한국아이비엠주식회사
(07326) 서울시 영등포구 국제금융로10
서울국제금융센터 (Three IFC)

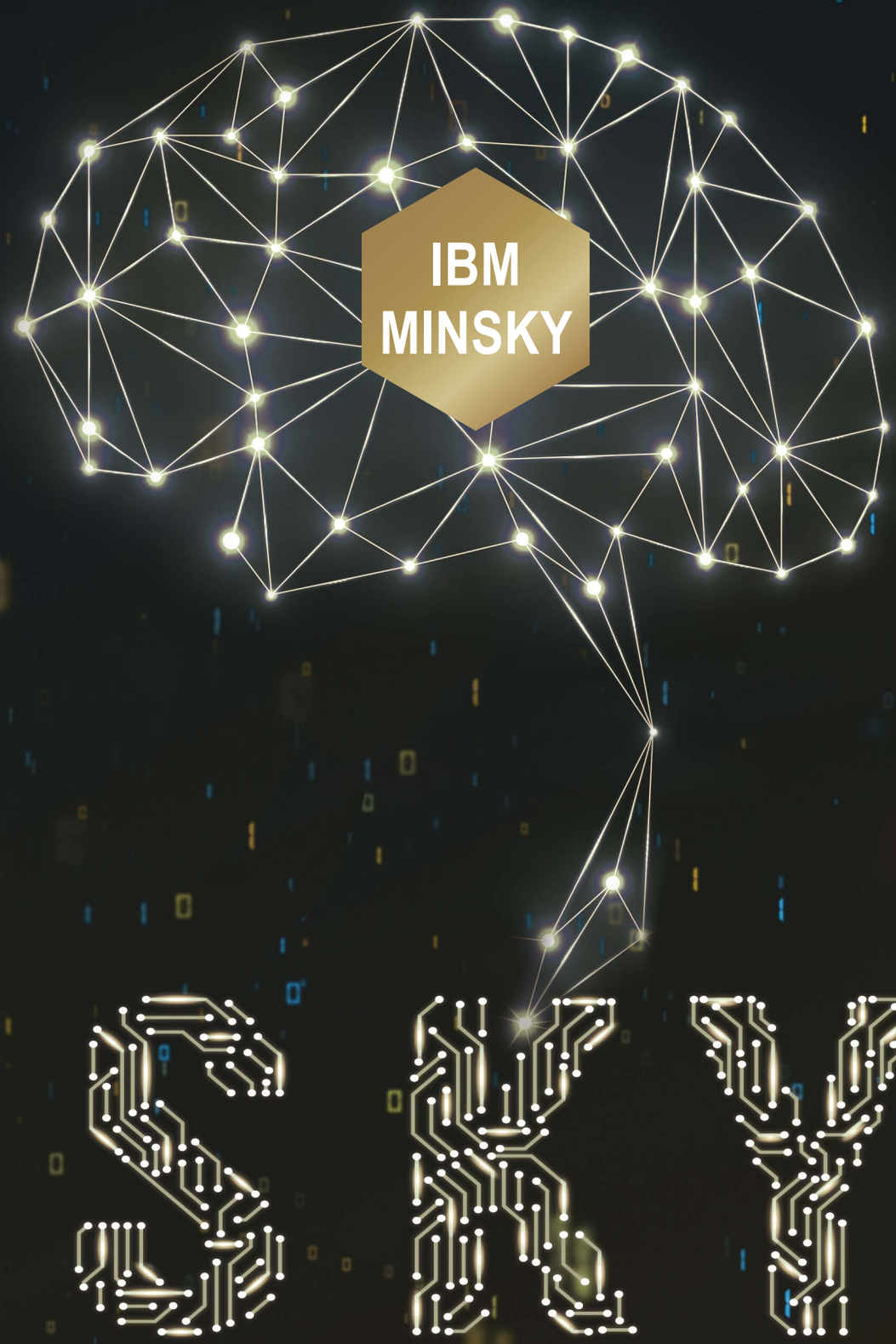
TEL : (02) 3781-7900
www.ibm.com/kr

2017년 2월

Printed in Korea

NVIDIA의 최신 GPU기술을 탑재한

IBM 딥러닝 서버 솔루션



딥러닝,
‘인공지능의 봄’을 알리다

딥러닝(Deep Learning)이란?

인공 신경망을 기반으로 한 머신러닝 방법론 중 하나로,
인간의 두뇌가 수많은 데이터 속에서 패턴을 발견한 뒤 사물을 구분하는
정보처리 방식을 모방해 컴퓨터가 사람처럼 스스로 학습하여 판단하는 기술입니다.

인공지능의 핵심기술 그 중심에는 ‘딥러닝’이 있습니다

자율 주행차, 실시간 금융 사기 방지 및 신약 개발과 같은
새로운 산업의 발전은 예전과는 다른 차원의
인공지능 기술을 필요로 합니다.
이러한 기술의 핵심에 ‘딥러닝’이 있습니다.

업계 최초의 혁신!
IBM 딥러닝 서버 솔루션

IBM MINSKY

IBM Minsky는
인공지능에서부터 딥러닝, 첨단 빅데이터 분석
그리고 연산 집약적인 작업을
더욱 빠르고 효율적으로 처리해 줍니다.

IBM 딥러닝 서버 솔루션

최신, 최고의 GPU
PASCAL P100

- NVIDIA의 최신 PASCAL 아키텍처 P100 GPU 장착
- 딥러닝을 위한 Half- precision 성능 **21TFLOPS**
- 기존의 **3배**에 달하는 GPU메모리 대역폭

쉽고 빠른 딥러닝 프레임워크 제공
PowerAI

- IBM의 **딥러닝 소프트웨어 툴킷** PowerAI 제공
- **CAFFE, Torch, TensorFlow** 등 주요 딥러닝 프레임워크를 최적화하여 패키지로 제공

IBM Minsky 4대 장점

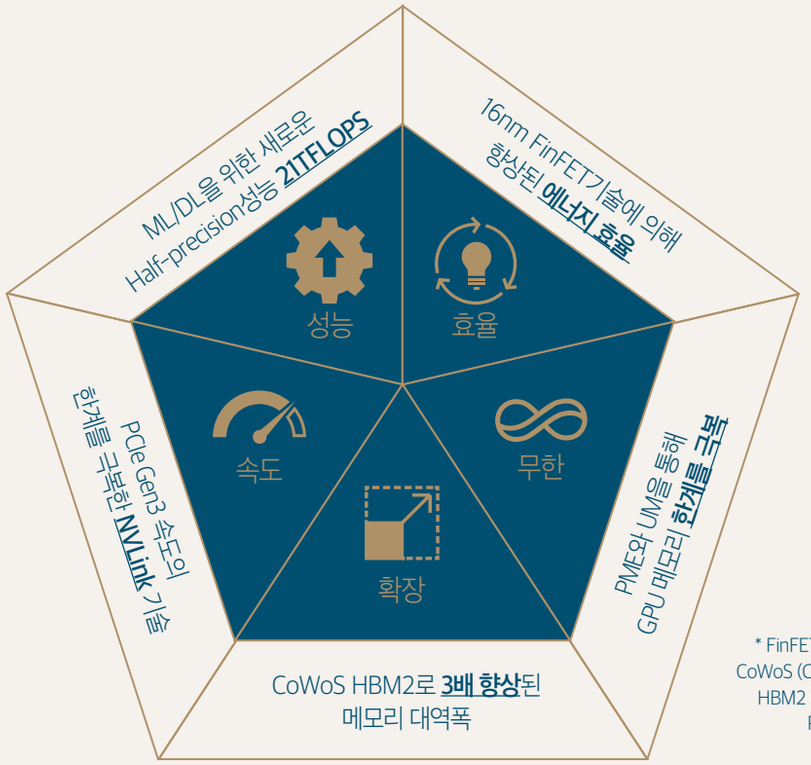
신기술에 의한
기존 문제의 해결

- **Unified Memory**로 GPU 메모리 한계 극복, P2P 문제 해결
- **NVLink** 기술로 GPU-CPU간 병목 해결, 획기적 성능 향상

진정한 오픈 아키텍처
OpenPOWER 플랫폼

- POWER 아키텍처 공개에 의한 진정한 오픈 아키텍처
- **IBM / Mellanox / NVIDIA** 협업을 통한 굳건한 GPU 솔루션 로드맵
- **NVIDIA-IBM Acceleration Lab** 지원

그 무엇도 따라갈 수 없는 NVIDIA PASCAL P100 아키텍처의 신기술 Big 5



* FinFET (Fin Field Effect Transistor)
CoWoS (Chip-on Wafer-on-Substrate)
HBM2 (High Bandwidth Memory 2)
PME (Page Migration Engine)
UM (Unified Memory)

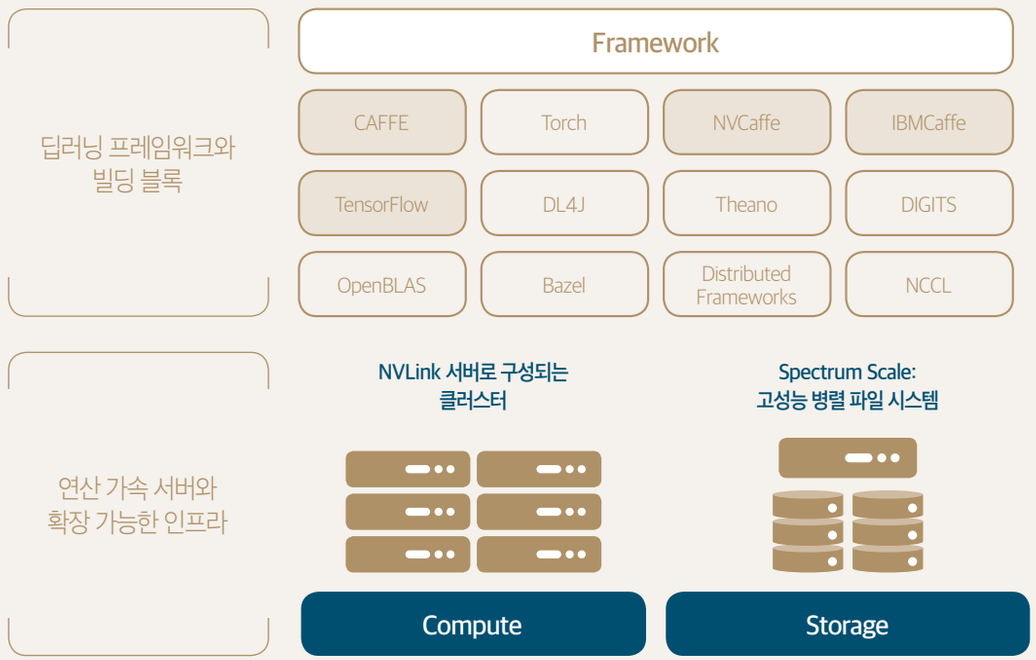
▷Tesla P100 vs. 기존 GPU 사양 비교

Tesla Products	Tesla K40	Tesla K80	Tesla M40	Tesla P100 (NVLink)
GPU / Form Factor	KeplerGK110 / PCIe	KeplerGK210 / PCIe	MaxwellGM200 / PCIe	PascalGP100 / SXM2
Stream Processors	2880	2 * 2496	3072	3584
Base Clock	745 MHz	562 MHz	948 MHz	1328 MHz
GPU Boost Clock	810/875 MHz	875 MHz	1114 MHz	1480 MHz
FP16 TFLOPs[1]	4.29	8.74	6.84	**21.2**
FP32 TFLOPs[1]	4.29	8.74	6.84	10.6
FP64 TFLOPs[1]	1.43	2.91	0.21	5.3
Memory Interface	384-bit GDDR5	384-bit GDDR5	384-bit GDDR5	4096-bit HBM2
Memory Bandwidth	288 GB/s	480 GB/s	288 GB/s	732 GB/s

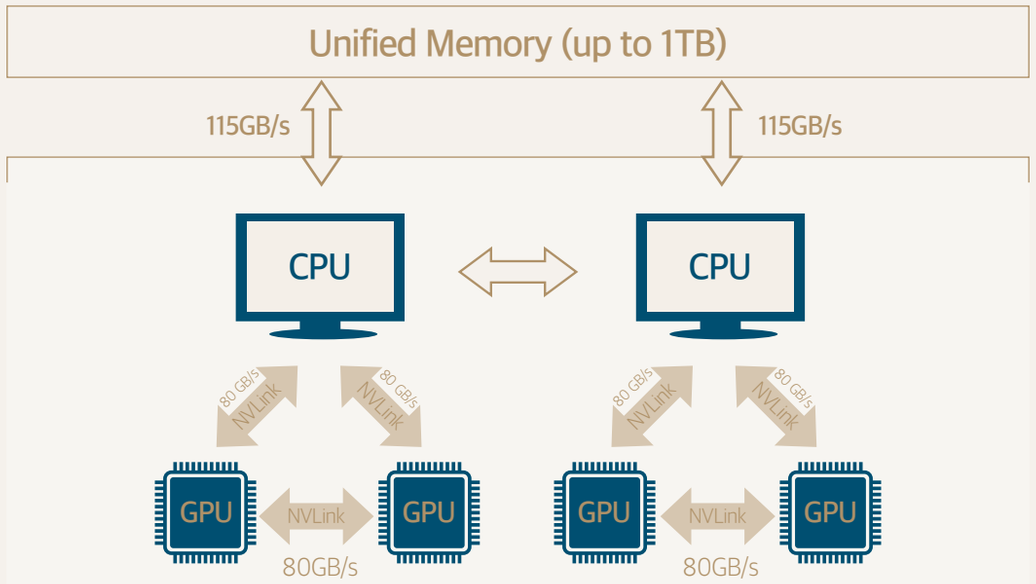
[Source] https://devblogs.nvidia.com/parallelforall/inside-pascal/
http://www.anandtech.com/show/8729/nvidia-launches-tesla-k80-gk210-gpu
http://www.anandtech.com/show/10222/nvidia-announces-tesla-p100-accelerator-pascal-power-for-hpc

딥러닝 프레임워크를 사용하는 가장 쉽고 빠른 방법, PowerAI

PowerAI는
오픈소스 딥러닝 프레임워크를 IBM Minsky에 최적화하여
제공하기에 누구나 쉽고 빠르게 설치하도록 지원합니다.



GPU를 'full peer' 로 취급하여 P2P 문제 해결

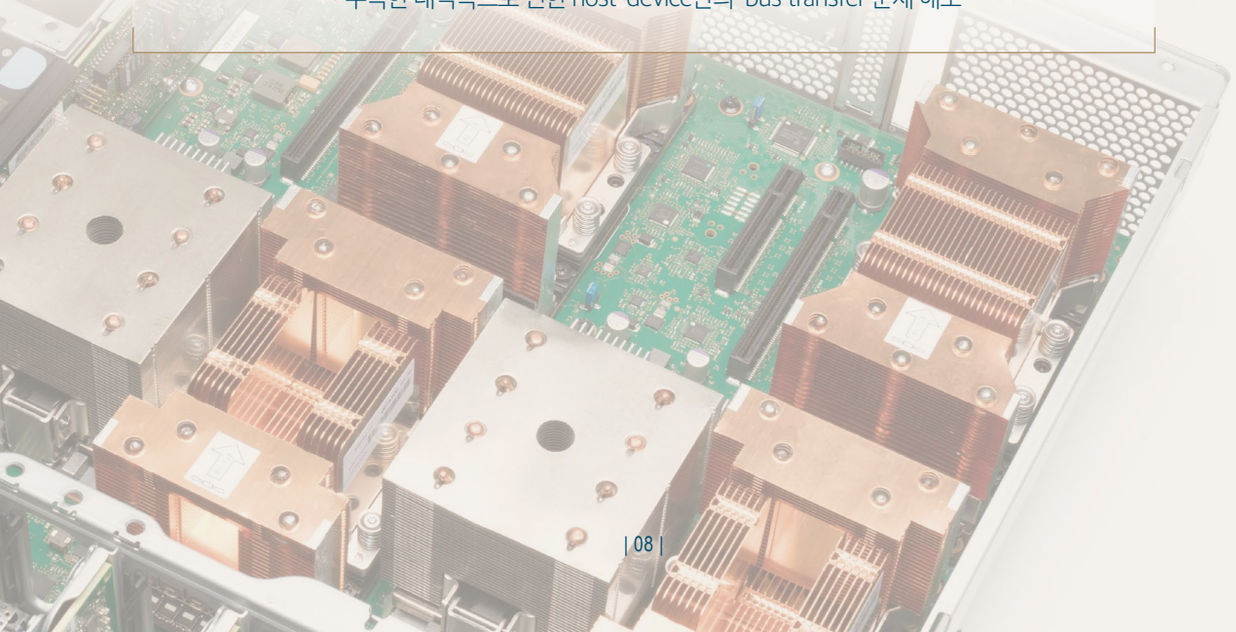


☑ Minksy는 '두껍고도 수평적으로' (both fat and flat) 설계된 시스템

- 어느 link에서도 data 병목이 생기지 않도록 설계
- GPU에서도 CPU처럼 시스템 메모리를 취급 (시스템 메모리 최대 1TB)

☑ 보편적 업무와 알고리즘에 잘 맞는 구조

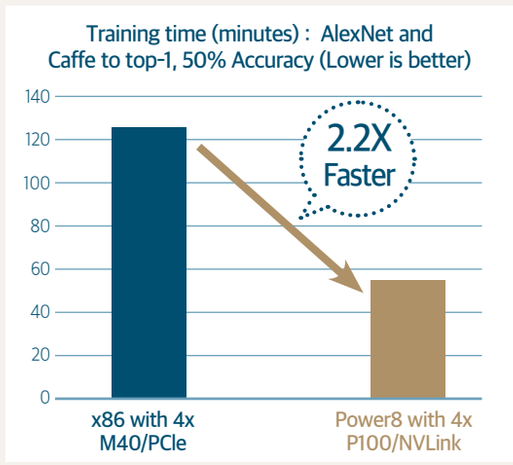
- Startup/teardown시 폭발적인 성능
- 두 GPU간의 안정적 transfer
- 부족한 대역폭으로 인한 host-device간의 bus transfer 문제 해소

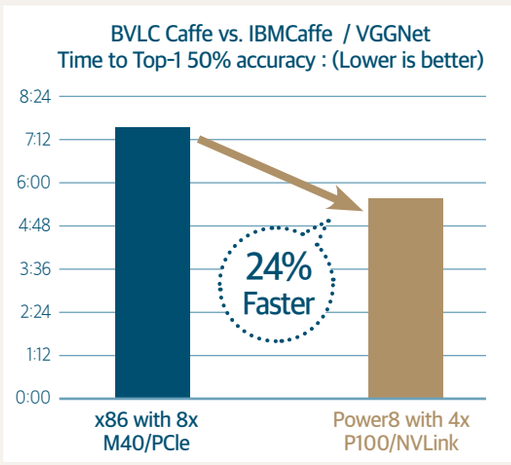


기존 GPU 컴퓨팅의 병목을 해결하는 NVLink

- 세계 유일 GPU-GPU 뿐만 아니라 GPU-CPU도 NVLink로 연결 가능
- 기존 PCIe Gen3 대비 약 2.5배 대역폭 제공

▷ 딥러닝 프레임워크 성능 테스트 결과



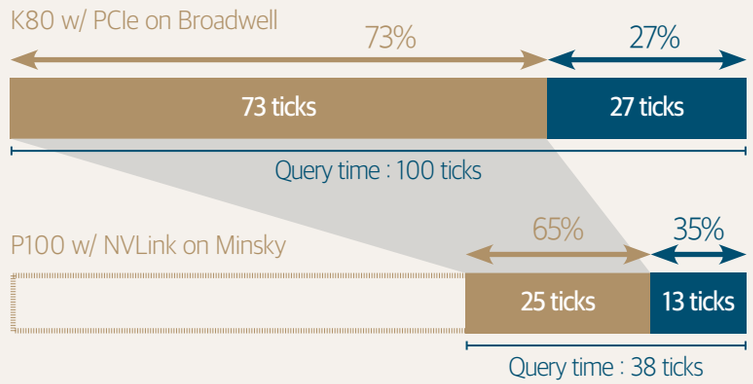


* IBM S822LC 20-cores 2.86GHz 512GB memory / **4 NVIDIA Tesla P100 GPUs** / Ubuntu 16.04 / CUDA 8.0.44 / cuDNN 5.1 / IBMCaffe 1.0.0-rc3 / Imagenet Data

* Intel Broadwell E5-2640v4 20-core 2.6 GHz 512GB memory / **8 NVIDIA Tesla M40 GPUs** / Ubuntu 16.04 / CUDA 8.0.44 / cuDNN 5.1 / BVLC Caffe 1.0.0-rc3 / Imagenet Data

▷ NVLink vs. PCIe Gen3 - GPU DB의 query 테스트 결과

NVLink로 data copy시 **속도 3배 향상** (73tick▷25tick)

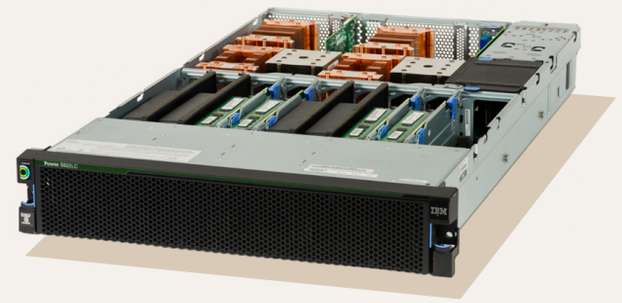


Data Transfer
GPU Compute

전체 속도 **2.6배 향상**

- 전체 소요 시간 감축 수치 : **62 tick (1 tick = 0.01 sec)**
- Data Transfer에서의 감축 : 48 ticks 전체 감소치의 77%
- GPU 계산에서의 감축 : 14 ticks 전체 감소치의 23%

IBM Minsky 세부 규격 및 하드웨어 구조



IBM Power8 CPU와 NVIDIA P100 GPU의 조합

- 최신 Pascal 아키텍처의 P100 4장 장착
- 양방향 40+40GB/sec의 대역폭을 가지는 NVLink를 통해 GPU-GPU는 물론, CPU-GPU도 연결
- 물리적 core 1개당 8개의 HW thread (SMT-8)를 가지는 Power8 프로세서
- 2U 공간 안에 강력한 GPU 컴퓨팅 파워를 압축하여 성능 대비 상면적 및 전력 소비량에서 월등한 이점

IBM Minsky(IBM Power System S822LC for HPC) 개요	
시스템 구성(8335-GTB)	
마이크로프로세서	8코어 3.25GHz Power8 프로세서 카드 2개 또는 10코어 2.86GHz Power8 프로세서 카드 2개
L2(Level 2) 캐시	코어당 512KB L2 캐시
L3(Level 3) 캐시	코어당 8MB L3 캐시
L4(Level 4) 캐시	소켓당 최대 64MB
메모리 최소/최대	4GB, 8GB, 16GB, 32GB DDR4 모듈, 128GB ~ 1TB 총 메모리
프로세서-메모리 대역폭	소켓당 115GB/초, 시스템당 230GB/초(SCM에서 L4 캐시까지의 최대 지속 메모리 대역폭) 소켓당 170GB/초, 시스템당 340GB/초(L4 캐시에서 DIMM까지 최대 피크 메모리 대역폭)
스토리지 및 입출력(I/O)	
표준 백플레인	하드 디스크 드라이브(HDD) 또는 솔리드 스테이트 디스크(SDD)를 위한 SFF(small form factor) 베이 2개
미디어 베이	해당 없음
RAID 옵션	통합 PCIe 어댑터에서 하드웨어 RAID 지원
어댑터 슬롯	PCIe Gen3 슬롯 3개: x16 PCIe Gen3 2개, x8 PCIe Gen3 1개. 모두 CAPI 지원
I/O 대역폭	64GBps
GPU 액셀러레이터	최대 4개의 NVIDIA Tesla P100(NVLink GPU)
전원, RAS, 시스템 소프트웨어, 물리적 특성과 보증	
전원	200V ~ 240V
RAS 기능	• 프로세서 명령 재시도 • Chipkill 메모리 • 결함 모니터링 기능이 있는 서비스 프로세서 • 핫플러그 및 이중 전원/냉각 팬(GPU 설치 시 전원 이중화 없음) • 선택 동적 펌웨어 업데이트 • ECC L2 캐시, L3 캐시 • 핫스왑 가능 디스크 베이
운영 체제*	Linux on POWER
시스템 크기	441.5W x 86H x 822D mm