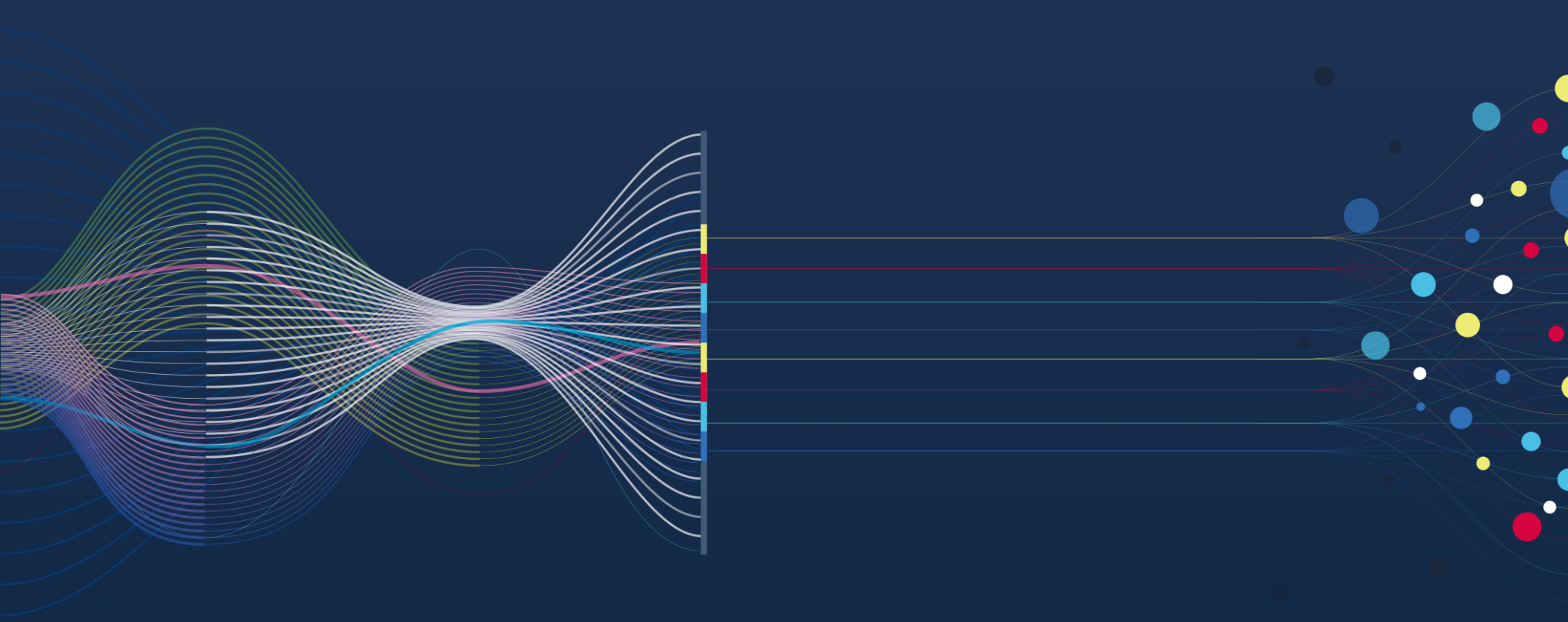
세상을 이해하는 또 하나의 눈

Insight-BigAI

빅데이터 기반 AI 서비스 운영 플랫폼



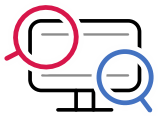
ASIANA IDT

당신 기업의 지금까지 AI 서비스 구축, 만족하십니까?

귀사의 빅데이터 기반 AI 서비스 구축 Issue 이렇진 않습니까?

- AI 분석을 위한 환경구성 라이선스 도입비용이 너무 높지 않습니까?
- AI 기반 분석 및 운영 시점에 지속적인 유지보수비가 발생하지 않습니까?

- 분석을 위한 Infra 집중으로 AI를 활용한 분석을 시작도 못하고 있지 않습니까?
- AI/분석전문가없이 업무담당자가AI기반모델을직접만들고운영환경에바로적용가능합니까?



- 개별 OSS에서 제공하는 다양한 모니터링 도구들의 활용이 가능합니까?
- Agile한 지원과 비즈니스 환경에 맞는 Customizing이 가능합니까?

ASIANA IDT

AI 서비스 구축 Issue,
새로운 관점의 시각에서 [Insight-BigAI] 해결해 드리겠습니다.

귀사에 최적화된 AI 서비스를 위한
Insight-BigAI 만의 Solution

- AI 서비스의 분석 및 인프라를 위한 합리적 도입 비용을 제안해 드리겠습니다.
- AI 서비스 운영 및 유지보수비 비용을 절감해 드리겠습니다.

- AI Infra 구성은 쉽게, 모델/분석 등 AI기반 Contents에 집중 할 수 있게 해드리겠습니다
- 사용자 기반 AutoML 구성으로 분석시 손쉽게 직접 모델 생성이 가능하게 해드리겠습니다.

- 단일화된 통합 모니터링 UI 제공으로 각각의 AI 서비스 상태 확인을 도와 드리겠습니다.
- 빅데이터 수집, 기계학습, 모델링, AI 서비스 까지 고객 맞춤 Custom 서비스를 지원해 드리겠습니다.

ASIANA IDT

빅데이터 기반 AI 서비스 구축 성공을 위한
Insight-BigAI의 3가지 제안

Insight-BigAI는 3가지 Suite을 통해
합리적이고 전문적이며 고객 기반 편의성을 제공하는
맞춤형 빅데이터 기반 AI 서비스 환경을 제공합니다.

BIGDATA PLATFORM
Suite

Open Hadoop기반
대용량 데이터 처리

MACHINE LEARNING
Suite

AI 모델링 프로세스
전처리 및 자동화

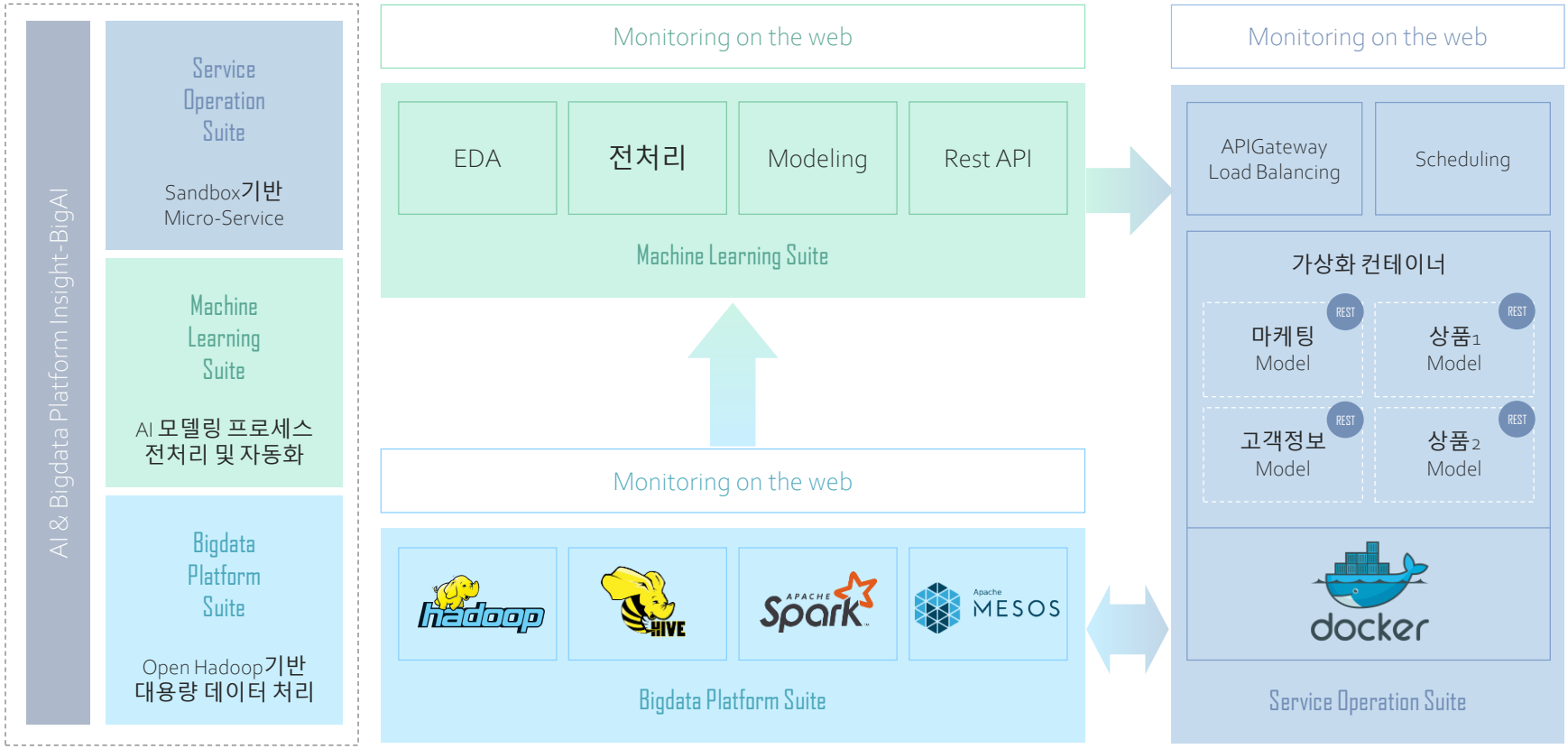
SERVICE OPERATION
Suite

Sandbox기반
Micro-Service

ASIANA IDT

대용량 데이터 저장 및 관리와 지능형 서비스 개발 / 안정적 운영을 지원하는 [빅데이터 기반 AI 서비스 운영 플랫폼] Insight-BigAI

Insight-BigAI는 데이터관리, 인공지능, 서비스운영을 담당하는 3개의 Suite으로 구성되는 서비스 플랫폼입니다.



ASIANA IDT

Suite별 기능 특성화 및 전문화를 통해 사용자가 요구하는 서비스 제공

Insight-BigAI Suite간의 기술 관계를 통해 Business에 적용 할 수 있는 서비스 제안

Insight-BigAI Product area

Service Operation Suite

- 하둡 core 아키텍쳐
- Hadoop eco system 통합
- 통합 모니터링
- 데이터 검색 및 접근

Machine Learning Suite

- 마이크로 서비스 아키텍쳐 (Netflix 검증 OSS)
- 도커 컨테이너 기반 독립 서비스 환경
- 서비스 로드 벨런싱
- AI API서비스 제공 'Insight-BigAI Machine Learning Suite'
- 통합 운영 모니터링 제공

Bigdata Platform Suite

- 데이터 소스 관리 (file / DB / Hadoop (from 'Insight-BigAI Bigdata Platform Suite'))
- 데이터 EDA 및 전처리
- Tensorflow 기반 기계학습 알고리즘 적용
- 기계학습 Hyper parameter 최적화
- API 서비스 배포 'Insight-BigAI Service Operation Suite'
- 설명가능 AI(XAI), 지속학습 (TBD)

Bigdata analysis, service development, service operation (management)

AI model development, service development, service operation (management)

Business Solution area

Aviation AI Service

- 화물/여객 RM
- NOTAM 메시지 분석
- 운송장 AI 인지
- 개인화 노선 추천
- 기내판매 면세품 추천

...

Financial AI services

- 보험 지급 자동 심사
- AML
- TBML
- 자산배분

...

ASIANA IDT

제안 I BIGDATA PLATFORM Suite을 통한 합리적이고 경제적인 빅데이터 Infra 제공

외산 솔루션대비 충분한 경제성 제공

구분	외산 솔루션 (Cloudera)	Insight-BigAI (Bigdata Suite)
비용	– 초기 도입비용이 높음 – 유지보수 및 시스템 개선 비용이 높음	– 초기 도입비용이 낮음 – 유지보수 및 기능 추가 비용 낮음
라이선스 정책	– 1 Node 당 년 $8,000 (최소 구성 : 10 Node 필수[마스터 3 + 데이터 7]) – 매년 라이선스 비용 발생 – 추가 Node 당 동일 라이선스 비용 발생	– 플랫폼 구성 1Set 기준 100,000 천원 (기본 구성 : 10 Node[마스터 3 + 노드 7]) – 유지 보수비 : 라이센스 15% – 최초 구축시 1회 발생 – 추가 Node 당 라이선스 20,000 천원 추가
최초 구성비용 (5년 TCO)	$8,000($1당 환율 1,170) x 10 node x 5 = 468,200 천원 ※ 최소 구성을 위해서 10 node 필요, (93,600천원) 발생 매년 비용	100,000 천원 + (100,000천원 x 0.15) x 4 [유지보수비] = 160,000 천원 ※ 최초 구축 시 1회 발생

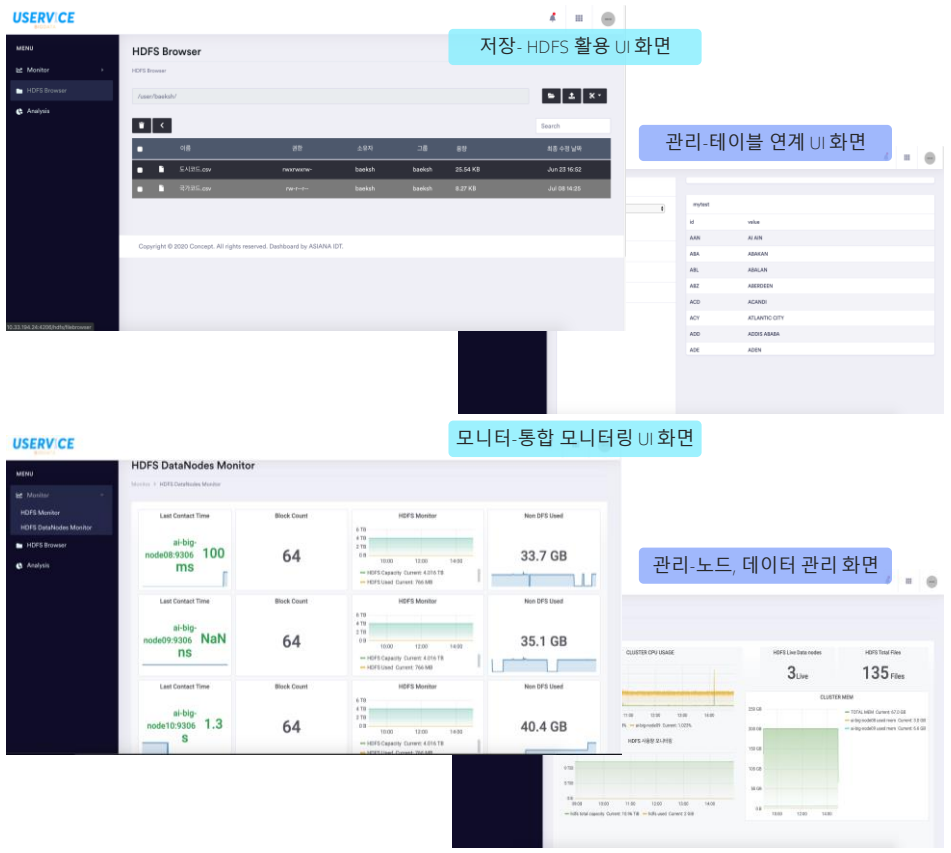
Insight-BigAI Bigdata Platform Suite을 통한 저비용 고효율의 Bigdata Platform 구성 현실화

제안 I BIGDATA PLATFORM Suite을 통한 합리적이고 경제적인 빅데이터 Infra 제공

Bigdata Platform Suite에 구성된 주요 핵심 기술을 통해 효율적 환경 구성 제공

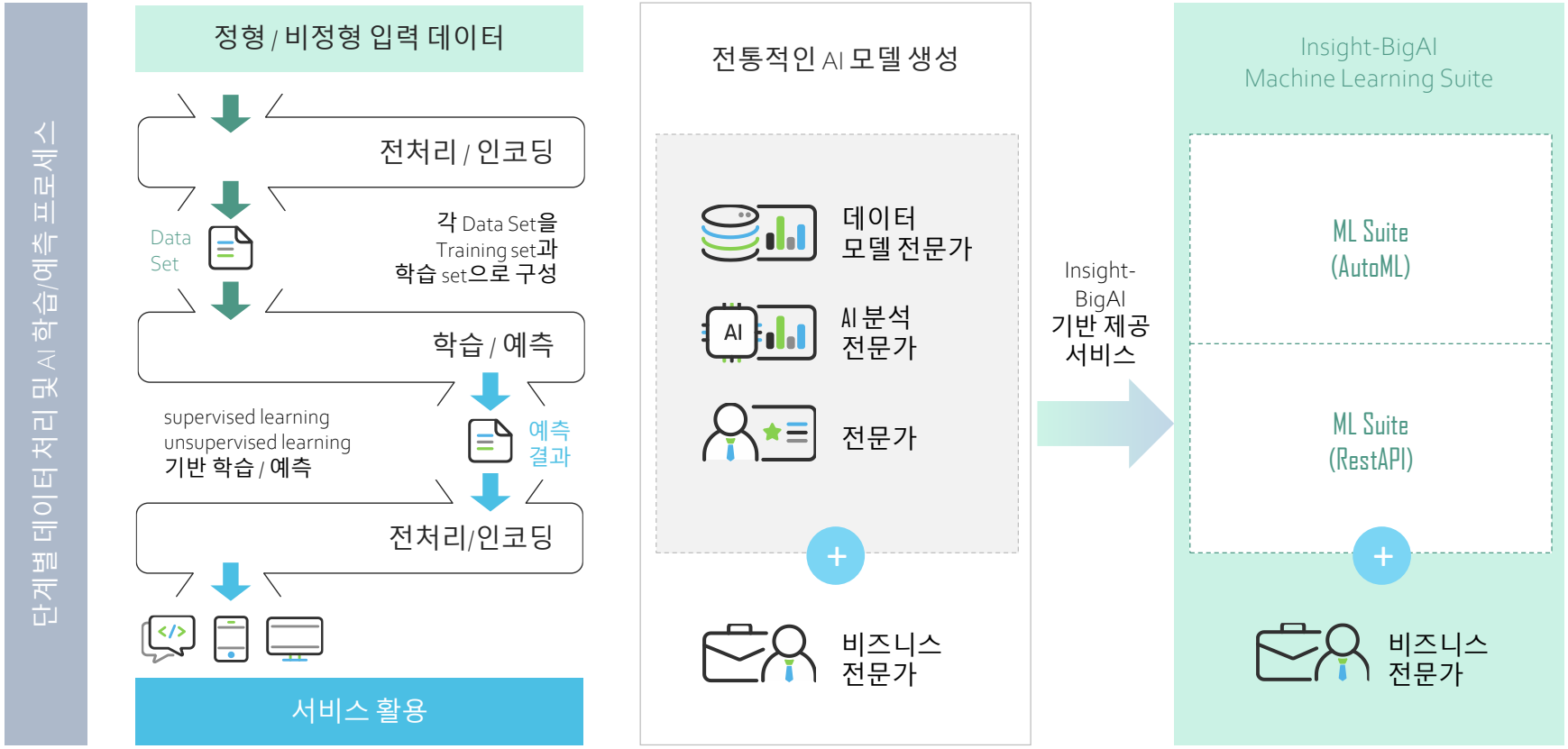
구분	기능	세부요소
데이터 저장	Hadoop Core	– HDFS (Hadoop Scalable, Flexibility, Reliability) – MESOS(enabling fault-tolerant and elastic distributed systems to easily be built and run effectively.)
데이터 관리	Apache Spark	– Spark SQL (Seamlessly mix SQL, Uniform Data Access,) – Spark Streaming (Fault Tolerance, batch and interactive queries)
	Apache Hive	– HiveQL (Familiar SQL-like Interface) – Hcatalog (Shared Data Structures) – Data mining
	Sqoop	– Transfer data – Keep table data and Hadoop in sync – Export generated or backed-up data from Hadoop – Load data into Hadoop's data Hive or HBase
데이터 모니터링	UI/UX	– Web GUI (Quickly discover relevant data) – Web interface (Save and iterate across teams) – Monitoring (intelligent event assistance)

Insight-BigAI (Bigdata Suite) 주요 화면



제안 2 MACHINE LEARNING Suite을 통한 전문적인 AI 모델링과 즉시 업무활용이 가능한 AI 서비스 구현

각 분야 전문가의 전문 지식 없이도 비즈니스 담당자만으로 다양한 AI 예측 서비스 구현 가능



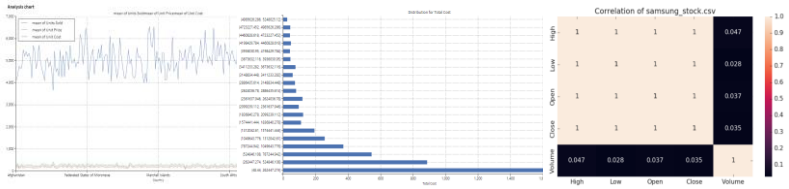
제안 2 MACHINE LEARNING Suite을 통한 전문적인 AI 모델링과 즉시 업무활용이 가능한 AI 서비스 구현

데이터 전처리에서부터 기계학습 모델 생성, 배포까지 전체적인 Lifecycle 제공

01 Dataset source 연계

- File (csv, xlsx), DB (MariaDB 외 확장 가능), HDFS (Bigdata Platform Suite연동) 다양한 Data source의 연동
- 복수 Data source의 병합
- API 를 통합 외부 Data source의 연계 (확장 개발)

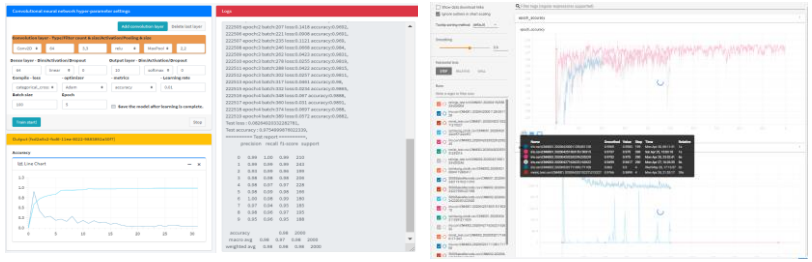
02 EDA 및 전처리



EDA
- Plot analysis
- Distribution
- Correlation
- Outlier analysis

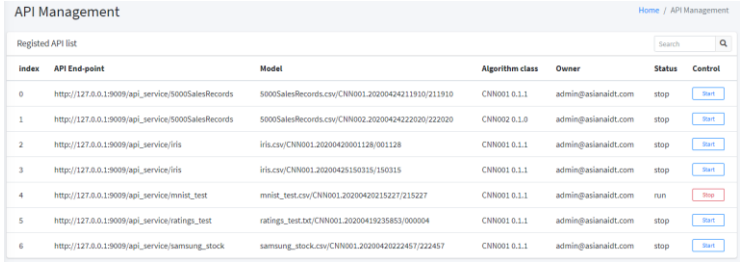
Pre-process
- Rename/Remove column
- Remove outliers/null
- Substitute null/fixed value
- Categorization
- Normalization
- Type cast
- Transform data shape
- Text embedding

03 Machine Learning



- CNN알고리즘 Hyper-parameter 설정
- 학습진행 중 Accuracy Chart 표출
- Hyper-parameter 자동 설정 (Hyperband알고리즘 적용)
- Tensorboard에 의한 성능 확인

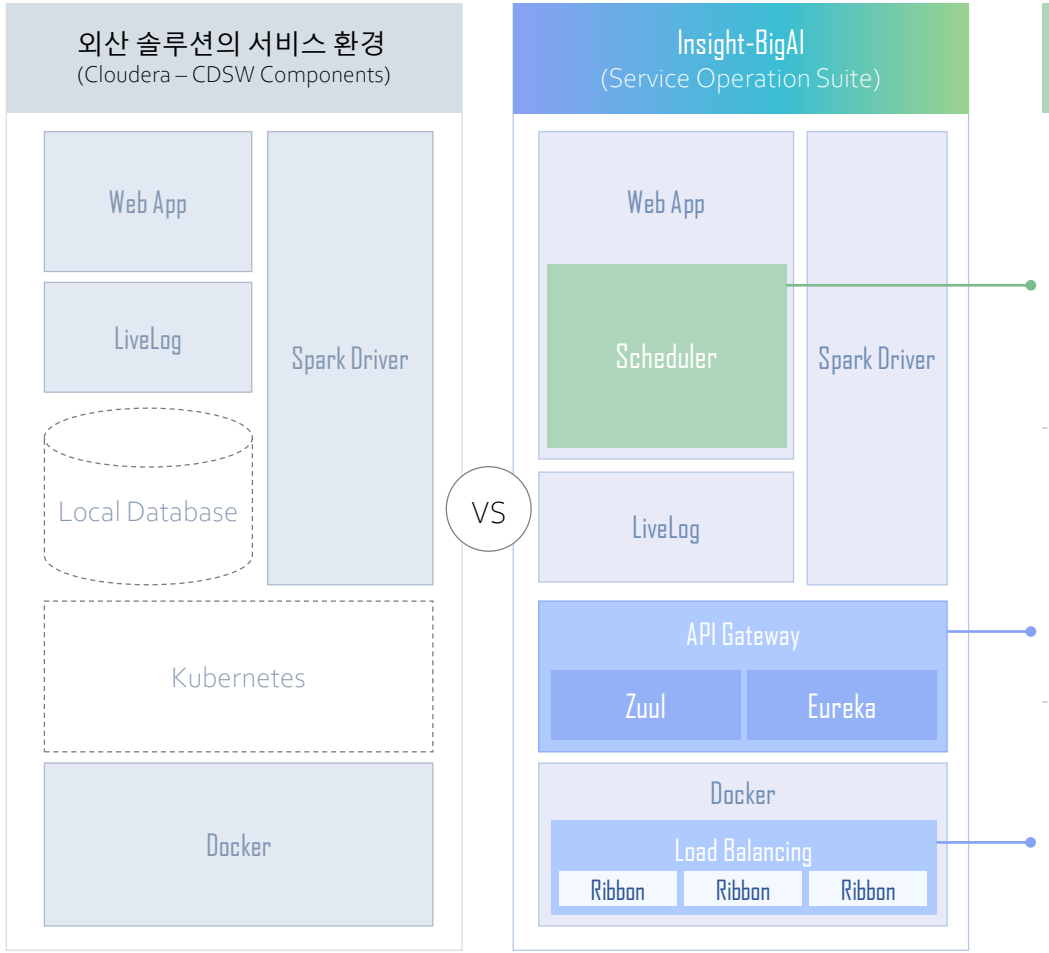
04 Model의 API Service 배포



- 모델의 Rest API 서비스 바로 배포
- Rest API의 인증 제공 (계정 or UUID 인증)
- API의 Model One-click 교체/업그레이드
- API의 Query data 이력(Lifelong data) 자동 저장 / 관리

제안 3 SERVICE OPERATION Suite 기반 즉시 활용 가능한 가상화 운영 환경 지원

Sandbox 기반의 개별 가상 환경 및 자동화된 분산처리로 서비스 관리의 편의성과 극대화된 효율 제공



Service Operation Suite 주요 특징

- 특정 시간/주기에 서비스 및 Event의 예약(반복/비반복) 작업(Job)의 자동 실행 제공
- scheduler주기, 병렬실행 및 직렬 실행, 단계별 스케줄 설정이 가능하도록 제공
- 실행 상태 모니터링 및 History / log 제공으로 적극적인 Issue 대응 가능

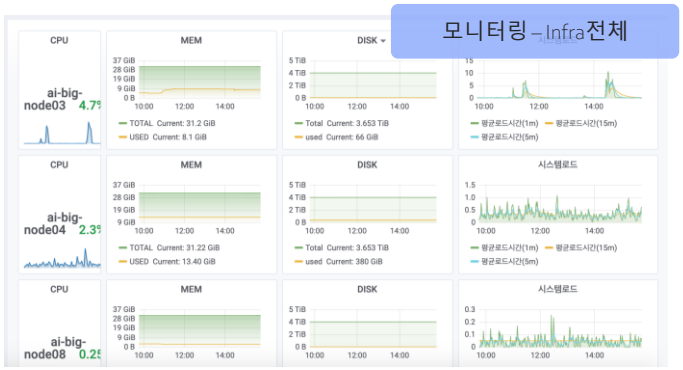
- 서비스 확장성을 구조적으로 제공하여 편리한 스케일링, 업그레이드, 유지보수 가능
- Netflix의 검증된 MSA모델 및 알고리즘 기반 구성으로 안정적이고 효율적 운영 가능
- 작은 단위의 독립적 서비스 운영으로 확장성과 안정성 극대화 제공

- HW 기반 Load Balancing대비 합리적 가격 및 유연한 구조 제공
- 자동화된 server 목록 확장 및 Issue 발생시 Retry를 통한 server 복구 기능 제공
- Peer to Peer 호출 기반으로 NW Hop 감소로 부하를 줄여 안정적 서비스 제공

표시는 Insight-BigAI Components**에 사용하지 않는 기능**

제안 3 SERVICE OPERATION Suite 기반
즉시 활용 가능한 가상화 운영 환경 지원

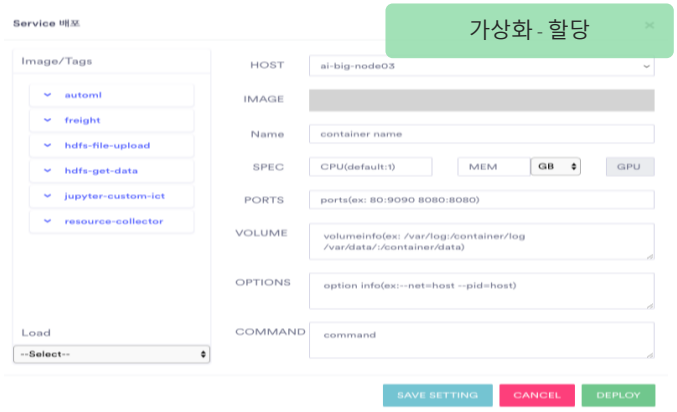
사용자에게 다양한 통합 모니터링을 제공하여 효율적 운영이 가능 할 수 있는 환경 제공



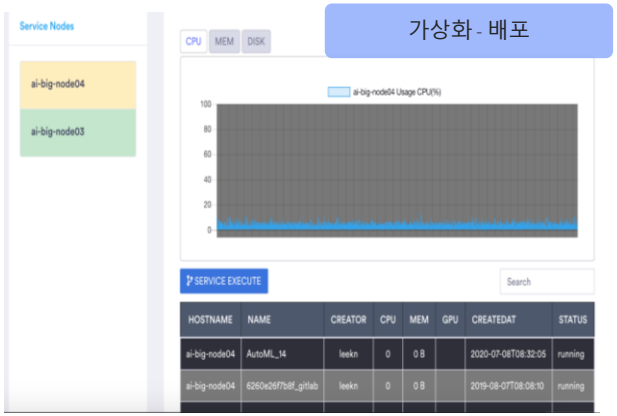
– Insight-BigAI에서 관리하는 전체 Infra node의 상태 모니터링

– Cluster 분산처리 node의 상태를 모니터링

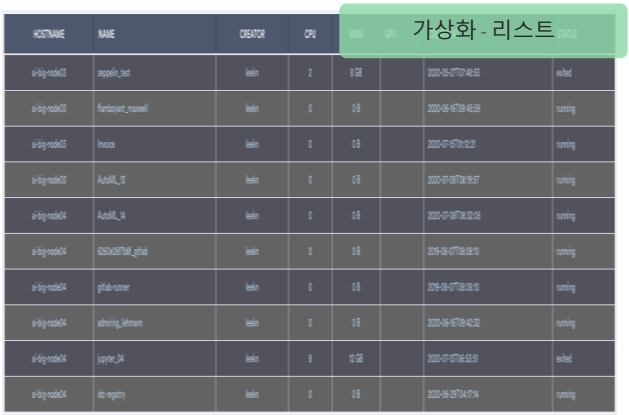
– Scheduler 관리 및 상태 모니터링/History/Log 제공



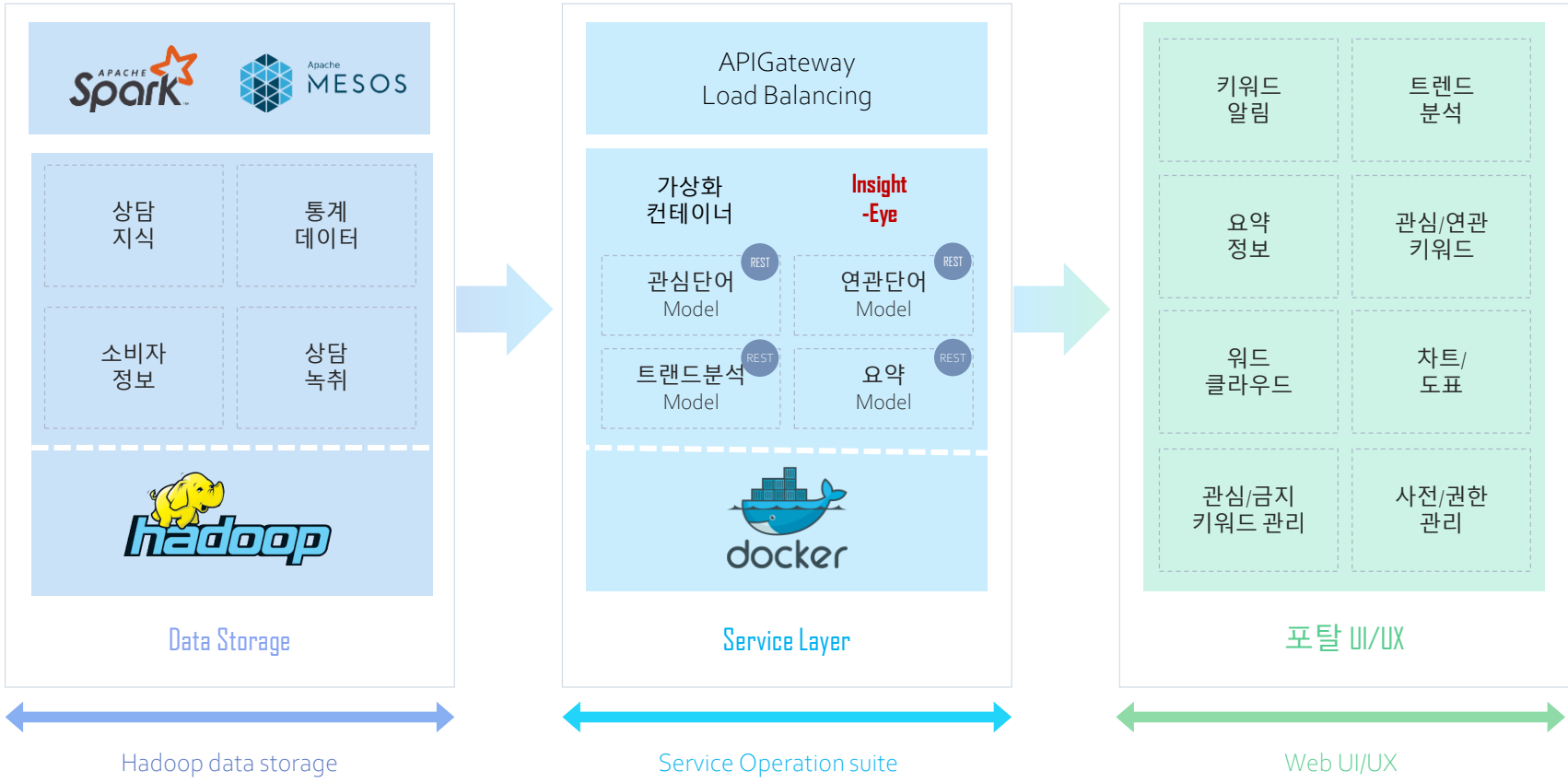
– 가상화 이미지 설정 및 서비스 관리 제공



– 배포된 서비스 통합 모니터링 기능 제공



– 노드에 배포된 서비스 이미지틑리스트 관리 기능 제공

별첨 #1 공공기관 : AI 빅데이터기반 민원시스템 구축

Suite간 기능적 관계도