GPU를 활용한 HPE AI 및 HPC 전략

일시: 2021. 10. 21(목), 10:30 ~ 11:30

플랫폼: 디지털데일리(DD튜브)

GPU를 활용한 HPE AI 및 HPC 전략

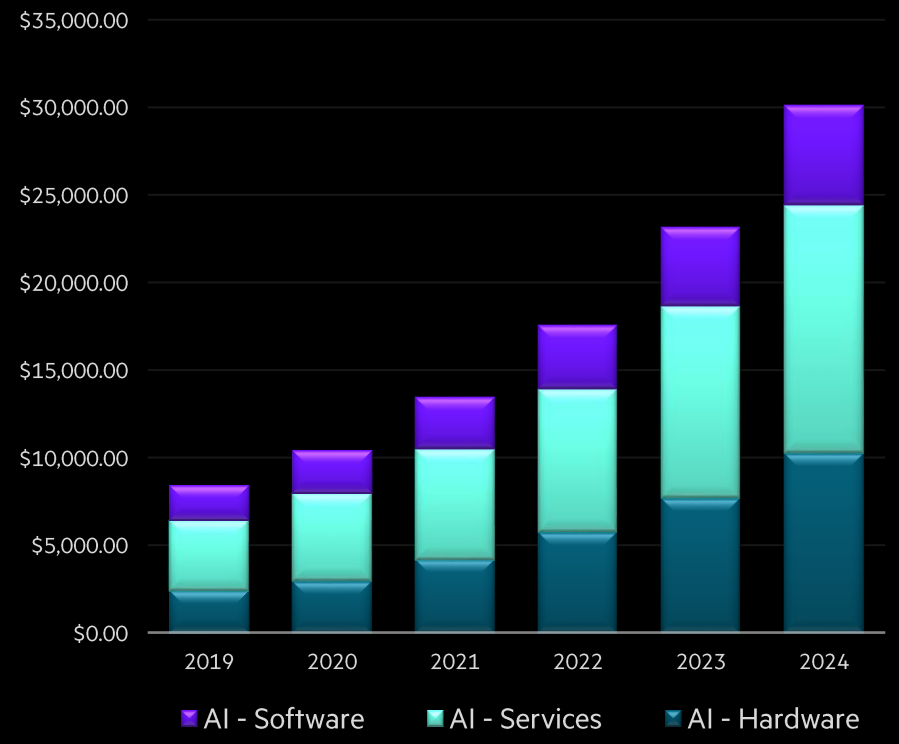
엄현필 부장 | Hewlett Packard Enterprise

AI TRENDS

CAGR ’20-’24

29%

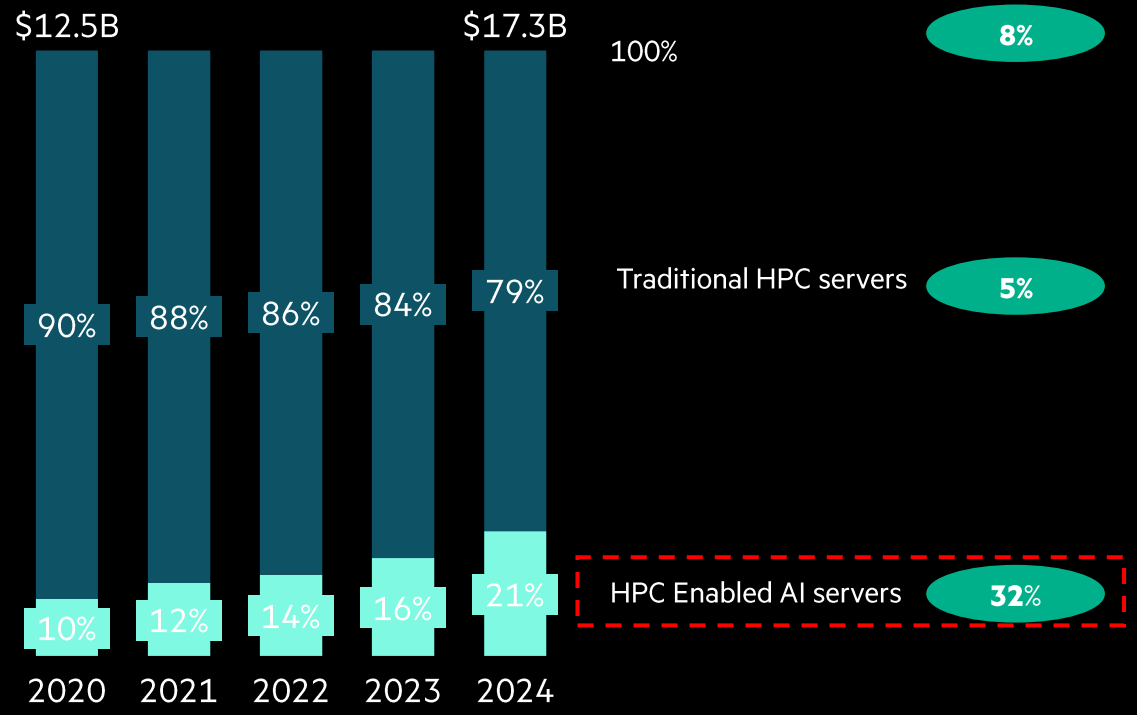
AI at the Edge Spending (WW)



* 고객은 AI를 비즈니스 프로세스(자체 모델 또는 ISV)에 통합하려는 의도가 보임

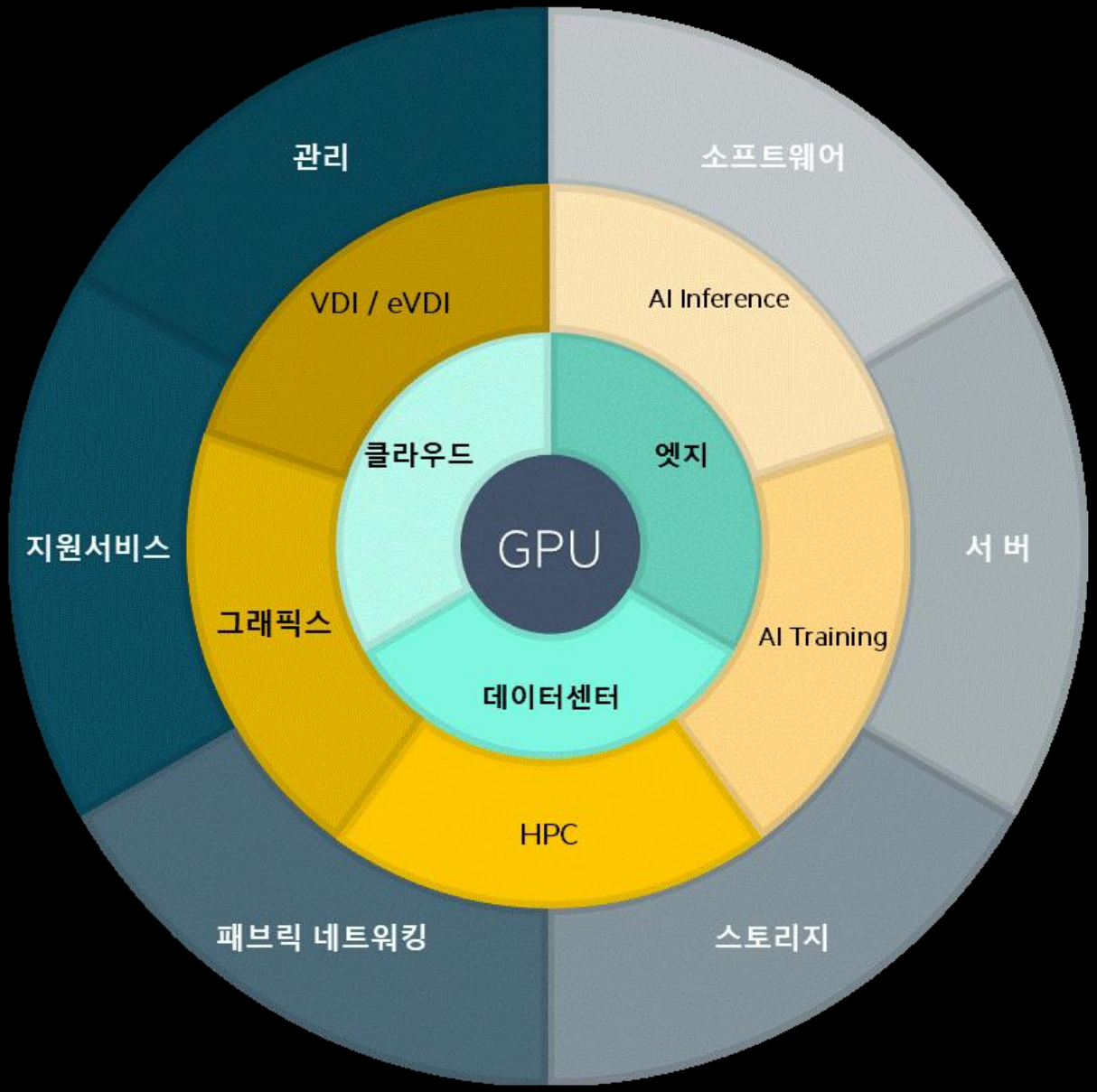
HPC server market size[1], ($B) (WWeC)

CAGR ’20-’24



* 고객은 최적화된 AI 모델 제공으로 생산성을 높여가고 있는 것을 예측할 수 있음

HPE의 GPU를 활용한 인프라 및 전략



700+ GPU-ACCELERATED HPC 어플리케이션

ARTIFICIAL INTELLIGENCE

44 apps

Including:
- Caffe2
- MXNet
- Tensorflow

COMP. FINANCE

16 apps

Including:
- O-Quant Options Pricing
- MUREX
- MISYS

CLIMATE & WEATHER

5 apps

Including:
- Cosmos
- Gales
- WRF

DATA SCI. & ANALYTICS

20 apps

Including:
- Anaconda
- H20
- OmniSci

FEDERAL & DEFENSE

15 apps

Including:
- ArcGIS Pro
- EVNI
- SocetGXP

LIFE SCIENCES

107 apps

Including:
- Amber
- LAMMPS
- GROMACS
- NAMD
- Relion
- VASP

MFG, CAD, & CAE

140 apps

Including:
- Ansys Fluent
- Abaqus SIMULIA
- AutoCAD
- CST Studio Suite

MEDIA & ENT.

167 apps

Including:
- DaVinci Resolve
- Premiere Pro CC
- Redshift Renderer

MEDICAL IMAGING

8 apps

Including:
- aidoc
- PowerGrid
- RadiAnt

OIL & GAS

19 apps

Including:
- Echelon
- RTM
- SPECFEM3D

RESEARCH: HER and SC

45 apps

Including:
- Chroma
- GTC
- MILC
- QUDA
- XGC

SAFETY, TOOLS & OTHER

16 apps

Including:
- Cyllance
- FaceControl
- Bright Cluster Manager
- HPCtoolkit

GPU를 활용한 국내 고객 적용 사례 및 워크로드

서비스 제공자

- 포털 게임사 & 클라우드 서비스사

1. 이미지/비디오 검색/분류
2. 음성 인식
3. 자연어 처리
4. 기사 추천
5. 언어 번역
6. 빅데이터 분석

생명과학

- 국가 의료 기관 / 대학 및 대형 병원

1. 영상 진단
2. 환자 데이터 분석 및 위험도 측정
3. 챗봇을 통한 의료 상담

통신/미디어/쇼핑

- Telco 3사 및 대형 쇼핑 / 미디어

1. 상품 추천
2. 이미지 검색
3. 자연어 처리
4. 실시간 번역
5. 콜센터

공공

- 국가정부기관 및 공공기관

1. 지능형 CCTV
2. 지능형 범죄 예방
3. 공공 서비스 챗봇
4. 학교 연구 프로젝트

제조

- 국내 Enterprise 전자 & 제조회사 및 자동차회사

1. 자율 주행
2. 스마트 팩토리
3. 영상 분석을 통한 결함 분석
4. 빅데이터 분석

금융

- 국/내외 메이저 은행 및 보험 / 증권사

1. 부정 거래 방지
2. 로보 어드바이저 (추천, 거래 등)
3. 챗봇 / 콜센터
4. 빅데이터 분석

GPU 워크플로우를 지원하는 시스템 권고

카테고리	AI Training	AI Inference	HPC	그래픽스	VDI / eVDI
CPU	2.4GHz 이상의 CPU 48 코어 이상의 CPU	2.4GHz 이상의 CPU 24 코어 이상의 CPU	2.5GHz 이상의 CPU 가속기 당 4~8 코어	2.6GHz 이상의 CPU 12 코어 이상의 CPU	OS : 사용자 : 코어 Win7 : 5 : 1 Win10 : 3 : 1 / 2 : 1
MEMORY	2.5x Accelerator Memory/CPU 2 DIMMs/Ch, Dual Rank, Best 1 DIMM/Ch, Dual Rank, Better 2 DIMMs/Ch, Single Rank, Good	2.5x Accelerator Memory/CPU 2 DIMMs/Ch, Dual Rank, Best 1 DIMM/Ch, Dual Rank, Better 2 DIMMs/Ch, Single Rank, Good	2.5x Accelerator Memory/CPU 2 DIMMs/Ch, Dual Rank, Best 1 DIMM/Ch, Dual Rank, Better 2 DIMMs/Ch, Single Rank, Good	1-2x Accelerator Memory 2 DIMMs/Ch, Dual Rank, Best 1 DIMM/Ch, Dual Rank, Better 2 DIMMs/Ch, Single Rank, Good	**2-16GB/사용자/VM** 2 DIMMs/Ch, Dual Rank, Best 1 DIMM/Ch, Dual Rank, Better 2 DIMMs/Ch, Single Rank, Good
FABRIC	50-200 GbE 1x/2xGPU @200GbE, Best 1x/2xGPU @100GbE, Better 1x/GPU@50GbE, Good	25-100 GbE 1x / 소켓	100-200 GbE 1-2x / 시스템	10 -25 GbE 1x / 시스템	10-25 GbE 1x / 시스템
STORAGE	NVMe, Best SSD, Better SAS/SATA, Good	NVMe, Best SSD, Better SAS/SATA, Good	NVMe, Best SSD, Better SAS/SATA, Good	NVMe, Best SSD, Better SAS/SATA, Good	NVMe, Best SSD, Better SAS/SATA, Good

다양한 워크로드를 위한 PCIE GPU

딥러닝 트레이닝

과학적 연구

데이터 분석

최고 성능의 컴퓨팅AI, HPC, 데이터 처리

Fastest Compute, FP64
Up to 7 MIG instances

랭귀지 프로세싱

대화형 AI

추천 시스템

AI 추론 및 메인 스트림 컴퓨팅

Versatile Mainstream Compute
FP64, Up to 4 MIG instances

엣지 AI

엣지 비디오

모바일 클라우드 게임

소규모 데이터 센터 및 엣지 추론

High density Video & Graphics
Compact & Versatile

가상화 워크스테이션

화상회의

4K 클라우드 게임

메인 스트림의 AI 그래픽과 비디오

4K Cloud Gaming
Graphics, Video with AI

클라우드 렌더링

클라우드 XR

옴니버스

최고 성능의 그래픽과 시각화 컴퓨팅

Fastest RT Graphics
Largest render models

가상화 데스크탑

트랜스코딩

초고집적 가상화 데스크탑

4K Resolution
Max # of encode/decode streams

A100
250W & 300W | 40G & 80G
2-slot FHFL | NVLINK

A30
165W | 24GB
2-slot FHFL | NVLINK

A16
250W | 4 x 16GB
2-slot FHFL

A40
300W | 48GB
2-slot FHFL | NVLINK

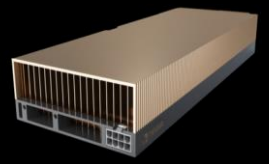
컴퓨트

그래픽스

어플리케이션과 ACCELERATOR OFFERINGS BY APPLICATION

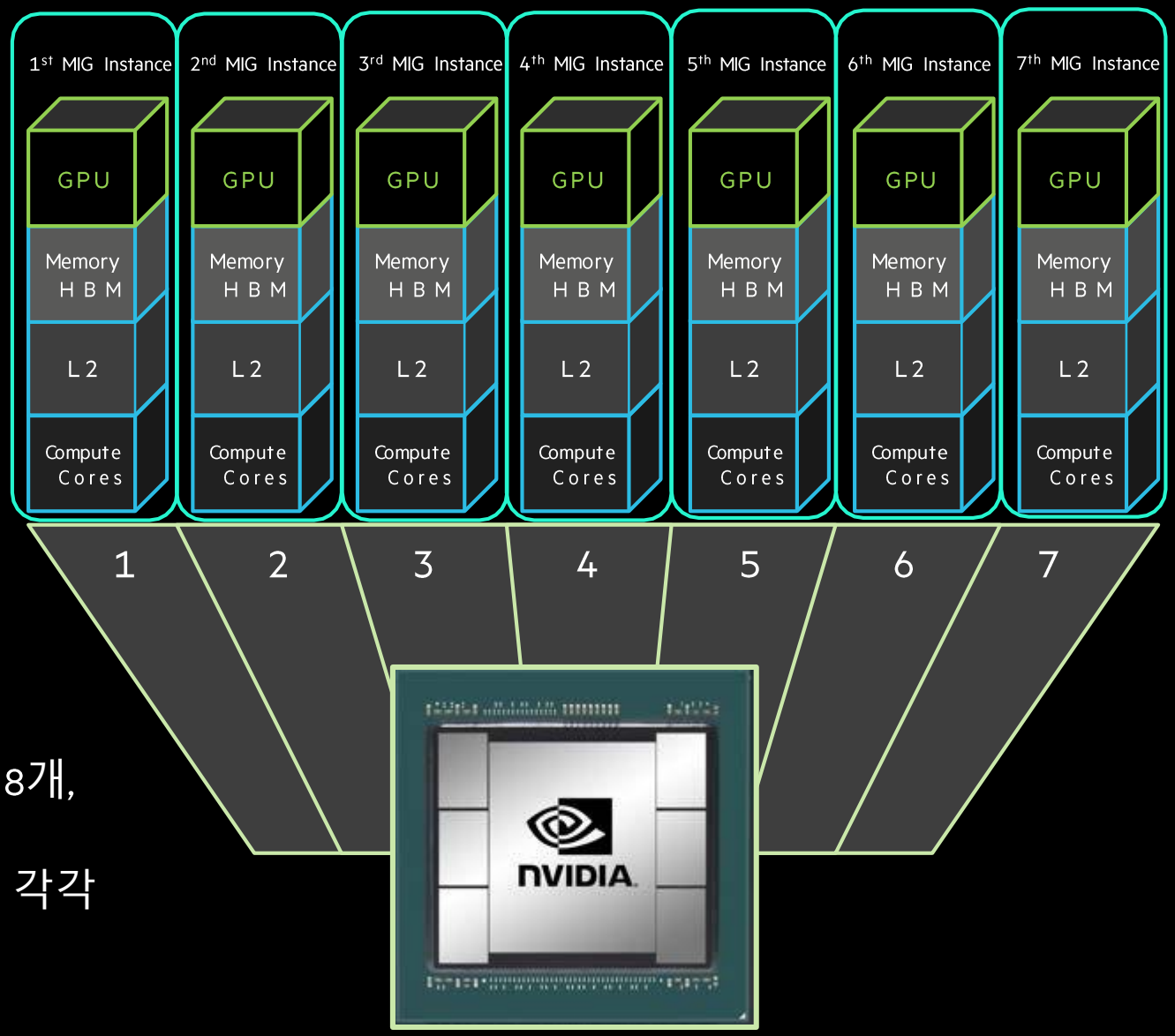
카테고리	워크로드	NVIDIA A100 SXM | PCIe	NVIDIA A40	NVIDIA A30	NVIDIA A16
		하이 퍼포먼스 컴퓨팅(HPC)	고성능 그래픽	메인스트림 컴퓨팅	주요 그래픽과 VDI를 위한 저전력 및 최적화
		서버당 GPU의 권고 수량			
딥러닝 트레이닝 데이터 분석	가장 빠른 모델 훈련과 분석에 사용	SXM PCIe 4-8개의 GPU 80GB 이상의 파라미터 모델(DLRM, GPT-2)			
딥러닝 추론	배치 작업과 실시간 추론에 사용	SXM PCIe Multi-Instance GPU (MIG)를 활용한 1-2개의 GPU 80GB 이상의 대형 배치사이즈 모델 (RNN-T)		Multi-Instance GPU (MIG)를 활용한 2-4개의 GPU	4-8개의 GPU
High Performance Computing(HPC) & AI	고등 교육 연구 및 과학 컴퓨팅 센터용으로 사용	SXM Multi-Instance GPU (MIG)를 활용한 1-4개의 GPU		Multi-Instance GPU (MIG)를 활용한 2-4개의 GPU	
렌더 팜	배치 작업과 실시간 렌더링에 사용		4-8개의 GPU		
그래픽스	전문가용 VDI를 위한 최고의 그래픽 성능을 위해 사용		고성능 가상 워크스테이션을 위한 2-4개의 GPU		보급형/ 고성능 가상 워크스테이션을 위한 2-8 GPU
클라우드 게임	4K 해상도와 안드로이드에 사용		4-8개의 GPU (4K 해상도)		1-2개 GPU (안드로이드) 4-8개의 GPU (4K 해상도)
엔터프라이즈 가속화	그래픽, 머신러닝, 딥러닝, 분석, 트레이닝 및 추론을 포함한 혼합 워크로드를 위해 사용	PCIe 컴퓨트 워크로드를 위한 Multi-Instance GPU (MIG) 활용 1-2개의 GPU	고성능 그래픽 워크로드를 위한 1-2개의 GPU	컴퓨트 워크로드를 위한 Multi-Instance GPU (MIG) 활용 1-2개의 GPU	고성능 그래픽 워크로드와 컴퓨트 워크로드를 위한 1-4개 GPU
엣지 가속	다양한 사례 및 배포 장소를 위해 사용	PCIe Multi-Instance GPU (MIG)를 활용한 1-2개의 GPU	고성능 그래픽 워크로드와 AR / VR을 위한 1-4개의 GPU	Multi-Instance GPU (MIG)를 활용한 1-2개의 GPU	추론과 비디오 워크로드를 위한 1-8개 GPU

워크로드별 GPU TRANSTIONS

워크로드 타입	운영중인 GPU	전환될 GPU	소프트웨어
딥러닝 트레이닝, 추론, HPC, AI, 데이터 사이언스	NVIDIA V100, V100S, P100, T4	NVIDIA A30 NVIDIA A100 	**NVIDIA AI Enterprise** NVIDIA Virtual Compute (vCS)
중 / 대형 가상화 워크스테이션	RTX 8000, RTX 6000, RTX 4000, T4, NVIDIA P100, P40, M60	NVIDIA A40 	NVIDIA RTX Virtual Workstation (vWS)
사무 생산성, 스트리밍 비디오, 소형 가상 워크스테이션	NVIDIA RTX 4000, T4, P6, P4, M10, M6, GRID K1, K2	NVIDIA A16 	NVIDIA Virtual PC (vPC) NVIDIA RTX Virtual Workstation (vWS)

MIG (MULTI INSTANCE GPU)

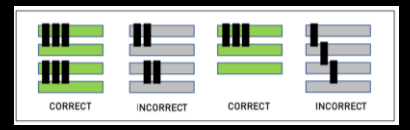
- A100 GPU, 최대 7개 GPU Slice로 분할
- A30 GPU, 최대 4개 GPU Slice로 분할
- 목적: GPU 사용률 극대화
- 사용 사례
  - A100 GPU 1장 미만 워크로드
  - A100 GPU 여러 개를 활용한 사용자 할당
  - 경량 학습, 추론, 개발, 일부 HPC
- 혜택
  - 분할된 MIG, H/W 독립, QoS 보장
  - Apollo 6500 Gen10 Plus 기준 SXM4 타입의 A100 8개, PCIe 타입 10개 지원
  - A100, A30 PCIe Type GPU는 10개 장착 가능하며 각각 최대 70개와 40개의 GPU Instance 생성 가능



PCIE GPU를 연결하는 GPU-TO-GPU BRIDGING

NVIDIA A100 40GB PCIe GPU

- NVIDIA Ampere NVLink 2x2 Bridge
- 최대 NVLink 대역폭 = 600GB/s
- 인접한 두 GPU 간의 빠른 메모리 공유 가능
- 8개의 A100 GPU를 가지고 2개씩 연결하여 4개의 연결된 GPU 쌍을 제공 가능
- A100 GPU의 각 쌍당 3개의 브리지 필요

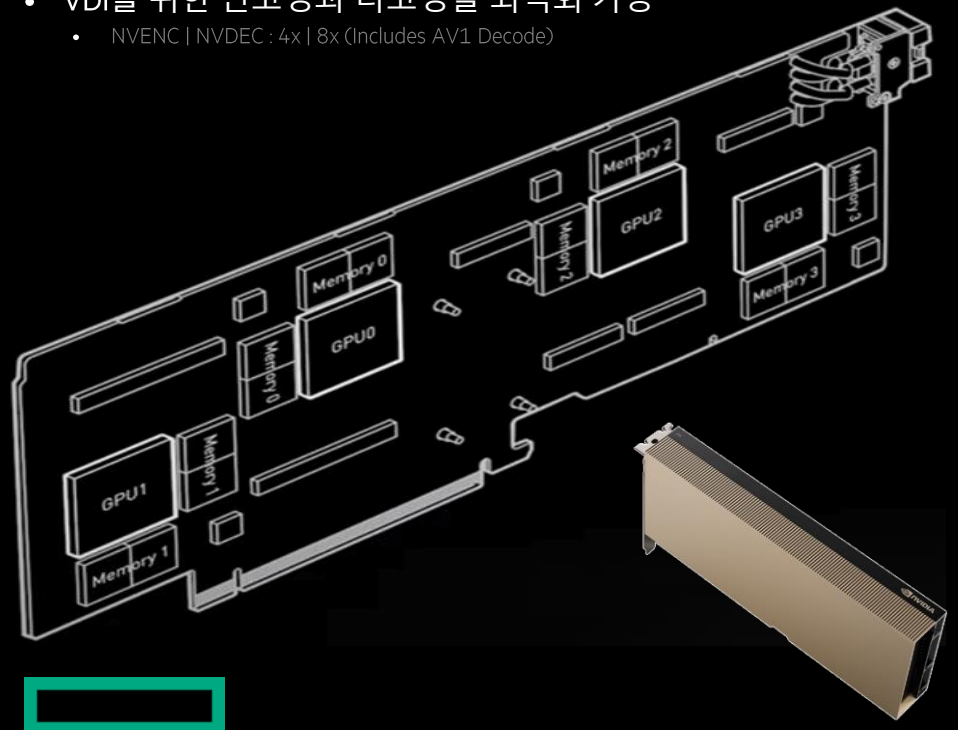


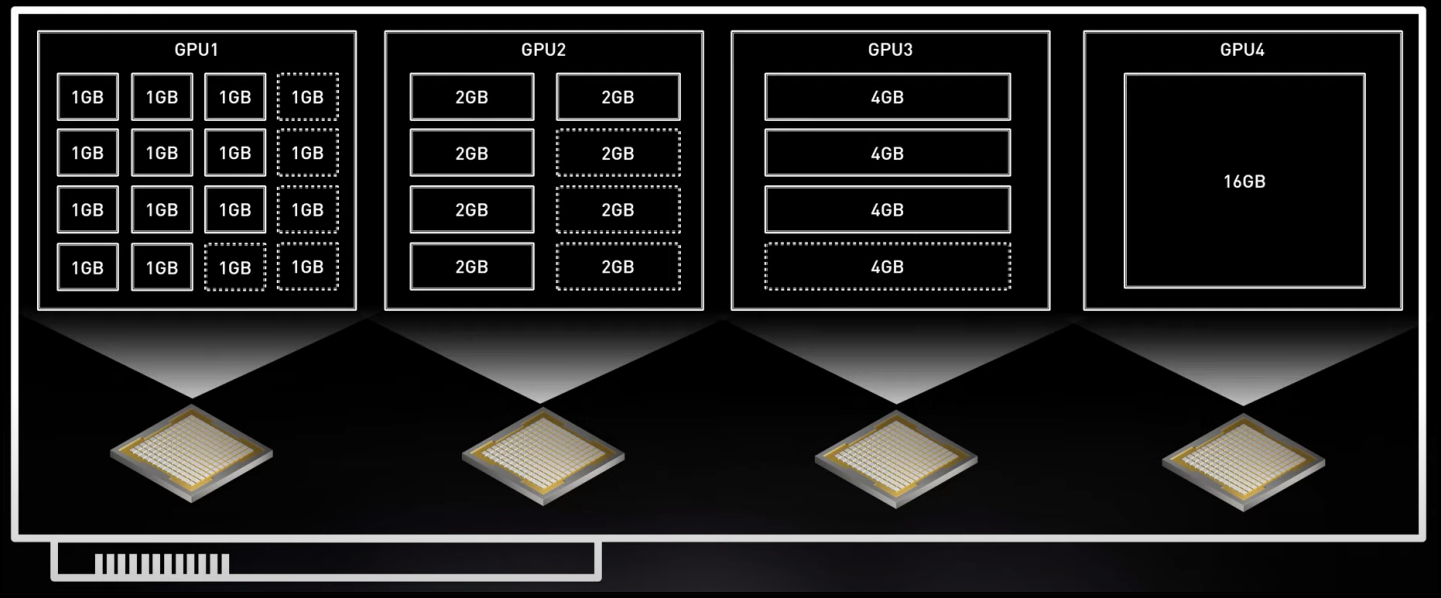
NVIDIA A40 48GB PCIe GPU

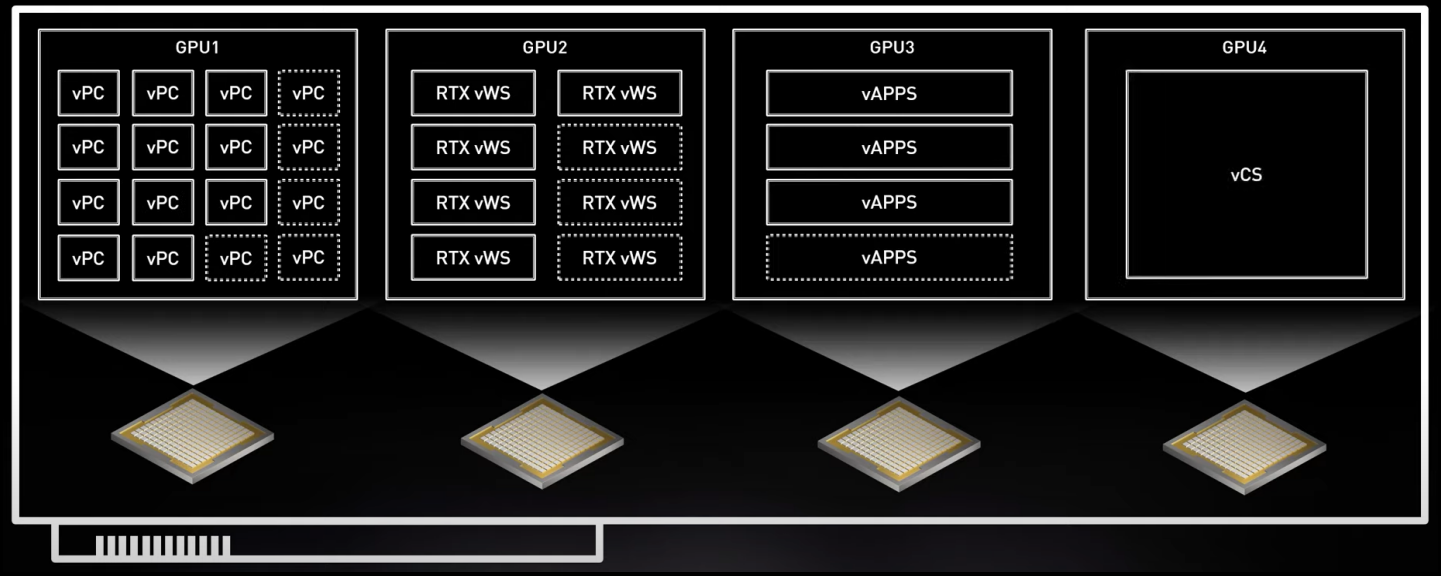
- NVIDIA Ampere NVLink 2x2 Bridge
- 최대 NVLink 대역폭 = 200GB/s
- 인접한 두 GPU 간의 빠른 메모리 공유 가능
- 4개의 A40 GPU를 가지고 2개씩 연결하여 2개의 연결된 GPU 쌍을 제공 가능
- A100 GPU의 각 쌍당 1개의 브리지 필요

NVIDIA AMPERE GPU – A16

- A16에는 총 4개의 개별 GPU로 구성
- 각 GPU는 16GB 메모리가 별도로 존재
  - 64 GB GDDR6 with Error Correcting Code (ECC) (16 GB per GPU)
- 4개의 GPU가 Mellanox PCIe Switch로 연결
  - 4x 232 GB/s
- GPU별 구성을 워크로드에 맞게 분리하여 구성 가능
- 사용자에 맞는 워크로드에 맞춰 Profile을 구성
- VDI를 위한 인코딩과 디코딩을 최적화 가능
  - NVENC | NVDEC : 4x | 8x (Includes AV1 Decode)







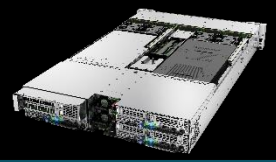
GPU를 활용할 수 있는 고객 워크로드 및 HPE 플랫폼

시장 및 워크로드	지원 범주	GPU 종류	HPE의 적용 서버 플랫폼
1. HPC : • 시뮬레이션 • 통합된 데이터 세트	공공 기관의 컴퓨팅 센터나 국가 연구소 및 Enterprise HPC 워크로드, 시뮬레이션 및 AI (Training/Inference) 및 데이터 사이언스를 지원	4x SXM NVLink (GPU) 4x OAM Accelerators (GPU/ASIC) 2-4 PCIE Accelerators (GPU/ASIC/FPGA)	Apollo 6500 (XL270d/XL645d) Apollo 2000 (XL190r/XL220n) DL380 / 385 Gen10 Plus & v2 Cray Supercomputer
2. AI : • 딥러닝 / 머신러닝 중심 학습 • 추론(Inference)	로컬 또는 분산 교육에 최적화된 트레이닝 목적의 클러스터 구축. 데이터센터, 엣지, 니어 엣지의 추론용을 지원.	8x SXM NVLink (GPU) 8x OAM Accelerators (GPU/ASIC) 1-4 lower power PCIE (GPU/FPGA)	Apollo 6500 (XL270d/XL675d) Apollo 2000, DL380 / 385 Gen10 Plus & v2 EL1000/4000/8000 Cray Supercomputer
3. 데이터센터 그래픽스 : • AR / VR • 렌더링	복합적인 워크로드, 원격 그래픽, AI, 트레이닝 및 추론, 비디오 분석, AR / VR, 애니메이션, 스트리밍 서비스에 지원.	2-4x Mid-High End Graphics (GPU) 2-4x High-Eng Compute (GPU) + Virtual software	Apollo 6500 (XL645d) Apollo 2000 (XL190r/XL220n) DL380 / 385 Gen10 Plus & v2 EL8000 e910 Synergy
4. Enterprise 고객 : • VDI / eVDI • 워크로드 가속화	NVIDIA GRID (vAPPS, vPC, vDWS, vCS)를 적용한 사용자 당 TCO를 최적화하여 가상 데스크톱 환경에 지원. GPU를 사용하여 혼합 워크로드를 처리	2-4 Dense VDI specific (GPU) 2-8 Low power multi purpose (GPU)	Apollo 2000 (XL190r/XL220n) DL380 / 385 Gen10 Plus & v2, ML350 Synergy
5. Edge : • 엣지 / 니어 엣지 • 스마트 팩토리 • 스마트 팜	니어엣지 (데이터 센터) / 에지 (강력한 환경) 솔루션. 다양한 고객 사례를 통한 고객 지원.(비디오 디코딩 / 분석 / Telco 회사(vRAN)	1-8x Low-mid range (GPU/FPGA) 1-4x High end (GPU)	Apollo 2000 (XL190r/XL220n) DL360 / 380 / 385 Gen10 Plus & v2 EL8000/4000/1000

HPE 의 GPU를 활용한 HPC & AI 서버 포트폴리오

데이터센터 트레이닝 → 대규모 인메모리 → 소규모 데이터센터 트레이닝 & 추론 / 니어 엣지 ⇔ 엣지 트레이닝 & 추론 → 엣지 추론



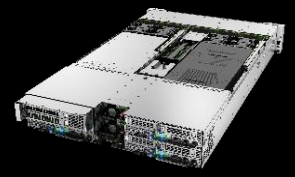




플랫폼	Apollo 6500 Gen10 & Plus	Superdome Flex	ProLiant DL380 / DL385 Gen10 Plus & v2	Apollo 2000 Gen10 & Gen10 Plus(XL190r/XL290)	Edgeline EL8000 e910/e920	Edgeline EL4000 m710x	Edgeline EL1000 m710x
워크로드	• 대량의 데이터세트 트레이닝 • 복합적 트레이닝 워크로드	• 금융 모델 • 스트리밍 분석 • 데이터베이스 검색 및 분석	• 소규모 데이터세트 트레이닝 및 추론 • 소규모 스케일 트레이닝 워크로드	• 데이터센터 / 니어 엣지 추론 & 소규모 트레이닝 • 비디오 / 데이터 분석 • 고성능 & 고집적 멀티 노드 워크로드	• 엣지 실시간 트레이닝 & 추론 • Telco – RAN AR/VR • 취약 환경 워크로드	• 엣지 실시간 추론 • 비디오 / 데이터 / 센서 분석 • 취약 환경 워크로드	• 엣지 실시간 추론 • 비디오 / 데이터 / 센서 분석 • 이동식, 무선, 취약환경 워크로드
GOOD	8-12x T4(XL270d) 8x A10(XL645d) 16x A10(XL675d)	4x RTX 6/8000	8x T4 6x A10	2-8x T4	1U: 1x T4 / CLX CPU	1x T4	1x T4
BETTER	8x RTX6/8000(XL270d) 4x A40(XL645d) 10x A40(XL675d)	4x V100S-32GB PCIE	6x RTX A4000 3x A30/A40	4-8x RTX 4000 2x RTX 8000	2U: 4x T4 2U: 1x RTX 6000	2x T4	2x T4
BEST	8x V100-32GB SXM2 4x A100 80GB SXM4 8x A100 80GB SXM4		3x A100-40GB-PCIE	2x V100S 32GB PCIE 2x A100-40GB-PCIE	2U: 1x V100S 32GB PCIE 1x A100-40GB-PCIE	4x T4	

트레이닝 워크로드 ⇔ 추론 워크로드

※ **Roadmap에 따라** GPU와 서버는 선택적일 수 있음

HPE 의 GPU를 활용한 GRAPHIC 서버 포트폴리오

	워크스테이션 그래픽	엔터프라이즈 그래픽			렌더팜	대규모 인메모리
						
플랫폼	ProLiant ML350 Gen10	ProLiant DL380 / DL385 Gen10 Plus & v2	Apollo 2000 Gen10 & Gen10 Plus(XL190r/XL290)	ProLiant SY480 Gen10 & Gen10 Plus	Apollo 6500 Gen10 & Plus	Superdome Flex
워크로드	• 의료 영상 및 이미지, DCC, Direct Attach • 비디오 / 데이터 사이언스 워크로드	• 레이 트레이싱, DCC, 비디오 프로세싱 • 의료 이미지 & 3D 모델링	• 레이 트레이싱, DCC, 비디오 프로세싱, Oil & Gas • 고성능 & 고집적 멀티노드 워크로드	• Oil & Gas, 디자인 & 제조 • 고성능 & 고집적 멀티노드 워크로드	• 실시간 레이 트레이싱 • 고성능 렌더링 & 시뮬레이션	• 스케일아웃, 단일 NUMA 노드 워크로드
GOOD	4x RTX 4000	4x RTX 4000	4-8x RTX 4000	4x RTX 4000	4x RTX 6000 1-10x A40 / 1-16x A10	4 socket, 4x RTX 6/8000
BETTER	2x RTX 6/8000 + NVLink Bridge	3x RTX 6000	1-2x RTX 8000	2x RTX 6000	1-10x A40 / 1-16x A10	8 socket, 8x RTX 6/8000
BEST	4x RTX 6/8000 + NVLink Bridge	3x RTX 8000 / A30 /A40 6x A10	2-4x RTX 8000	2x RTX 8000 2x A40 / 4x A10	10x A40 / 16x A10	16 socket, 16x RTX 6/8000

사무실 → 데이터센터

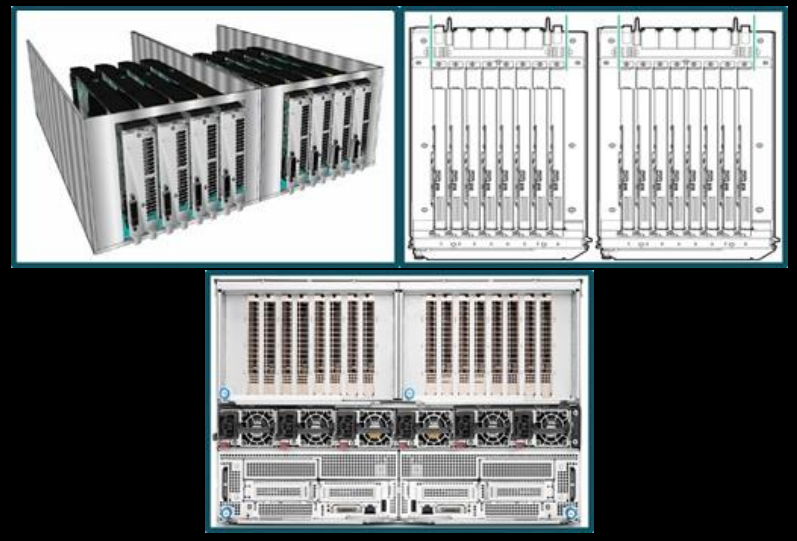
※ Roadmap에 따라 GPU와 서버는 선택적일 수 있음

멀티 워크로드를 위한 시스템 – APOLLO 6500 GEN10 PLUS

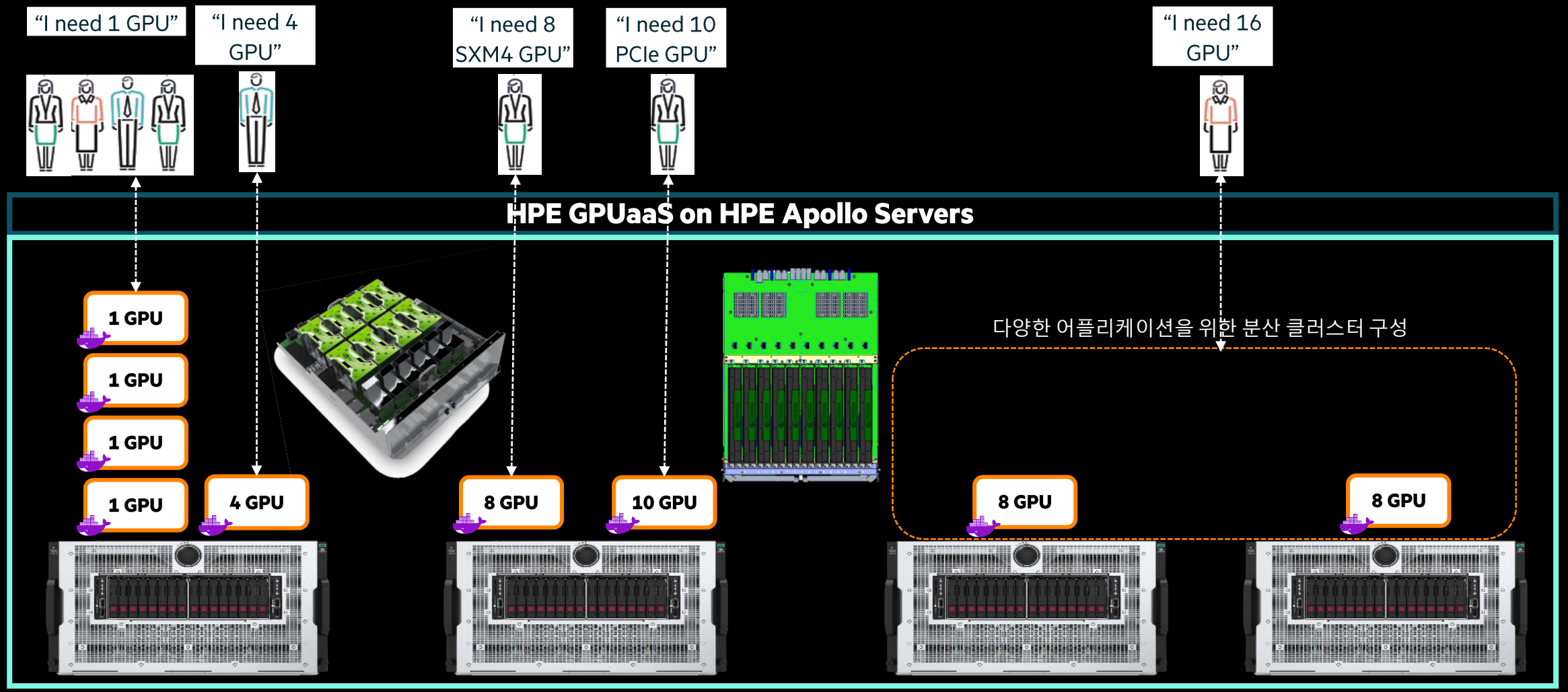
NVIDIA 4 (XL645d x 2ea) / 8 (XL675d) SXM A100 GPU

10 Double-width or 16 Single-width PCIe GPU (XL675d)

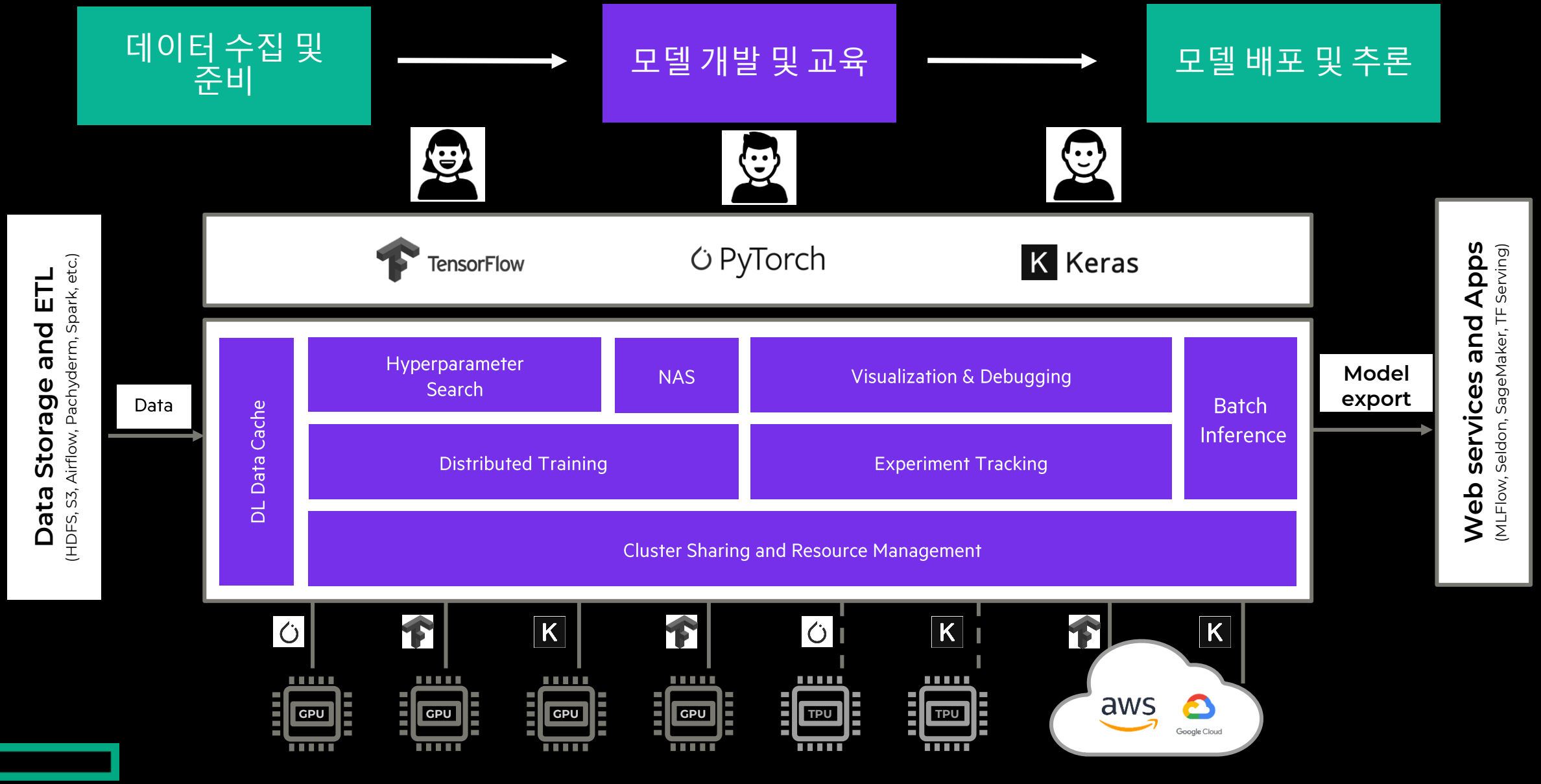
4 Double-width or 8 Single-width PCIe GPU per Node (XL645d)



APOLLO 6500 GEN10 PLUS를 활용한 멀티 워크로드 배포 방법



트레이닝에 최적화 된 DETERMINED AI TRAINING PLATFORM



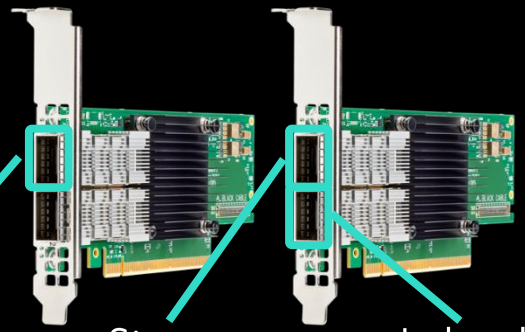
HPE AI / HPC 소프트웨어 포트폴리오

어플리케이션과 소프트웨어 개발 에코시스템	개발 환경	HPE Cray Programming Environment • C/C++, Fortran, UPC, R, Python • Compiling Environment	Intel® Parallel Studio XE (w/Intel MPI)	AOCC with AMD ROCm	
			NVIDIA HPC SDK	NVIDIA GPU Cloud	GNU Compilers
	디버그와 성능	• Debuggers • Performance analysis and optimization tools • Code parallelization assistant	Arm® Forge Professional	TotalView™ by Perforce	Vampir
	MPI	• HPE Cray MPI	HPE Message Passing Interface (MPI)	Open MPI	Mellanox® HPC-X™
워크로드 관리와 오케스트레이션		Altair® PBS Professional®	Slurm®	Kubernetes®	Containers: Docker®, Singularity
원격 시각화		NICE DCV and EnginFrame			
스토리지 파일시스템		Cray ClusterStor E1000 Storage Solution (Lustre-based)			
데이터 관리		HPE Data Management Framework (DMF)			
시스템 관리		HPE Cray supercomputer software • HPE Cray System Management	HPE Performance Cluster Manager	Bright Cluster Manager®	
패브릭 소프트웨어		• HPE Slingshot fabric manager	Mellanox® Unified Fabric Manager™	Intel® Omni-Path Fabric Software	
운영체제		• HPE Cray Operating System	SUSE® Linux® Enterprise Server	Red Hat® Enterprise Linux	

● HPE Apollo, HPE ProLiant DL, HPE SGI ● HPE Cray supercomputer ● Cray ClusterStor

GPU DIRECT RDMA & GPU DIRECT STORAGE를 위한 INFINIBAND 구성

Compute
(InfiniBand)

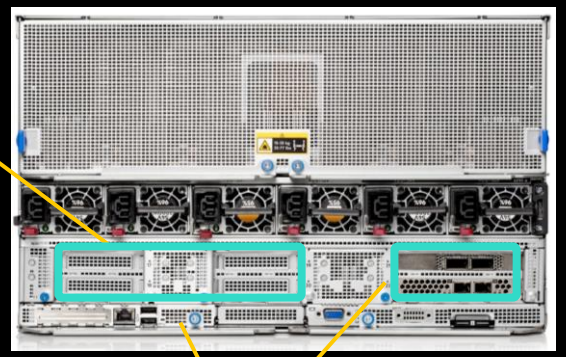


Storage
(InfiniBand)

Storage
(InfiniBand)

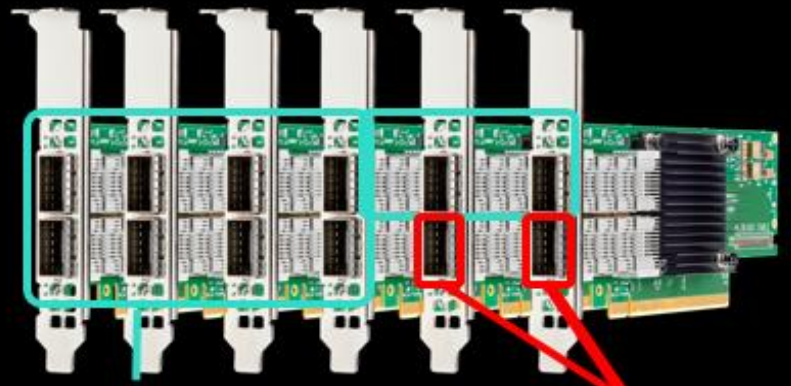
In-band
(100GbE)

4x Mellanox ConnectX-6 1-port VPI
Compute Fabric (InfiniBand - HDR)



2x Mellanox ConnectX-6 2-port VPI
Storage Fabric (InfiniBand - HDR)
In-band management (Ethernet 100Gb)

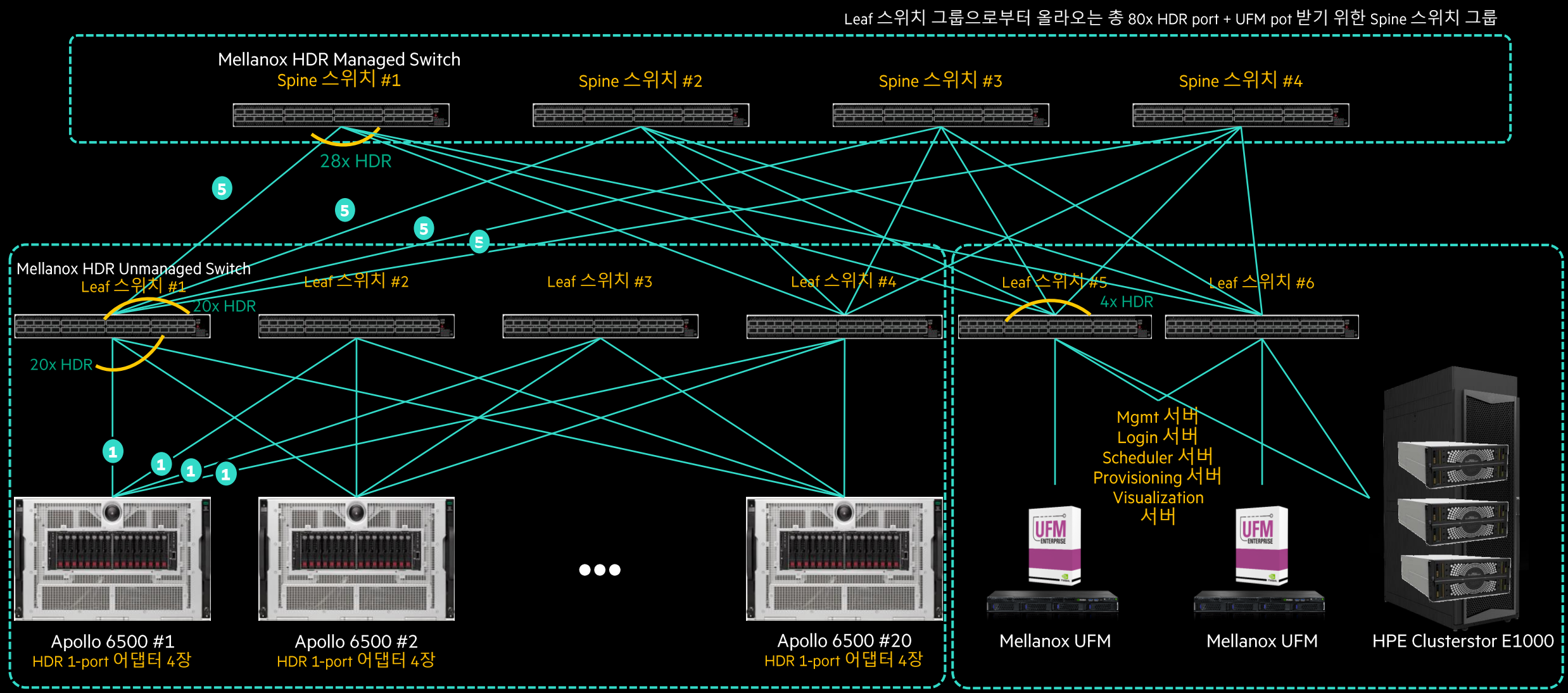
iLO management port
Out-of-band management (1GbE)



InfiniBand(HDR)

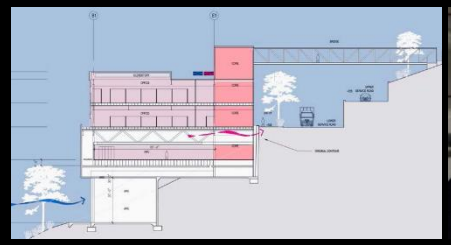
Ethernet(100Gb)

HPE SUPERPOD 20대 CLUSTER 및 STORAGE ARCHITECTURE



AI / HPC USE CASE

FAST CLUSTER WITH NVIDIA OF THE WORLD



Cluster Special Feature

- 3072 CPU only nodes
- 1536 CPU + GPU nodes
- AMD Milan EPYC 7763 CPU
- NVIDIA A100 GPU 6144ea
- Slingshot interconnect
- 35PB All-Flash ClusterStor E1000-F
- Cooling : Direct liquid cooling
- 4 exaflops for AI
- Size : $100M
- Delivery : 2020 ~ 2021
- Support: 7,000 users and 700 project

GPU Programming Models

- CUDA: MILC, Chroma, HACC ...
- CUDA FORTRAN: Quantum ESPRESSO, StarLord (AMREX)
- OpenACC: VASP, E3SM, MPAS, GTC, XGC ...
- KoKKos: LAMMPA, PELE, Chroma ...
- Raja: SW4

THE FASTEST SUPERCOMPUTER IN 2023

CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

Hewlett Packard Enterprise

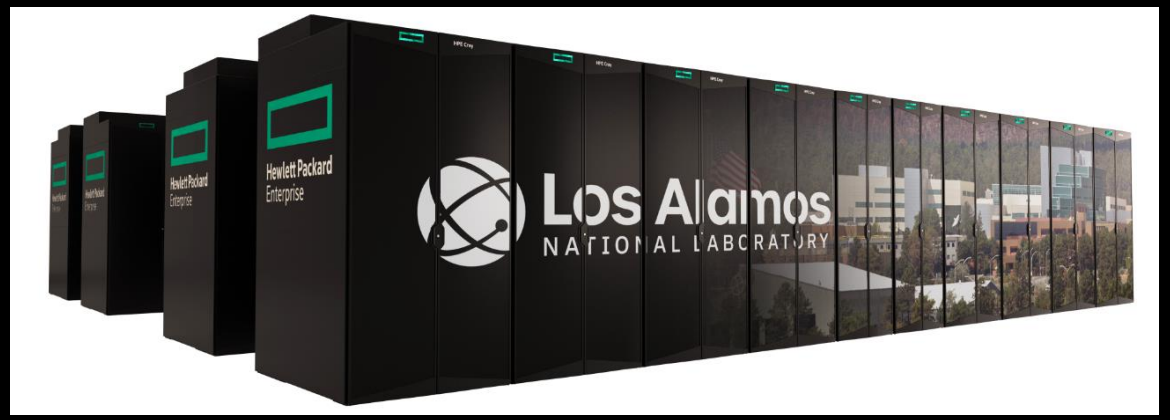
NVIDIA

- 20 exaflops for AI (7x faster than Selene supercomputer).
- Built on HPE Cray EX supercomputer
- Online in 2023

Los Alamos
NATIONAL LABORATORY

Hewlett Packard Enterprise

NVIDIA



- First United States customer to receive Grace
- HPE as the system provider
- Target delivery early 2023

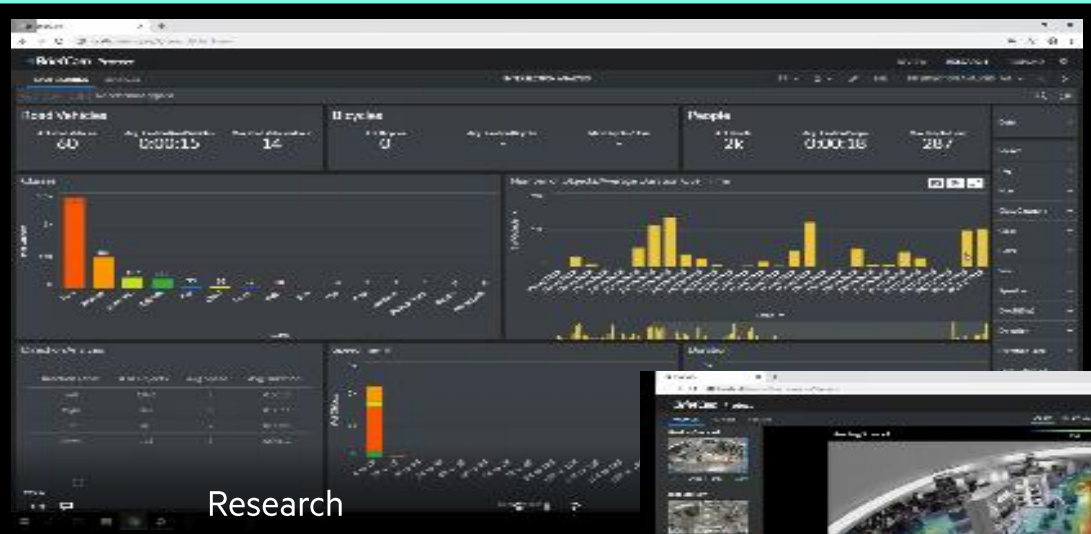
GRAPHICS USE CASE

RENDERING CASE

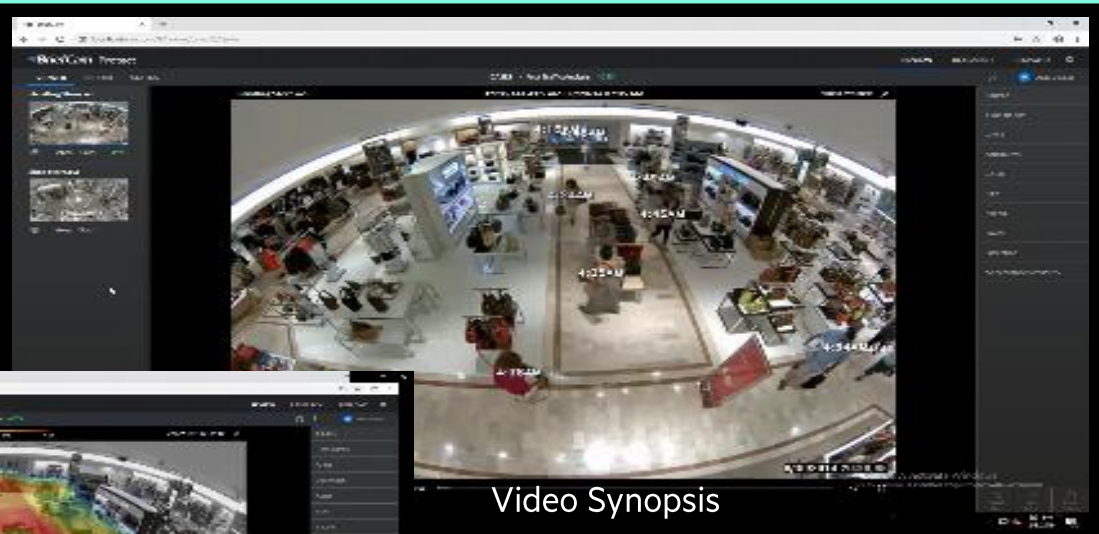
Ray Tracing & DLSS

Ray Tracing & DLSS

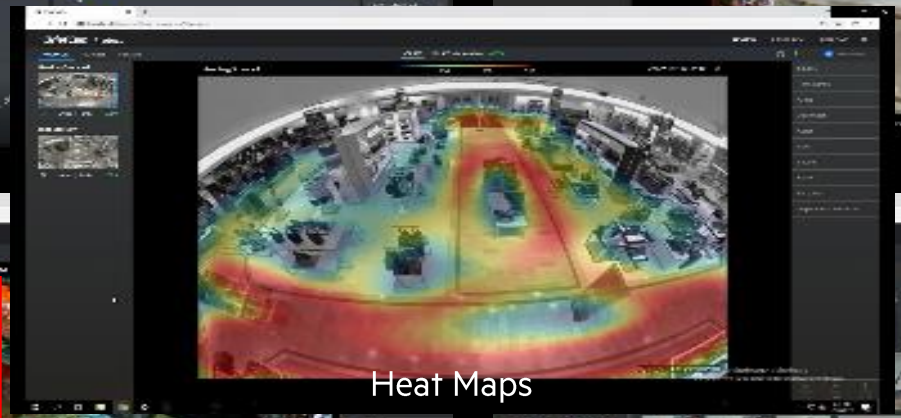
VIDEO CASE



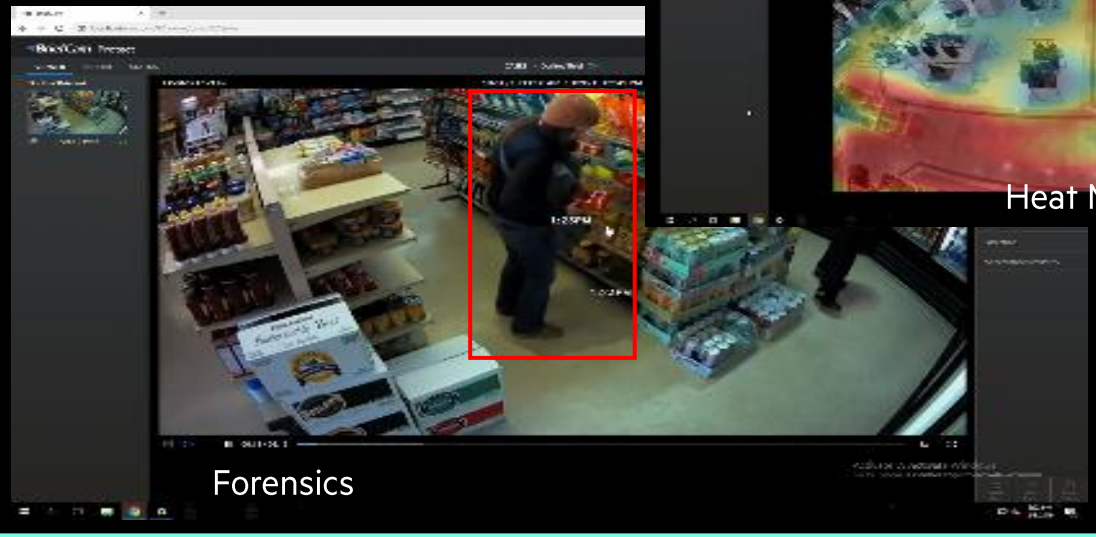
Research



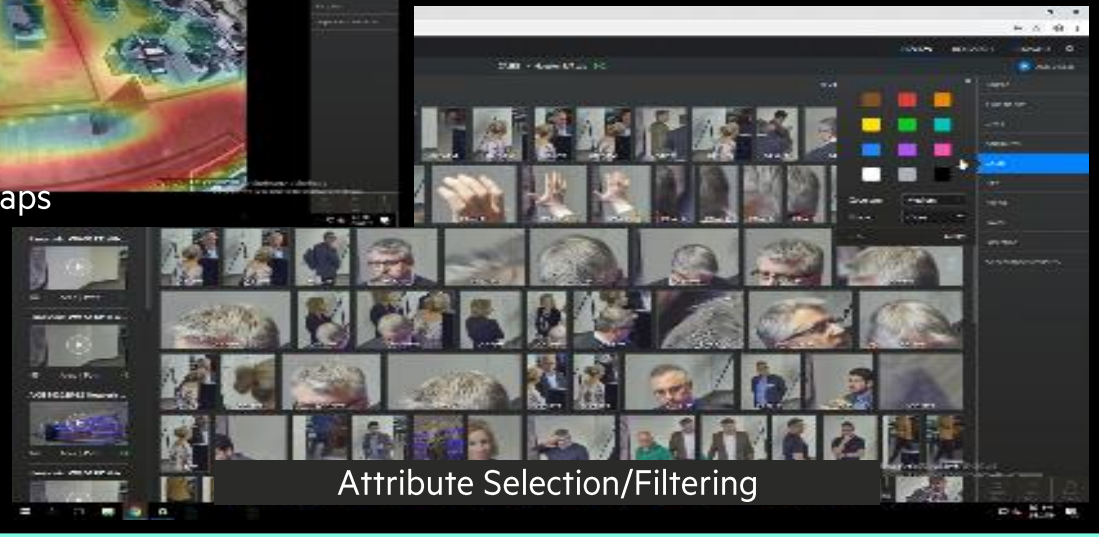
Video Synopsis



Heat Maps

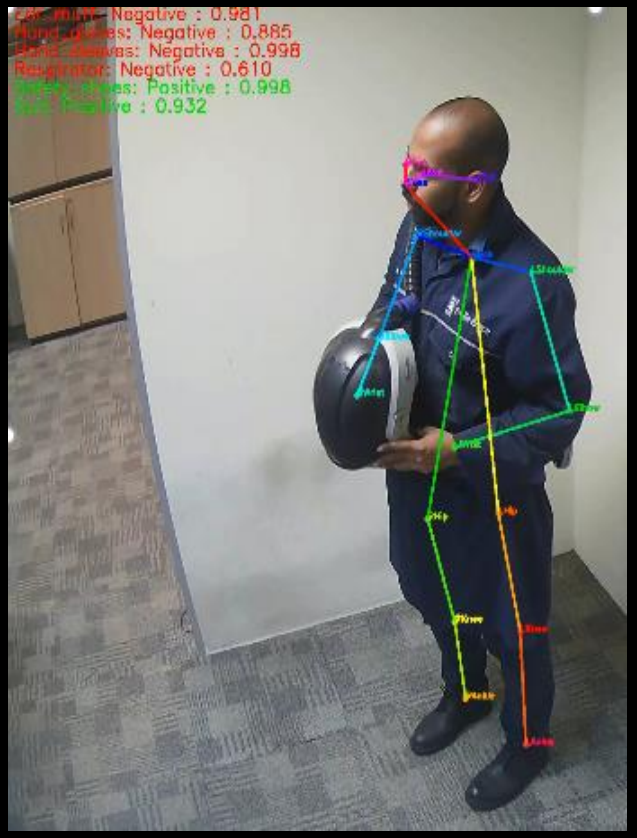


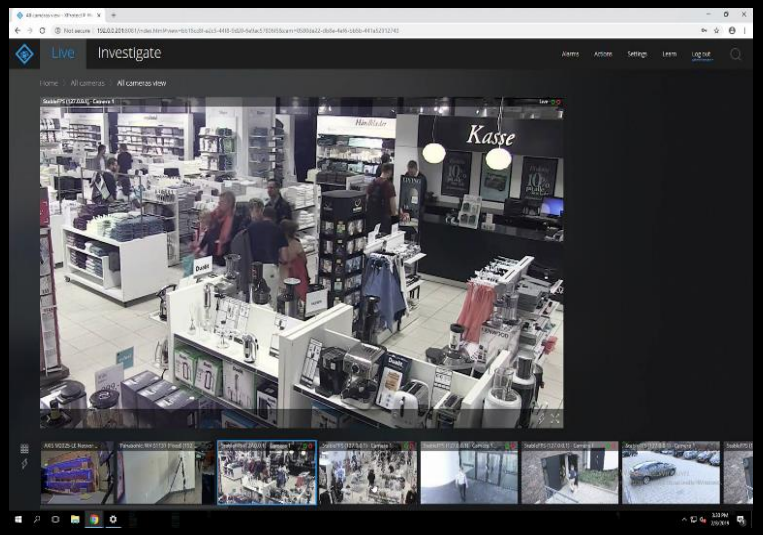
Forensics

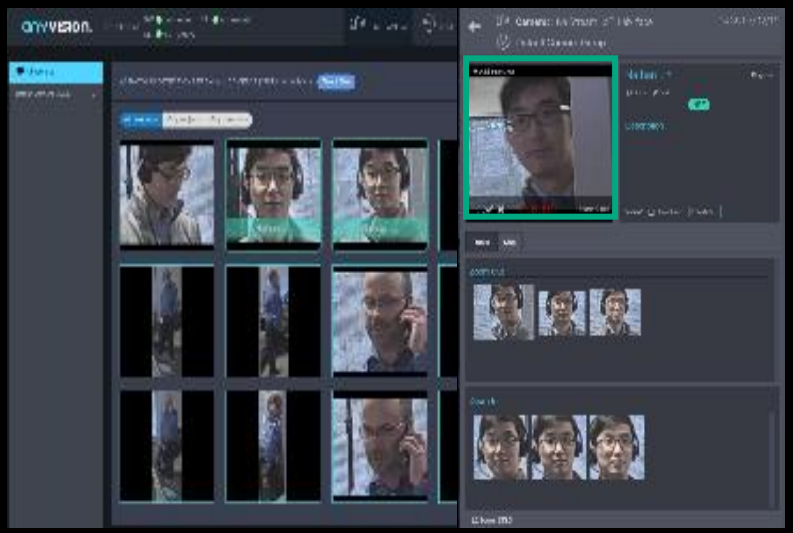


Attribute Selection/Filtering

CCTV CASE


Negative : 0.981
Negative : 0.885
Negative : 0.998
Negative : 0.610
Positive : 0.998
Positive : 0.932


Live
Investigate
Kasse


onvision

EDGE USE CASE

작업자의 자동화 및 가시적 원격 지원

현장 공장 작업자와 제어실간에 시각적으로 안내되는 실시간 대화 세션을 활성화

기술 적용 부분

- RealWear HMT-1 (HMT-1z1) Helmet
  - Class 1, Division 1 rating
  - Hands-free device
  - Noise-cancelling microphone
  - Camera
- HPE MyRoom VRG
  - Live-share complete images, video, audio and data

장점

- 작업자의 안전화, 작업 효율(hands-free)
- 예방 정비에 대한 활성화
- 원격에서의 가이드 제공
- 위급 상황 시에 대한 Push Notification(화학물질노출)
- 장비 상태 점검을 위한 증강 현실 제공
- 생산성 / 효율성 / 직원 참여 향상
- 시간과 비용의 절약
- 기술적인 학습 향상

작업자의 자동화 및 가시적 원격 지원

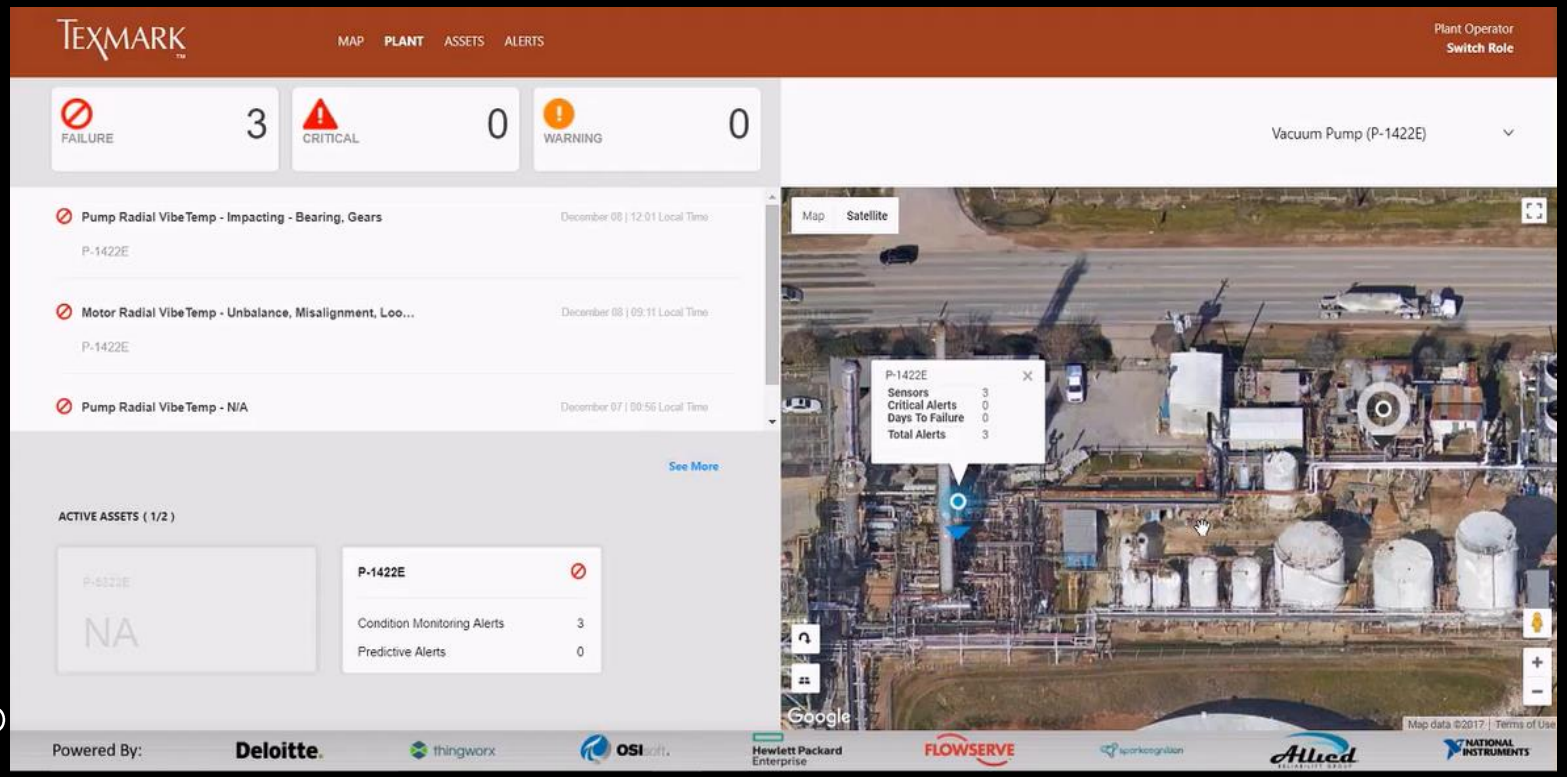
현장 공장 작업자와 제어실간에 시각적으로 안내되는 실시간 대화 세션을 활성화

기술 적용 부분

- RealWear HMT-1 (HMT-1z1) Helmet
  - Class 1, Division 1 rating
  - Hands-free device
  - Noise-cancelling microphone
  - Camera
- HPE MyRoom VRG
  - Live-share complete images, video, audio and data

장점

- 작업자의 안전화, 작업 효율(hands-free)
- 예방 정비에 대한 활성화
- 원격에서의 가이드 제공
- 위급 상황 시에 대한 Push Notification(화학물질노출)
- 장비 상태 점검을 위한 증강 현실 제공
- 생산성 / 효율성 / 직원 참여 향상
- 시간과 비용의 절약
- 기술적인 학습 향상

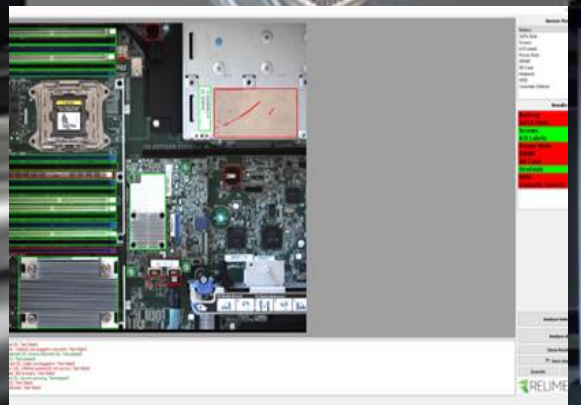


Hewlett Packard
Enterprise

21x

빠른 속도

HPE 서버 제품의 품질 보증을 위해 EDGE COMPUTING AI와 결합된 고해상도 카메라를 사용하여 HPE 제조는 클라우드에서 EDGE 컴퓨팅으로 전환하고 합격 / 불합격 시간을 21 초에서 1 초로 단축



Hewlett Packard Enterprise | NVIDIA

GPU를 활용한
HPE AI 및 HPC 전략

THANK YOU