

Part II

CIO Summit 2024

스스로 탐구하는 AGI와 GPU 인프라 이야기



2월 22일, 효성인포메이션시스템은 전자신문 CIO Summit 2024에서 ‘스스로 탐구하는 AGI와 GPU인프라 이야기’를 주제로 AI를 넘어선 AGI 시대에 더욱 중요해진 GPU 시스템, 그리고 AI 업무에 적합한 고속 데이터 처리를 위한 인프라에 대해 소개했다.

AGI 시대, 최적의 인프라 도입 요건

생성형 AI로 인해 기업들의 AI 비즈니스가 활발해지고 있으며, 거대언어 모델(LLM) 개발 열풍까지 불고 있다. 그리고 AI를 넘어 AGI가 등장하며 AI 비즈니스에도 변화가 있을 것으로 예상되고 있다.

AGI는 인간과 유사한 지능과 스스로 학습할 수 있는 능력을 갖춘 AI로, 학습이 가속화될수록 더욱 많은 데이터가 생성될 것으로 예상된다. AGI의 등장은 더 큰 데이터 세트와 복잡한 AI 모델이 필요함을 의미하며, GPU시스템 및 데이터 처리 효율이 매우 중요해질 수밖에 없다는 것을 시사한다.

기업들이 자사에 최적화된 AI 인프라를 도입하기 위해서는 어떤 요건이 필요할까?

첫째, 최신 기술이 적용된 GPU 인프라를 사용해야 한다. 모든 기업이 고비용에 최신 성능의 GPU를 도입할 수는 없다. 효성인포메이션시스템은 글로벌 전문 파트너인 Supermicro와 함께 기업별 환경에 적합한 GPU를 제안 및 공급하고 있다.

둘째, GPU 시스템 사용 효율을 고려해야 한다. 컨테이너 환경에서 GPU 자원의 분할 가상화 솔루션을 도입해야 하며, 데이터 과학자가 AI 개발 업무에 집중하도록 손쉬운 AI 개발 환경을 제공해야 한다.

마지막으로 AI 최적화 시스템을 설계해야 한다. GPU의 고성능을 최대로 발휘하기 위해 서버와 네트워크단의 최적의 설계가 전제되어야 한다. 또한 업무 단위로 적절한 GPU 인프라를 효율적으로 배치해주는 컨설팅 및 기술지원이 뒷받침되어야 한다.

AI 데이터 처리를 위한 완벽한 솔루션

효성인포메이션시스템은 AI 데이터의 고속 처리를 위해 초고성능 병렬 파일 스토리지 HCSF를 제공하고 있다.

AI 업무는 랜덤 입출력과 빈번한 Small I/O, 메타데이터 오버헤드가 발생하기 때문에 IOPS 성능이 매우 중요하다. HCSF는 타사 대비 7.5~225배의 높은 IOPS 성능을 나타내, AI 업무를 위한 저장장치로서 완벽한 성능을 증명하고 있다. 또한, HCSF는 단일 통합 파일시스템으로 다양한 멀티 프로토콜을 지원해 AI 업무 모든 단계의 애플리케이션 용도에 맞춰 저장자원으로 활용된다.

HCSF는 증설 및 확장할수록 성능이 선형적으로 증가하는 AI 데이터 인프라이다. 메타데이터를 처리하는 시스템이 노드별로 분산되어 있기 때문에 데이터 처리에 있어 병목이 발생하지 않는다. 따라서 확장 시 성능저하 이슈가 전혀 없는 솔루션으로, AI 인프라가 충족해야 할 모든 요구 사항을 갖춘 고성능 솔루션이다.

효성인포메이션시스템은 고성능 GPU 서버 및 HPC 서버를 기반으로 기업들이 AI 업무를 위한 연산 인프라를 구축할 수 있도록 지원하며, 고성능 데이터 처리에 탁월한 AI 스토리지를 제공한다. 특히 AI/ML Ops, GPU DB 등 전문 파트너와 협업을 통해 기업 환경에 최적화된 AI 솔루션을 제공한다.