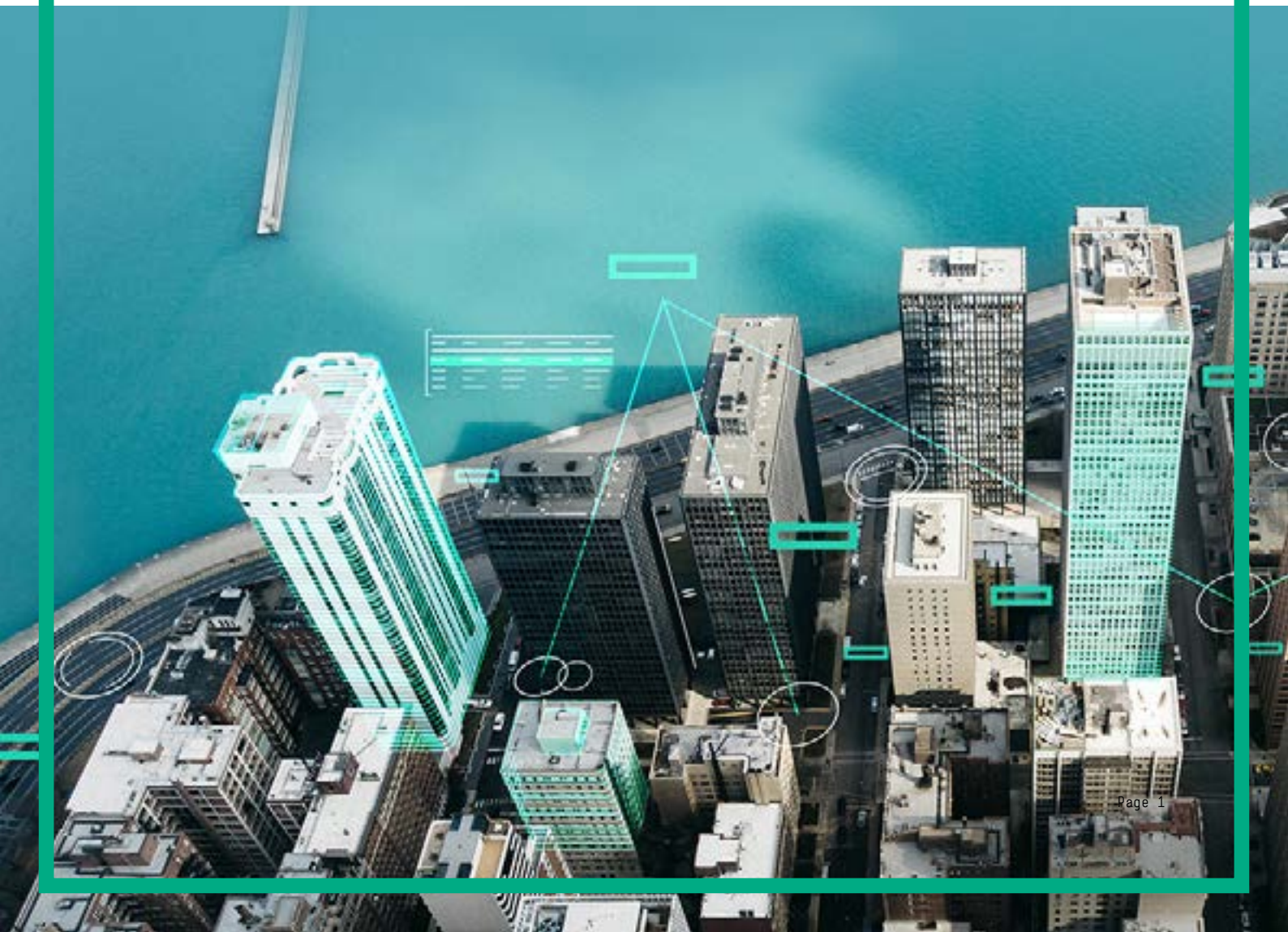


# HPE-NVIDIA AI Bundle Offering

AI Training 및 Inference 전용구성





# INDEX

인공지능의 미래 및 HPE HPC/AI 연구여정 .....Page 3

HPE Edge-to-Core AI 통합 모델 구성도 .....Page 4

## HPE-NVIDIA AI 오퍼링 - AI Training 용도

Deep Learning + Big Data 전용구성 .....Page 5

WEKA I.O 알아보기 .....Page 6

Deep Learning 전용 구성 .....Page 8

NVIDIA GPU V100 소개 .....Page 9

## HPE-NVIDIA AI 오퍼링 - AI Inference 용도

Edge Compute + Data Center 전용 구성 .....Page 11

HPE Converged Edgeline System .....Page 12

NVIDIA GPU T4 소개 .....Page 13

## HPE AI 최적화 제품 포트폴리오

HPE Apollo 6500 Gen10 .....Page 16

HPE ProLiant ML350 Gen10 .....Page 17

HPE ProLiant DL380 Gen10 .....Page 18

HPE Edgeline 컨버지드 에지 시스템 및 HPE Intelligent 게이트웨이 .....Page 19

HPE Pointnext AI 컨설팅 .....Page 20

## 인공지능의 미래

인공지능은 지난 50여 년의 연구 과정을 거치면서 우리가 살고, 일하고, 사람들과 관계를 맺는 방식을 변화시키고 있습니다. 인공지능 기술은 인간의 직각, 추론, 학습능력 등을 컴퓨팅 파워를 이용하여 구현하는 기술로, 자율주행, 지능형 금융서비스, 의료진단, 법률서비스 지원, 게임, 기사작성, 지능형 감시 시스템 등 다양한 산업분야에서 널리 활용되고 있습니다.



**자율주행**  
~ 2027



**리테일 자동화**  
~ 2031



**외과수술**  
~ 2043



**베스트셀러 작품**  
~ 2049



**수학연구**  
~ 2060



**노동자동화**  
~ 2140

## HPE의 HPC/AI 연구 여정



### The mission to Mars

NASA와의 협업으로 Spaceborne Computer 우주선 탑재, 우주연구 프로젝트에 기여



### Seagate

하드디스크 품질관리를 위해 엣지-투-코어 (edge-to-core) 기반의 머신비전 시시스템 구축



### WINNING

Against the World's Best Poker Players

### AI Poker

카네기멜론과의 협업으로 인공지능 포커플레이어 탄생, 예측 불가능한 환경에서 인공지능 기술 구현



### HPE Manufacture in Switzerland

HPE서버 제품관리를 위해 엣지-투-코어 (edge-to-core) 기반의 머신비전 시시스템 구축



### DNZE

독일의 뇌행성 뇌질환 센터와의 협업으로 메모리 기반 컴퓨팅을 통해 알츠하이머 치료개선 기여



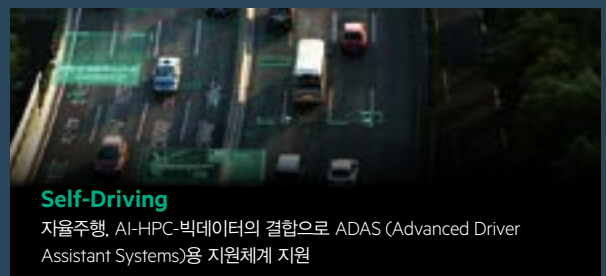
### International Airport

비디오 분석과 감시를 활용한 안전 강화 사례



### F1

Mercedes-AMG Petronas Motorsport팀과의 파트너십으로 레이싱 시뮬레이션을 통해 초당 퍼포먼스 성능향상



### Self-Driving

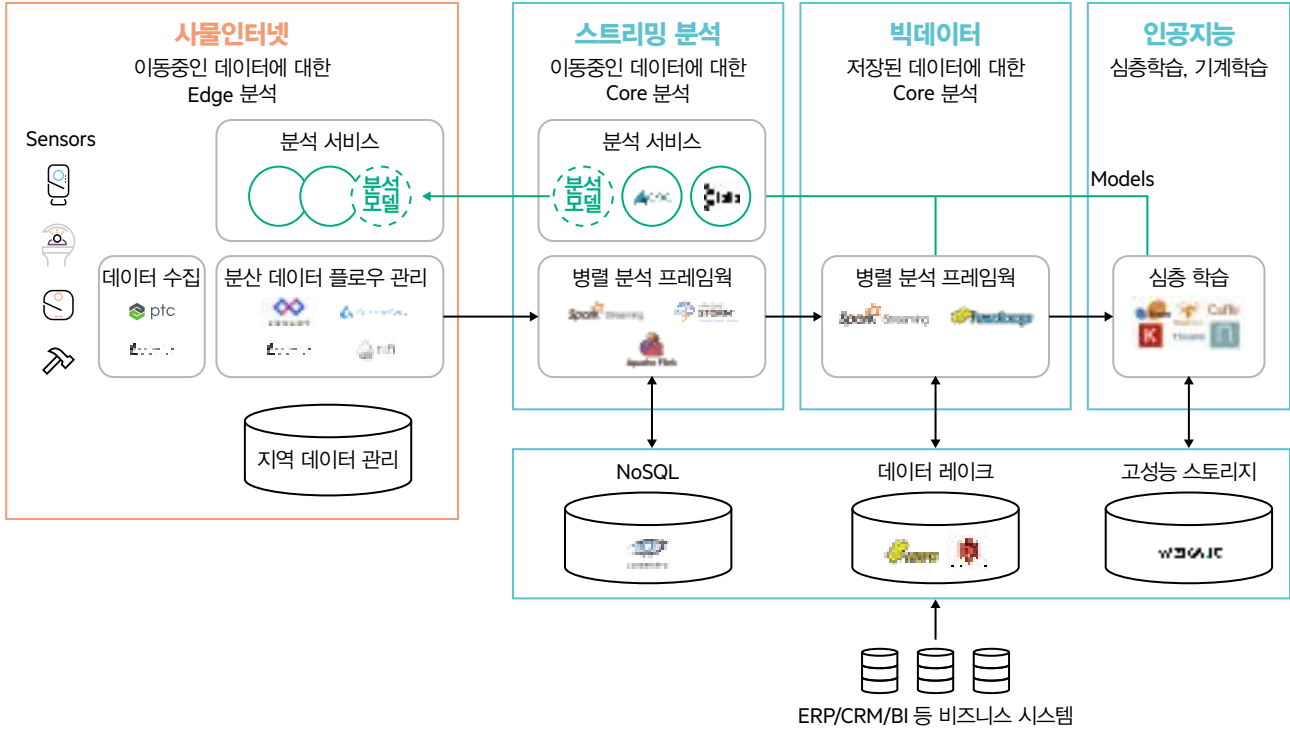
자율주행, AI-HPC-빅데이터의 결합으로 ADAS (Advanced Driver Assistant Systems)용 지원체계 지원

# HPE Edge-to-Core AI 통합모델 구성도

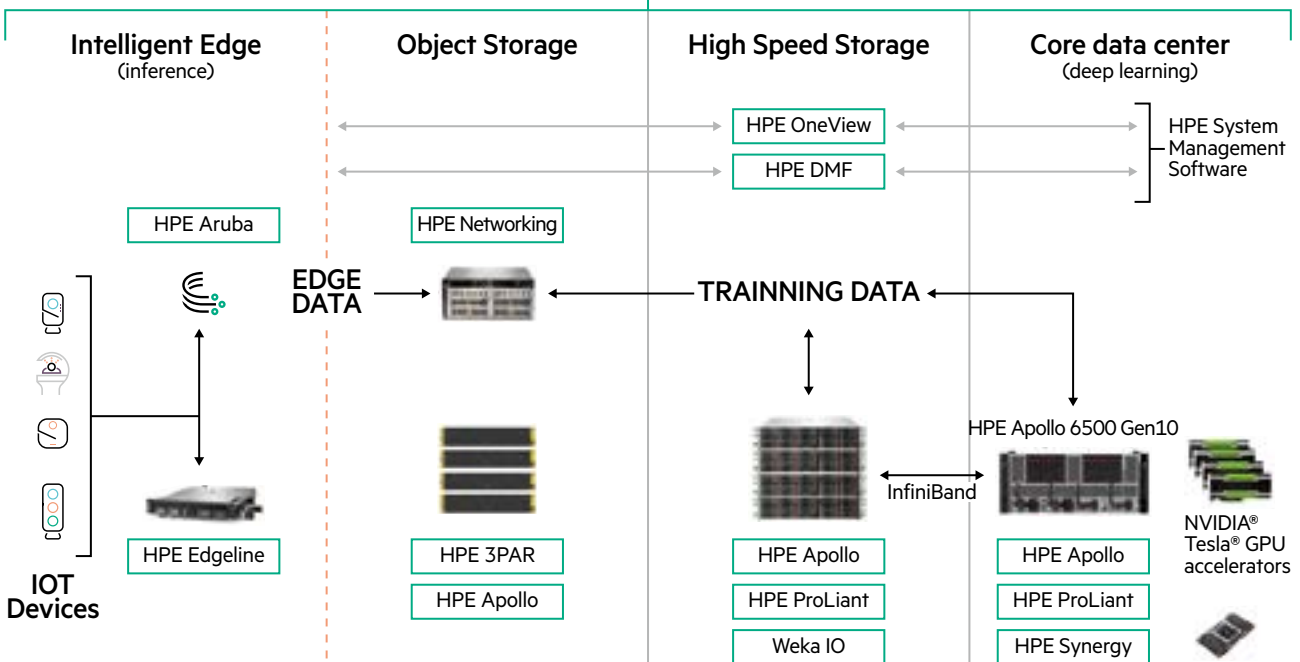
## 서비스와 솔루션

### 데이터 과학 도구

데이터 플로우 설계, 데이터 과학자용 워크벤치, 데이터 모델 관리, 어플리케이션 개발



### HPE Pointnext

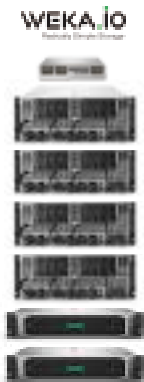


# HPE-NVIDIA AI 오퍼링 AI Training 용도

## Deep Learning Compute + Big Data 전용구성

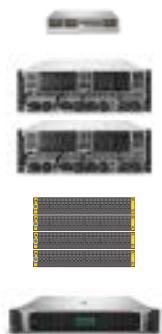
빅데이터에서 시작하여 인공지능 학습까지 연결되는 최적화된 구성입니다.  
고객사의 학습모델에 맞추어, 컴퓨팅 및 데이터의 처리 성능에 따른 3단계의 스케일별 구성을 제공합니다.

### Large Scale



구성	수량	제품 사양 (노드당)
Server	4	HPE Apollo 6500 Gen10 Servers 2x Intel Gold 6242 (2.8GHz/16c), 512GB of Memory (최대 3TB확장), 4x 2TB NVMe SSDs, 2x 480GB SATA MU M.2 SSDs
GPU	32	NVIDIA Tesla V100 32GB SXM2 GPUs
Switch	1	Mellanox 100Gbe/IB
Storage System		WEKA.io on Apollo 2000 or DL360 Gen10 50TB Usable, 30GB/s   2.5M IOPs, 8 Servers with (4) 3.2TB NVMe per server in 4u

### Medium Scale



구성	수량	제품 사양 (노드당)
Server	4	HPE Apollo 6500 Gen10 Servers 2x Intel Gold 6242 (2.8GHz/16c), 512GB of Memory (최대 3TB확장), 4x 2TB NVMe SSDs, 2x 480GB SATA MU M.2 SSDs
GPU	16	NVIDIA Tesla V100 32GB SXM2 GPUs
Switch	1	Mellanox 100Gbe
Storage System		3PAR 8400 2N (3.84TB x 24, 16Gb FC x 12ports) Usable 56 TB (RAID-6 6D+2P) SAN Switch SN3000B 24ports/12ports x 2ea

### Entry Scale



구성	수량	제품 사양 (노드당)
Server	1	HPE Apollo 6500 Gen10 Servers 2x Intel Gold 6242 (2.8GHz/16c), 512GB of Memory (최대 3TB확장), 4x 2TB NVMe SSDs, 2x 480GB SATA MU M.2 SSDs
GPU	8	NVIDIA Tesla V100 32GB SXM2 GPUs
Storage System		Apollo 4510 Gen10 (3.84TB SSD x 32ea) Over Usable 50TB

# Weka.IO 알아보기

## 차세대 플래시 스토리지 솔루션 - Weka.IO

오늘날 스토리지 기술들은 확장성과 성능측면에서 장단점이 있기 때문에 특수 목적으로 설계된 스토리지 아키텍처와 최신 워크로드 사이에 갭이 나타나고 있습니다. 이러한 새로운 워크로드를 처리하기 위해 혁신적인 스토리지 기술을 사용할 수 있습니다. WekaIO Matrix는 컴퓨팅 및 데이터 집약형 워크로드를 처리하기 위해 플래시에 최적화된 차세대 병렬식 파일 시스템을 제공하여 병렬식 파일 시스템의 취약점을 극복합니다. 이는 NVMe 플래시의 성능과 클라우드의 확장성, 그리고 NAS의 단순성을 하나로 결합한 시스템입니다.



WekaIO는 HPC 및 AI 환경 내에서 차세대 스토리지 워크로드에 필요한 대역폭과 IOPs 및 지연 시간을 제공합니다. 전송 속도를 테스트한 결과, 올 플래시 NAS 솔루션이 1~1.5GB/초인 데 반해, WekaIO는 GPU 서버당 6GB/초 이상의 빠른 속도를 기록했습니다. WekaIO Matrix는 GPU 노드가 I/O 중심이 아닌 컴퓨팅 중심이 되도록 보장합니다.

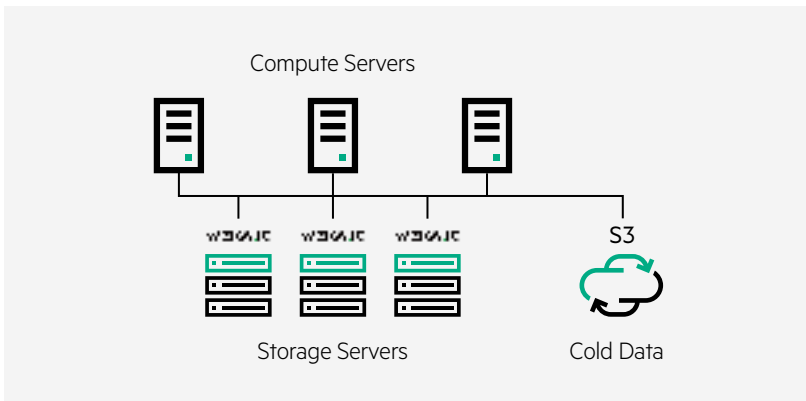
## Weka.IO 주요특징

- Weka.IO Matrix는 네이티브 NVMe를 기반으로 하는 완전 병렬식 및 완전 분산형의 공유 파일 시스템으로, NVMe-oF(NVMe-over-Fabrics)과 동급의 기술로 볼 수 있습니다.
- Weka.IO Matrix는 온프레미스 및 클라우드 배포를 지원하여 두 환경 간에 통합 네임스페이스를 제공합니다. 고객들은 두 환경에서 동시에 작업을 실행한 다음, 결과를 동기화할 수 있습니다.
- Weka.IO는 NVMe 데이터를 객체 스토리지에 티어링할 수 있도록 지원합니다. 각 SKU 구독은 고객에게 1TB의 객체 스토리지 티어링 자격을 부여합니다.
- 플래시, 고성능 및 낮은 지연 시간을 고려해 설계된 파일 시스템입니다.
- 분산형 메타 데이터 아키텍처를 통해 수십억 개의 파일과 엑사바이트급 용량으로 확장이 가능합니다.
- 이레이저 코딩, 데이터 보호 체계 및 고속 재빌드를 통해 장애 복원력을 제공합니다.
- 오버플로우 또는 용량 스토리지를 위해 S3를 통한 객체 스토리지 티어링이 제공됩니다.

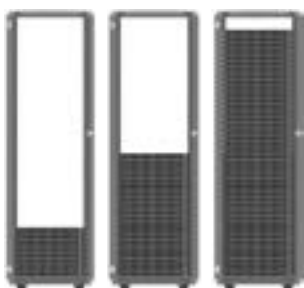
애플리케이션	가상 머신
POSIX 파일 시스템	
분산형 메타데이터	
동시 스냅샷	데이터 압축/중복 제거
암호화	N+4 이레이저 코딩
멀티 프로토콜 S3, SMB, HDFS, NFS	최적화된 네트워크 스택 (IB, 이더넷)

## Weka.IO 구성

- WekaIO Matrix is loaded on a dedicated servers
- S3 interface to tier overflow/cold data to object/cloud storage



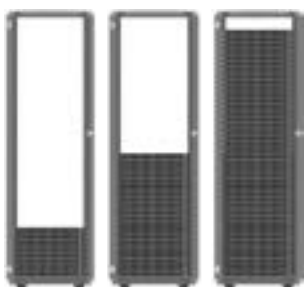
### DL360 Gen10 구성시



Entry Midrange High End

Type	Node Count	Raw Capacity (TB)	Usable Capacity (TB)	Read 4K IOPS (Millions)	Read 1M Bandwidth (GB/s)
Entry	8	102.4	54.6	2.5	30
Midrange	20	512	364.1	8.7	160
High End	40	1280	910.2	17.4	320

### Apollo 2000 Gen10 구성시



Entry Midrange High End

Type	Node Count	Raw Capacity (TB)	Usable Capacity (TB)	Read 4K IOPS (Millions)	Read 1M Bandwidth (GB/s)
Entry	8	102.4	54.6	2.5	30
Midrange	20	512	364.1	8.7	160
High End	40	1280	910.2	17.4	320



# Deep Learning 전용구성

## AI Training 용도- Deep Learning Compute Only (단위 시스템 별)

HPE는 고객의 다양한 딥러닝 요구사항에 맞춘 최적화된 단위시스템을 제공합니다.

### Large Scale



구성	수량	제품 사양 (노드당)
Server	1	HPE Apollo 6500 Gen10 Servers 2x Intel Gold 6242 (2.8GHz/16c), 512GB of Memory (최대 3TB확장), 4x 2TB NVMe SSDs, 2x 480GB SATA MU M.2 SSDs
GPU	8	NVIDIA Tesla V100 32GB SXM2 GPUs

### Medium Scale



구성	수량	제품 사양 (노드당)
Server	1	HPE DL385 Gen10 Servers 2x Intel Gold 6248 (2.5GHz/20c), 384GB of Memory (최대 3TB확장), 4x 2TB NVMe SSDs, 2x 480GB SATA MU M.2 SSDs
GPU	4	GV100 32GB NVLink

### Large Scale

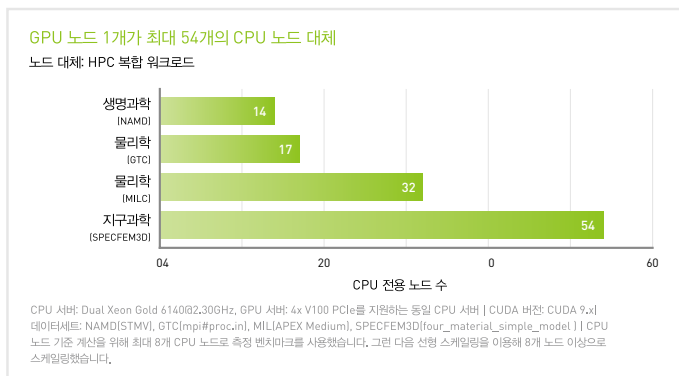
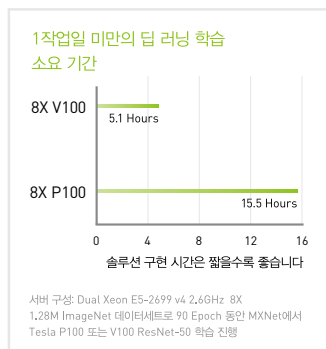
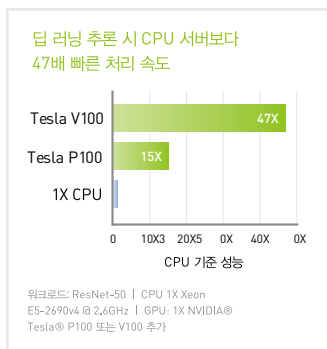


구성	수량	제품 사양 (노드당)
Server	1	HPE DL380 Gen10 Servers 2x Gold 6248 (2.5GHz/20c), 384GB of Memory, 2x960GB SATA 6G MU SSDs, 2x240GB SATA 6G RI SSDs for Boot
GPU	2	NVIDIA Tesla V100 32GB PCIe GPUs

# NVIDIA TESLA V100 GPU 가속기

## HPC 및 AI 전용 데이터센터 GPU

NVIDIA® Tesla® V100은 지금까지 개발된 데이터센터 GPU 중 세계에서 가장 앞선 데이터센터 GPU로서 AI, HPC 및 그래픽을 가속합니다. 최신 GPU 아키텍처인 NVIDIA Volta로 지원되는 Tesla V100은 단일 GPU로 최대 100개의 CPU에 성능을 제공하여 불가능의 영역에 머물러있던 도전 과제를 해결할 수 있도록 데이터 과학자, 연구자 및 엔지니어를 지원합니다.



## 사양



	Tesla V100 PCIe	Tesla V100 SXM2
GPU 아키텍처	NVIDIA Volta	
NVIDIA Tensor 코어	640개	
NVIDIA CUDA® 코어	5,120개	
다중정밀도 성능	7 TFLOPS	7.8 TFLOPS
단정밀도 성능	14 TFLOPS	15.7 TFLOPS
Tensor 성능	112 TFLOPS	125 TFLOPS
GPU 메모리	32GB /16GB HBM2	
메모리 대역폭	900GB/초	
ECC	지원	
상호연결 대역폭	32GB/초	300GB/초
시스템 인터페이스	PCIe Gen3	NVIDIA NVLink
폼 팩터	PCIe 전체 높이/길이	SXM2
최대 전력 소비	250 W3	00 W
열 솔루션	패시브	
Compute APIs	CUDA, DirectCompute, OpenCL™, OpenACC	

## 획기적인 혁신



### Volta 아키텍처

통합 아키텍처 내에서 CUDA 코어와 Tensor 코어를 페어링함으로써 Tesla V100 GPU가 탑재된 단일 서버는 수백 대의 범용 CPU 서버를 대체하는 성능으로 종래의 HPC와 딥 러닝을 수행할 수 있습니다.



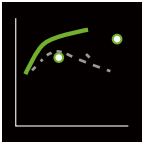
### Tensor 코어

640개의 Tensor 코어가 탑재된 Tesla V100은 125 teraFLOPS의 딥 러닝 성능을 제공합니다. 이는 NVIDIA Pascal™ GPU와 비교해 DL 학습의 경우 12X Tensor FLOPS, DL 추론의 경우 6X Tensor FLOPS에 상응합니다.



### 차세대 NVLink

Tesla V100에 탑재된 NVIDIA NVLink는 이전 세대 대비 2배 빠른 처리 속도를 지원합니다. 최대 8대의 Tesla V100 가속기를 최대 300GB/s의 속도로 상호 연결할 수 있어 단일 서버에서 구현 가능한 최고의 애플리케이션 성능이 발휘됩니다.



### 최대 효율 모드

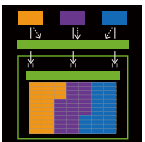
새로운 최대 효율 모드를 이용해 데이터센터는 기존 전력 예산 범위 내에서 랙당 최대 40% 향상된 컴퓨팅 성능을 달성할 수 있습니다.

본 모드에서 Tesla V100은 최고 처리 효율로 가동되며 절반의 전력 소비로 최대 80%의 성능을 제공합니다.



### HBM2

향상된 900GB/s의 기본 대역폭과 95%의 DRAM 사용 효율을 결합한 Tesla V100은 STREAM 측정기준으로 Pascal GPU보다 1.5배 향상된 메모리 대역폭을 제공합니다. 현재 Tesla V100은 표준 16GB 제품 메모리의 2배인 32GB 구성을 지원합니다.



### 프로그래밍 기능

Tesla V100은 처음부터 철저하게 프로그래밍 기능 단순화를 위해 설계되었습니다. 새로운 독립 스레드 스케줄링 기능은 소규모 작업들 간 리소스 공유를 통해 미세 동기화를 지원하며 GPU 사용률을 향상시킵니다.

# HPE-NVIDIA AI 오퍼링 AI Inference 용도

## Edge Compute + Data Center 전용구성

엣지(Edge)에서 생성되는 다양한 데이터를 수집하고 분석할수 있도록 구성된 AI Inference용 구성입니다.

### Large Scale



구성	수량	제품 사양 (노드당)
Server	1	HPE Apollo 6500 Gen10 Servers 2x Intel Gold 6242 (2.8GHz/16c), 512GB of Memory (최대 3TB확장), 4x 2TB NVMe SSDs, 2x 480GB SATA MU M.2 SSDs
GPU	8	NVIDIA Tesla T4

### Medium Scale



구성	수량	제품 사양 (노드당)
Server	1	HPE DL380 Gen10 Servers 2x Gold 6248 (2.5GHz/20c). 384GB of Memory, 2x960GB SATA 6G MU SSDs, 2x240GB SATA 6G RI SSDs for Boot
GPU	4	NVIDIA Tesla T4

### Edge-large Scale



구성	수량	제품 사양 (노드당)
Server	1	HPE Edgeline EL4000 4x10GbE 2xQSFP 2xm510 Xeon D-1548 2.0GHz 8-core, 128GB, 2x240GB SATA, 4x 1TB NVMe x4 Lanes Read Intensive M.2 2280
GPU	4	NVIDIA Tesla T4

### Edge-medium Scale



구성	수량	제품 사양 (노드당)
Server	1	<b>HPE Edgeline EL1000 10GbE 2xSFP</b> 1xm510 Xeon D-1587 1.7GHz 16-core, 128GB, 240GB SATA, 2x 1TB NVMe x4 Lanes Read Intensive M.2 2280
GPU	1	NVIDIA Tesla T4

# HPE Converged Edgeline System

## 운영기술과 엔터프라이즈급 IT를 결합한 edge to cloud 환경을 위한 새로운 제품 카테고리

### 엔터프라이즈를 cloud에서 에지로 확장

오늘날, 성공적인 조직은 에지로의 확장에 투자하고 있습니다. "에지"는 데이터센터와 클라우드가 아닌 그 어느곳이 될 수 있습니다. 예를들어 공장, 바다의 오일링, 공항 또는 사업장도 에지가 될 수 있습니다. 우리는 에지에서 "Things"를 통해 다양한 데이터를 모을 수 있습니다. "Things"는 자동차, 드릴, 펌프, 카메라 등 모든 기기를 의미합니다. 이러한 데이터를 신속하고 효율적으로 수집, 분석 및 실행함으로써 조직은 운영을 최적화하고 직원의 경험을 재정의하며 고객 만족도를 높이고 비즈니스 모델을 차별화 할 수 있습니다.

HPE Edgeline 컨버지드 에지 시스템은 혹독한 환경에서 사용하는 경우를 염두에 둔 컴팩트한 크기의 견고한 시스템으로, 에지에서 데이터 센터급 기능을 제공하여 IoT 데이터로부터 즉각적인 통찰력을 확보할 수 있습니다. 이를 통해 기업의 실시간 의사결정이 가능해지므로 비즈니스 운영에 가치를 더하고 결과적으로 성과를 촉진할 수 있습니다. 또한, 대기시간, 대역폭, 비용, 보안, 복제, 손상, 규제 준수 문제를 해결할 수 있어 비용 절감과 성공의 세가지 핵심 요소인 즉각적인 통찰력, 비용 절감, 신속한 조치가 가능해집니다.

Deep / Far Edge



고성능 딥 에지 컴퓨팅		EL4000
엔트리 딥 에지 컴퓨팅	 	EL1000 EL300
고성능 인텔리전트 게이트웨이		GL20
엔트리 인텔리전트 게이트웨이		GL10

#### Intel Xeon D/E3 최대 16코어 x 4

- 메모리 최대 128GB x 4,
- 작동 온도 0-55°C

#### Intel Xeon D/E3 최대 16코어

- 메모리 최대 128GB
- 작동 온도 0-55°C
- Intel Core i5 듀얼 코어
- 메모리 최대 32GB
- 작동 온도 -30~70°C, IP50 충격 · 진동

#### Intel Core i5 듀얼 코어

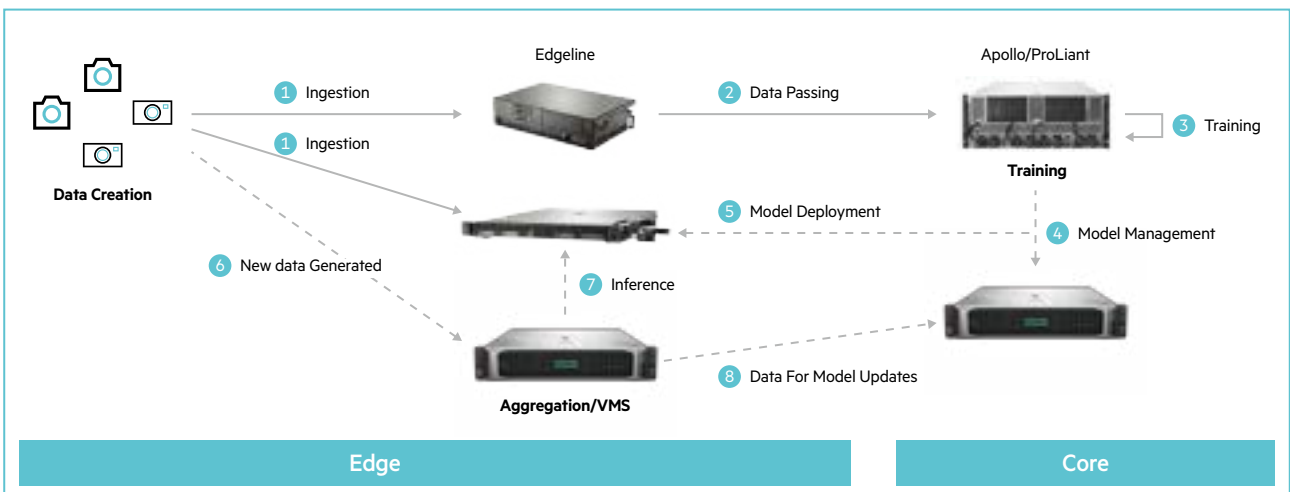
- 메모리 8GB, SSD 64GB
- 작동 온도 -20-60 °C, IP40 충격 · 진동
- 3G, LTE, WiFi / BT

#### Intel Core i5 듀얼 코어

- 메모리 8GB, SSD 64GB
- 작동 온도 -20-60 °C, IP40 충격 · 진동
- 3G, LTE, WiFi / BT

Near Edge

## Intelligent Video Analytics Data Flow



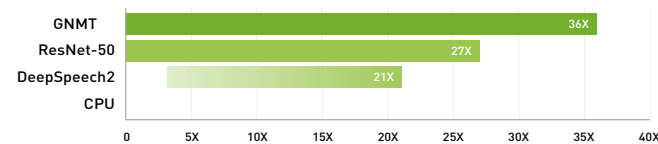


# NVIDIA T4 TENSOR CORE GPU

## 더욱 강력한 스케일아웃 AI 트레이닝 및 추론

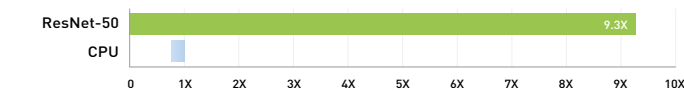
세계 최고 성능을 자랑하는 스케일 아웃 가속 그래픽 카드인 NVIDIA® T4 GPU로 서버의 속도를 높이십시오. NVIDIA Turing™ 텐서 코어가 탑재된 로우 프로파일 70W 설계로 획기적인 다중 정밀도 성능을 구현하여 다양한 최신 애플리케이션을 가속화할 수 있습니다. 에너지 효율이 뛰어난 70W의 소형 PCIe 폼팩터를 갖춘 이 고성능 GPU는 스케일아웃 서버에 최적화되어 있으며 최첨단 AI 구현을 위해 설계되었습니다.

### 추론 성능



NVIDIA T4 GPU 1개  
와 XEON GOLD 6140 CPU  
탑재 서버의 비교

### 트레이닝 성능



듀얼 NVIDIA T4 GPU  
와 듀얼 소켓  
XEON GOLD 6140 CPU  
탑재 서버의 비교



### GPU 아키텍처 NVIDIA TURING

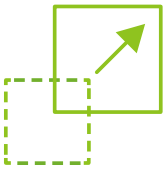
NVIDIA Turing 텐서 코어 수	320
NVIDIA CUDA 코어 수	2,560
단정밀도	8.1TFLOPS
혼합 정밀도(FP16/FP32)	65TFLOPS
INT8	130TOPS
INT4	260TOPS
GPU 메모리	16GB GDDR6 300GB/s
ECC	지원
인터커넥트 대역폭	32GB/sec
시스템 인터페이스	x16 PCIe Gen3
폼팩터	로우 프로파일 PCIe
열 솔루션	수동적
컴퓨팅 API	CUDA, NVIDIA TensorRT™, ONNX

## 데이터 센터 가속화를 실현하는 스케일아웃 성능

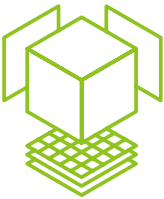


**70W의 소형 폼팩터 설계**로 스케일아웃 서버에 최적화된 T4는 CPU 대비 에너지 효율이 50배 높으며 그 결과 운영 비용을 크게 절감할 수 있습니다.

NVIDIA의 추론 플랫폼은 지난 2년간 10배 넘게 효율성을 높이면서 분산형 AI 트레이닝 및 추론을 위한 에너지 효율 솔루션으로서 최고의 자리를 지키고 있습니다.



**NVIDIA T4 데이터 센터 GPU**는 분산 컴퓨팅 환경에 이상적인 범용 가속 그래픽 카드입니다. 획기적인 다중 정밀도 성능으로 딥 러닝 및 기계 학습의 트레이닝과 추론, 비디오 트랜스코딩, 가상 데스크톱 속도를 대폭 향상합니다. 모든 AI 프레임워크와 네트워크 유형을 지원하는 T4는 성능과 효율성을 크게 높여 대규모 배포의 효율성을 극대화합니다.



**Turing 텐서 코어 기술**은 시용 다중 정밀도 컴퓨팅과 더불어 FP32부터 FP16, INT8 및 INT4 정밀도에 이르기까지 획기적인 성능을 제공합니다.

트레이닝에서는 CPU 대비 최대 9.3배, 추론에서는 최대 36배나 높은 성능을 발휘합니다.



## 고성능 컴퓨팅 및 인공지능/빅데이터를 위한 고집적 최적화 솔루션

HPE Apollo 시스템 제품군은 빅데이터, 분석, 개체 스토리지 및 고성능 컴퓨팅(HPC) 워크로드에 랙 스케일의 효율적인 컴퓨팅, 스토리지, 네트워킹, 전력 및 냉각 솔루션을 제공하도록 설계되었습니다. 이러한 랙 스케일 효율성을 바탕으로 다음과 같이 탁월한 비즈니스 이점을 제공하고 있습니다.

- 특정 워크로드에 최적화된 시스템에서 확장성, 성능 및 효율성을 적정 수준으로 조정
- 수개월이 걸리던 구현 시간을 단 며칠로 단축함으로써 가치 창출을 가속화
- 스케일업 및 스케일아웃 솔루션에 유연한 아키텍처를 제공
- 자본 지출(CAPEX) 및 운영 비용(OPEX)을 대폭 절감
- 파트너, 통합업체 및 ISV(독립 서비스 공급업체)가 참여하는 확장된 에코시스템을 활용
- 모든 것을 갖춘 서비스 및 지원 기능을 통해 안심하고 사용

HPE Apollo 시스템 포트폴리오에서 스케일아웃 컴퓨팅을 담당하는 솔루션으로는 하이퍼스케일과 범용 스케일아웃 컴퓨팅용 Apollo 2000 시스템, 빅데이터 분석 및 개체 스토리지용 Apollo 4000 시스템 제품군, HPC 및 슈퍼컴퓨터용 Apollo 6000 및 SGI 8600 수냉 시스템이 있습니다. 이렇게 포괄적인 제품군 덕분에 규모에 관계 없이 모든 기업들이 고밀도 서버 스토리지, 관리 및 랙 스케일 효율성을 경험할 수 있게 되었습니다.

또한, 단계적 접근 방식을 통해 데이터 기반 기업을 위한 빅데이터, 개체 스토리지 및 HPC를 활용하고 싶어하는 고객들이 간편하면서도 합리적으로 솔루션을 구현할 수 있도록 시작점을 제공하고 있습니다.

시 포트폴리오





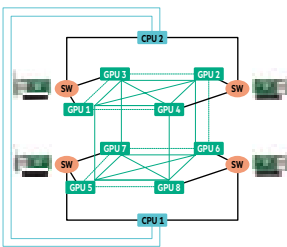
# HPE AI 최적화 제품 포트폴리오

## HPE Apollo 6500 Gen10

### 딥러닝과 HPE Apollo 6500 - 완벽한 조합

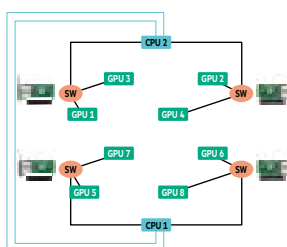
한층 다양한 애플리케이션 영역에 AI를 실현하는 딥러닝은 고성능 컴퓨팅에 기반하여 방대한 양의 데이터 내에서 패턴과 관계를 도출합니다. 지금까지, 기존 HPC시스템은 딥러닝이 요구하는 전례 없는 수준의 요건을 부응할 수 없었습니다. 따라서, 딥러닝의 이점은 사실상 이용되지 않은 상태로 남아 있습니다. HPE Apollo 6500 Gen10 시스템은 다양한 워크로드에 따른 토폴로지를 제공합니다.

**NVLink 2.0**



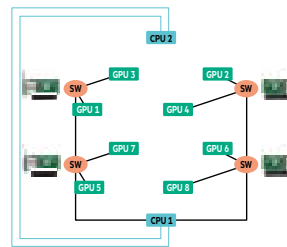
딥러닝/AI/HPC 애플리케이션을 위한 하이브리드 큐브 메쉬 기반 NVLink 2.0 토폴로지

**PCIe 4:1**



HPC 애플리케이션을 위한 가장 전통적인 4:1 토폴로지

**PCIe 8:1**



딥러닝 트레이닝을 위한 8:1 토폴로지 (GPUDirect)



<b>그래픽 처리 장치</b>	서버당 최대 8개의 GPU, 최대 350 와트 GPU 지원, NVLink 2.0 또는 표준 PCIe를 이용한 고속 전용 GPU간 통신 옵션 제공
<b>지원 가속기</b>	NVLink 및 PCIe 구성의 NVIDIA® Tesla V100, T4, Quadro RTX 6000, NEC Vector Engine IOB, AMD MI25
<b>어댑터</b>	4가지 고속 패브릭 어댑터 (이더넷, Intel Omni-Path Architecture, InfiniBand EDR, 향후 InfiniBand HDR)
<b>프로세서</b>	1세대 및 2세대 Intel® Xeon® 프로세서 Scalable Family 탑재
<b>메모리</b>	24개 2933MT/s HPE SmartMemory 지원 (최대 3TB)
<b>스토리지</b>	최대 16개의 SFF 장치; 최대 4 NVMe SSD 전면 액세스 스토리지를 탑재한 핫플러그 SAS/SATA SSD
<b>네트워크 옵션</b>	<ul style="list-style-type: none"> <li>• 4개의 1GbE 단일 포트 RJ-45</li> <li>• InfiniBand EDR 또는 Intel OPA 또는 고속 이더넷을 지원하는 4개의 16PCI Express 슬롯, 최대 100 Gbps</li> </ul>
<b>I/O 슬롯</b>	<ul style="list-style-type: none"> <li>• 1x 16FHHL PCIe Gen3 시스템 보드 모듈</li> <li>• GPU 모듈의 4 x 16PCIe 로우 프로파일 Gen3</li> </ul>
<b>전원공급장치</b>	최대 4개의 2200W 80 Plus Platinum (2 + 2)
<b>랙</b>	표준 1075 mm 깊이 랙에서 사용 가능

# HPE ProLiant ML350 Gen10

## 가장 강력하고 유연한 ProLiant 2소켓 타워서버



지원 프로세서	1세대 및 2세대 Intel® Xeon® 프로세서 Scalable Family 탑재
프로세서 수	2
프로세서당 코어	4/6/8/10/12/14/16/18/20/22/24/26/28
최대 프로세서 성능/캐시	3.8GHz / 38.50 MB
최대 내장 메모리	24개 2933MT/s HPE SmartMemory지원 (최대 3TB)
스토리지 컨트롤러	HPE Smart Array S100i/E208i/P408i/P816시리즈
스토리지 드라이브 베이	최대 24 SFF HDD/SSD 지원 또는 최대 12 LFF HDD/SSD 최대 8 NVMe PCIe SSD 지원
최대 내장 스토리지 용량	368.64TB
GPU 지원	최대 Double Wide GPU 4개 장착 가능 최대 Single Wide GPU 4개 장착 가능
네트워킹 포트/FlexibleLOM	4 x 1GbE 온보드 / FlexibleLOM 슬롯 및 PCIe NIC 지원 (1/10/25/100GbE)
폼 팩터	타워형 (4U), 랙형 (5U)
시스템 LOM	UEFI Class 2 / Legacy BIOS
시스템 관리	HPE iLO 5, SUM, RESTful Interface Tool, UEFI
보증기간 (연 기준: parts/labor/onsite)	3/3/3

# HPE ProLiant DL380 Gen10

## 멀티 워크로드 컴퓨팅을 위한 업계 1위의 데이터센터 표준 서버



SFF 새시 전면



LFF 새시 전면



SFF 새시 후면



LFF 새시 후면

지원 프로세서	1세대 및 2세대 Intel® Xeon® 프로세서 Scalable Family 탑재
프로세서 수	2
프로세서당 코어	4/6/8/10/12/14/16/18/20/22/24/26/28
최대 프로세서 성능/캐시	3.8GHz / 38.50 MB
최대 내장 메모리	24개 2933MT/s HPE SmartMemory지원 (최대 3TB)
퍼시스턴트 메모리	12개 16GB 2666MT/s NVDIMM지원 (최대 192GB) 12개 128/256/512GB 2666MT/s DC Persistent Memo지원 (최대 6TB)
스토리지 컨트롤러	HPE Smart Array S100i/E208i/P408i/P816시리즈
스토리지 드라이브 베이	최대 24+6 SFF 또는 12+4+3 LFF+2 SFF HDD/SSD지원 또는 최대 20 NVMe PCIe SSD 2개의 M.2 추가 지원 (PCIe Riser)
최대 내장 스토리지 용량	460.8TB
GPU 지원	최대 Double Wide GPU 3개 장착 가능 최대 Single Wide GPU 5개 장착 가능
네트워킹 포트/FlexibleLOM	4 x 1GbE 온보드 / FlexibleLOM 슬롯 및 PCIe NIC 지원 (1Gb, 10Gb, 25Gb, 40Gb, 50Gb, 100Gb)
폼 팩터	랙형 (2U)
시스템 LOM	UEFI Class 2 / Legacy BIOS
시스템 관리	HPE iLO 5, SUM, RESTful Interface Tool, UEFI
보증기간 (연 기준: parts/labor/onsite)	3/3/3

# HPE Edgeline 컨버지드 에지 시스템 및 HPE Intelligent 게이트웨이

## OT 및 엔터프라이즈 IT 컨버전스를 구현하는 고성능 에지 시스템



**GL10 Gateway**



**GL20 Gateway**

<b>컴퓨팅</b>	Intel E3826 듀얼 코어 Atom 1.46GHz 및 내장형 GPU 4GB DDR3L-1067 SO-DIMM (8GB CTO 옵션)	Intel 4300U 듀얼 코어 i5 1.9GHz 및 Intel HD 4400 그래픽 8GB DDR3L-1600/1333 SO-DIMM
<b>I/O</b>	GbE 이더넷 포트 1개 표준 크기 미니 PCIe 확장 슬롯 1개, ½ 크기 미니 PCIe 확장 슬롯 1개 Wifi, BT, 3G 용 Sim 슬롯	GbE 이더넷 포트 1개 표준 크기 미니 PCIe 확장 슬롯 3개, ½ 크기 미니 PCIe 확장 슬롯 1개 Wi-Fi, BT, 3G 및 LTE 용 Sim 슬롯
<b>스토리지</b>	2.5" SATA HDD 베이 1개 32GB SATA SSD (표준) 미니 PCIe 확장 슬롯을 통해 mSATA 지원	2.5" SATA HDD 베이 1개 64GB SATA SSD (표준) 미니 PCIe 확장 슬롯을 통해 mSATA 지원
<b>환경</b>	사용 온도 범위: -20C ~ 60 C	사용 온도 범위: -20C ~ 60 C
<b>충격 및 진동</b>	3Grms, IEC 60068-2-64, 무작위, 5~500Hz, 1시간/축 30G, IEC 60068-2-27 반정현, 11ms 내진	3Grms, IEC 60068-2-64, 무작위, 5~500Hz, 1시간/축 30G, IEC 60068-2-27 반정현, 11ms 내진
<b>크기</b>	가로 13.9cm, 세로 11.7cm, 높이 3.6cm DIN 및 벽면 장착용 범용 마운팅 키트	가로 26.5cm, 세로 13.3cm, 높이 7.5cm DIN 및 벽면 장착용 범용 마운팅 키트
<b>전원</b>	12V DC 입력 전압, 범용 AC 전원 키트 (미국, 영국, 유럽, 일본)	9~36V DC 입력 전압, 범용 AC 전원 키트 (미국, 영국, 유럽, 일본)



**Edgeline EL300**



**Edgeline EL1000**




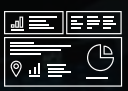


**Edgeline EL4000**

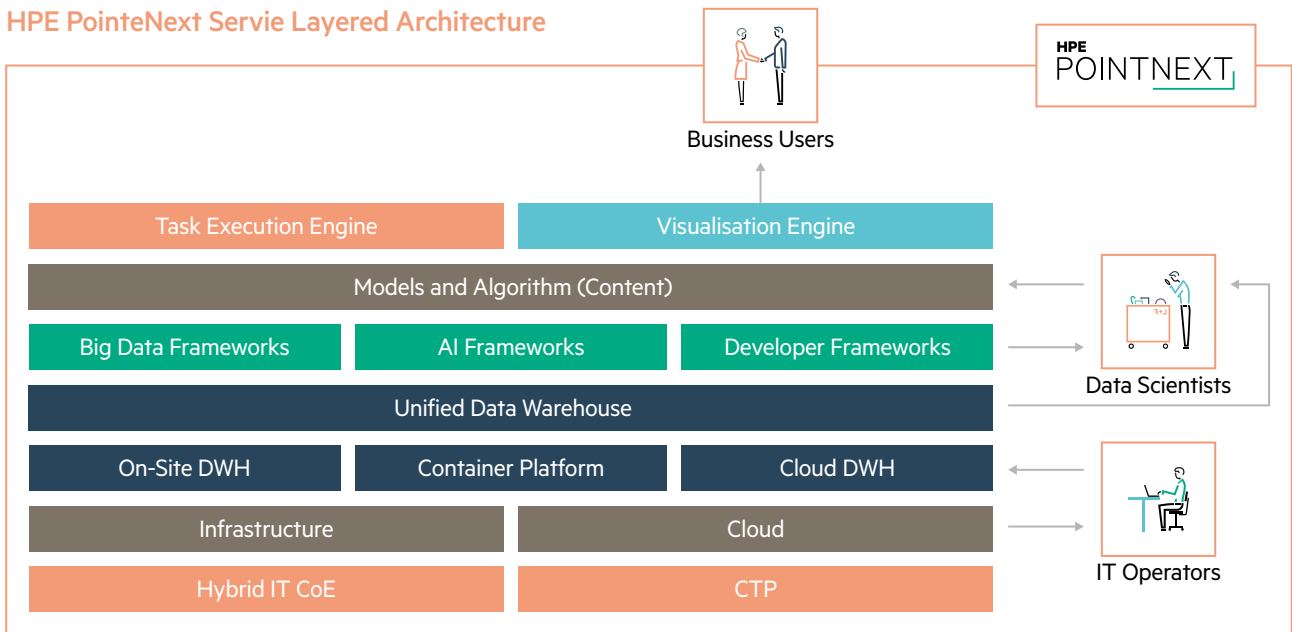
<b>컴퓨팅</b>	<ul style="list-style-type: none"> <li>Intel 7300U 듀얼 코어 i5 2.6GHz</li> </ul>	<ul style="list-style-type: none"> <li>HPE m510 (Intel Xeon D 8C/16C 또는 m710x (Intel Xeon E3 4C + workstation GPU) 1개</li> </ul>	<ul style="list-style-type: none"> <li>HPE m510 (Intel Xeon D 8C/16C 또는 m710x (Intel Xeon E3 4C + workstation GPU) 최대 4개</li> </ul>
<b>I/O</b>	<ul style="list-style-type: none"> <li>최대 6개의 1GbE 포트 (TSN)</li> <li>Daughter card 1개 옵션 (4포트 1GbE TSN, CAN Bus, Serial, GPIO)</li> <li>WiFi, BT, LTE 용 Sim 슬롯</li> </ul>	<ul style="list-style-type: none"> <li>최대 PCIe Full-Height Half-Length (FHHL) 카드 2개</li> <li>PXI/PXle 하이브리드 슬롯 2개</li> <li>내부 미니 PCIe 슬롯 2개</li> <li>최대 2개의 1GbE/10GbE 포트</li> <li>WiFi, BT, LTE 용 Sim 슬롯</li> </ul>	<ul style="list-style-type: none"> <li>최대 PCIe Full-Height Half-Length (FHHL) 카드 4개</li> <li>PXIe 슬롯 4개</li> <li>10 GbE SFP+ 또는 4x10 GbE QSFP+ 포트 2개</li> </ul>
<b>스토리지</b>	<ul style="list-style-type: none"> <li>M.2 SATA SSD 또는 2.5" SFF SSD 1개</li> <li>M.2 NVMe SSD 1개</li> </ul>	<ul style="list-style-type: none"> <li>2.5" SFF HDD 또는 SSD 2개 (최대 4TB)</li> <li>온 보드 서버 카트리리지 SSD</li> </ul>	<ul style="list-style-type: none"> <li>온 보드 서버 카트리리지 SSD</li> </ul>
<b>환경</b>	<ul style="list-style-type: none"> <li>사용 온도 범위: -30 ~ 70° C</li> </ul>	<ul style="list-style-type: none"> <li>사용 온도 범위: 0 ~ 55° C</li> </ul>	<ul style="list-style-type: none"> <li>사용 온도 범위: 10 ~ 55° C</li> </ul>
<b>충격 진동 범위</b>	<ul style="list-style-type: none"> <li>3Grms, IEC-600068-2-64, 5Hz ~ 500Hz, 1분</li> <li>30G, IEC-600068-2-27 반정현, 11ms 내진</li> </ul>	<ul style="list-style-type: none"> <li>3Grms, 무작위, IEC-60068-2-64, 5Hz ~ 500Hz, 1시간/축</li> <li>30G, IEC-600068-2-27 반정현, 11ms 내진</li> </ul>	<ul style="list-style-type: none"> <li>3Grms, 무작위, IEC-60068-2-64, 5Hz ~ 500Hz, 1시간/축</li> <li>30G, IEC-600068-2-27 반정현, 11ms 내진</li> </ul>
<b>크기</b>	<ul style="list-style-type: none"> <li>가로 232mm, 세로 100mm, 높이 200.5mm</li> </ul>	<ul style="list-style-type: none"> <li>가로 35.1cm, 세로 23.3cm, 높이 8.8cm</li> </ul>	<ul style="list-style-type: none"> <li>가로 43.5cm, 세로 61.5cm, 높이 4.3cm, 1U</li> </ul>
<b>전원</b>	<ul style="list-style-type: none"> <li>Typical: 80W</li> <li>12VDC</li> </ul>	<ul style="list-style-type: none"> <li>Typical: 100~150 W</li> <li>95~265VAC 또는 -48VDC</li> </ul>	<ul style="list-style-type: none"> <li>Typical: 400~600 W,</li> <li>95~265VAC 또는 -48VDC</li> </ul>
<b>관리</b>	<ul style="list-style-type: none"> <li>HPE iSM</li> <li>HPE EIM</li> </ul>	<ul style="list-style-type: none"> <li>HPE iLO 4</li> </ul>	<ul style="list-style-type: none"> <li>HPE iLO 4</li> </ul>
<b>메모리</b>	<ul style="list-style-type: none"> <li>시스템당 최대 32 GB</li> </ul>	<ul style="list-style-type: none"> <li>시스템당 최대 128 GB</li> </ul>	<ul style="list-style-type: none"> <li>시스템당 최대 512 GB</li> </ul>

# HPE Pointnext AI 컨설팅

HPE PointNext는 Deep Learning을 위한 SW Stack의 구축 및 지원부터 IoT를 위한 Edge Inference 프로젝트까지 지원하고 있습니다. GPU 전문 엔지니어, Data Scientists, Docker SW 엔지니어 등 전문 인력을 통한 AI 비즈니스 구현이 가능합니다.

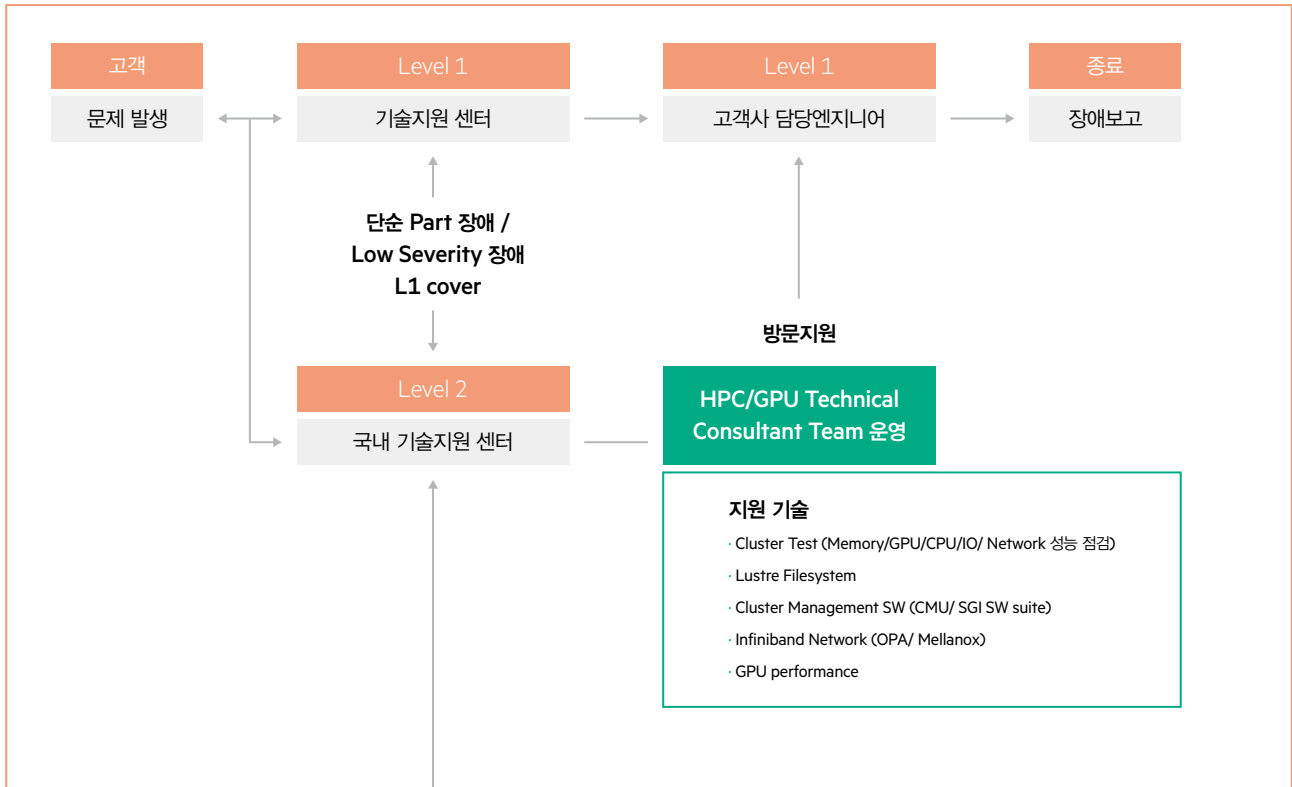
Data Sources	Data Management	Data Sources	Insight
			
<ul style="list-style-type: none"> <li>· Multiple formats</li> <li>· Multiple Sources</li> <li>· Multiple Standards</li> <li>· Security</li> </ul>	<ul style="list-style-type: none"> <li>· Data Platform</li> <li>· Data Lake</li> <li>· On Premise</li> <li>· Hybrid Cloud Integration</li> </ul>	<ul style="list-style-type: none"> <li>· Multiple Machine Learning Libraries</li> <li>· Multiple AI trained models</li> <li>· Multiple AI languages</li> </ul>	<ul style="list-style-type: none"> <li>· Business Use Cases</li> <li>· Integrated Analysis</li> <li>· 360 view of data</li> <li>· CDO Dashboards</li> </ul>

## HPE PointNext Service Layered Architecture

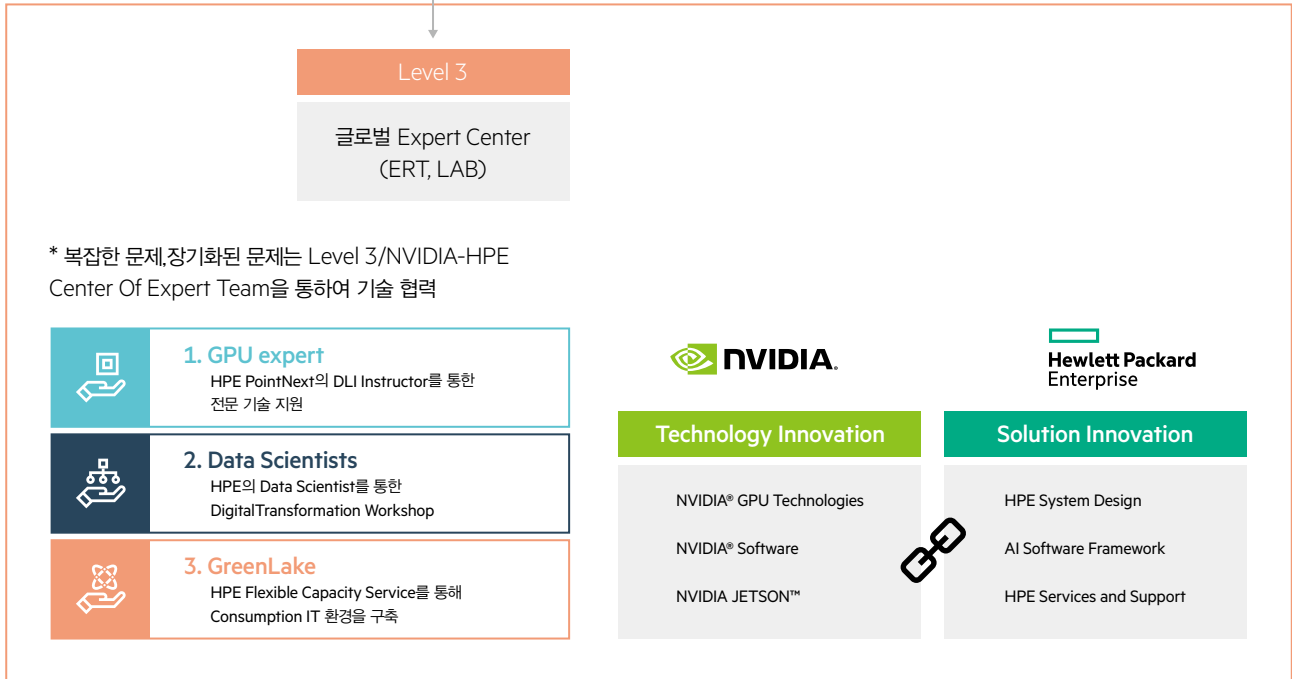


# HPE PointNext - HPC/AI지원 프로세스

## 국내 기술지원



## Global 기술지원





Sign up for updates

★ Rate this document

---

## 휴렛팩커드 엔터프라이즈

제품정보 및 구입안내: 080-703-0700

©Copyright 2018 Hewlett Packard Enterprise Development LP.

본 문서에 게재된 정보는 사전고지 없이 변경될 수 있습니다. HPE 제품 및 서비스에 대한 보증은 오직 해당 제품 및 서비스에 첨부된 보증서 내용에만 상응합니다. 본 문서의 어떠한 부분도 추가적인 보증으로 유추될 수 없으며, HPE는 본 문서의 기술적 오류나 편집 상 오류, 또는 누락에 대한 책임을 지지 않습니다.

© 2018 NVIDIA Corporation. All rights reserved. NVIDIA, NVIDIA 로고, NVIDIA Turing, CUDA 및 TensorRT는 미국 및 기타 국가에 소재한 NVIDIA Corporation의 상표 및/또는 등록 상표입니다. 그 밖에 모든 상표 및 저작권은 각 소유권자의 재산입니다. DEC18

