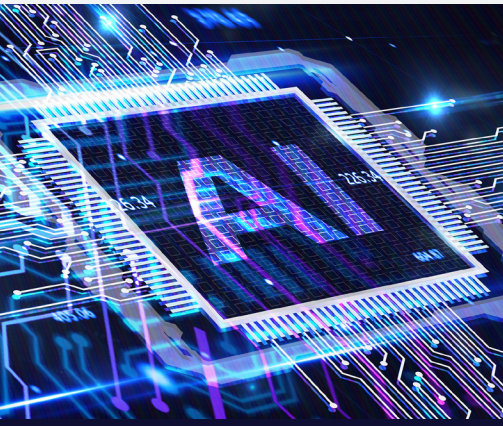


핵심 요약



Dell과 관련하여
조사에서 확인된 사항:



가장 광범위한 Generative
AI 포트폴리오* 보유



MLPerf® 결과에 따르면,
NVIDIA H100 SXM5 GPU
4개를 탑재한 Dell
PowerEdge XE8640 서버는
9가지 범주에서 제출된 모든
GPU 4개 서버 가운데 가장
높은 AI 처리량을 달성했다.



AI에 맞춤화된 광범위한
전문 서비스 포트폴리오

Dell AI 포트폴리오로 AI 워크로드의 당면 과제 해결

Dell AI 포트폴리오와 HPE의 유사 오퍼링 비교

AI 도입은 기업과 데이터 센터 IT 직원에게 다음과 같은 새로운 당면 과제를 안겨준다.

- 교육 또는 외부 채용을 통해 기존 인력에 존재하던 기술 격차 해소
- 기업 데이터의 품질, 수량, 위치 및 현재 상태를 포함하여 AI의 데이터 준비 요구 사항 이해
- 기업 목표를 진단하여 어떤 AI 모델 및 구현이 이점을 제공할지 현명하게 판단
- 계획된 AI 시스템의 컴퓨팅, 네트워킹 및 스토리지 요구 사항을 진단하고 이러한 시스템 확보

Dell Technologies와 HPE 같은 인프라스트럭처 공급업체는 AI 지원 포트폴리오를 통해 AI 수명주기 전반을 포괄하는 통합 솔루션을 제공한다. 본 연구에서는 Dell과 HPE AI 포트폴리오 모두에 대해 공개된 정보를 살펴보고 MLPerf®의 벤치마크 테스트를 조사했다. Dell은 고성능 서버 및 스토리지 옵션, 검증된 솔루션, 계획부터 운영까지 프로세스를 안내하는 전문 서비스를 통해 기업이 AI를 적극적으로 수용할 수 있도록 지원할 준비가 되어 있는 것으로 확인되었다.

AI 모델 벤치마크 성능: MLPerf 결과 비교

공개된 MLPerf® 벤치마크 테스트에 따르면 Dell AI 포트폴리오의 오퍼링은 AI 워크로드에 일관된 강력한 성능을 제공한다. MLPerf®는 여러 AI 모델에서 학습과 추론 모두에 대한 성능을 테스트한다. 이 보고서에 언급된 데이터는 2023년 11월 MLCommons® 웹사이트에 게시된 MLPerf® v3.1 추론 데이터 센터 결과를 기반으로 한다.¹ GPU 8개 및 GPU 4개 서버를 모두 비교했으며, 본 요약에 각 범주의 결과를 하나씩 포함했다. 자세한 내용은 [여기](#)에서 보고서 전문을 확인할 수 있다.

*Dell 분석 기준, 2023년 8월. Dell Technologies는 워크스테이션 PC(모바일 및 고정형)부터 HPC(High Performance Computing) 서버, 데이터 스토리지, 클라우드 네이티브 소프트웨어 정의 인프라스트럭처, 네트워킹 스위치, 데이터 보호, HCI 및 서비스에 이르기까지 AI 워크로드를 지원하도록 엔지니어링된 솔루션을 제공합니다. 데이터 출처: Robert McNeal, "Dell, VMware and NVIDIA Bring AI to Your Data" (2024년 1월 17일 열람), <https://www.dell.com/en-us/blog/dell-vmware-and-nvidia-bring-ai-to-your-data/>.

1. MLCommons, "MLPerf Inference: Datacenter Benchmark Suite Results", 2023년 12월 12일 열람, <https://mlcommons.org/en/inference-datacenter-31/>.

GPU 8개 서버에 대한 MLPerf 성능 비교

GPU 8개 서버에 대한 MLPerf® v3.1 결과에서 NVIDIA SXM5 H100 GPU를 탑재한 Dell PowerEdge XE9680은 NVIDIA SXM4 A100 GPU를 탑재한 HPE ProLiant XL675d Gen10 Plus보다 최대 4.25배 뛰어난 성능을 보였다(그림 1 참조).

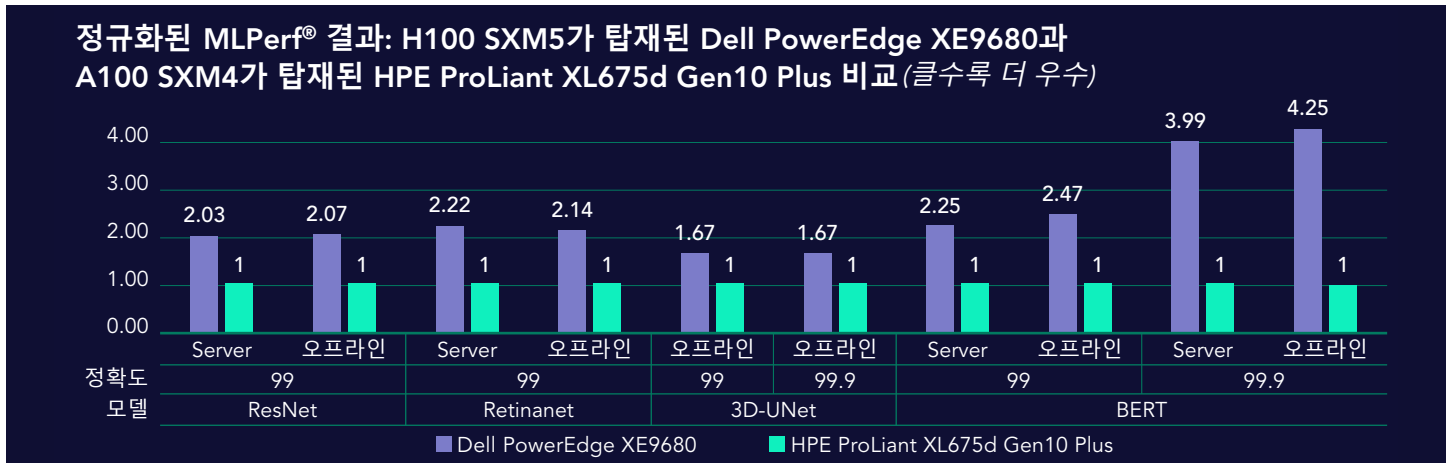


그림 1: 2023년 11월 29일 현재 Dell PowerEdge XE9680 및 HPE ProLiant XL675d Gen10 Plus에 대해 공개된 MLPerf® 결과. Dell 시스템은 NVIDIA H100 GPU를 사용하지만 HPE 시스템의 GPU는 한 세대 이전 버전이다. 출처: MLCommons®의 데이터를 사용하는 Principled Technologies.*

GPU 4개 서버에 대한 MLPerf 성능 비교

NVIDIA H100 SXM5 GPU 4개를 탑재한 Dell PowerEdge XE8640 서버는 9가지 범주에서 제출된 모든 GPU 4개 서버 가운데 가장 높은 AI 처리량을 달성했다. 그림 2에서 볼 수 있듯, MLPerf® 벤치마크에서 HPE ProLiant DL380a 서버와 비교하여 최대 2.07배 높은 점수를 받았다.

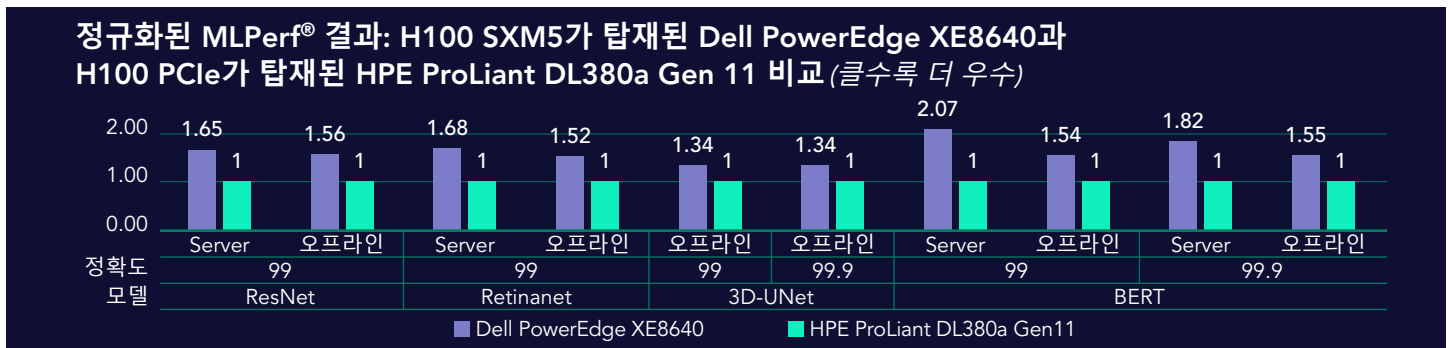


그림 2: 2023년 11월 29일 현재 Dell PowerEdge XE8640 및 HPE ProLiant DL380a Gen11에 대해 공개된 MLPerf® 결과. Dell 시스템은 NVIDIA H100 SXM 폼 팩터를 사용하고, HPE 시스템은 덜 강력한 PCIe 폼 팩터를 사용한다. 출처: MLCommons®의 데이터를 사용하는 Principled Technologies.*

요약: AI 성공에 기여하는 추가 요인

AI 모델이 성공하려면 고성능 서버 이상의 비결이 필요하다. AI 솔루션을 계획, 준비, 배포 및 관리하려면 비정형 데이터용 스토리지와 전문 서비스도 고려해야 한다. Dell 포트폴리오는 파일 스토리지용 PowerScale Series와 오브젝트 스토리지용 Elastic Cloud Storage 또는 ObjectScale 스토리지를 갖춘 AI 데이터 세트용 스토리지를 제공한다.

조직은 AI를 위한 Dell의 전문 서비스 및 컨설팅 서비스의 이점을 누릴 수 있다. 이 서비스에는 HPE에서 제공하지 않는 데이터 준비와 같은 일부 서비스도 포함된다. 또한 Dell 포트폴리오에는 Validated Designs for AI가 포함되어 있어 AI 설계 및 배포의 불확실성을 해소한다.

*v3.1 Inference Closed의 검증된 MLPerf. 점수.

<https://mlcommons.org/benchmarks/inference-datacenter/> 2023년 12월 5일, 항목 3.1-0069, 3.1-0085, 3.1-0067 및 3.1-0084에서 발췌. MLPerf. 이름과 로고는 MLCommons의 등록 및 미등록 상표. 미국 및 다른 국가에 있는 협회. All rights reserved. 무단 사용을 엄격히 금지합니다. 자세한 내용은 www.mlcommons.org를 참조하십시오.

▶ 이 보고서의 원본(영문) 보기:
<https://facts.pt/PiSgyY7>

보고서 보기 ▶



Facts matter.®

Principled Technologies는 Principled Technologies, Inc.의 등록 상표입니다. 다른 모든 제품 이름은 해당 소유주의 상표입니다. 자세한 내용은 보고서를 검토하십시오.