

[IDC Future Enterprise Summit 2024]

미래 AI환경을 위한 데이터 레이크하우스 전략

권동수 전문위원

HS효성인포메이션시스템 데이터사업팀

2024년 11월 27일

Agenda

1. AI 트렌드

2. AI 활용을 위한 데이터 레이크 구축 사례

1. AI 트렌드

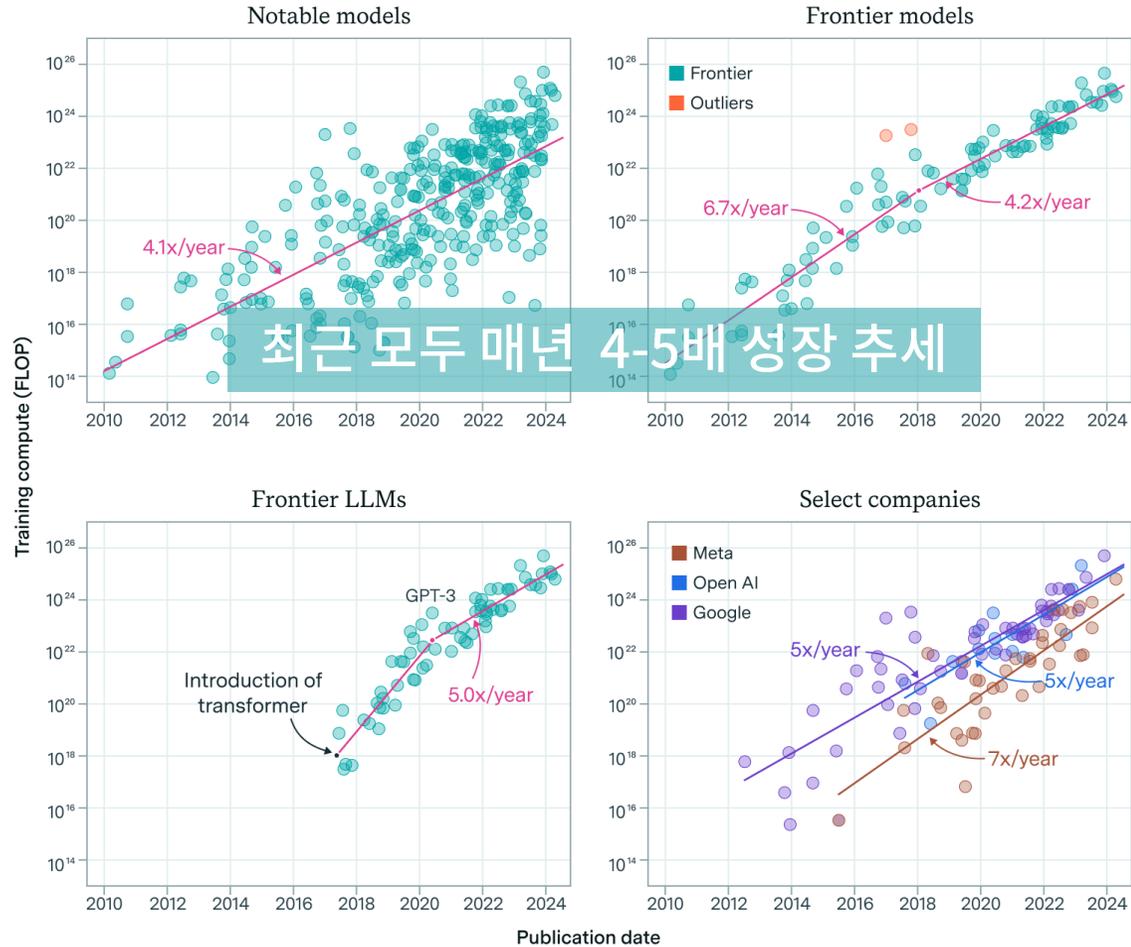
1. ANI(약인공지능), AGI(일반인공지능) and ASI(초인공지능)
2. 생성형AI 도입 시 고려사항
3. 통합 AI PLATFORM
4. AI를 위한 현대적 데이터 아키텍처



ANI(약인공지능), AGI(일반인공지능) and ASI(초인공지능)

Summary of compute trends in AI

EPOCH AI



AI Artificial Intelligence

ANI

2024년
(Artificial Narrow Intelligence)

- 음성 인식, 이미지 처리, 자연어 처리, 추천 시스템, 자율주행 등

AGI

(Artificial General Intelligence)

- 다양한 문제 해결, 과학적 발견, 창의적 작업 등

ASI

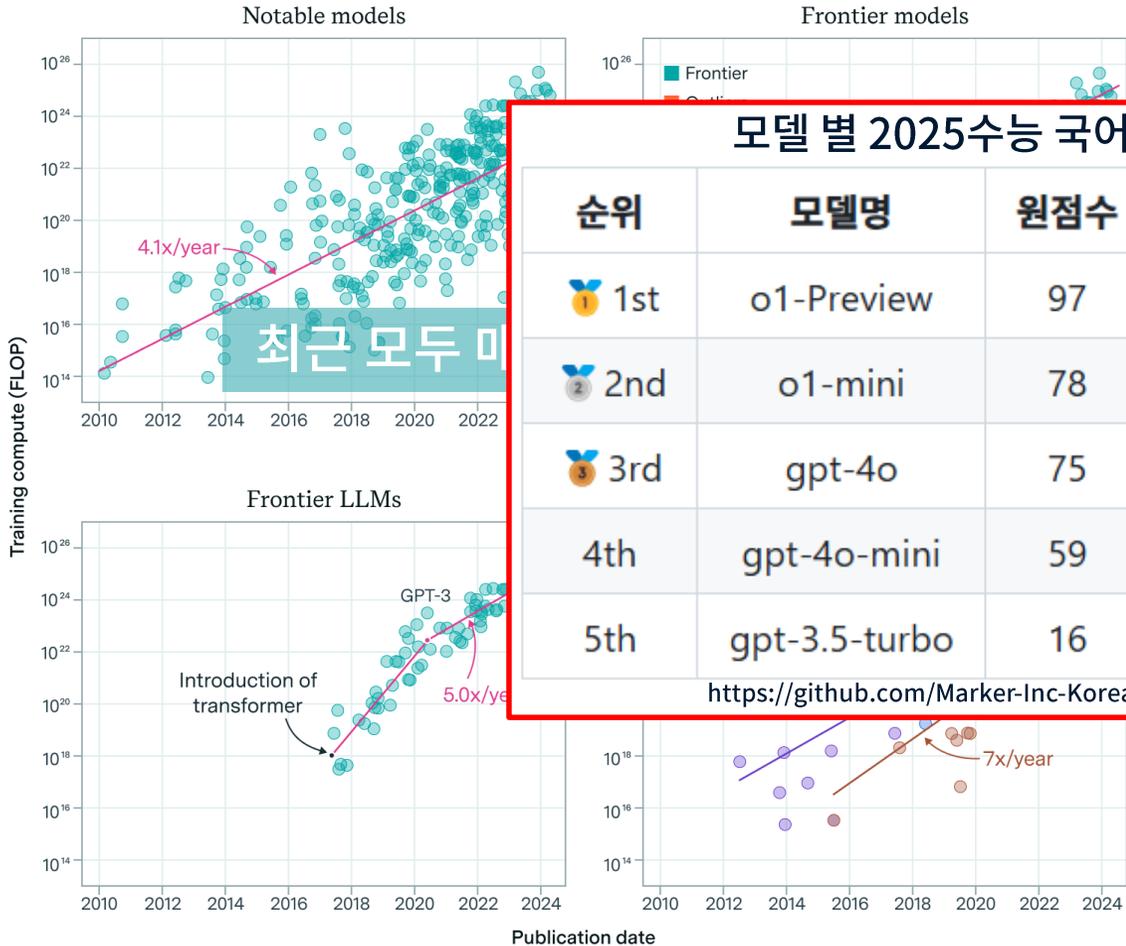
(Artificial Super Intelligence)

- 인간의 능력을 초월하는 모든 분야 활동 영역

ANI(약인공지능), AGI(일반인공지능) and ASI(초인공지능)

Summary of compute trends in AI

EPOCH AI



모델 별 2025수능 국어시험 테스트 결과

순위	모델명	원점수	추정 등급컷(2025.11.18기준)
1st	o1-Preview	97	1등급
2nd	o1-mini	78	4등급
3rd	gpt-4o	75	4등급
4th	gpt-4o-mini	59	5등급
5th	gpt-3.5-turbo	16	8등급

<https://github.com/Marker-Inc-Korea/Korean-SAT-LLM-Leaderboard>

AI

Intelligence

AGI

(Artificial General Intelligence)
 다양한 문제 해결, 과학적 발견, 창의적 작업 등

ASI

(Artificial Super Intelligence)
 • 인간의 능력을 초월하는 모든 분야 활동 영역

(Artificial Narrow Intelligence)
 • 음성 인식, 이미지 처리, 자연어 처리, 추천 시스템, 자율주행 등

비즈니스에 성공적인 생성형 AI를 도입하고 궁극적으로 지속가능 성장을 위해 기업이 고려할 5가지 중요한 사항

#1 목표 파악 : 맞춤형, 확장성, 성과 지향적 생성형 AI

#2 핵심 엔터프라이즈 표준과 구체적인 기술 지침 결합

#3 하이브리드 환경에서 생성형 AI 지원

#4 기업비즈니스의 이해와 기술 역량을 갖춘 AI 파트너 선택

#5 고성능·로딩 최적화·AI 어플라이언스 사용

통합 AI PLATFORM

1. AI 트렌드



고속 네트워크

200Gb/400Gb Ethernet & InfiniBand
연산 및 저장자원 간 성능 최적화를 위한 초고속 네트워크

GPU 서버

HGX, H100, H200, B200 등
NVIDIA가 공인한 슈퍼마이크로의 최신 GPU서버

초고성능 병렬 파일시스템

Hitachi Content Software for File
GPU환경, AI워크로드에 최적화된 NVMe 기반 초고성능 분산 병렬파일시스템

오브젝트 스토리지

Hitachi Content Platform
확장성이 뛰어난 오브젝트 스토리지와 정책기반 자동화 티어링으로 효율성 향상

AI 인프라는 HS효성에 맡기고,
‘비즈니스 혁신’에만 집중하세요.

01

사전 설계·검증된 통합 AI 인프라 플랫폼 제공

02

데이터 생애주기를 위해 데이터레이크 구성

03

컨설팅부터 설계, 구축, 운영까지 통합 수행

04

IDC 공인 국내 하이엔드/고성능 스토리지 시장 10년 연속 1위 기업

05

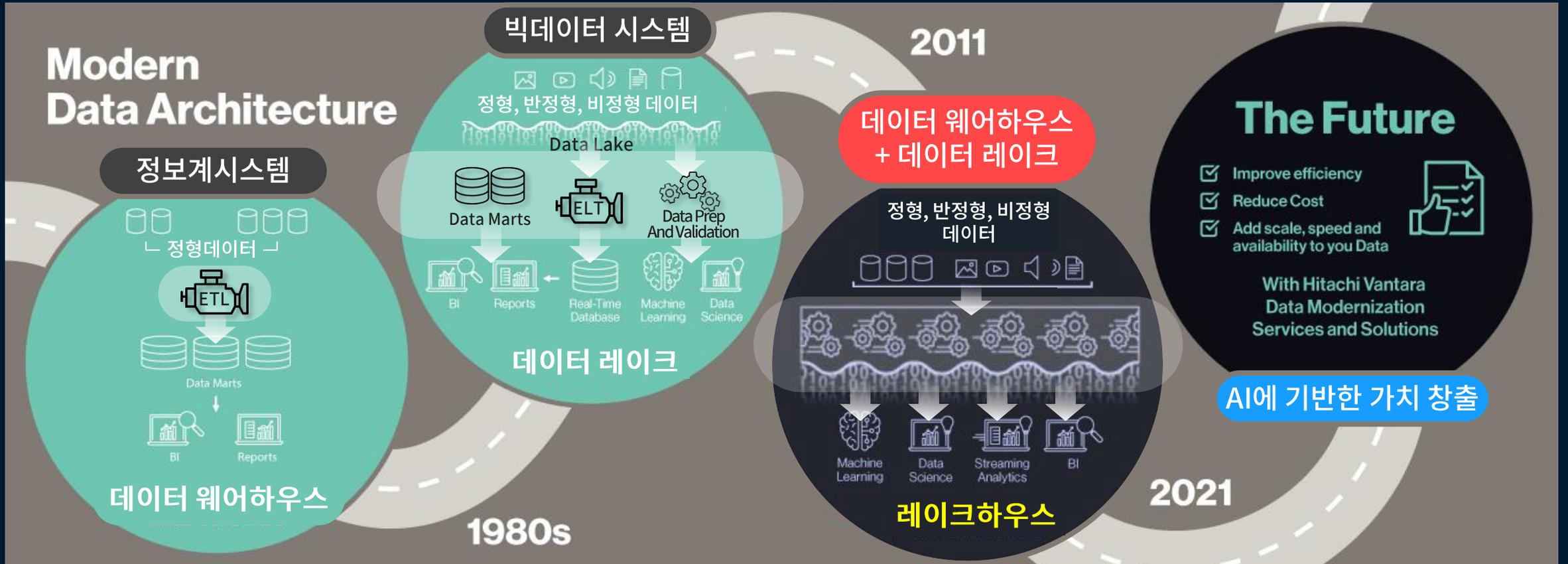
파트너 에코시스템과 긴밀한 협업으로 최적의 AI Ops 제공

06

LLM, AI/HPC 프로젝트 다수 수행

AI를 위한 현대적 데이터 아키텍처

- 데이터 레이크의 다양하고 방대한 데이터를 통해 인사이트 도출
- 새롭게 도출된 인사이트를 통한 가치 창출



Source : Hitachi Vantara

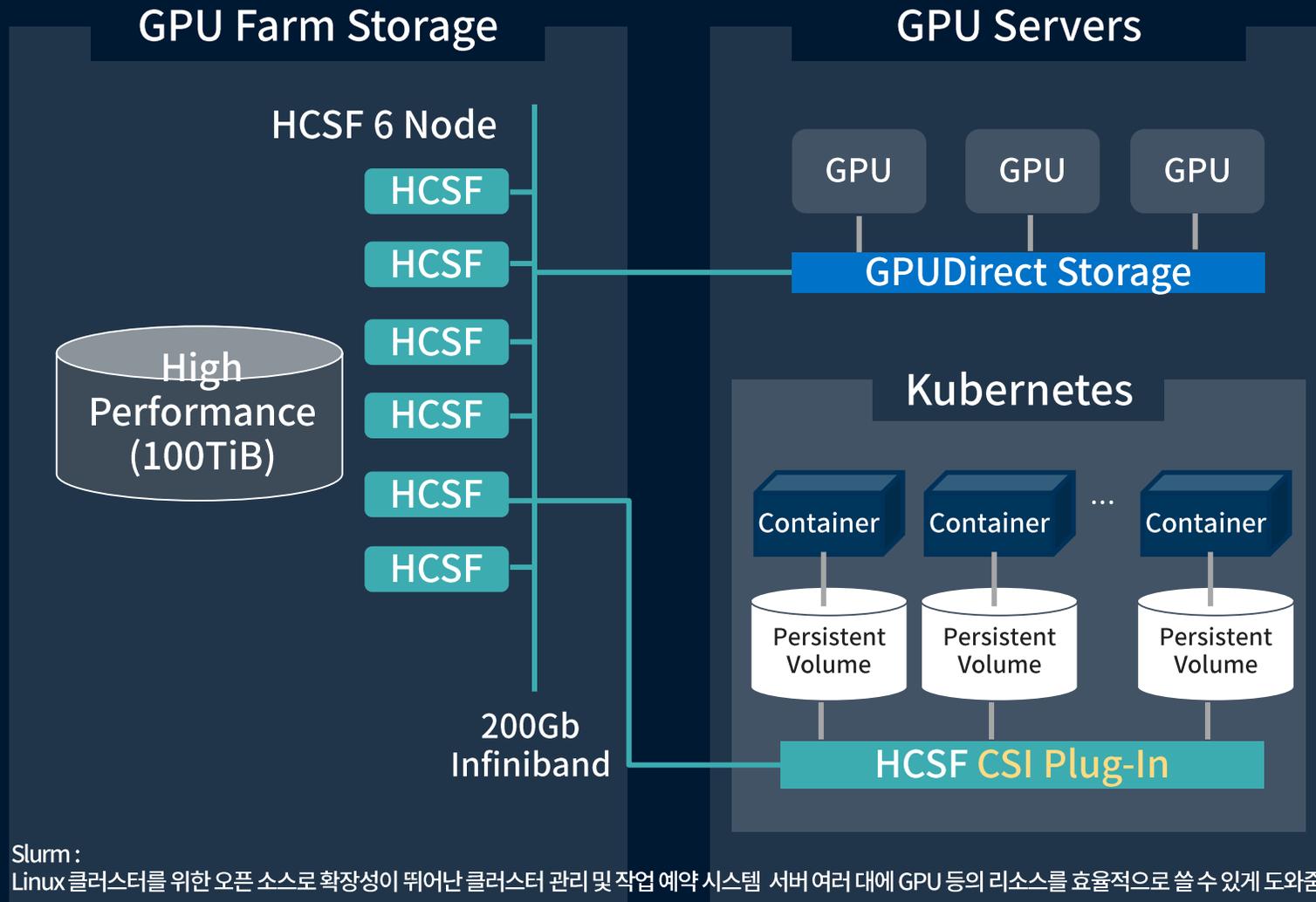
2. AI 활용을 위한 데이터 레이크 구축 사례

1. 사례 1. 국내 그룹 SI 기업
2. 사례 2. 국내 AI 서비스 기업
3. 사례 3. 국내 R&D기업
4. 사례 4. 국내 금융기관



사례 1. 국내 그룹 SI 기업

2. AI 활용을 위한 데이터 레이크 구축 사례



사업 목적

- sLLM 모델 개발 업무를 위한 AI/HPC 인프라 구축

구축 내용

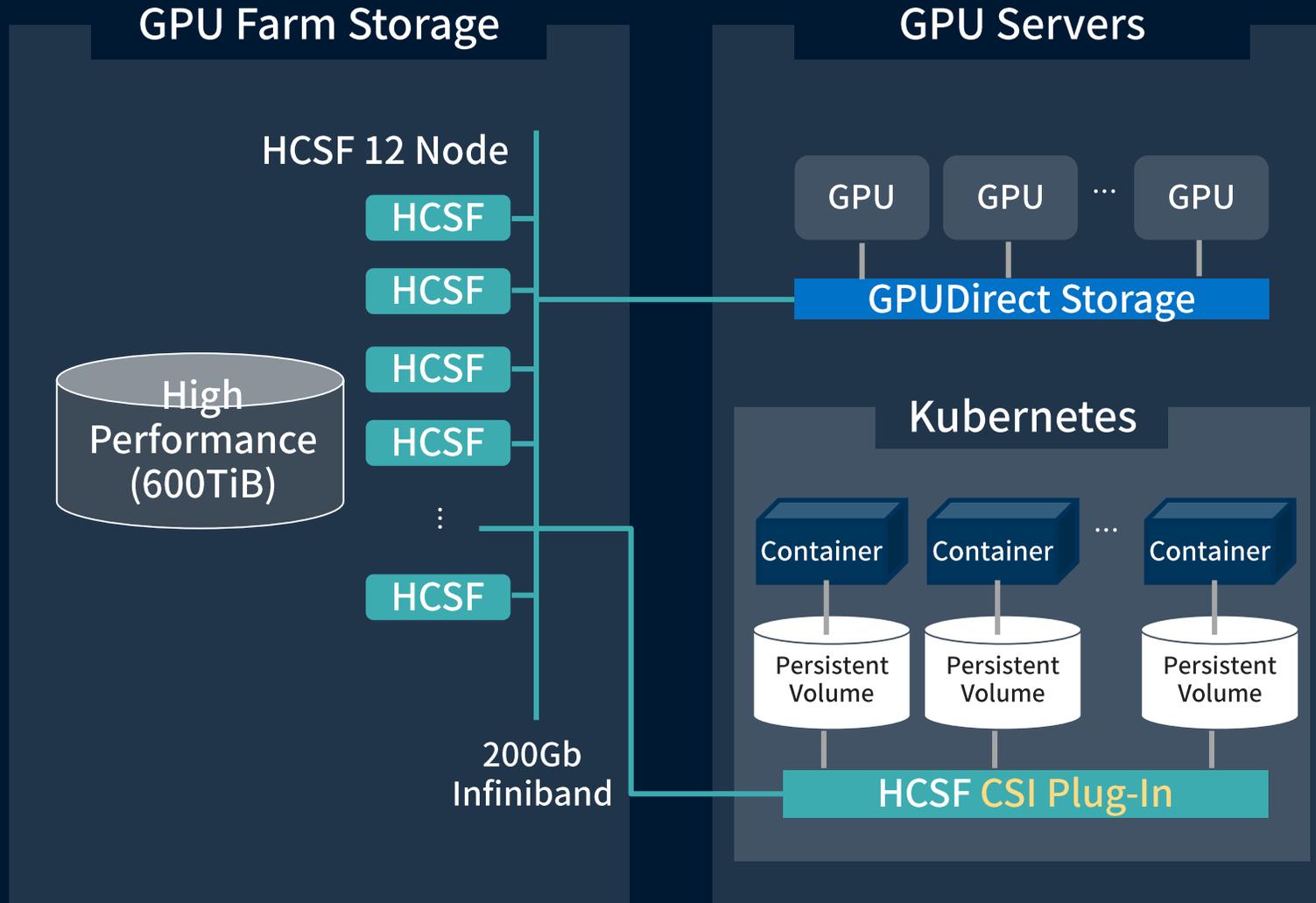
- 최신 Supermicro GPU 서버, 고성능 스토리지 HCSF, 고속 네트워크, Slurm AI HPC 인프라 구축
- Kubernetes 연계(CSI) 기능을 제공하여 볼륨 운영 최적화/효율화 지원

도입 효과

- MLPerf 테스트 검증된 최고 성능 GPU 서버와 GPUDirect Storage 적용으로 GPU 연산 성능 최적화
- 확장에 따른 성능 저하 이슈 없는 무제한 급증 설 지원 저장 자원 제안
- AI 플랫폼 구축 경험 있는 전문가로 구성된 전담 조직으로 AI 플랫폼 연계 기술 지원 보장

사례 2. 국내 AI 서비스 기업

2. AI 활용을 위한 데이터 레이크 구축 사례



사업 목적

- AI 교육 및 자사 AI 학습 모델 테스트 환경을 위한 프라이빗 클라우드 기반 HPC 인프라 구축

구축 내용

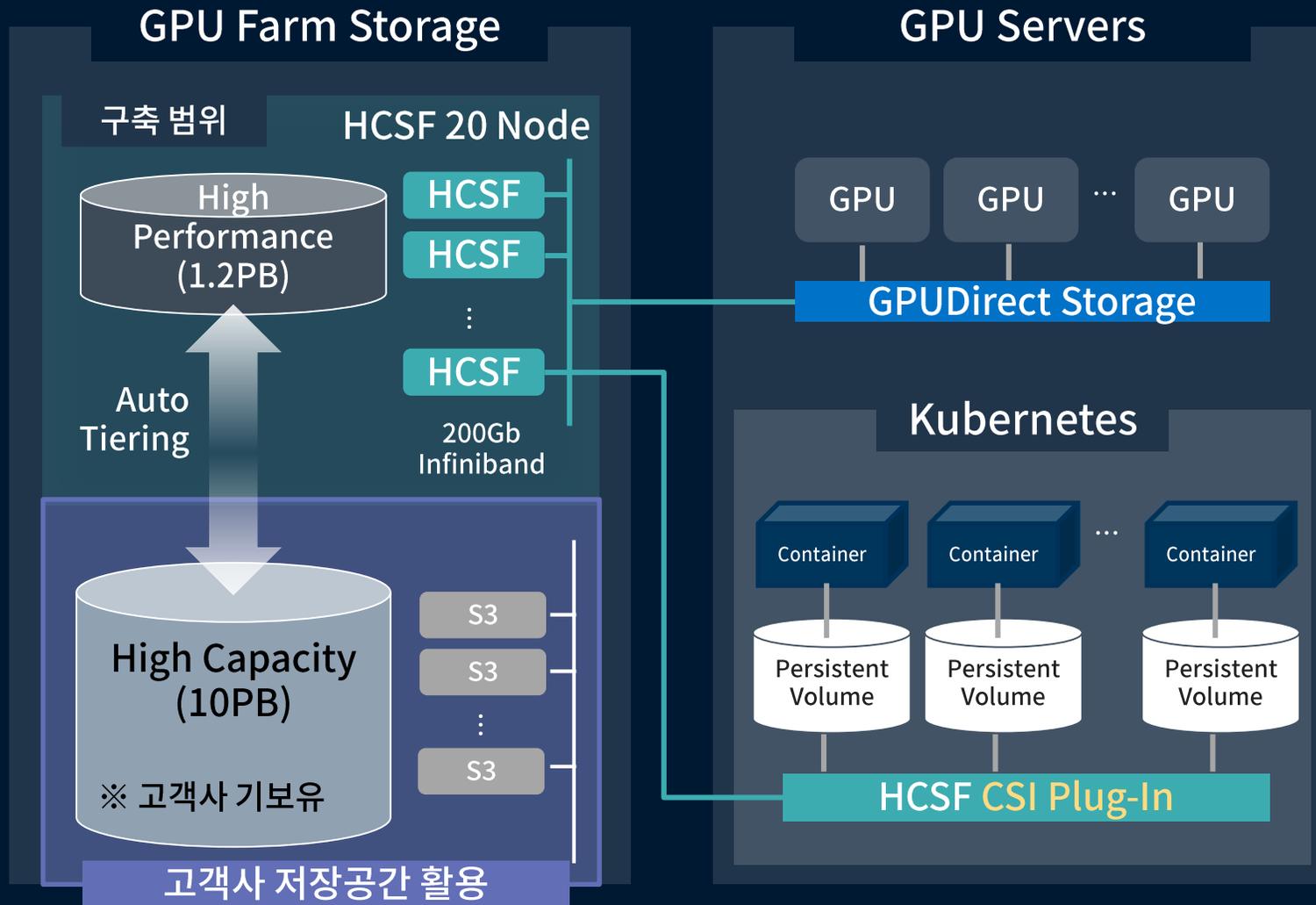
- GPU 환경 구현을 위한 고성능 스토리지 HCSF 구축
- GPU 성능 최적화 GDS 지원 스토리지 구축
- Kubernetes 연계(CSI) 기능을 제공하여 볼륨 운영 최적화/효율화 지원

도입 효과

- 다수 개발자/분석가/교육생 공동 운영을 위한 GPU 전용 스토리지 인프라 확보
- 향후 요구 사항 변화 (데이터 증가, 사용자 증가 등)에 유연한 대응 확보
- GPU 서버 인프라 고려한 스토리지 최적화

사례 3. 국내 R&D기업

2. AI 활용을 위한 데이터 레이크 구축 사례



사업 목적

- 자연어처리알고리즘 학습 성능 개선을 위한 GPU 서버와 전용 스토리지를 도입

구축 내용

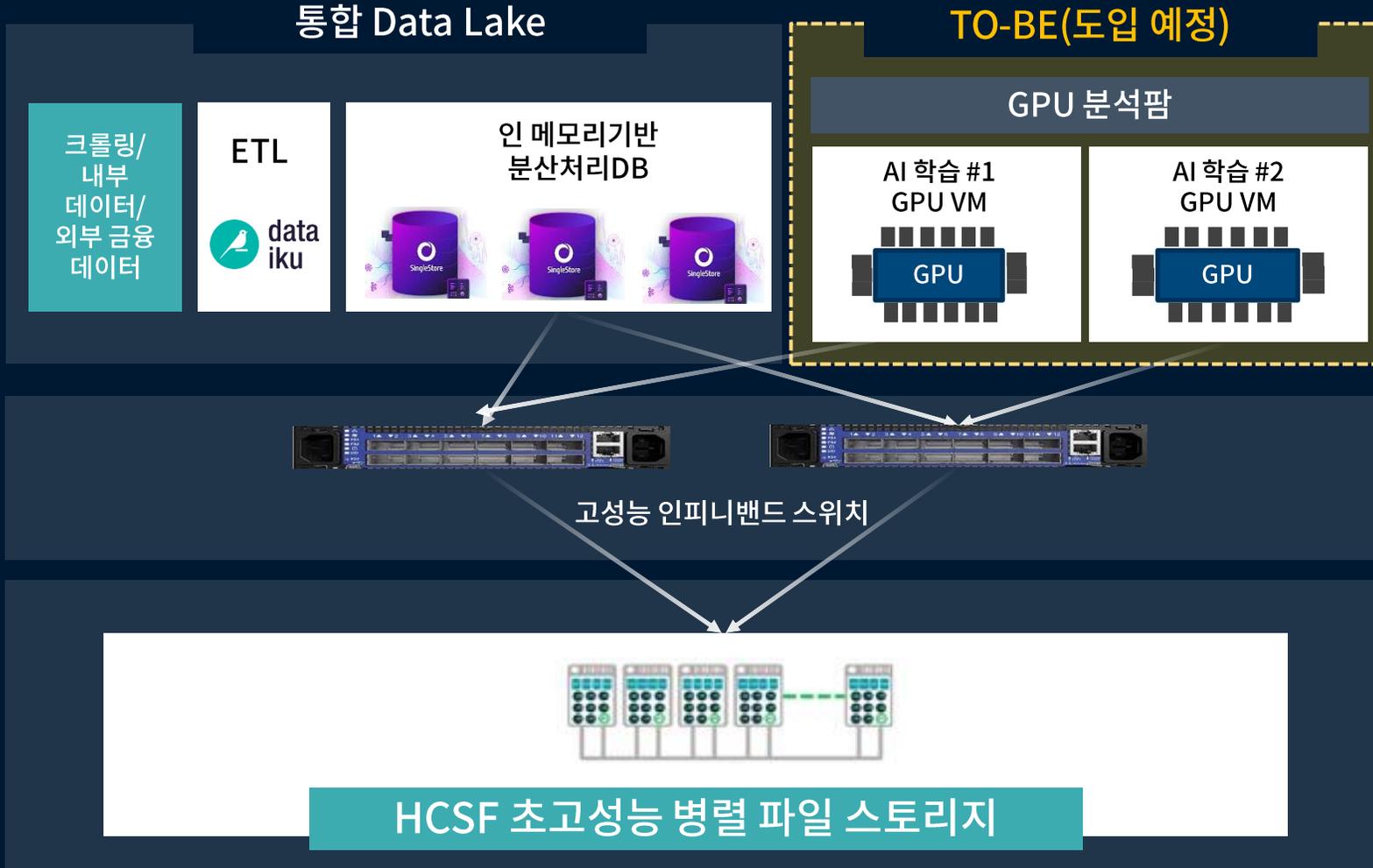
- GPU 환경 구현을 위한 고성능 스토리지 HCSF 구축
- GPU 성능 최적화 GDS 지원 스토리지 구축
- Kubernetes 연계(CSI) 기능을 제공하여 볼륨 운영 최적화/효율화 지원
- 고객사 S3 스토리지와 Tiering 구성

도입 효과

- 자연어처리 기계 학습 연산 성능 향상
- 서버 인프라 구성 고려한 스토리지 최적화
- 자체 S3 스토리지 활용성 확보를 통한 비용 효율적인 구성

사례 4. 국내 금융기관

2. AI 활용을 위한 데이터 레이크 구축 사례



사업 목적

- 외부기관데이터/국내금융기관데이터 통합으로 효과적인 데이터 분석 및 분석 체계 확립

구축 내용

- 정형/반정형 데이터 분석을 위한 데이터 레이크 구축
- 차년도 예정 사업인 GPU 기반 딥러닝 분석 환경으로 확장 가능한 아키텍처 수립

도입 효과

- 통합 플랫폼 (CPU → + GPU) 을 이용해 관리 포인트 최소화 및 확장 유연성 제공
→ 단 2명의 소수 운영자로 관리가 가능해짐
- 고성능 단일 데이터 레이크 저장소 구축 운영
→ 여러 대의 업무 별 스토리지를 하나로 통합
- 고성능 쿼리 분석 엔진 도입



HPC 요구 사항 확인 및 PoC/BMT 필수!!!
IOPS, 처리량 요구 성능 검증



HPC 시스템 데이터 레이크 구축 경험 필수!!!
대용량의 파일 처리, 안정적인 확장



국내외 실 사례를 통한 국내 기술력(인력) 필수!!!
장애 지원, 신규 AI 솔루션 연계 지원

감사합니다.

