

IBM Power System IC922

AI 모델 가동을 위해 설계된 가속 추론 서버

주요 특징

- 우수한 IO 및 메모리 대역폭을 통해 데이터에서 빠르게 인사이트 확보
- AI에 최적화된 하드웨어 및 소프트웨어 스택 지원
- 모듈식으로 비즈니스 요구에 맞게 확장 가능
- AI 여정에 필요한 엔터프라이즈급 보안 및 지원 제공

조직의 AI 여정이 진행됨에 따라 AI 이니셔티브는 훈련 단계에서 추론 단계로 전환됩니다. AI를 구현하려면 단계마다 다양한 인프라 요구사항이 충족되어야 하므로 다목적 AI 서버가 훈련과 추론 둘 다에 가장 적합한 것은 아닐 수 있습니다. 조직은 IT 인프라를 평가하여 AI의 모든 단계에 적합한지 확인해야 합니다. 그렇지 않으면 AI 프로젝트가 실패로 돌아갈 위험이 있습니다.

AI 추론을 위해 설계된 IBM Power System IC922는 훈련된 AI 모델에서 비즈니스 인사이트를 확보하는 데 필요한, 컴퓨팅 집약적이고 Latency가 짧은 인프라를 제공합니다. POWER9을 기반으로 하는 Power IC922는 진보된 형태의 인터커넥트(PCIe 4세대, OpenCAPI)를 제공하는데, 이를 통해 데이터 처리량을 향상시키고 Latency를 단축할 수 있습니다. 가속화된 Power IC922는 최대 6개의 NVIDIA® T4 GPU를 지원합니다.

추론을 수행하려면 AI 모델이 스토리지 또는 메모리에 쉽고 빠르게 액세스할 수 있어야 합니다. 풍부한 용량의 스토리지를 갖춘 Power IC922는 소켓당 170GB/s의 최대 피크 메모리 대역폭을 제공하며 2U 시스템에서 최대 24개의 SAS/SATA 드라이브와 결합됩니다.

AI는 복잡한 워크로드이므로 통합된 AI 스택이 필요합니다. Power IC922를 사용하면, AI에 최적화된 POWER9 하드웨어 및 소프트웨어를 활용할 수 있으며 Power AC922 훈련 플랫폼과 통합할 수 있습니다. 이를테면, Power AC922에서 모델을 훈련시킨 다음, 해당 모델을 Power IC922로 넘겨 추론 작업을 수행하도록 할 수 있습니다.

또한, Power IC922는 모듈식으로 설계되어 있기 때문에 조직은 온프레미스이건 프라이빗 클라우드 환경이건 상관없이 비즈니스 요구에 맞게 인프라를 확장하고 엔터프라이즈급 지원을 받을 수 있습니다.



IBM Power System IC922 서버

Power System IC922(9183-22X) 요약

시스템 구성	
마이크로프로세서	POWER9 2개 - 12코어, 16코어 또는 20코어
레벨 2(L2) 캐시	512K
레벨 3(L3) 캐시	10MB
RAM(메모리)	최대 2TB - 제공된 16개 또는 32개 DDR4 RDIMM 슬롯을 통해 메모리 대역폭 최대화
내부 디스크 스토리지	SFF(2.5") 드라이브 베이 24개
프로세서-메모리 대역폭	시스템당 340GB/s
L2-L3 캐시 대역폭	7TB/s 온칩 대역폭
어댑터 슬롯	<ul style="list-style-type: none"> • NVIDIA T4 GPU 가속기 최대 6개 • PCIe x16 3.0 FHFL 슬롯 2개(2배폭 가속기 지원) • PCIe x16 4.0 LP 슬롯 2개 • PCIe x8 3.0 FHFL 슬롯 2개(실질적으로 x16) • PCIe x8 3.0 FHFL 슬롯 2개 • PCIe x16 3.0 LP 슬롯 2개
표준 기능	
I/O 포트	<ul style="list-style-type: none"> • USB 3.0 2개, 1Gb 이더넷/VGA 2개 • BMC 전용 1Gb 1개
POWER 하이퍼바이저	베어메탈
RAS 기능	<ul style="list-style-type: none"> • 프로세서 명령 재시도 • 선택형 동적 펌웨어 업데이트 • Chipkill 메모리 • ECC L2 캐시, L3 캐시 • 고장 모니터링이 포함된 서비스 프로세서 • 동시 유지보수 디스크 • 이중 핫플러그 전원 • 이중 핫플러그 팬 • 고객 설치 및 복구 • 간소화된 Op 패널 • 인랙(In-rack) 시스템 서비스
운영 체제	Red Hat Enterprise Linux 7.6-alt
전원 요구사항	220V
시스템 크기	<ul style="list-style-type: none"> • 너비: 436mm(17.2인치) • 깊이: 820mm(32.3인치) • 높이: 86mm(3.4인치) • 무게: 30.9kg(68파운드)
보증	3년 제한 보증, 기타 모든 유닛(국가별로 다름)의 경우 CRU(고객 교체 가능 유닛), 익일(영업일 기준) 오전 9시 ~ 오후 5시 지원(휴일 제외), 보증 서비스 업그레이드 및 유지보수 이용 가능

왜 IBM인가?

IBM은 100년 넘게 기술 분야를 선도하면서 기업 고객들에게 비즈니스 데이터의 관리 및 마이닝을 지원하는 서비스를 제공해 왔습니다. 이제 IBM은 고객의 AI 도입을 지원하고 있습니다. IBM Research의 초기 머신러닝 시스템부터 Jeopardy 퀴즈쇼의 Watson에 이르기까지 IBM에게는 AI가 단순한 유행어가 아닙니다. IBM은 기업의 전체 AI 여정에 필요한 기술과 서비스를 유연하고 포괄적인 구성으로 제공합니다.

다음 단계

→

<https://www.ibm.com/marketplace/power-system-ic922>

자세한 정보

IBM Power System IC922에 대한 자세한 내용은 IBM 영업대표 또는 IBM 비즈니스파트너에게 문의하거나 <https://www.ibm.com/marketplace/power-system-ic922>에서 확인하세요.

또한 IBM Global Financing에서는 고객이 비즈니스 성장에 필요한 기술을 도입하는 데 도움이 되도록 다양한 결제 옵션을 제공합니다. IBM은 IT 제품 및 서비스의 도입부터 폐기까지 라이프사이클 전반을 관리하도록 지원합니다. 자세한 내용은 [ibm.com/financing](https://www.ibm.com/financing)에서 확인하세요.

© Copyright IBM Corporation 2020.

IBM, IBM 로고 및 ibm.com은 전세계 여러 국가에 등록된 International Business Machines Corp.의 상표입니다. 기타 제품 및 서비스 이름은 IBM 또는 타사의 상표입니다.

현재 IBM 상표 목록은 웹

<https://www.ibm.com/legal/us/en/copytrade.shtml>에 있습니다. 또한 본 문서에서 참조되는 타사의 상표는 https://www.ibm.com/legal/us/en/copytrade.shtml#section_4에 있습니다.

본 문서에는 IBM Corporation의 등록상표 및/또는 상표인, 다음 IBM 제품에 적용되는 정보가 포함되어 있습니다.

IBM® Power Systems™, POWER9™, OpenCAPI™



Linux는 미국 또는 기타 국가에서 사용되는 Linus Torvalds의 등록상표입니다.

IBM이 제시하는 방향 또는 의도에 관한 모든 언급은 특별한 통지 없이 변경될 수 있습니다.