX H100 시스템 클러스터를 쉽게 지원할 수 있습니다. 더 높은 스토리지 성능 요구 사항이 있는 대규모 배포의 경우 NetApp ONTAP 클러스터에 최대 12개의 HA 쌍(24개 노드)까지 추가 AFF 시스템을 추가할 수 있습니다. 이 솔루션에 설명된 FlexGroup 기술을 사용하면 24노드 클러스터가 단일 네임스페이스에서 79PB 이상과 최대 552GBps의 처리량을 제공할 수 있습니다. AFF A400, A250 및 C800과 같은 다른 NetApp 스토리지 시스템은 더 낮은 비용으로 더 작은 규모의 배포에 대해 더 낮은 성능 및/또는 더 큰 용량 옵션을 제공합니다. ONTAP 9는 혼합 모델 클러스터를 지원하므로 고객은 처음에는 작은 공간으로 시작한 후 용량과 성능 요구 사항이 증가함에 따라 클러스터에 더 많거나 더 큰 스토리지 시스템을 추가할 수 있습니다. 아래 표는 각 AFF 모델에서 지원되는 A100 및 H100 GPU의 수를 대략적으로 추정하는 것입니다.

NetApp 스토리지 시스템 크기 조정 지침

		Throughput ²	Raw capacity (typical ³ / max)	Connectivity	# NVIDIA A100 GPUs supported ⁴	# NVIDIA H100 GPUs supported ⁵
NetApp® AFF A1K	1 HA pair ¹	56 GB/s	368TB / 14.7PB	200 GbE	1-160	1-80
	12 HA pairs	672 GB/s	4.4PB / 176.4PB		1920	960
AFF A90	1 HA pair	46 GB/s	368TB / 6.6PB	200 GbE	1 – 128	1-64
	12 HA pairs	552 GB/s	4.4PB / 79.2PB		1536	768
AFF A70	1 HA pair	21 GB/s	368TB / 6.6PB	200 GbE	1-48	1-24
	12 HA pairs	252 GB/s	4.4PB / 79.2PB		576	288

NVIDIA DGX 시스템을 탑재한 NVA-1173 NetApp AIPOd - 결론 및 추가 정보

이 섹션에는 NVIDIA DGX 시스템을 탑재한 NetApp AIPOd 에 대한 추가 정보에 대한 참조가 포함되어 있습니다.

결론

DGX BasePOD 아키텍처는 동등하게 고급 스토리지 및 데이터 관리 기능을 요구하는 차세대 딥 러닝 플랫폼입니다. DGX BasePOD를 NetApp AFF 시스템과 결합하면 NetApp AIPOd 와 DGX 시스템 아키텍처를 거의 모든 규모로 구현할 수 있습니다. NetApp ONTAP 의 탁월한 클라우드 통합 및 소프트웨어 정의 기능과 결합된 AFF 성공적인 DL 프로젝트를 위해 에지, 코어 및 클라우드를 아우르는 광범위한 데이터 파이프라인을 지원합니다.

추가 정보

이 문서에 설명된 정보에 대해 자세히 알아보려면 다음 문서 및/또는 웹사이트를 참조하세요.

- NetApp ONTAP 데이터 관리 소프트웨어 - ONTAP 정보 라이브러리

["https://docs.netapp.com/us-en/ontap-family/"](https://docs.netapp.com/us-en/ontap-family/)

- NetApp AFF A90 스토리지 시스템-

<https://www.netapp.com/pdf.html?item=/media/7828-ds-3582-aff-a-series-ai-era.pdf>

- NetApp ONTAP RDMA 정보-

["https://docs.netapp.com/us-en/ontap/nfs-rdma/index.html"](https://docs.netapp.com/us-en/ontap/nfs-rdma/index.html)

- NetApp DataOps 툴킷

["https://github.com/NetApp/netapp-dataops-toolkit"](https://github.com/NetApp/netapp-dataops-toolkit)

- NetApp Trident

"개요"

- NetApp GPUDirect 스토리지 블로그-

["https://www.netapp.com/blog/ontap-reaches-171-gpudirect-storage/"](https://www.netapp.com/blog/ontap-reaches-171-gpudirect-storage/)

- NVIDIA DGX BasePOD

["https://www.nvidia.com/en-us/data-center/dgx-basepod/"](https://www.nvidia.com/en-us/data-center/dgx-basepod/)

- NVIDIA DGX H100 시스템

["https://www.nvidia.com/en-us/data-center/dgx-h100/"](https://www.nvidia.com/en-us/data-center/dgx-h100/)

- NVIDIA 네트워킹

["https://www.nvidia.com/en-us/networking/"](https://www.nvidia.com/en-us/networking/)

- NVIDIA Magnum IO™ GPUDirect® 스토리지

["https://docs.nvidia.com/gpudirect-storage"](https://docs.nvidia.com/gpudirect-storage)

- NVIDIA 베이스 명령

["https://www.nvidia.com/en-us/data-center/base-command/"](https://www.nvidia.com/en-us/data-center/base-command/)

- NVIDIA Base Command Manager

["https://www.nvidia.com/en-us/data-center/base-command/manager"](https://www.nvidia.com/en-us/data-center/base-command/manager)

- NVIDIA AI 엔터프라이즈

["https://www.nvidia.com/en-us/data-center/products/ai-enterprise/"](https://www.nvidia.com/en-us/data-center/products/ai-enterprise/)

감사의 말

이 문서는 NetApp Solutions 및 ONTAP Engineering 팀(David Arnette, Olga Kornievskaia, Dustin Fischer, Srikanth Kaligotla, Mohit Kumar, Raghuram Sudhaakar)의 작품입니다. 저자는 또한 지속적인 지원을 해주신 NVIDIA 와 NVIDIA DGX BasePOD 엔지니어링 팀에 감사드리고 싶습니다.

저작권 정보

Copyright © 2025 NetApp, Inc. All Rights Reserved. 미국에서 인쇄된 본 문서의 어떠한 부분도 저작권 소유자의 사전 서면 승인 없이는 어떠한 형식이나 수단(복사, 녹음, 녹화 또는 전자 검색 시스템에 저장하는 것을 비롯한 그래픽, 전자적 또는 기계적 방법)으로도 복제될 수 없습니다.

NetApp이 저작권을 가진 자료에 있는 소프트웨어에는 아래의 라이선스와 고지사항이 적용됩니다.

본 소프트웨어는 NetApp에 의해 '있는 그대로' 제공되며 상품성 및 특정 목적에의 적합성에 대한 명시적 또는 묵시적 보증을 포함하여(이에 제한되지 않음) 어떠한 보증도 하지 않습니다. NetApp은 대체품 또는 대체 서비스의 조달, 사용 불능, 데이터 손실, 이익 손실, 영업 중단을 포함하여(이에 국한되지 않음), 이 소프트웨어의 사용으로 인해 발생하는 모든 직접 및 간접 손해, 우발적 손해, 특별 손해, 징벌적 손해, 결과적 손해의 발생에 대하여 그 발생 이유, 책임론, 계약 여부, 엄격한 책임, 불법 행위(과실 또는 그렇지 않은 경우)와 관계없이 어떠한 책임도 지지 않으며, 이와 같은 손실의 발생 가능성이 통지되었다 하더라도 마찬가지입니다.

NetApp은 본 문서에 설명된 제품을 언제든지 예고 없이 변경할 권리를 보유합니다. NetApp은 NetApp의 명시적인 서면 동의를 받은 경우를 제외하고 본 문서에 설명된 제품을 사용하여 발생하는 어떠한 문제에도 책임을 지지 않습니다. 본 제품의 사용 또는 구매의 경우 NetApp에서는 어떠한 특허권, 상표권 또는 기타 지적 재산권이 적용되는 라이선스도 제공하지 않습니다.

본 설명서에 설명된 제품은 하나 이상의 미국 특허, 해외 특허 또는 출원 중인 특허로 보호됩니다.

제한적 권리 표시: 정부에 의한 사용, 복제 또는 공개에는 DFARS 252.227-7013(2014년 2월) 및 FAR 52.227-19(2007년 12월)의 기술 데이터-비상업적 품목에 대한 권리(Rights in Technical Data -Noncommercial Items) 조항의 하위 조항 (b)(3)에 설명된 제한사항이 적용됩니다.

여기에 포함된 데이터는 상업용 제품 및/또는 상업용 서비스(FAR 2.101에 정의)에 해당하며 NetApp, Inc.의 독점 자산입니다. 본 계약에 따라 제공되는 모든 NetApp 기술 데이터 및 컴퓨터 소프트웨어는 본질적으로 상업용이며 개인 비용만으로 개발되었습니다. 미국 정부는 데이터가 제공된 미국 계약과 관련하여 해당 계약을 지원하는 데에만 데이터에 대한 전 세계적으로 비독점적이고 양도할 수 없으며 재사용이 불가능하며 취소 불가능한 라이선스를 제한적으로 가집니다. 여기에 제공된 경우를 제외하고 NetApp, Inc.의 사전 서면 승인 없이는 이 데이터를 사용, 공개, 재생산, 수정, 수행 또는 표시할 수 없습니다. 미국 국방부에 대한 정부 라이선스는 DFARS 조항 252.227-7015(b)(2014년 2월)에 명시된 권한으로 제한됩니다.

상표 정보

NETAPP, NETAPP 로고 및 <http://www.netapp.com/TM>에 나열된 마크는 NetApp, Inc.의 상표입니다. 기타 회사 및 제품 이름은 해당 소유자의 상표일 수 있습니다.