

UCP AI 플랫폼 소개

효성인포메이션시스템

효성인포메이션시스템
클라우드사업팀

Contents

- 01 AI 시장 동향 및 특징
- 02 AI 플랫폼의 개선 방향
- 03 효성의 AI 플랫폼
- 04 효성의 AI 플랫폼 특장점
- 05 결론

1. AI시장 동향 및 특징

AI 시장 동향

한정된 분야 AI 일부 적용에서, 전통적 업무영역까지 AI 적극도입

인공지능 AI 국가전략 발표 (2019. 12) - '30년까지 최대 455조 원 경제효과 창출 목표

업종 구분 없이 전 산업군 AI기술 접목 시도 증가

1) 한국신용정보원
2) IDC 2019년 자료

AI 사업 특징

높은 성장률

국내시장

'25년까지 연평균 38.4% 성장,
10.51조 원의 시장 형성 전망¹⁾

기술/플랫폼 복잡성

AI 플랫폼은
다양한 S/W 및 H/W 조합,
비 정형화 및 비 표준화

분석 업무 효율 저하

다양한 모델 관리가
수작업으로 진행 되어 있어
업무 협업 및 자동화 개선 필요

2. AI 플랫폼의 개선 방향

도입 및 운영에 대한 고민

- 개발자, IT 관리자 간 협업의 어려움
- 비효율적 자원 운영
- 많은 수작업으로 시간지연

개선 요구 사항

- 업무 협업 환경
- 연산/저장 자원 및 비용효율
- AI/ML 모델 관리

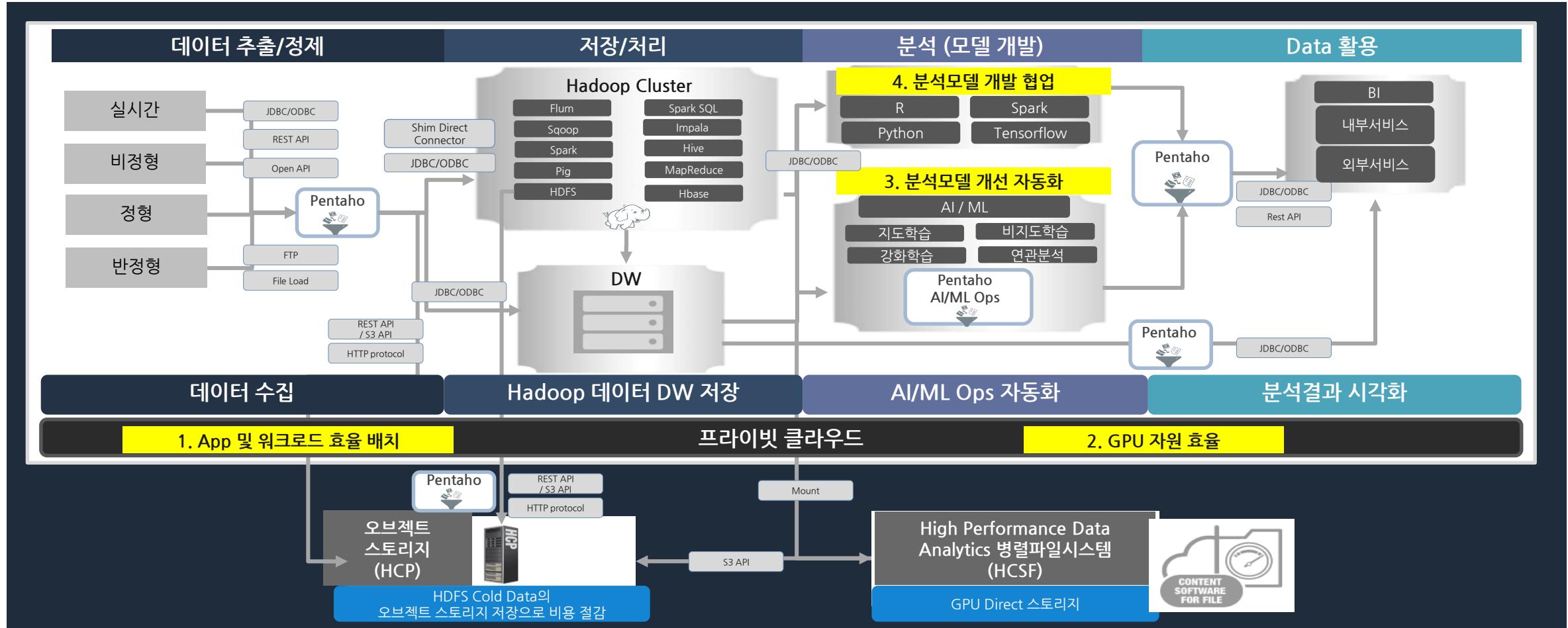
구현 방안

- 컨테이너 기반 모델개발 협업 환경 구현
- Remote GPU 구성 및 GPU 가상화
- 추출/처리/적재/ML 모델링의 자동화

AI분석 업무 프로세스 자동화 및 업무 효율을 보장하는 AI 분석 플랫폼 요구

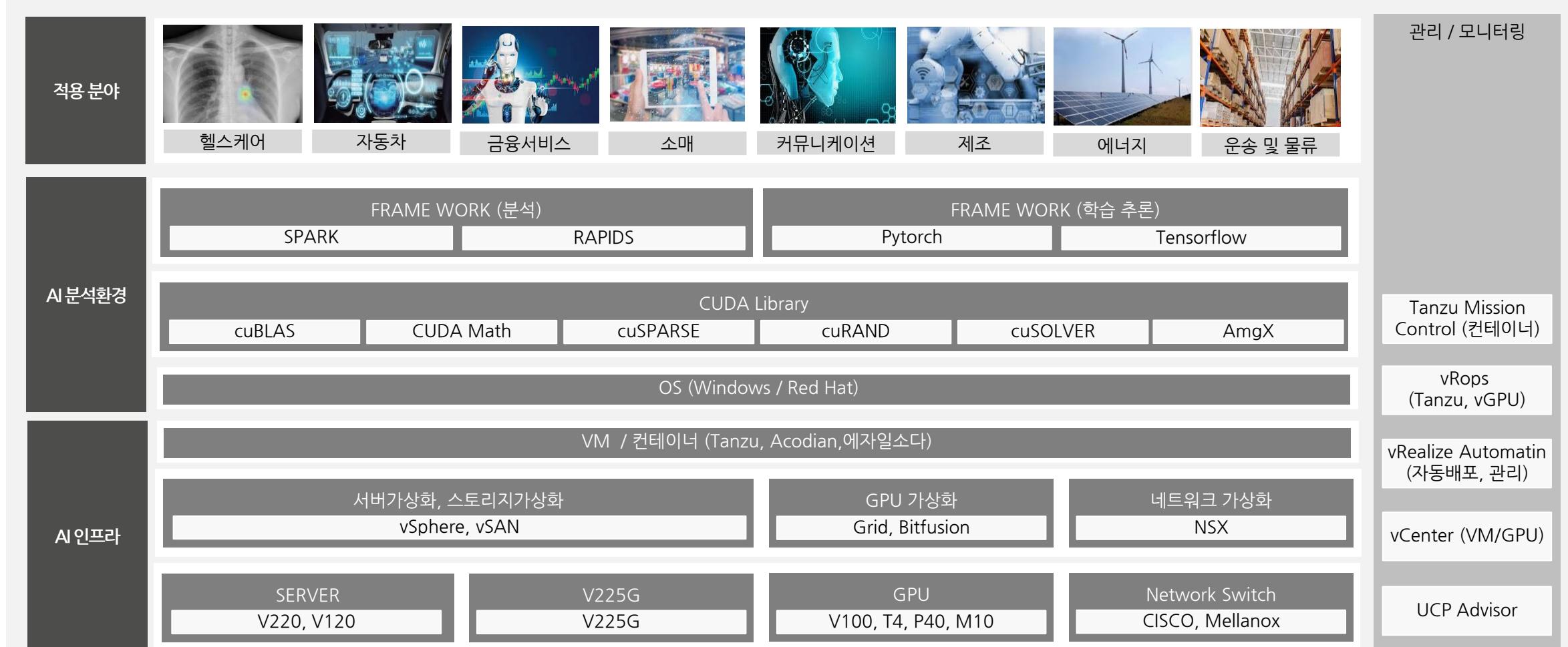
3. 효성의 AI분석 플랫폼

- 효성의 AI 분석 플랫폼은 데이터 수집, 저장, 분석 및 시각화를 위한 전체 Workflow를 End to End로 제공 가능
- DATA 분석의 가속화를 위한 GPU Direct 스토리지 제공으로 분석 업무 최적화



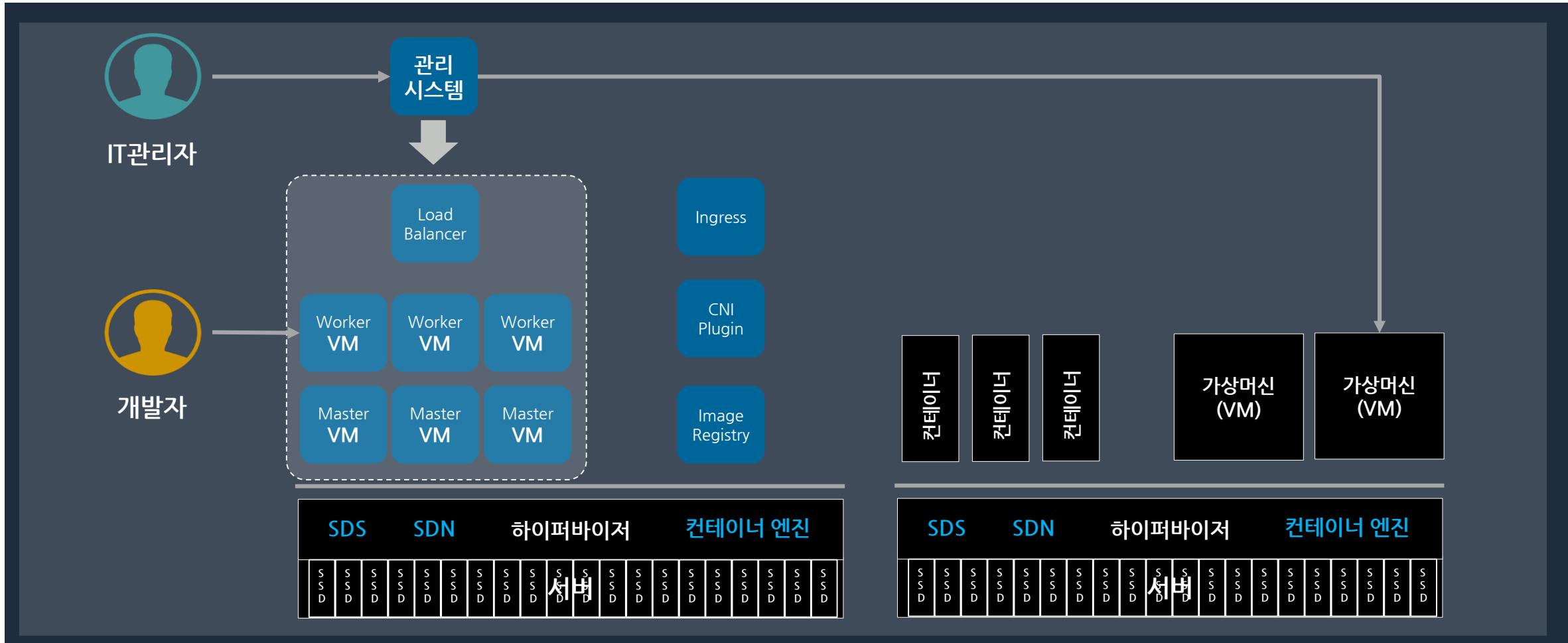
3. 효성의 AI분석 플랫폼

- AI 분석 플랫폼을 위한 하드웨어 (서버, 네트워크, 스토리지, GPU) 및 소프트웨어 (가상화 솔루션 및 컨테이너)를 Turn-key 제공



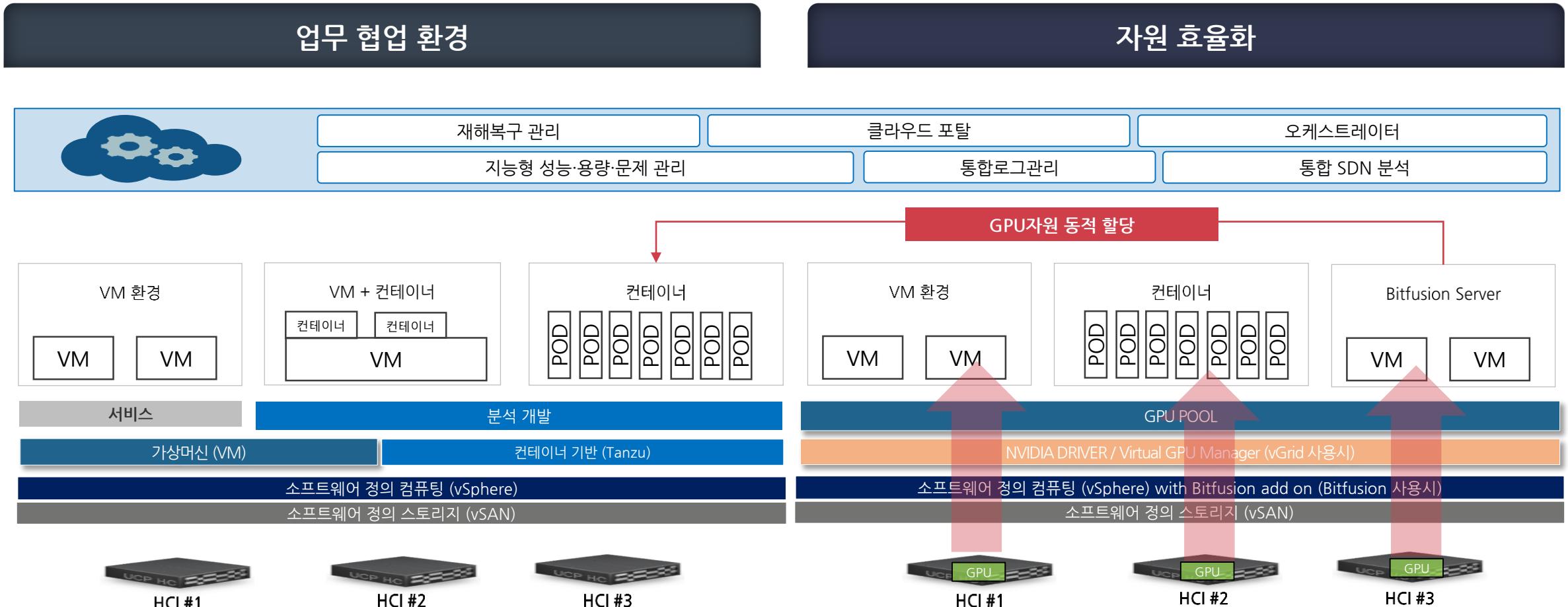
4. 특장점 - APP 및 워크로드의 효율적 배치

- 가상머신과 컨테이너를 동일 하이퍼바이저 Level에 구축할 수 있어, 효율적인 분석 업무 배포 및 관리가 가능



4. 특장점 - 통합 AI 워크로드 구현

- 통합 AI 워크로드 구현으로 업무 협업 환경에서 효율적인 자원 활용이 가능한 구성으로 설계 됨

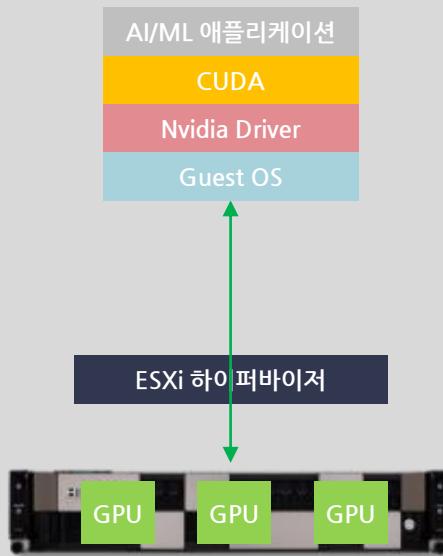


4. 특장점 - 자원 효율화 (GPU 가상화)

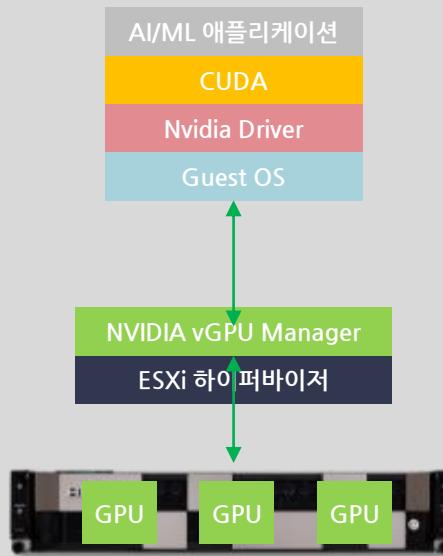
- 가상화 환경에서 GPU를 사용하는 방법은 3가지로 가상머신에 직접 할당하는 DirectPath I/O 방식과 GPU 리소스를 가상화하여 할당하는 vGPU와 Bitfusion 방식이 있음

GPU 활용방식

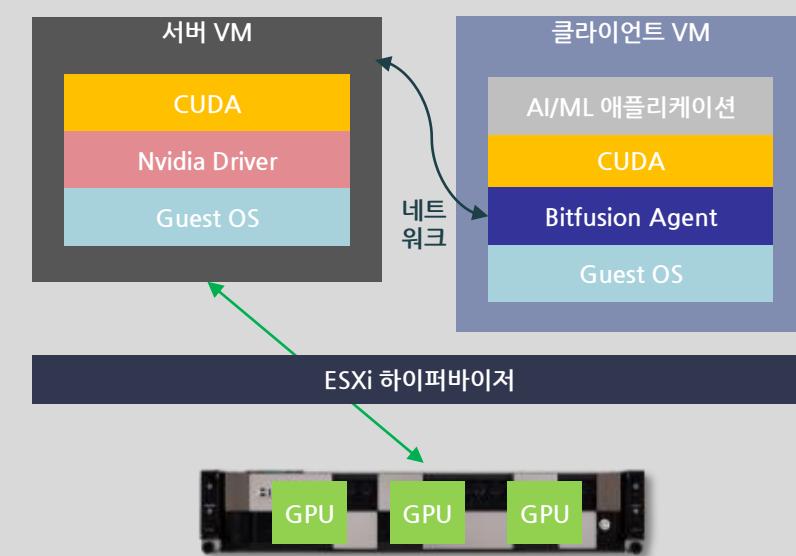
DirectPath I/O



vGPU



Bitfusion

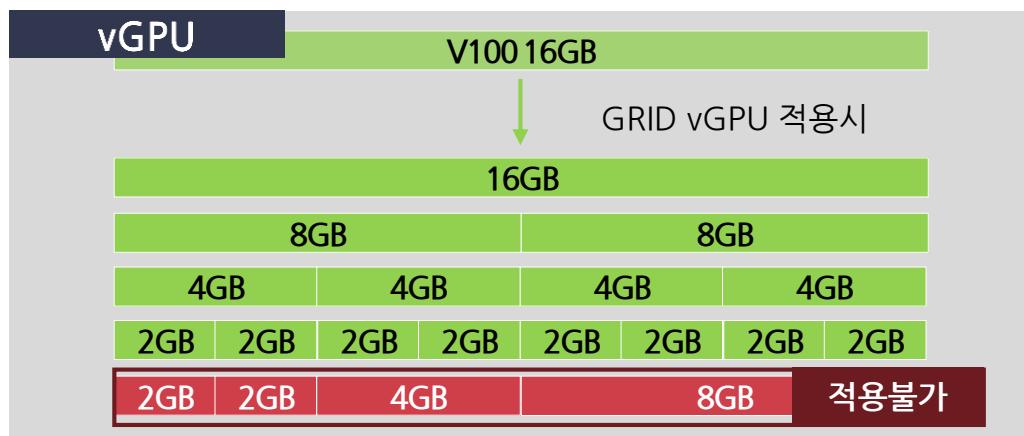


- (단일 애플리케이션) DirectPath I/O, vGRID, Bitfusion 순으로 성능 우수
- (전체 워크로드) 조각 GPU와 동적할당으로 인한 Bitfusion의 전체 이용률 기준 기대 성능 우수

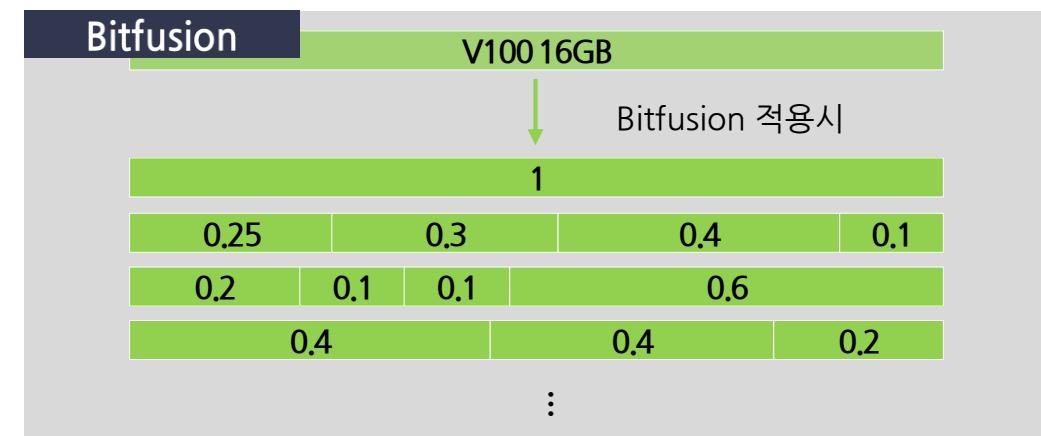
4. 특장점 - 자원 효율화 (GPU 가상화)

- GPU 리소스 활용 방식에 따라 vGPU와 Bitfusion 방식을 적용할 수 있고, 리소스의 동적 할당 및 리소스 배포면에서 Bitfusion이 활용도가 높음

No	항목	vGPU	Bitfusion	비고
1	GPU 자원 할당	할당 GPU Profile 사이즈 고정 (1GB, 2GB, 4GB 등) 동일 물리 GPU 상 다른 Profile 사이즈 적용 불가	자유로운 GPU 분할 (소수점 단위, MB 단위 할당)	비용효율, 유연성
2	GPU 자원 온라인 적용	VM 재 부팅 필요	온라인 적용 가능	관리효율
3	GPU 모니터링	vROps 라이선스 추가 구매 후, 별도 패키지 설치	vCenter에 모니터링 메뉴 포함	운영효율
4	GPU 자원 점유	GPU 할당 시 자원 상시 점유	Application 구동 시 GPU 리소스 사용 (GPU메모리 용량 보다 초과 할당 가능)	적은 GPU리소스 더 많은 사용자 할당



VS.



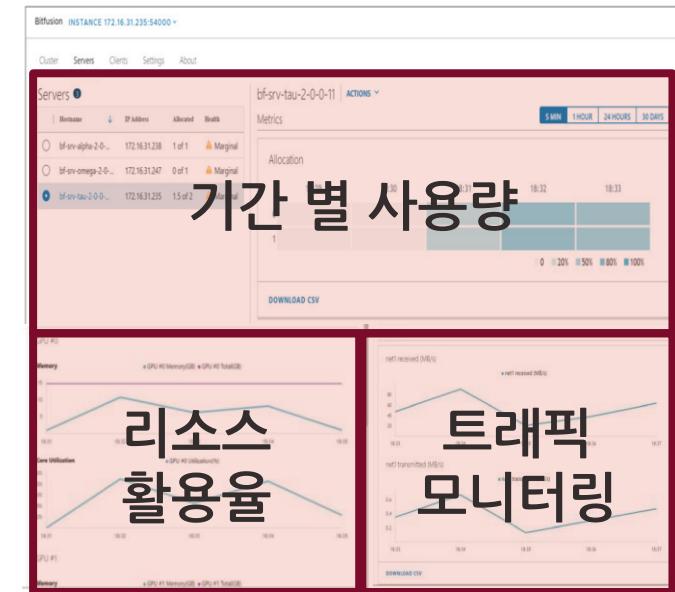
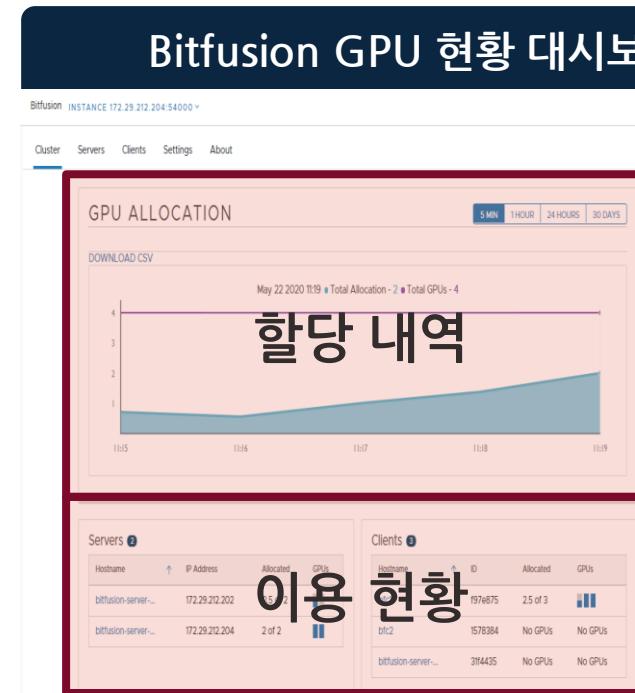
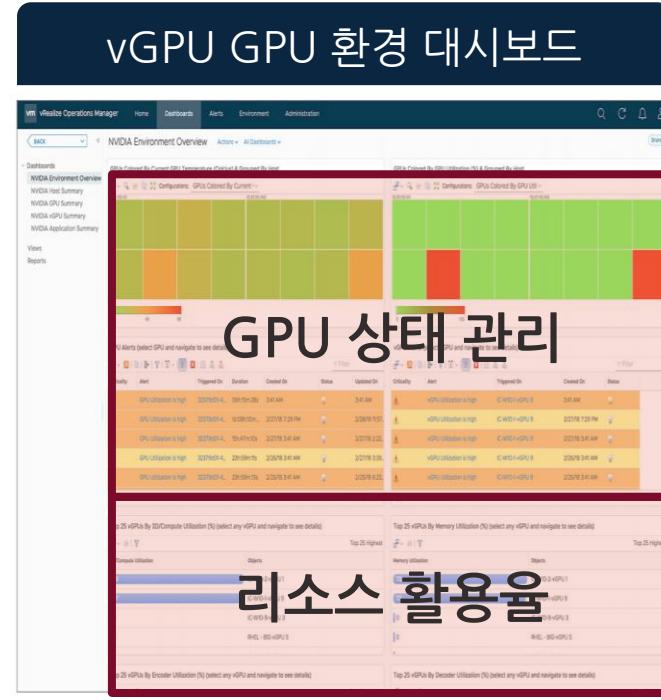
4. 특장점 - 자원 효율화 (GPU 자원 관리)

- AI 플랫폼 주요 리소스인 GPU에 대한 자원 관리 및 운영 모니터링 지원으로 효율적인 리소스 관리가 가능



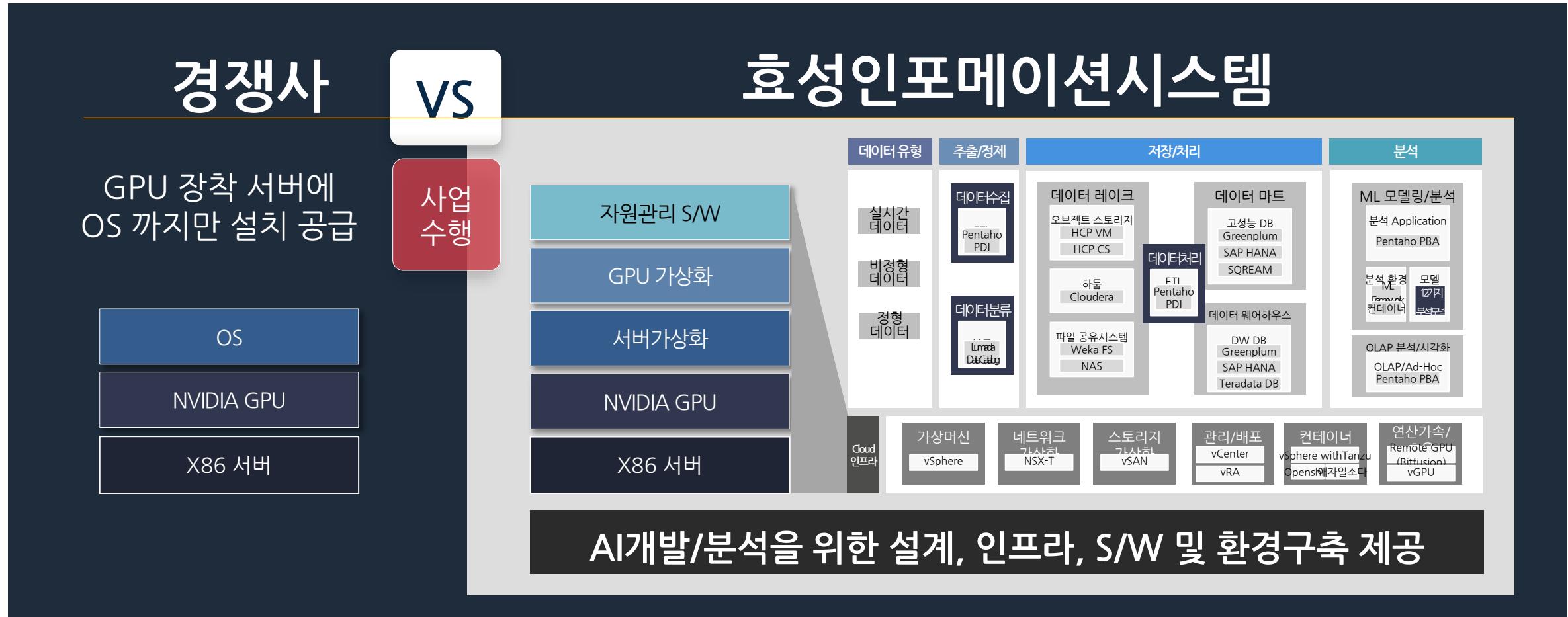
GPU 자원 관리

- 운영 GPU에 대한 사용률, 온도, 알림 정보 및 GPU 할당 내역을 관리하고, 트래픽을 모니터링



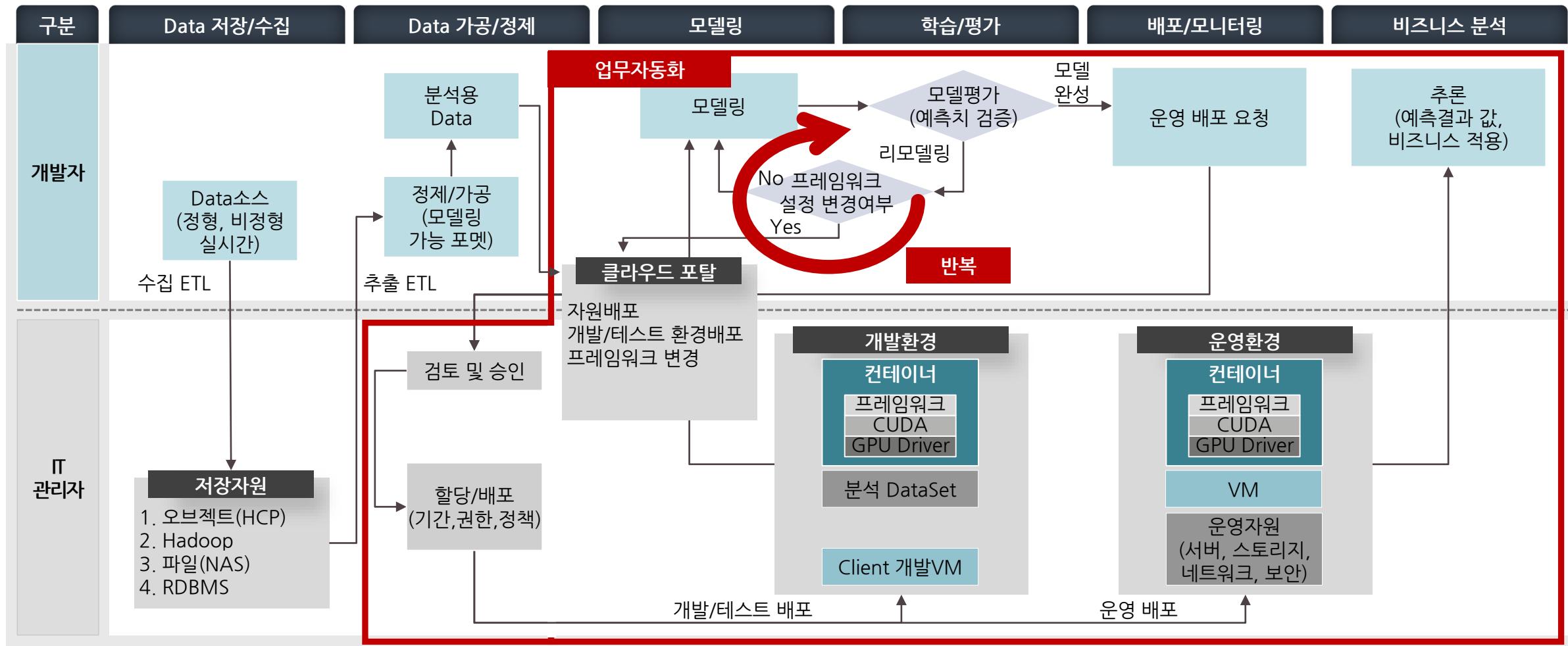
4. 특장점 - AI플랫폼 설계

- 효성은 경쟁사와 같이 단순 GPU 서버만 공급하는 것이 아니라, 전체 AI 분석 업무를 구성하는 전체 Step에 대해 설계 및 공급을 지원함



4. 특장점 - 프로세스 자동화

- AI 분석 업무에서 반복되는 훈련, 튜닝, 테스트 및 업데이트 구간을 자동화하여 개발에 집중할 수 있는 분석 환경 제공



4. 특장점 - 배포 자동화

- 사전 정의 된 개발환경 및 인프라를 적시 할당 받아 신속한 업무 구현

효성인포메이션시스템 AutoMationCenter - Service Broker

카탈로그 배포 승인

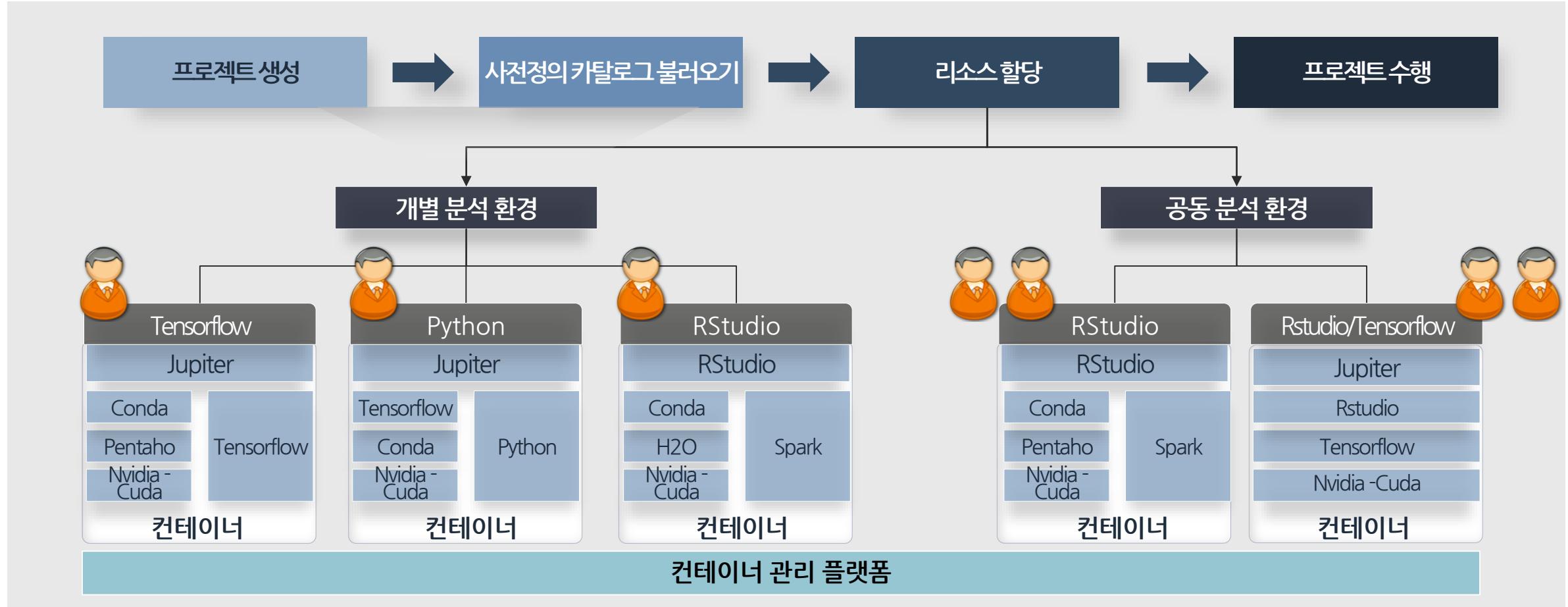
카탈로그 항목 10 개 항목 ▾

검색 정렬: 이름(오름차순)

부정적발 Cloud Assembly Blueprint R TensorFlow CUDA 10.1 V100 프로젝트: AI/ML TFT - Daej... 요청	빅데이터 Cloud Assembly Blueprint Python PyTorch CUDA 10.0 V100 프로젝트: AI/ML TFT - Daej... 요청	예지정비 Cloud Assembly Blueprint Python TensorFlow CUDA 11.0 V100 프로젝트: AI/ML TFT - Daej... 요청	마케팅분석 Cloud Assembly Blueprint Python PyTorch CUDA 11.0 V100 프로젝트: AI/ML TFT - Daej... 요청	이미지분석 Cloud Assembly Blueprint R TensorFlow CUDA 11.0 V100 프로젝트: AI/ML TFT - Daej... 요청	이상치탐지 Cloud Assembly Blueprint Python MXNet CUDA 10.0 V100 프로젝트: AI/ML TFT - Daej... 요청
텍스트분석 Cloud Assembly Blueprint R PyTorch CUDA 10.2 T4 프로젝트: AI/ML TFT - Daej... 요청	불량거래탐지 Cloud Assembly Blueprint Python PyTorch CUDA 11.0 T4 프로젝트: AI/ML TFT - Daej... 요청	금융시계열예측 Cloud Assembly Blueprint Python TensorFlow CUDA 11.0 V100 프로젝트: AI/ML TFT - Daej... 요청	부동산가격예측 Cloud Assembly Blueprint Python PyTorch CUDA 10.0 V100 프로젝트: AI/ML TFT - Daej... 요청		

4. 특장점 - 업무 협업 환경

- 개인 및 그룹 별 컨테이너 기반 작업 공간에서 (Sandbox) AI 모델 개발 및 분석 수행



5. 결론



플랫폼 최적화

- 기존 및 디지털 APP의 관리를 위한 **컨테이너/가상머신 운영 효율화**
- 네트워크 기반 GPU 가상화로 탄력적인 리소스 배포 및 통합으로 **효율적 GPU 관리**
- Remote GPU 구성 및 GPU 가상화 (GPU당 비용 50% 절감, 사용률 2~4배 향상)



프로세스 자동화

- 추출/처리/적재/ML모델링을 Workflow로 **통합 End to End 자동화**
- 전체 ML Workflow를 One-step으로 관리하여 예측모형의 Train, Tune, Test 및 **업데이트 자동화**



협업환경

- 사전정의 된 컨테이너 기반 GPU 가상화 환경 적시 배포로 **데이터 과학자와 IT 관리자간 협업 환경 제공**
- Cloud Portal에서 분석환경 요청 및 설계, 구성, 배포로 **개별 독립적 Sandbox 제공**

AI분석 업무 개선

Thank
you