



Hewlett Packard
Enterprise

엣지 투 클라우드 전환을 위한 최선의 선택,

HPE Ezmeral

빅데이터 구축 사례 소개

문 영 상 | 숭실대학교 정보과학대학원
(주) 성우HS 기술 자문

빅데이터의 요구사항 증대

빅데이터와 인공지능의 활용 측면

Why 빅데이터인가?

- 1 대규모의 컴퓨팅 파워를 필요로 하는 현업의 요구 사항 증대
- 2 대량의 데이터를 체계적으로 수집, 저장, 처리, 분석하기 위한 니즈 증대
- 3 데이터적측면에서 비정형 데이터 즉, 음성, 이미지, 동영상, 텍스트를 분석하기 위한 필요성 증대
- 4 통계 모델 중심의 분석이외의 다양한 분석 요구사항의 증가
- 5 인공지능 기반 서비스를 효과적으로 수행하기 위한 필요성 증대
- 6 빅데이터 중심의 서비스 창출을 통한 기업의 경쟁력 강화

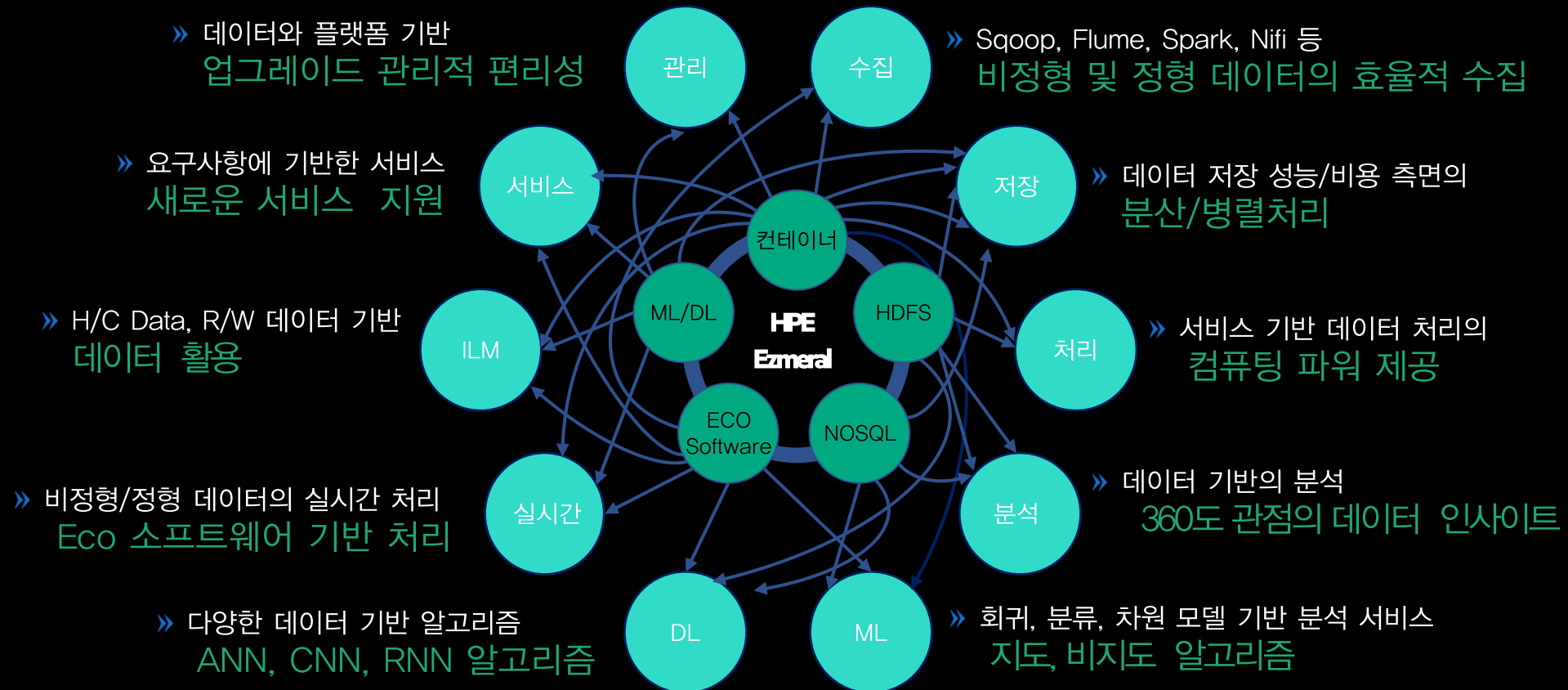
빅데이터

빅데이터 프로젝트 시 고려사항

- 1 플랫폼은? 하드웨어 자원, 소프트웨어 자원, ECO 소프트웨어 자원, 기존 IT자원의 통합
- 2 데이터는 ? 정형, 비정형(음성, 이미지, 동영상, 텍스트), 기타 데이터 (문서기반)등의 분석
- 3 비용은 ? 오픈소스 기반, 엔터프라이즈 서브 스크립션, 하이브리드 라이선스등)
- 4 업그레이드 및 유지보수 정책 (Major, Miner, Rolling 업그레이드)
- 5 인공지능 모델은? 머신러닝, 딥러닝, 통계 기반, 규칙 기반, 데이터 기반
- 6 통계와 AI 기반의 차이점은? 데이터 차원
- 7 인공지능의 출발점은? .AI의 출발점은 소프트웨어 중심, 라이브러리 중심, 데이터 중심
- 8 규칙 기반 모델과 통계적 모델의 모형은? Rule, Score, Statistics
- 9 인공지능 모델에 대한 평가? 성능/ 성과/ 관리 지표
- 10 효율적인 인공지능 모델을 구현하기 위해서? 신경망등 차원, GPU 할당, AI 등 소프트웨어의 적용
- 11 AI는 믿을 수 있는가? Black Box의 Explainable화에 대한 대응

HPE Ezmeral Data Fabric

Converged Bigdata Platform



Ezmeral

Ezmeral Converged Platform 구성

HPE E Z M E R A L

AI/ML and DATA ANALYTICS

For Data Science teams and ML Ops teams

For Data Analytics and DataOps teams

HPE Ezmeral ML Ops

HPE Ezmeral Data Fabric

CONTAINER MANAGEMENT

For Developers, and DevOps teams

HPE Ezmeral Container Platform

SECURITY

For Security and SecOps teams

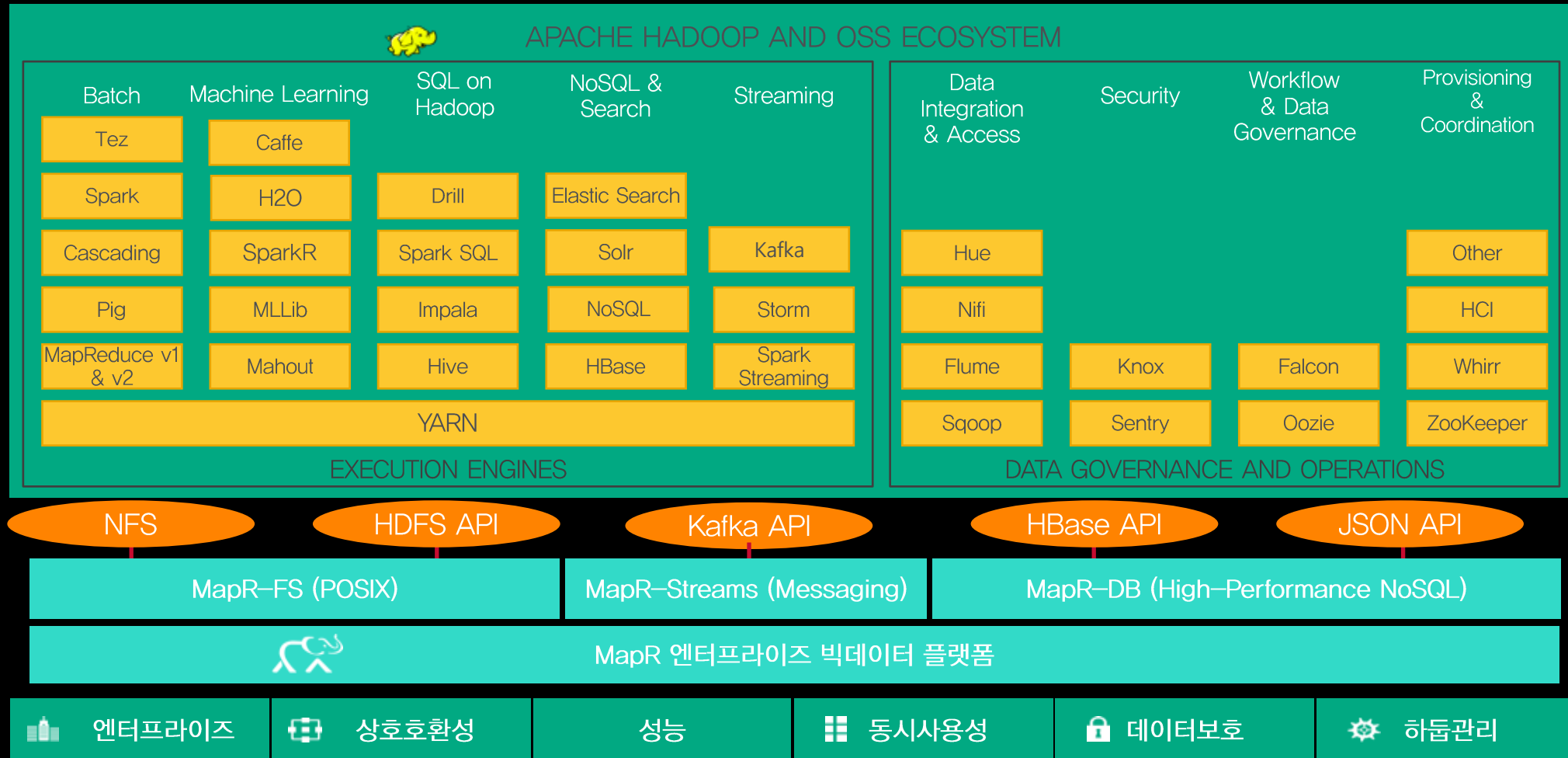
SPIFFE*

SPIRE*

*Open-Source projects

MapR data Frame work

MapR Data Fablic Structure

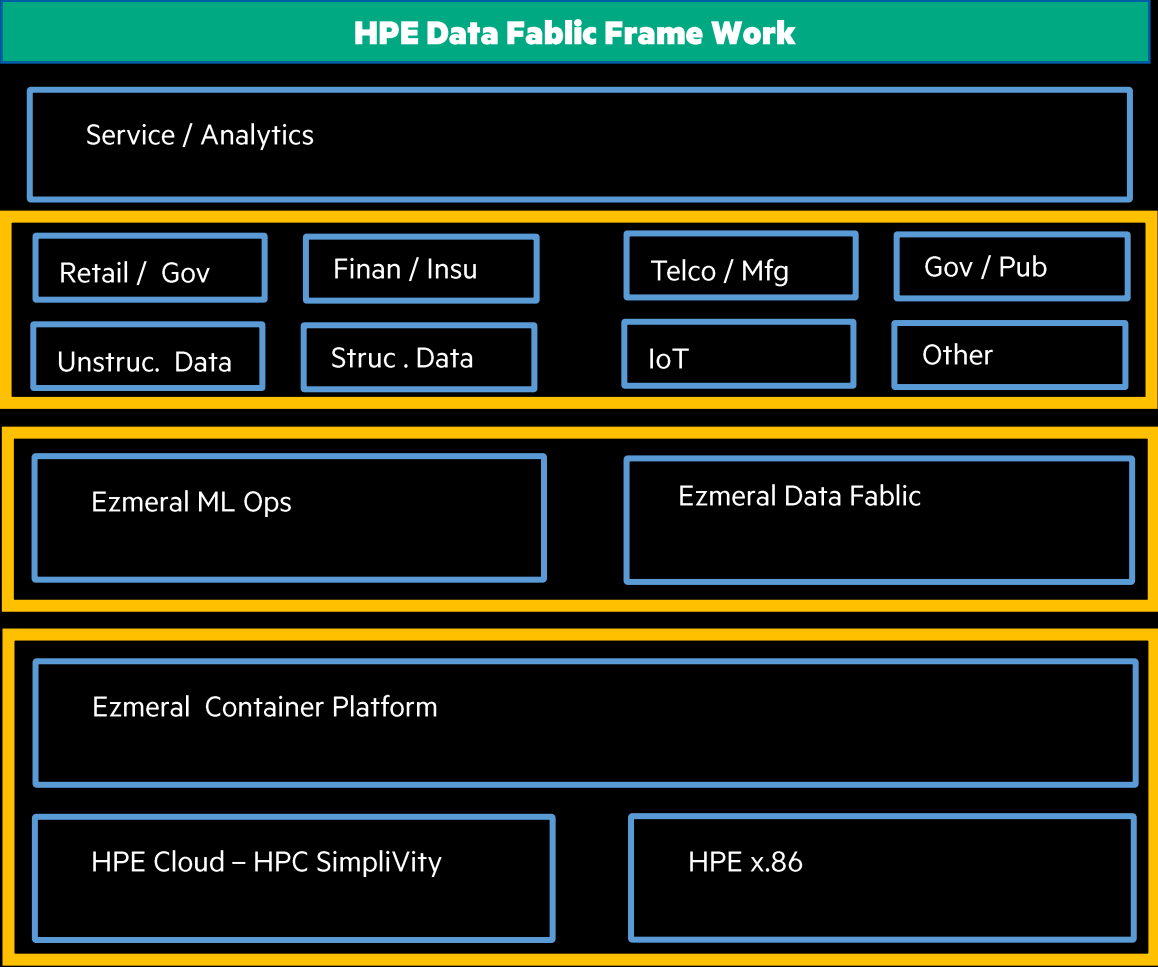


Include Software

Bigdata Fablic Eco Software

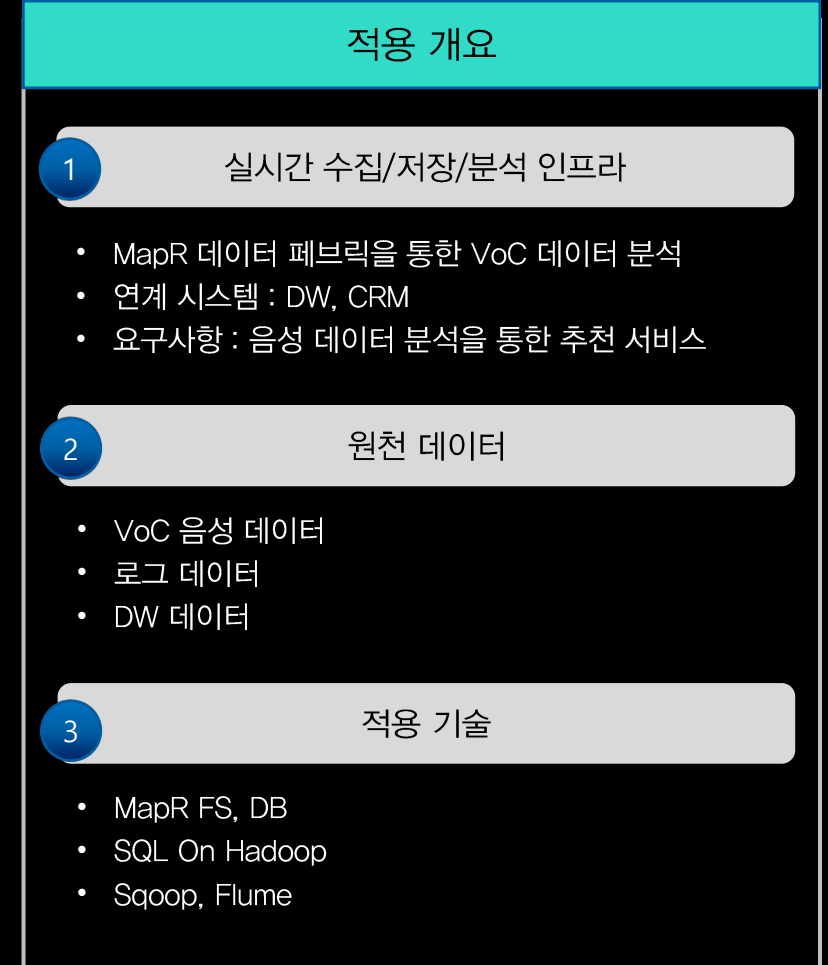
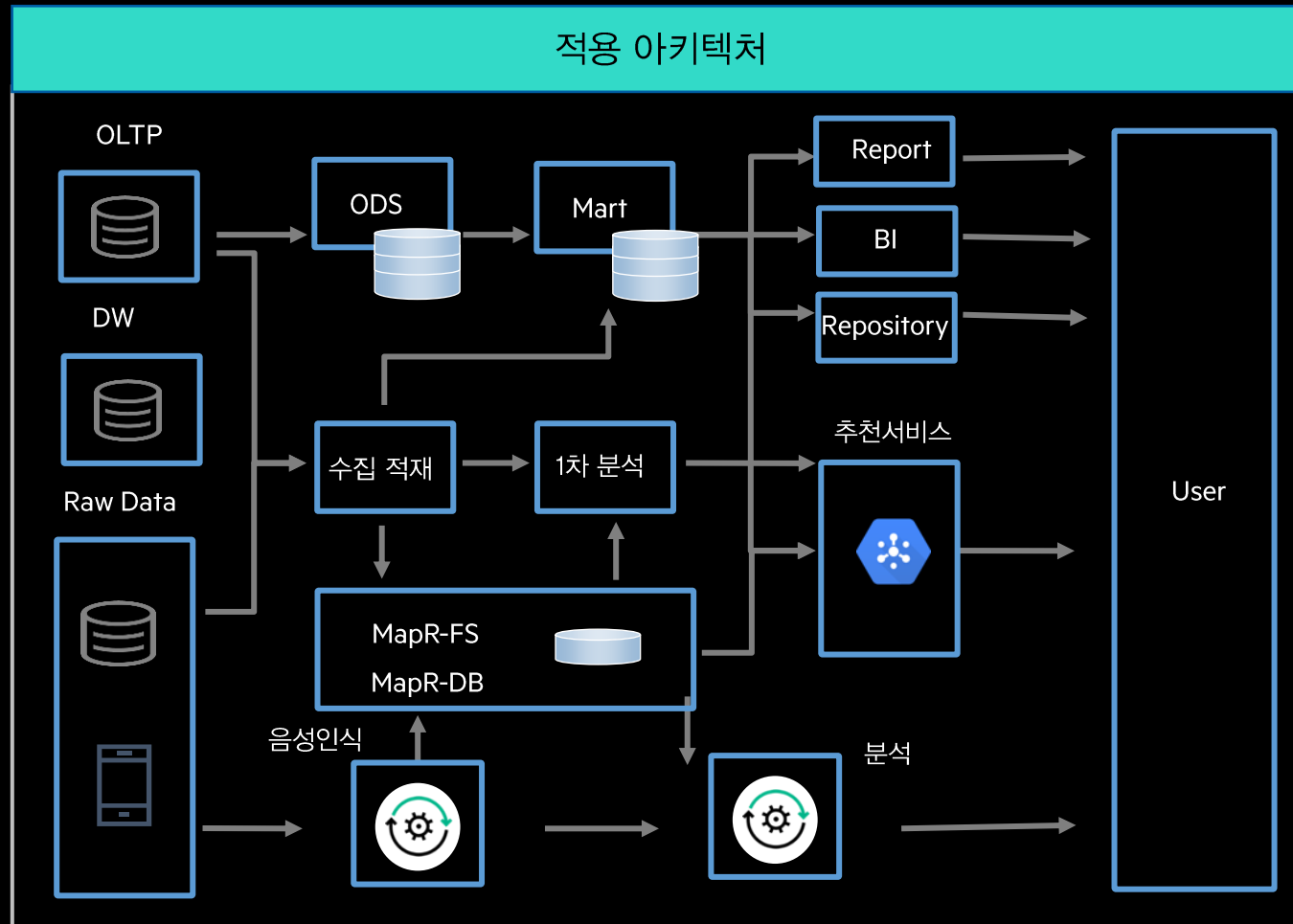
* MCS등 MapR Eco 제외

Include Eco	Eco Software
HDFS	MapR FS
DB (NoSQL)	MapR DB
SQL On Hadoop	Drill
	Impala
	Hive / Hive QL
	Pig
Cascade	Cascading
Data Transfer	Sqoop
	Flume
Real time Transfer	Spark
	Stome
Hcatalog	Hcatalog
Work Flow	Oozie
Development	Hue
ML/ DL	Mahout



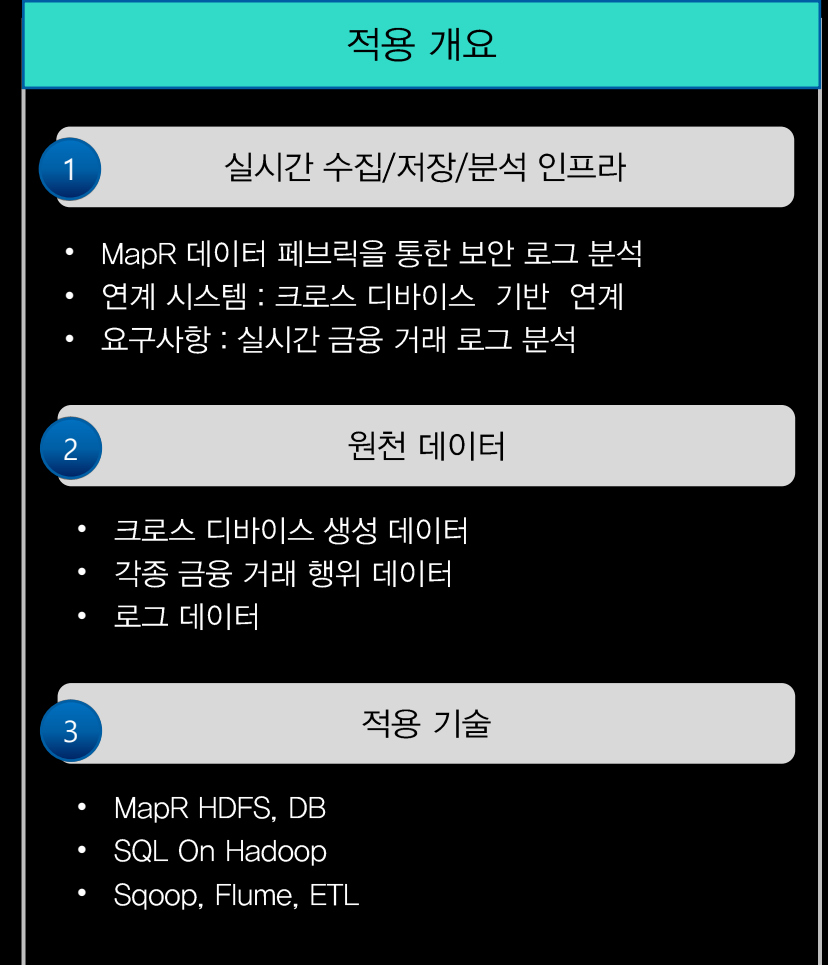
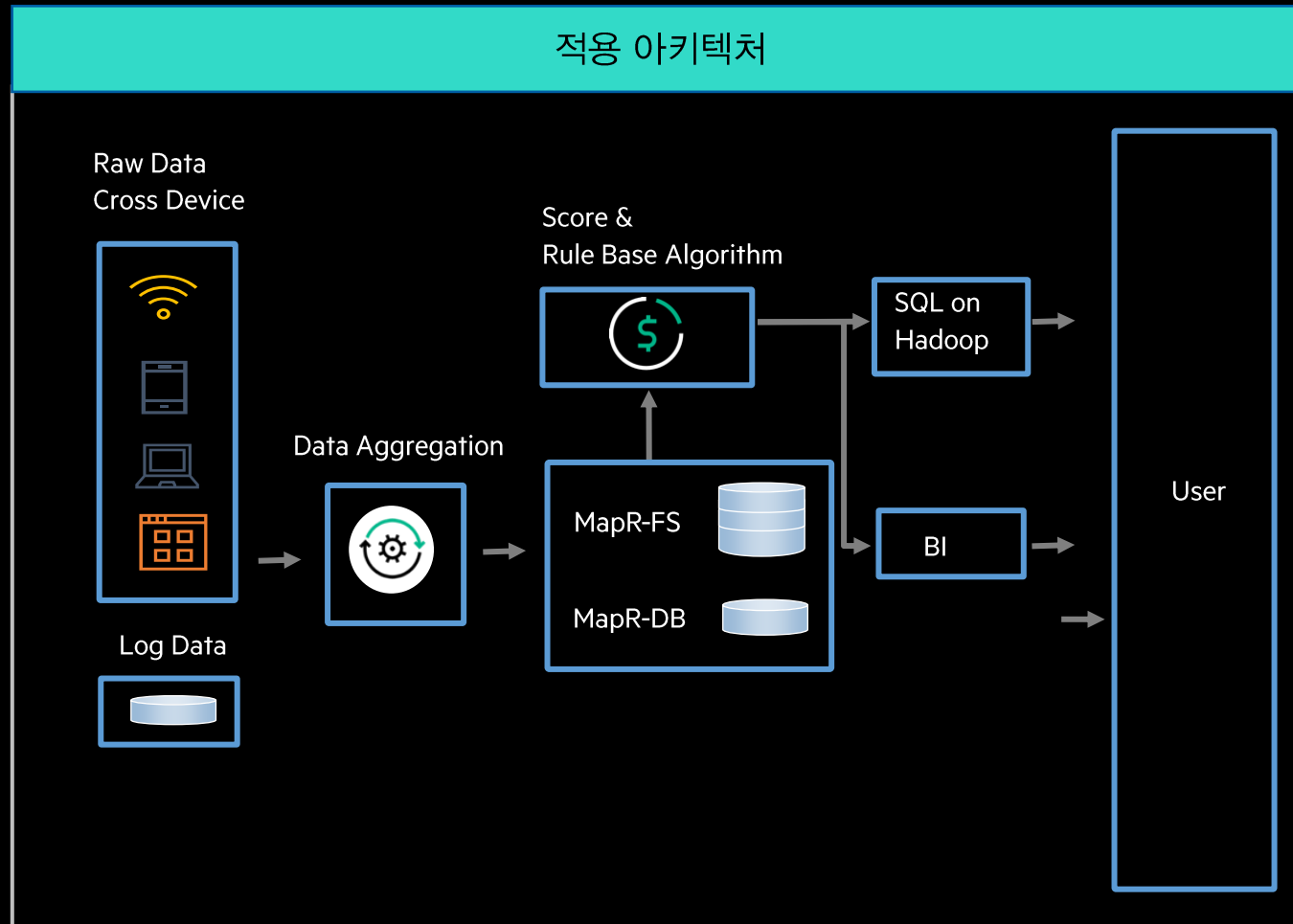
Use Case 1

금융: 음성데이터 분석을 통한 추천 서비스



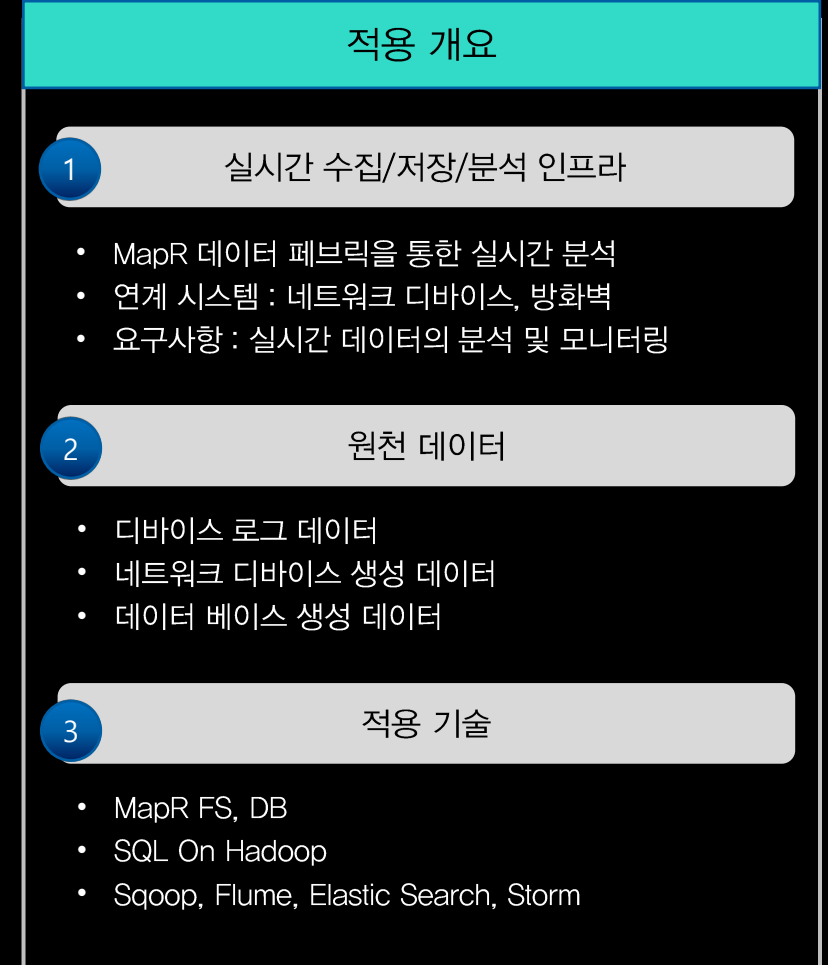
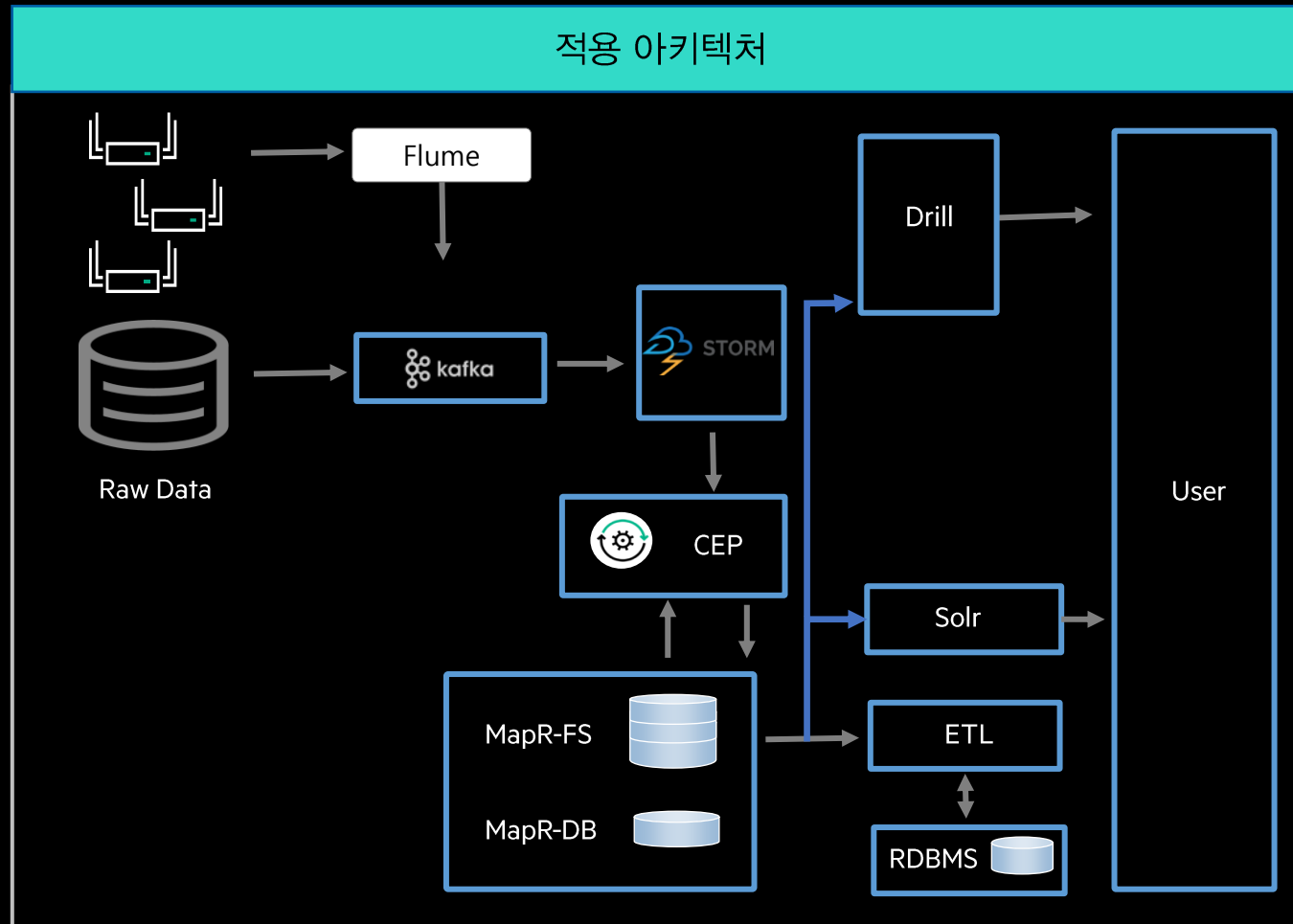
Use Case 2

금융: 보안 로그 분석



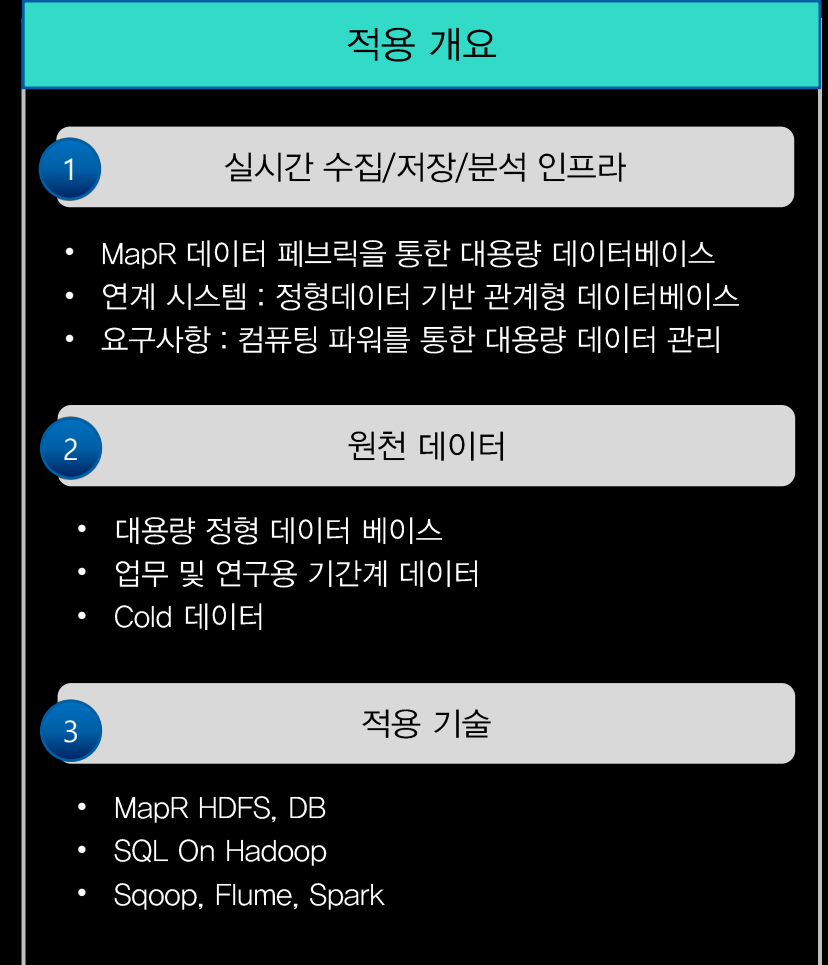
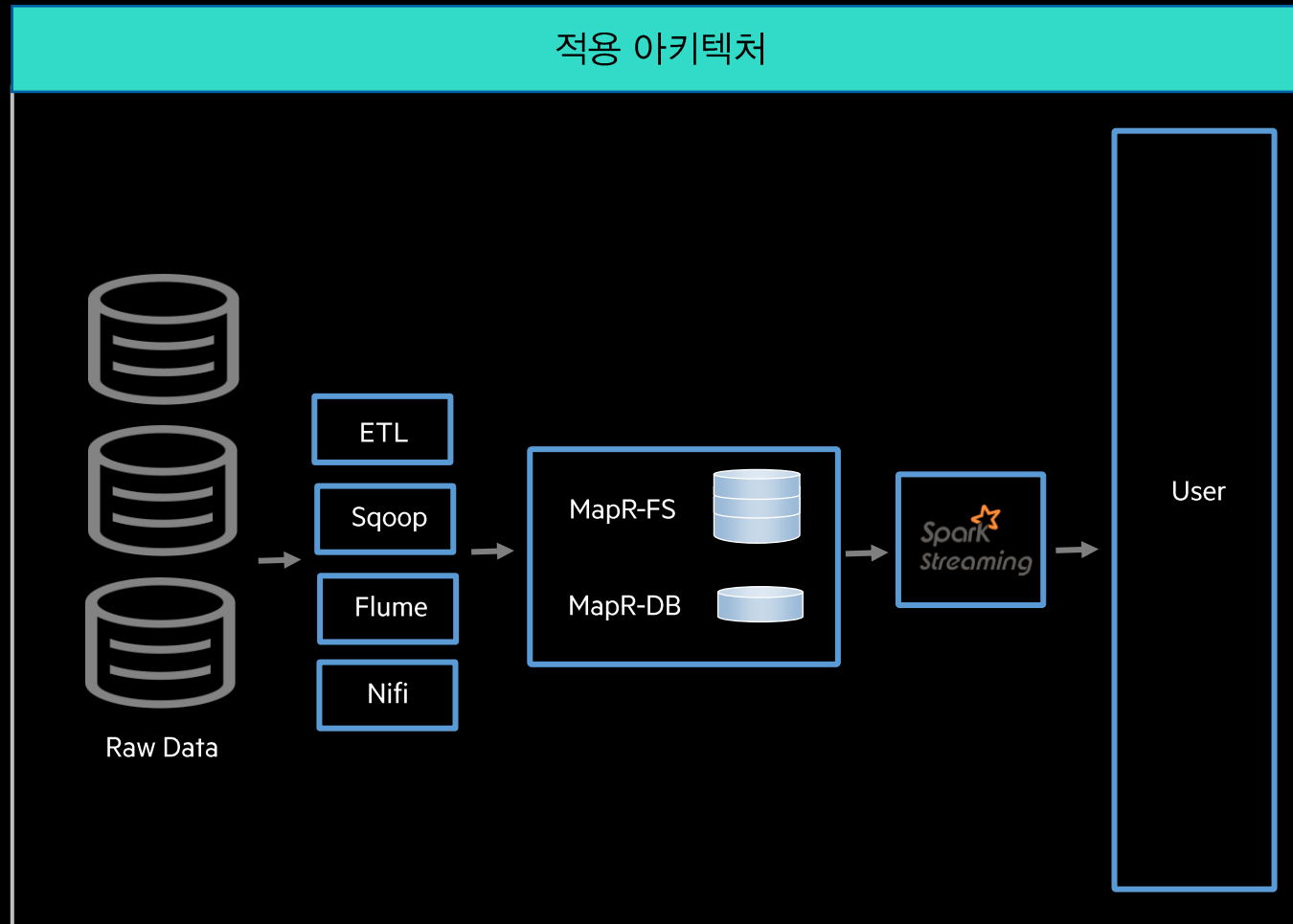
Use Case 3

고객 서비스: 각종 실시간 데이터 분석 모니터링



Use Case 4

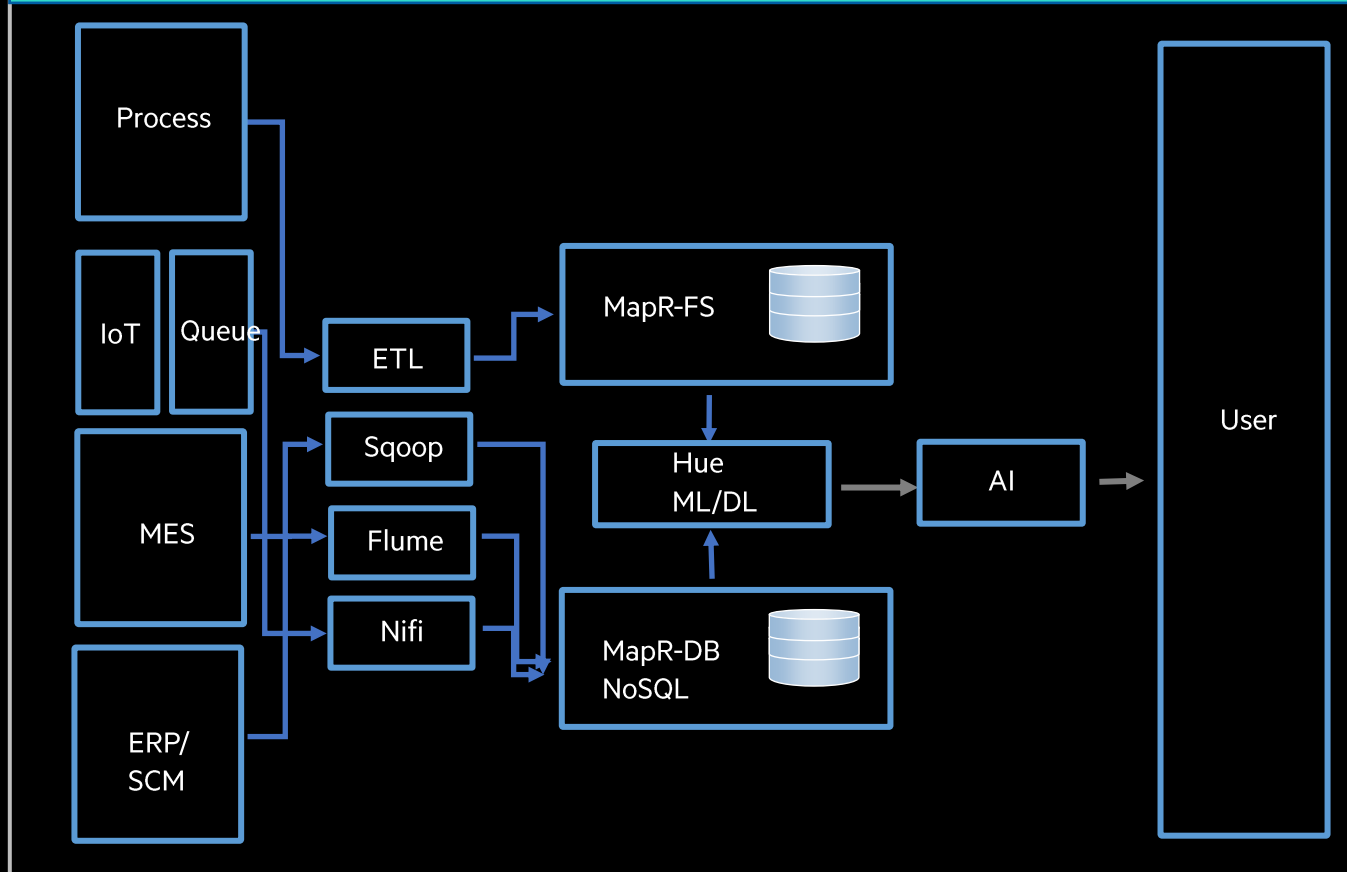
IT서비스: 대용량 데이터 베이스



Use Case 5

제조: 생산 공정 스마트 팩토리

적용 아키텍처



적용 개요

1

실시간 수집/저장/분석 인프라

- MapR 데이터 페브릭을 통한 스마트팩토리
- 연계 시스템 : 생산공정, 생산장비, MES 등 연계
- 요구사항 : 제조 공정 발생 데이터를 통한 수집 및 품질 예측 및 예지 보전

2

원천 데이터

- ERP 및 MES, SCM 데이터
- 생산공정 발생 데이터 및 디바이스 데이터
- 제조 장비의 로그 데이터

3

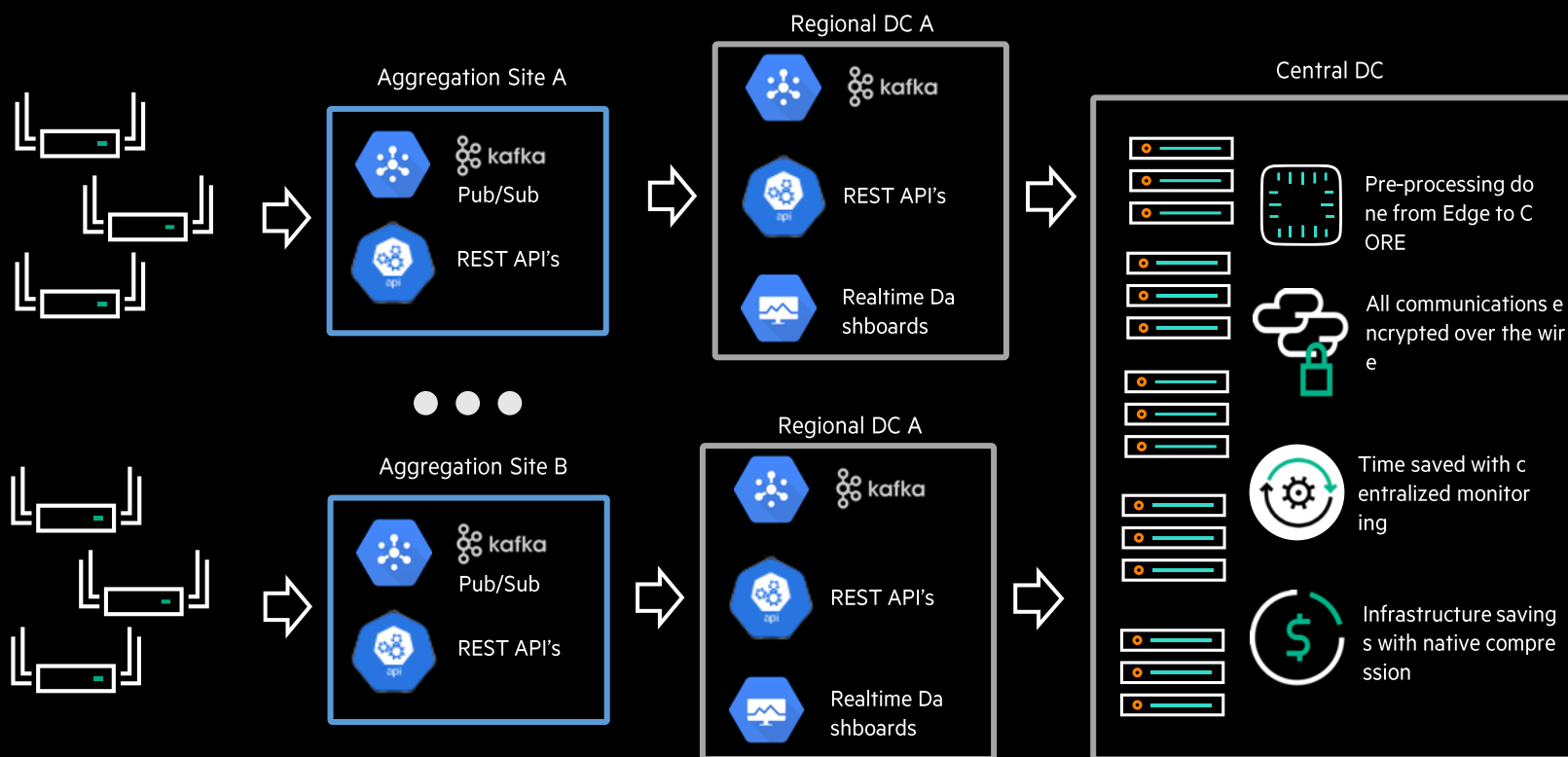
적용 기술

- MapR FS, DB, NoSQL
- SQL On Hadoop, Hue 기반 AI Engine
- Sqoop, Flume, Spark, Nifi, Kafka

Use Case 5

Overseas 통신: 실시간 데이터의 수집 및 분석

적용 아키텍처



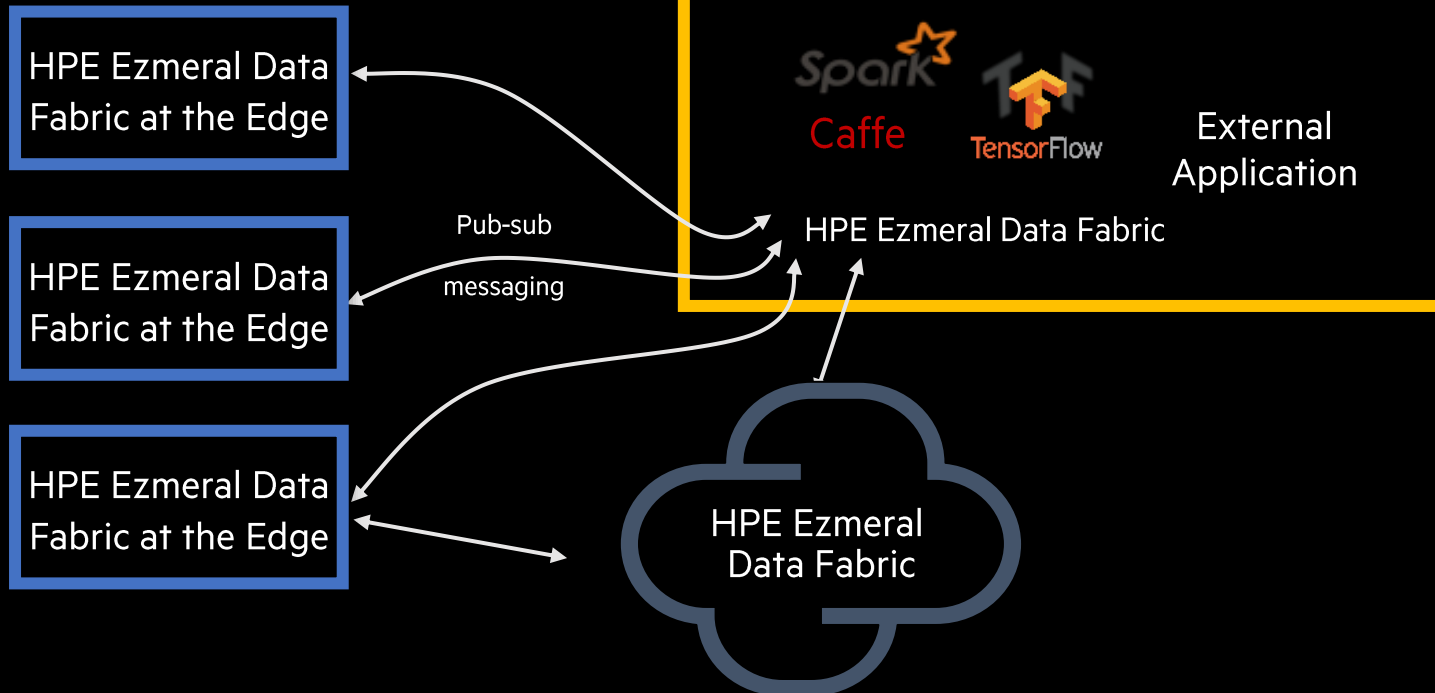
적용 개요

- 실시간 수집/저장/분석 인프라**
 - BI 기반 실시간 모니터링 및 분석
 - AD-Hoc 처리 및 분석
 - 요구사항 : 통신 디바이스 및 데이터를 통한 수집 및 레포팅 처리 시장의 단축
- 원천 데이터**
 - 통신 관련 실시간 데이터의 수집 및 저장
 - 디바이스 및 API 기반 실시간 데이터 수집
- 적용 기술**
 - MapR FS, DB, NoSQL
 - SQL On Hadoop, BI
 - API 기반, Spark, Kafka

Use Case 6

HPE Ezmeral Data Fabric Edge에 대해서

HPE Ezmeral Data Fabric at the Edge



Why HPE?

Ingest & integrate
from many APIs, sources

Full range of processing –
Analytics, ML, AI

Secure for multiple
tenants, data types

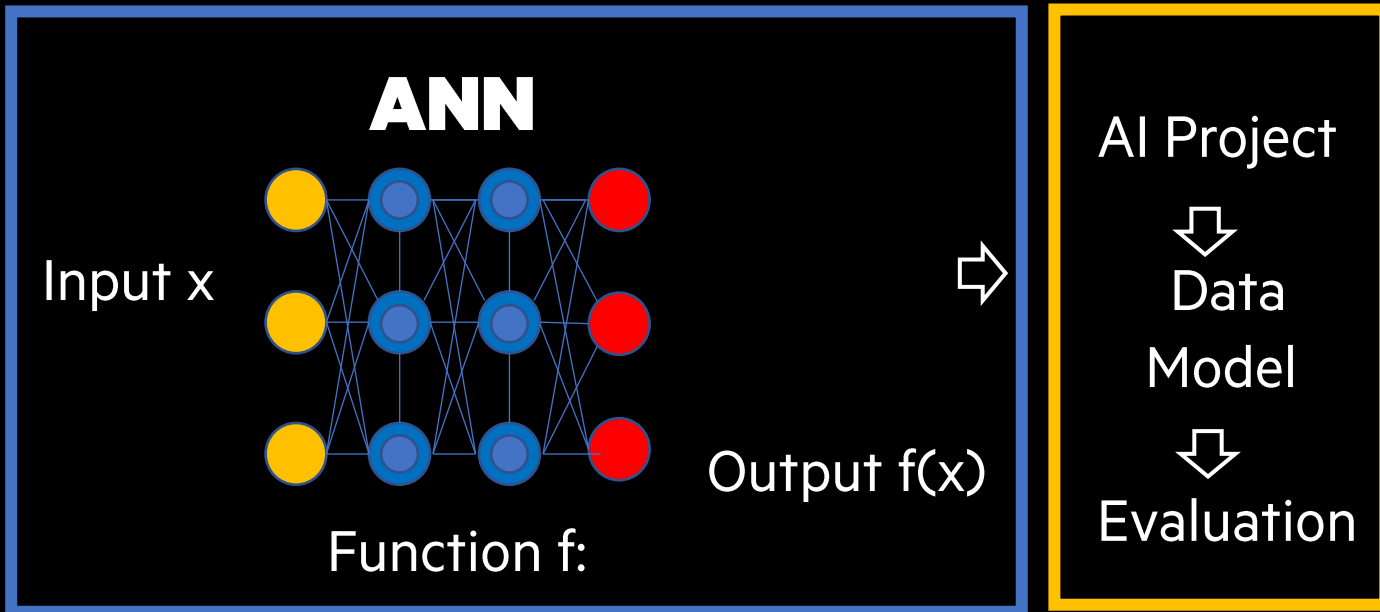
Governance, access,
catalog, control

Deliver insights into
operational apps

AI 알고리즘의 평가

데이터 기반 / 통계 모델 기반 / ML & DL 기반 AI 이슈

다양한 빅데이터 기반의 인공지능 프로젝트를 효과적으로 마무리 하는 방법은?



Bigdata기반 AI	
알고리즘	규칙기반 (Rule based Model)
	통계/계량 모형
구현	전통적 머신러닝 모형
	딥러닝 모형
기반	BigData Platform
	Eco Software
기반	AI 기반 Software
	AI Theory
기반	AI 기반 Library
	.AI
기반	Infrastructure
	Cloud or Onpromise

Thank you

Contact: ysmoon@ssu.ac.kr

