

# (주)모비젠 회사소개서

Make The Best For Smart Business

A New Experience  
In Big Data Analytics

[www.mobigen.com](http://www.mobigen.com)

TEL. (02) 538-9360 FAX. (02) 538-9369 E-MAIL. [sales@mobigen.com](mailto:sales@mobigen.com)

# 회사 소개

1. 기업 현황
2. 조직 및 인원 현황
3. 기업 인증 및 수상내역
4. 사업분야
5. 주요 레퍼런스

## 1.1. 일반현황 및 연혁

(주)모비젠은 데이터 기반의 기업 비즈니스 환경을 제공하는 빅데이터 전문 기업으로서 최초로 빅데이터 플랫폼을 출시했으며 “고객의 스마트 비즈니스를 위한 최상의 솔루션 제공으로 고객과 함께하는 최고의 ICT 파트너”를 추구하는 경영 철학을 가지고 있습니다.

### 일반현황 및 연혁

1. 법인명	(주)모비젠	2. 대표자	김태수
3. 사업의 종류	<ul style="list-style-type: none"> <li>빅데이터 플랫폼 솔루션 개발/판매 및 구축</li> <li>통신망관리 응용 솔루션 개발/판매 및 구축</li> <li>시스템 SW개발, 전산 시스템 납품 및 설치</li> </ul>		
4. 주소	서울특별시 송파구 법원로 128, C동 1610호		
5. 연락처	02-538-9360	FAX	02-538-9369
6. 회사 설립년월일	2000년 3월 21일		
7. 해당부문 사업기간	2000년 3월 21일 ~ 현재 (20년 이상)		

8. 자본금 및 최근 3년간 매출액 (단위:백만원)			
항 목	2019년	2020년	2021년
자 본 금	300	300	300
매 출 액	23,374	28,340	26,089

신용등급	평가기관	이크레더블	등급평가일	2022. 04. 12
	평가 목적	정기 평가	등급유효기간	2023. 04. 11
	결산기준일	2021. 12. 31	기업신용등급	<b>BBB+</b>

### 9. 주요연혁

#### 2010 ~

- 2021 인터넷대상 비즈니스부문 과학기술정보통신부 장관상
- 2020 DNA(Data, Network, AI)분야 혁신기업 선정(정보화진흥원)
- 2019 SKT New ICT Pride Awards 수상  
SK주식회사 동반성장상 수상
- 2017 제12회 대한민국 인터넷대상 인터넷 기술선도 장관상 수상 (과학기술정보통신부)
- 2016 강남구청 지방세 우수납부 기업선정
- 2015 글로벌 창조 소프트웨어(GCS)사업 주관기관 선정  
중소기업청 고성장 기업 선정
- 2014 정보통신공사사업 면허 취득  
Big Data 플랫폼 TM 지식경제부 2012년 신 소프트웨어
- 2012 상품 대상 일반 SW 부문 대상 수상 (지식 경제부)  
Big Data 플랫폼 TM GS 인증 획득(IRIS Enterprise DB)

#### 2000 ~

- 2009 기술보증기금 A+ 멤버스 기업인증 획득
- 2005 INNO-BIZ 기술혁신기업 선정 (중소기업청)  
품질경영 ISO 9001 인증 선정 (국제표준인증원)
- 2004 우수기술기업 선정 (기술보증기금)  
병역특례(전문연구요원) 지정업체 선정 (병무청)
- 2002 2003년 서울벤처대상 특별상 수상  
유망중소기업 선정 (중소기업진흥공단)
- 2001 정보통신 우수신기술 지정 (정보통신부: IT 인증 획득)  
벤처기업 확인 등록 (중소기업청 : 연구개발기업)
- 2000 모비젠 설립 / 기술연구소 설립

## 1.2. 매출 현황 (최근 3년)

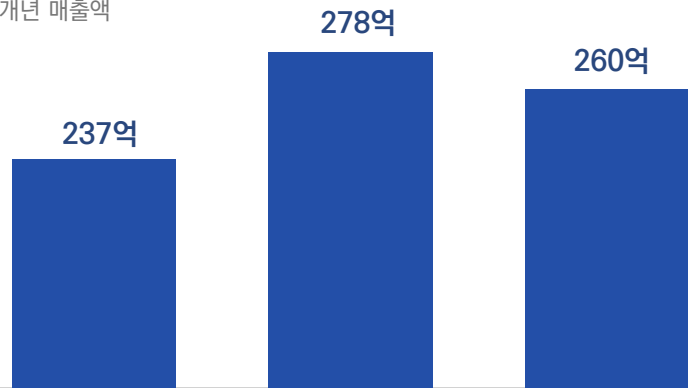
(주)모비젠은 2000년 3월 설립되어 투명한 경영을 통해 지속적인 성장과 이익을 창출하고 있으며, 최근 3년간 지속적으로 기업신용평가 등급 **BBB+** 이상을 유지하고 있는 안정되고 신뢰받는 기업입니다

### 경영 상태

(단위:백만원)

구분	2019년	2020년	2021년
매출액	23,374	27,820	26,089
총 자산	24,919	28,340	27,947
자기 자본	19,147	20,728	20,723
유동 부채	4,632	6,125	5,925
유동 자산	19,177	22,138	19,324
당기순이익	2,188	987	-213

● 최근 3개년 매출액



### 기업정보

기업제명	(주)모비젠 (MOBIGEN CO.,LTD.)		신용등급
사업자등록번호	220-81-81227	법인번호	110111-1912496
본사주소	(05854)서울특별시 송파구 법원로 128, 씨동 1610호-1621호(문정동, 에스케이이원지엘메트로시티)		<b>BBB+</b>
대표자	김태수		
전화번호	02-538-9360		현금흐름등급
산업분류	응용 소프트웨어 개발 및 공급업		
설립일	2000년 03월 20일		<b>B</b>
재무결산일	2021년 12월 31일	보고서 최종수정일	
신용등급 유효기간	2022년 04월 12일~2023년 04월 11일		

### 신용등급이력

결산기준일	정기평가			하반기평가			Watch 등급
	평가일	신용등급	현금흐름등급	평가일	신용등급	현금흐름등급	
2021/12/31	2022/04/12	BBB+	B	-	-	-	<b>정상</b> 2022년04월10일 현재
2020/12/31	2021/04/28	A-	B	2021/08/25	A-	B	
2019/12/31	2020/05/21	A-	B	2020/10/06	A-	B	

### 주요 재무 현황

(단위:백만원)

결산기준일	총자산	부채총계	자본총계	매출액	영업이익	당기순이익
2021/12/31	27,947	7,224	20,723	26,089	-1,149	-213
2020/12/31	28,340	7,611	20,728	27,820	828	987
2019/12/31	24,919	5,772	19,147	23,374	2,154	2,188

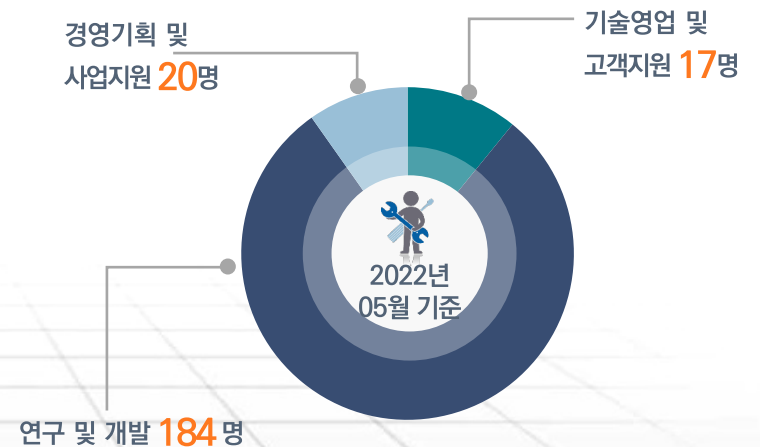
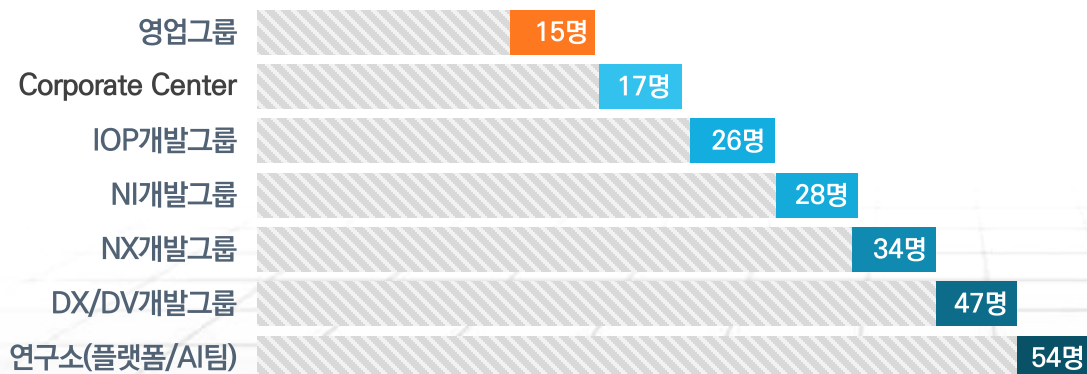
## 2.1. 조직 구성

(주)모비젠은 빅데이터 분야에서 20년 이상 축적한 경험과 기술력을 바탕으로 고객의 요구에 강한 책임감을 가진 검증 받은 기술인력을 보유하고 있으며, 전체 인력의 80% 이상이 연구개발 인력으로 구성된 기술 주도형 전문 SW 개발 기업입니다.

### 조직 구성 및 인원 현황



#### 그룹별 인원 현황



### 3.1. 기업 신인도

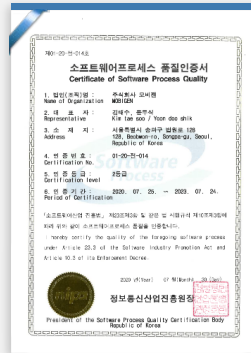
(주)모비젠은 시스템의 품질 관리 및 보안 관리를 위한 다양한 형태의 프로젝트를 수행해 왔으며, 국내 최초로 빅데이터 플랫폼을 출시한 기술력을 인정받은 빅데이터 시스템 구축 전문 기업으로 분산 컴퓨팅 빅데이터 플랫폼 특허 등 15개의 특허를 보유하고 있습니다.

#### 인증, 보유 특허 및 수상 내역

##### 주요 인증 / 수상 내역



ISO 9001



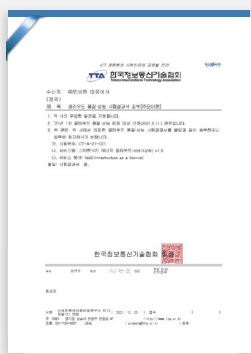
SP 2등급



GS 인증



이노비즈 인증



클라우드 품질 성능 확인서



신소프트웨어 대상

##### 인증 현황



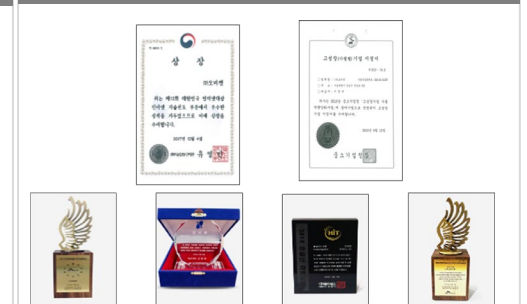
순번	인증서 종류
1	소프트웨어사업자 신고확인서 (B18-157819-005)
2	정보통신공사업 등록증 (제202175호)
3	엔지니어링사업자 신고증 (제 E-06-0001395 호)
4	직접생산확인증명서 (제2020-0507-00925호)
5	기술혁신형 중소기업(INNO-BIZ)확인서 (제 R6012-0502 호)
6	ISO 9001 (ISCQM-0470)

##### 특허 보유 현황



분산 컴퓨팅 빅데이터 플랫폼 특허 등 총 15건 기술 특허 보유

##### 수상 내역

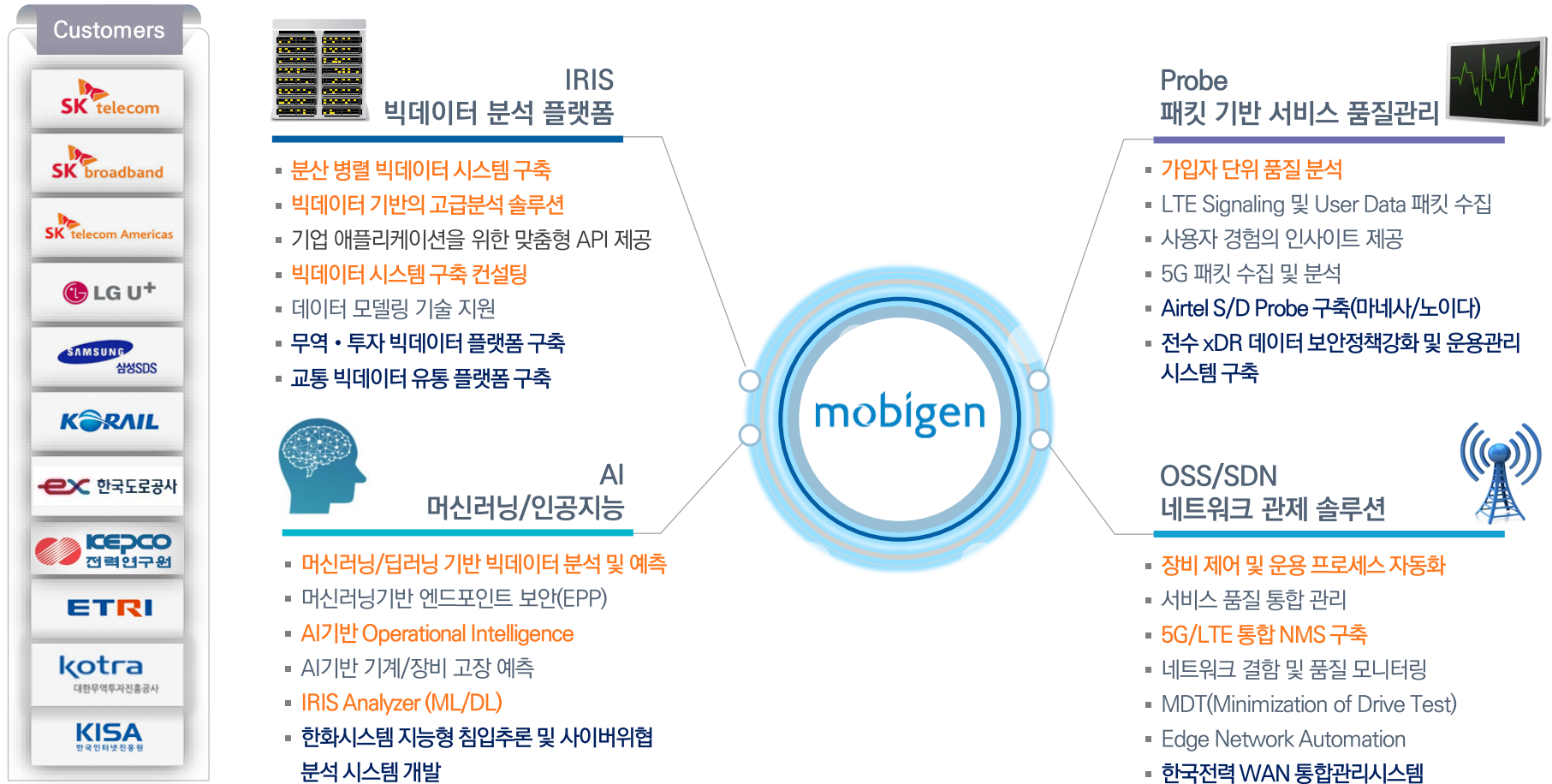


지식경제부 신 S/W 대상수상 등 다수의 수상 내역 보유

## 4.1. 주요 사업 분야

(주)모비젠은 각 그룹별 상호 유기적인 지원 체계를 기반으로 공공기관 프로젝트 및 다양한 사업진행을 성공적으로 수행하였으며, 데이터 기반의 비즈니스 환경을 제공하는 빅데이터 전문 기업입니다.

### 주요 사업 내용



## 5.1. 주요 레퍼런스

(주)모비젠은 다양한 빅데이터 분석 경험을 보유하고 있으며, M/L 기반의 분석/예측 시스템 개발 및 빅데이터 플랫폼 구축사업 등 다양한 빅데이터 플랫폼 구축사업 수행을 통해 고급분석, 기계학습, 딥러닝 등 인공지능 응용관련 풍부한 경험과 기술을 축적 보유하고 있습니다

### 빅데이터/AI 분야 주요 수행 사업

순번	사업 내용	사업기간	발주처	주요 적용기술	비고
1	무역투자 빅데이터 플랫폼 고도화	20.07~21.04	대한무역투자 진흥공사	빅데이터, 데이터품질관리, 분석	공공기관
2	통신설비 종합분석시스템 구축	20.07~21.01	한국철도공사	빅데이터, NMS, 데이터 품질관리	공공기관
3	국토교통 데이터 통합채널 구축	20.06~20.12	국토교통부	빅데이터, 데이터품질관리, 오픈API	공공기관
4	국가교통빅데이터 플랫폼 시스템 유지관리	20.05~20.12	한국도로공사	빅데이터, 데이터품질관리, 오픈API	공공기관
5	교통 빅데이터 플랫폼 구축	19.10~19.12	한국교통연구원	빅데이터, 데이터 품질관리, 데이터개방	일반기업
6	무역투자 빅데이터 분석 플랫폼 구축	19.06~20.03	대한무역투자 진흥공사	빅데이터, 데이터 품질관리	공공기관
7	19년 New Mobility TF 교통 예측 플랫폼 개발	19.03~19.12	SK텔레콤	빅데이터, 데이터 품질관리	일반기업
8	밸류링크유 해운 물류 빅데이터 분석 플랫폼 구축	19.05~19.10	밸류링크유	빅데이터 분석	일반기업
9	스마트 시민소통 · 정책결정지원 공공데이터 통합 플랫폼 구축	19.01~20.03	수원시청	빅데이터, 데이터 품질관리, 데이터개방	공공기관
10	오픈소스 기반 전력 빅데이터 플랫폼 및 분석/시각화 기능 개발	19.01~19.07	한전케이디엔	빅데이터, 데이터 품질관리	공공기관
11	New Mobility TF 교통 예측 플랫폼 개발	19.03~19.12	SK텔레콤	교통 데이터분석, 오픈API	일반기업



## 5.2. 주요 레퍼런스

(주)모비젠은 다양한 빅데이터 분석 경험을 보유하고 있으며, M/L 기반의 분석/예측 시스템 개발 및 빅데이터 플랫폼 구축사업 등 다양한 빅데이터 플랫폼 구축사업 수행을 통해 고급분석, 기계학습, 딥러닝 등 인공지능 응용관련 풍부한 경험과 기술을 축적 보유하고 있습니다

### 빅데이터/AI/SDN 분야 주요 수행 사업

순번	사업 내용	사업기간	발주처	주요 적용기술	비고
12	빅데이터 기반 RNMS 기능 고도화 통합 관리 시스템 (TANGO-O) RNMS RM기능 및 감시기능 고도화 개발	18.07~18.11	SK텔레콤	NMS	일반기업
13	국방광대역 통합망 T-SDN구축 U실험사업 (T-SDN 사전 실험 사업)	18.10~19.07	국방부	SDN	공공기관
14	빅데이터 네트워크 트래픽 분석 사업 (NaaS 플랫폼 개발)	18.07~18.12	SK텔레콤	SDN	일반기업
15	SNS 빅데이터 기반 재난안전사고 이슈 지능적 분석 기술 개발	18.06~19.01	국립재난안전연구원	AI	공공기관
16	스타트업을 위한 스마트시티 개방형 플랫폼 구축	18.07~18.10	YM솔루션	Cloud	공공기관 (전력연구원)
17	스마트시티 보안관제 시스템 개발	18.06~19.05	한국전력공사 전력연구원	AI, ML	공공기관
18	하남선 LTE-R 구축사업(NMS)	18.06~20.12	SK텔레콤	빅데이터, NMS	일반기업
19	지능형침입추론 및 사이버위협 분석시스템 개발	18.03~19.07	한화시스템	AI, ML	일반기업
20	사이버위협 정보 공유 자동화 시스템 공급	18.03~18.04	금융보안원	빅데이터, 보안	일반기업
21	공항철도 차세대 통합정보시스템 구축 (전사 정보분석체계 시스템 부문)	18.01~19.03	케이씨씨 정보통신	빅데이터, 분석	공항철도 (공공기관)
22	다중센서 분산처리 및 GEOINT 자료 통합처리 시스템 구축(분산처리 부문)	16.12~19.07	로딕스	빅데이터	국방과학 연구소 (공공기관)



A New Experience  
In Big Data Analytics

# iS IRIS

지능형 빅데이터 분석 플랫폼

## 1.1. IRIS 지능형 빅데이터 분석 플랫폼 제품 라인업

### IRIS ETL

(데이터 수집 및 연동)

- 다양한 종류의 외부 시스템으로부터 정형, 비정형, 반정형 데이터의 수집, 추출 및 변형, 적재 기능 제공

### IRIS Enterprise DB

(데이터 저장 및 분석)

- 빅데이터 분석을 위한 병렬분산 구조
- In-Memory, SSD, HDD 동시사용 Hybrid 구조
- 자유로운 Scale-Out 시스템 구조
- DBMS 형태의 빅데이터 관리 기법
- 초당 200만 레코드 이상 인덱싱 처리 및 검색 속도 보장

### IRIS Analyzer

(AI 기반 고급 분석)

- 원시 빅데이터의 직접 접근을 통한 데이터 추출 및 분석
- 분석가 관점의 지능형 분석 도구
- 비 전문가의 데이터 분석 업무 지원
- 데이터 분석 표준 프로세스 제공
- 자체적인 분석 및 시각화 도구 제공

### IRIS Discovery Service

(미들웨어 분석 서비스)

- 동일한 명령어로 다양한 데이터소스에 저장된 데이터 검색 및 분석 환경 제공
- 분산 환경에서 그대로 사용 할 수 있는 DSL(Domain-Specific Language) 지원 서비스
- RESTful API를 제공

### IRIS Hadoop

(데이터 저장 및 분석)

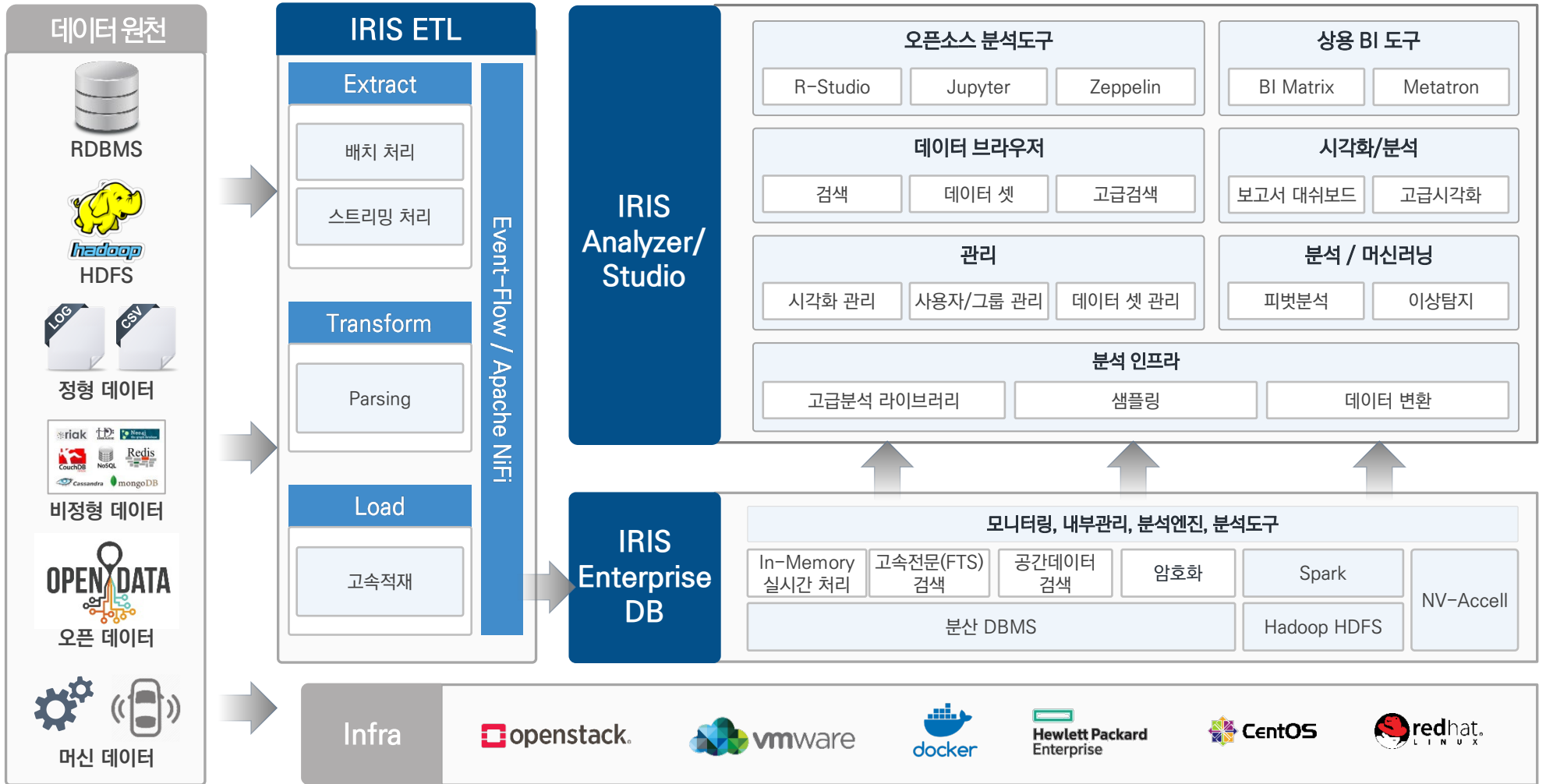
- 빅데이터 분석을 위한 병렬분산 구조
- 관리자, 개발자, 분석가를 위한 통합 UI 제공 (망분리 환경 고려)
- 고객맞춤 기업용 Hadoop 패키지

### IRIS Studio

(사용자 정의 보고서)

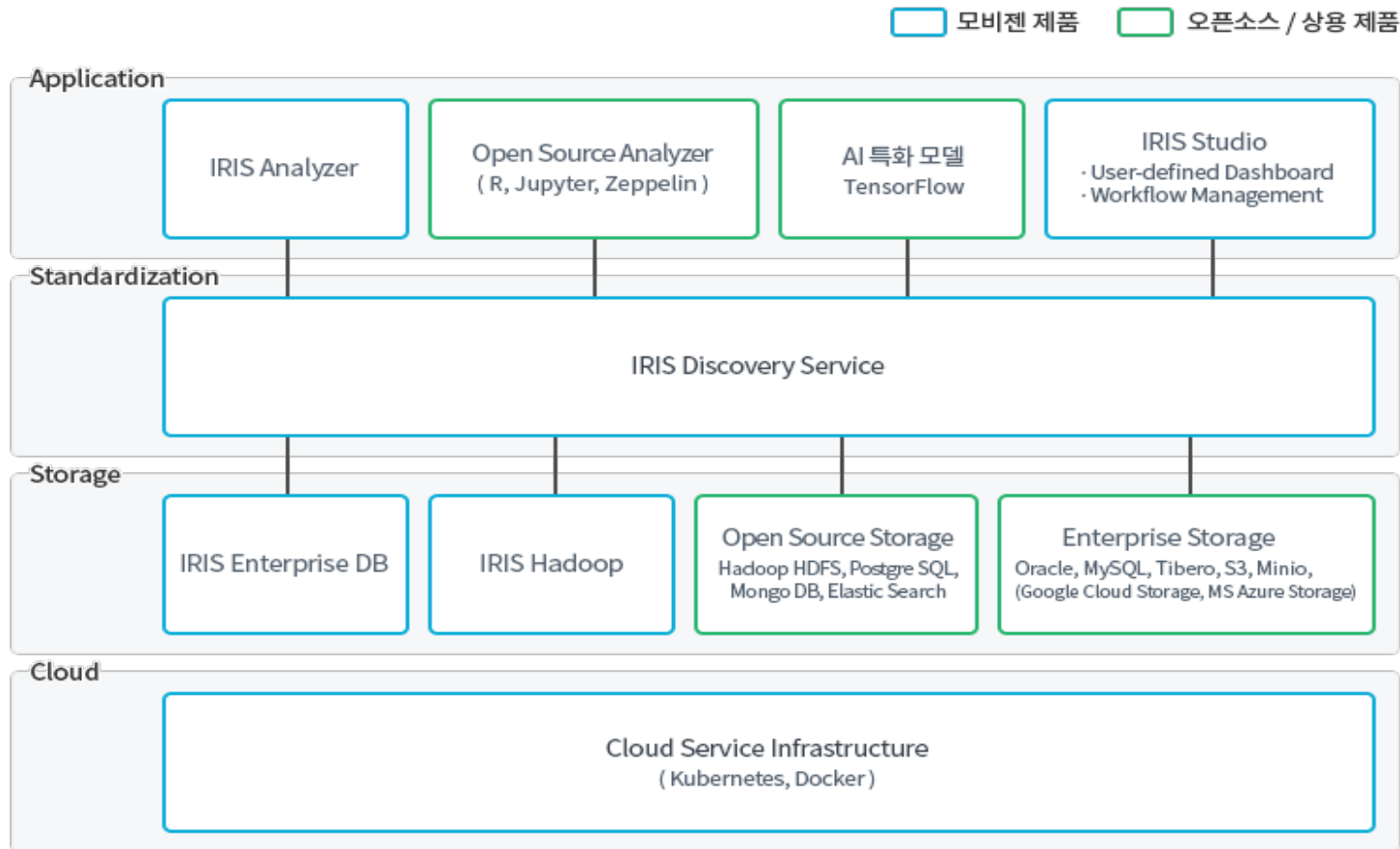
- Web 기반 시각화 도구 제공
- 데이터 분석화면/보고서 제작 도구 제공
- 데이터소스, 탐색 범위, 보고서 갱신 주기 등을 동적으로 보여주는 대쉬보드 활용

## 1.2. IRIS 지능형 빅데이터 분석 플랫폼 제품 구성도



## 1.3. IRIS 지능형 빅데이터 분석 플랫폼 아키텍처

클라우드 기반의 지능형 빅데이터 플랫폼인 아이리스(IRIS)는 빅데이터의 수집, 저장, 가공, 분산처리, 분석, 시각화, 공유를 통합적으로 지원하는 "빅데이터 End to End 처리 플랫폼" 으로 빅데이터 파이프라인 인 "수집-저장-분석-시각화-공유" 모든 과정을 통합한 빅데이터 통합 분석 및 시각화 플랫폼입니다.



mobigen



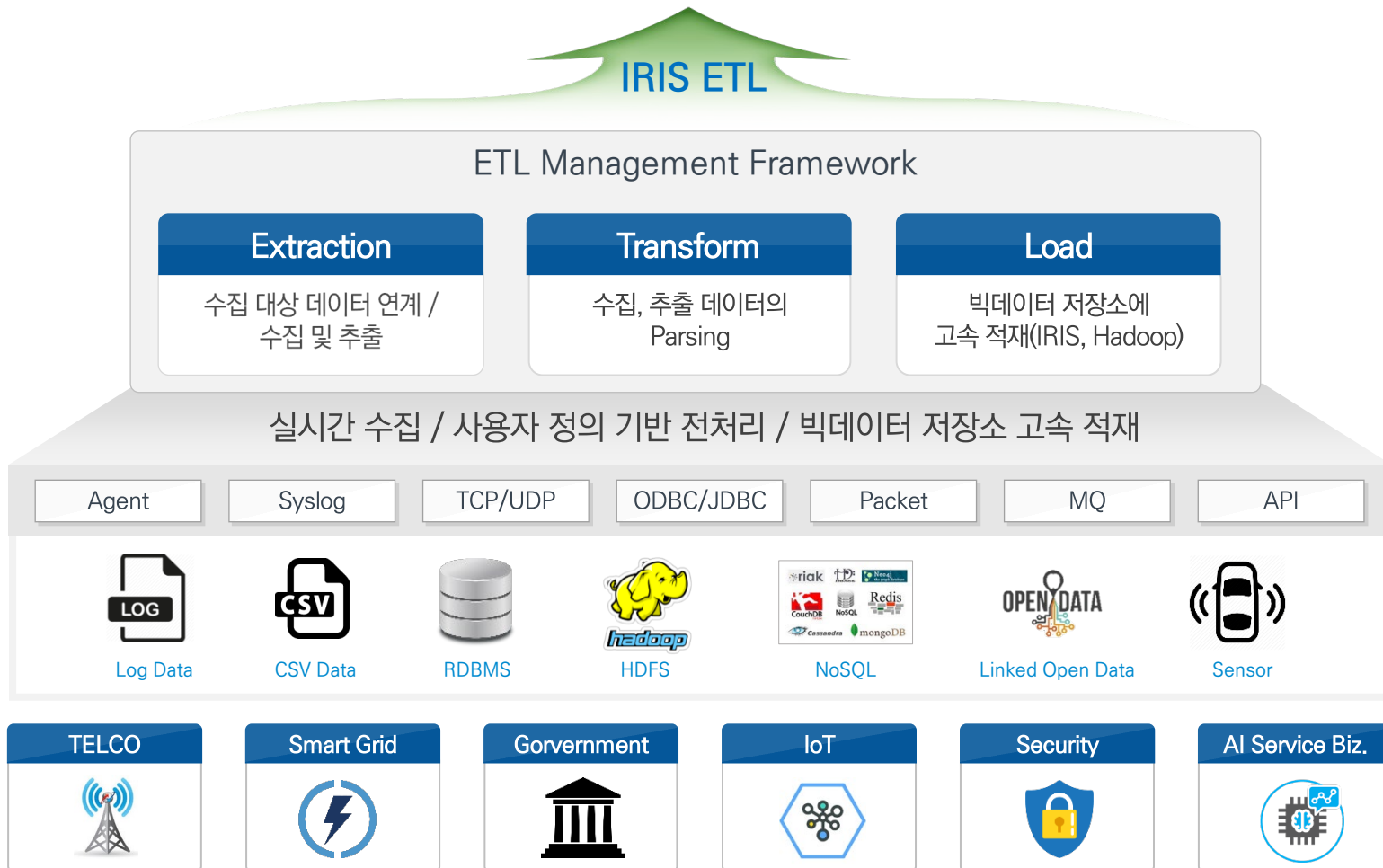
**IRIS ETL**

Extendable Big Data ETL System for IRIS

A New Experience  
In Big Data Analytics

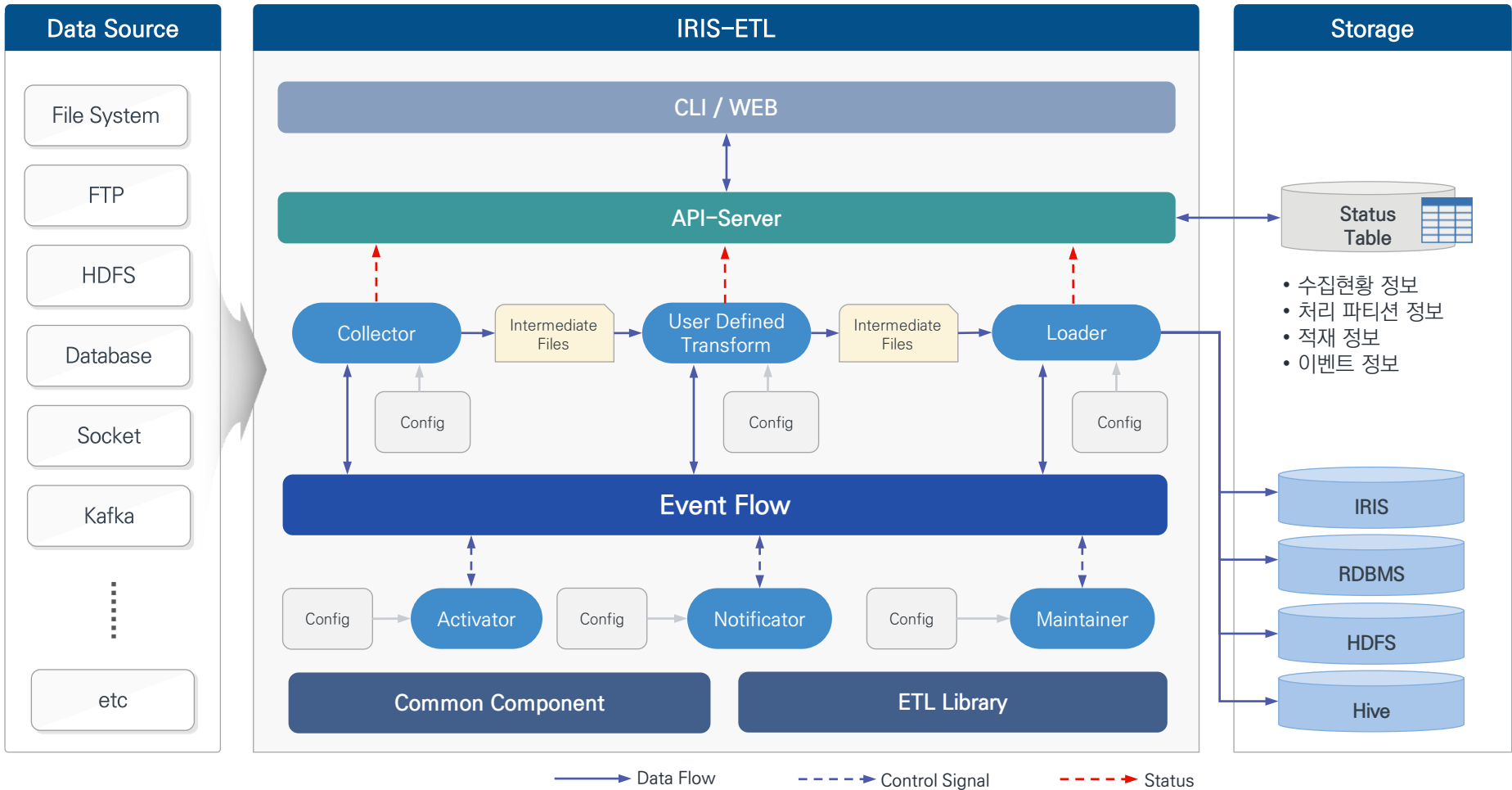
# IRIS ETL 시스템

다양한 종류의 외부 시스템으로부터 정형, 반정형, 비정형 데이터를 수집 · 추출 및 변형 · 적재



# IRIS ETL Architecture

데이터 처리 프로세스의 유연한 확장과 신속한 장애복구 대응이 가능한 EDA 기반 프레임워크





# IRIS ETL 시스템 특징

## 빅데이터 시스템 구축을 위한 기간 및 경제성 확보

### 확장성

- 가변적 분산 수집 구조로 수집 대상 시스템의 확장 대응
- 이중화 및 고가용성(HA)

### 안정성

- 대용량 데이터 고속 처리
- 룰 기반 재처리 자동화를 통한 완벽한 손실 방지

### 운영성

- ETL 프로세스의 컴포넌트 기반 독립적 관리 (재사용성)
- 다양한 데이터 형식 / 프로토콜 지원 (유연성)
- 실시간 / 배치처리 지원 (주기성)

### Distributed Data Integration System for IRIS

#### 프레임워크 기반 ETL

- 수집 / 변환 / 적재 과정의 ETL 프로세스 모듈화
- 룰 기반 파서 적용
- 모듈화된 ETL 프로세스 통합 관리
- 프로세스의 Scale-Out 구성

#### Standard Stream I/O

- 커널 레벨의 안정적인 고속 I/O 처리
- 표준 스트림(STDIN/OUT/ERR) 데이터 채널 기반 이벤트 처리
- 프로세스 데몬의 Kill 방지

IRIS ETL

#### Easy DevOps

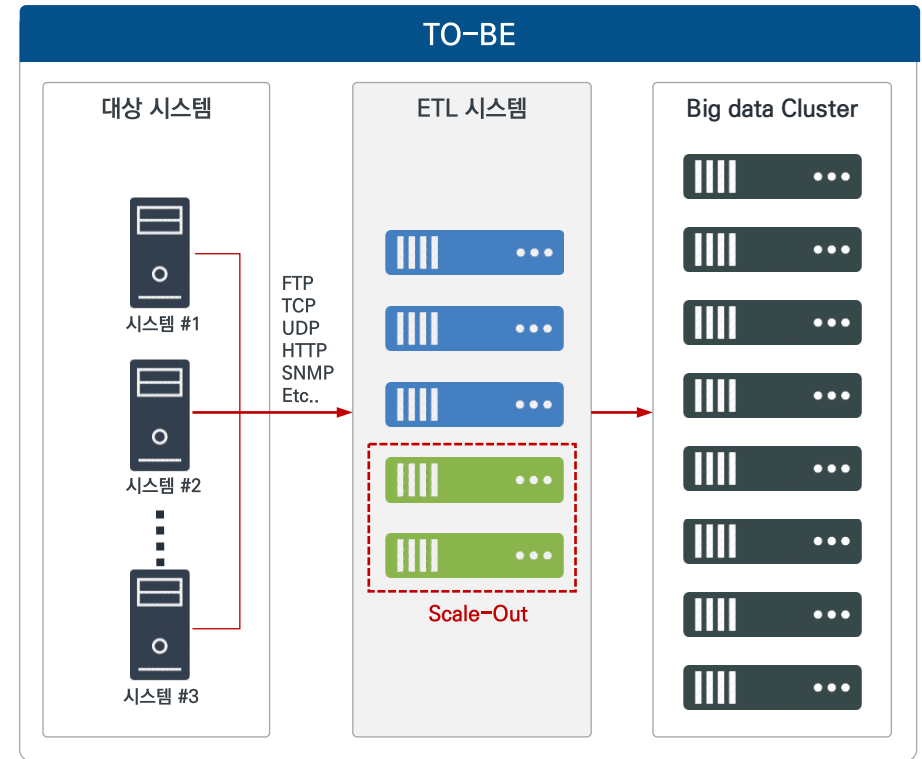
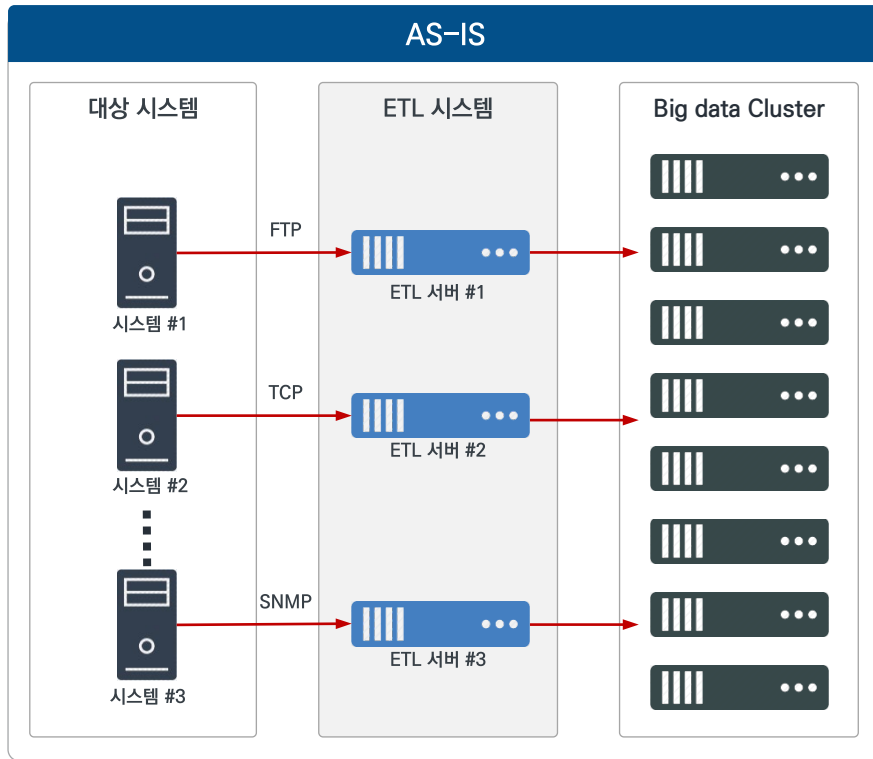
- ETL Job 개발 / 테스트 / 배포
- 다수 프로세스의 실행 / 중지 / 모니터링
- 재사용 가능한 컴포넌트 구조
- 신속한 장애 대응

#### Security

- ETL 설정 정보 암호화를 통한 시스템의 위 변조 방지

## 실시간 분산 수집

### 수집 대상 데이터(시스템)의 구성형태 / 처리용량 / 발생 속도에 따라 가변적 구성 및 확장성 제공



- 수집 대상 시스템에 맞게 별도의 수집서버 증설 적용
- 수집 데이터량 증가시 별도의 수집서버 증설
- 수집 서버의 이중화 기능 별도 구현
- 시스템 구성 복잡도 증가 및 운용 복잡도 증가

- 가변적 분산 수집 구조
- 수집 대상 및 수집 용량에 맞는 가변적 확장성(Scale-Out) 제공
- 빅데이터 전처리를 통한 META Data 생성 지원
- 수집 기능 이중화 제공 및 운용 안정성, 고가용성(H/A) 제공

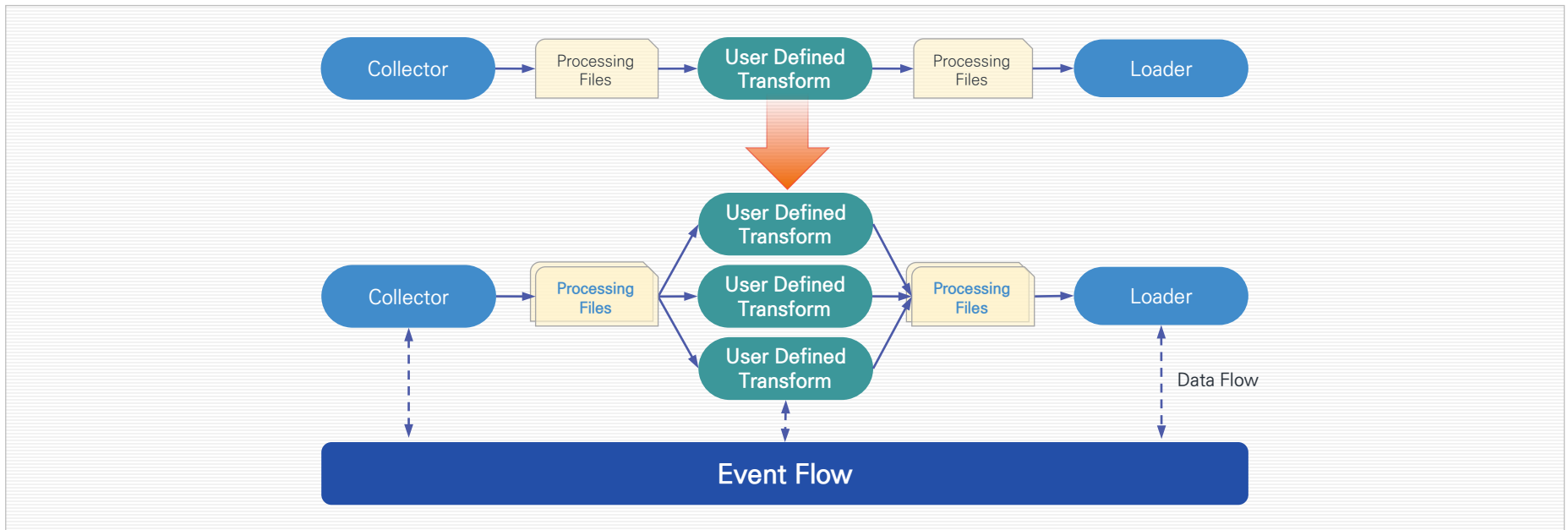
# ETL 프로세스의 Scale-Out 분산 확장

## 프로세스 노드의 I/O 관리

- 시스템 커널 레벨의 표준 스트림 데이터 신호 처리
- MQ(Message Queue) 기반의 ETL 프로세스 이벤트 관리
- ETL 이벤트 처리 상태 저장 및 복구

## 프로세스의 분산 확장

- 다른 프로세스의 출력을 여러 프로세스로 분산 가능
- Sharing mode, Broadcast mode 지원



# ETL 확장 구성 모델

ETL 서버의 프로세스 I/O 관리, 부하 분산 및 데이터 중복 제어를 위한 Flow Control 기반 분산처리

## Sharing Model

**쉐어링 1:N 모델**

Collector 1 → Event Flow (4, 3, 2, 1) → Parser 1 (3), Parser 2 (4)

**쉐어링 N:N 모델**

Collector 1 → Event Flow (2, 1) → Parser 1 (3), Parser 2 (4)  
 Collector 2 → Event Flow (4, 3) → Parser 1 (3), Parser 2 (4)

[예시] Collector(out:1,2,3,4) → (in:1,3)Parser1, (in:4,2)Parser2

- 큐 메시지(처리 이벤트)를 중복 없이 다음 프로세스로 전달
- 같은 큐 내에서 입/출력 순서 보장
- 서로 다른 큐 내에서의 입/출력 순서는 처리 순서에 의해서 결정
- 주로 여러 이벤트를 동일한 여러 작업으로 분산하여 수행하거나 자동 절체 (Fail Over)로 수행할 때 사용

## Broadcasting Model

**브로드캐스트 1:N 모델**

Parser 1 → Event Flow (2, 1) → IRIS Loader (2), HDFS Loader (2)

**브로드캐스트 N:N 모델**

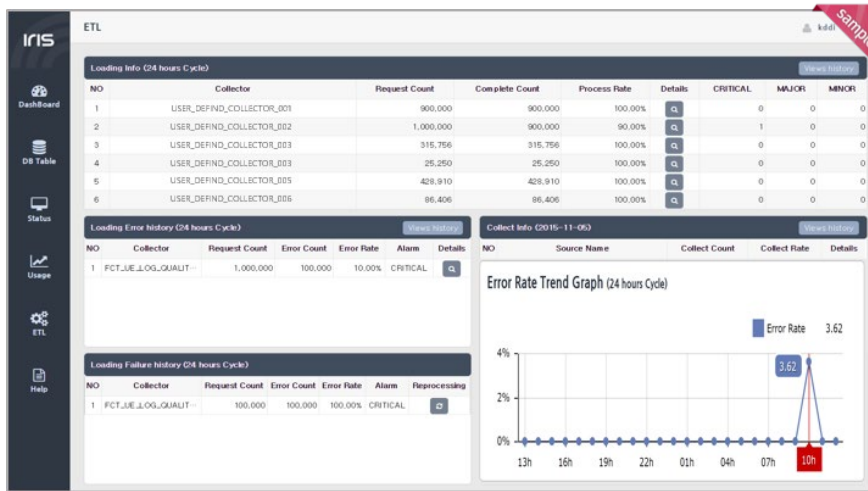
Parser 1 → Event Flow (1) → IRIS Loader (2), HDFS Loader (2)  
 Parser 2 → Event Flow (2) → IRIS Loader (2), HDFS Loader (2)

[예시] Parser(out:1,2) → (in:1,2)IRIS\_Loader, (in:1,2)HDFS\_Loader

- 큐 메시지(처리 이벤트)를 연결된 모든 노드(프로세스)로 중복되어 전달
- 같은 큐 내에서 입/출력 순서 보장
- Sharing 모드와 Broadcast 모드를 혼합하여 사용 가능
- 주로 하나의 이벤트를 종류가 다른 여러 작업으로 나누어 병렬 수행할 때 사용

# ETL 모니터링

## ETL 처리 전 과정의 상태 확인 및 이력 정보 조회 (Web & CLI 지원)



### Collector 별 데이터 적재 상태 확인

- 적재 요청 및 완료 레코드 수
- 수집기 별 진행률
- 적재 실패 알람 정보

### 과거 이력 조회

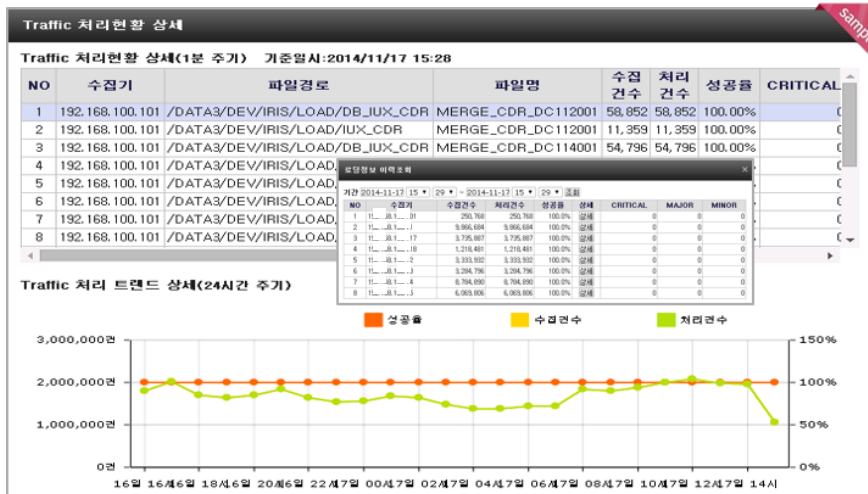
- 기간 별 ETL 처리 결과 이력 조회
- 특정 수집기의 상세정보 확인

### ETL 처리 상태의 트렌드 모니터링

- 시계열 차트로 적재상태 정보 제공
- 요청 건에 대한 처리비율 확인

### 시스템 레벨 모니터링 (CLI)

- 모든 프로세스의 상태 정보 확인
- 수집 파일 목록 / 수집 이력 및 경로 / 수집 파일 상세 정보
- ETL 프로세스의 이벤트 입출력 내역
- 시스템 로그 관리





# IRIS Enterprise DB

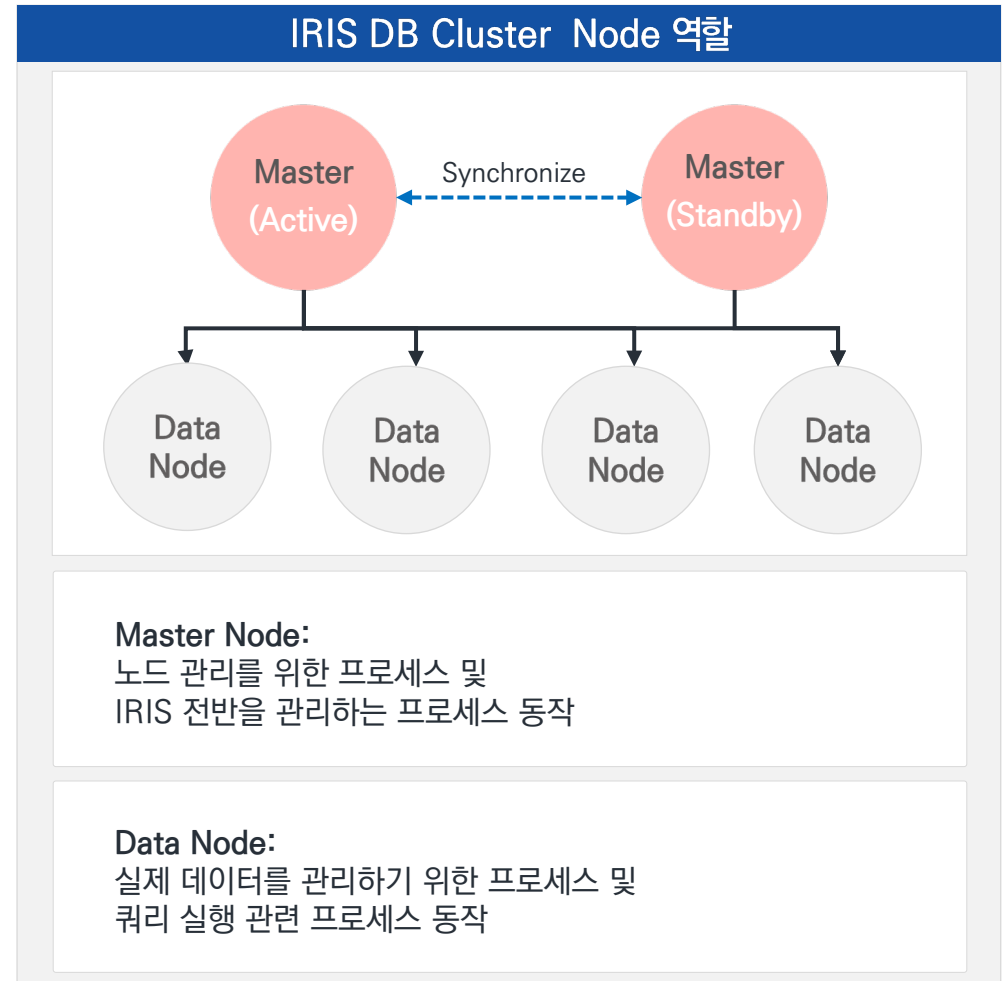
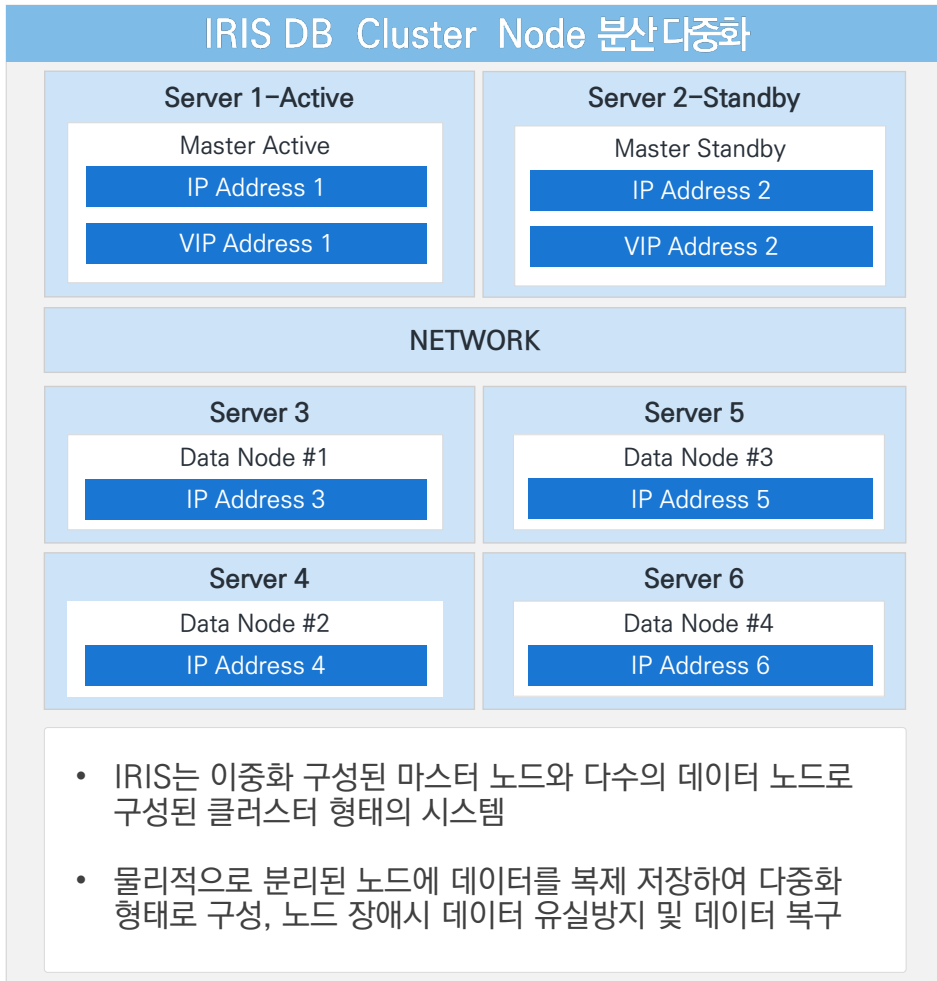
대용량 빅데이터 실시간 처리를 위한 분산 DBMS

A New Experience  
In Big Data Analytics

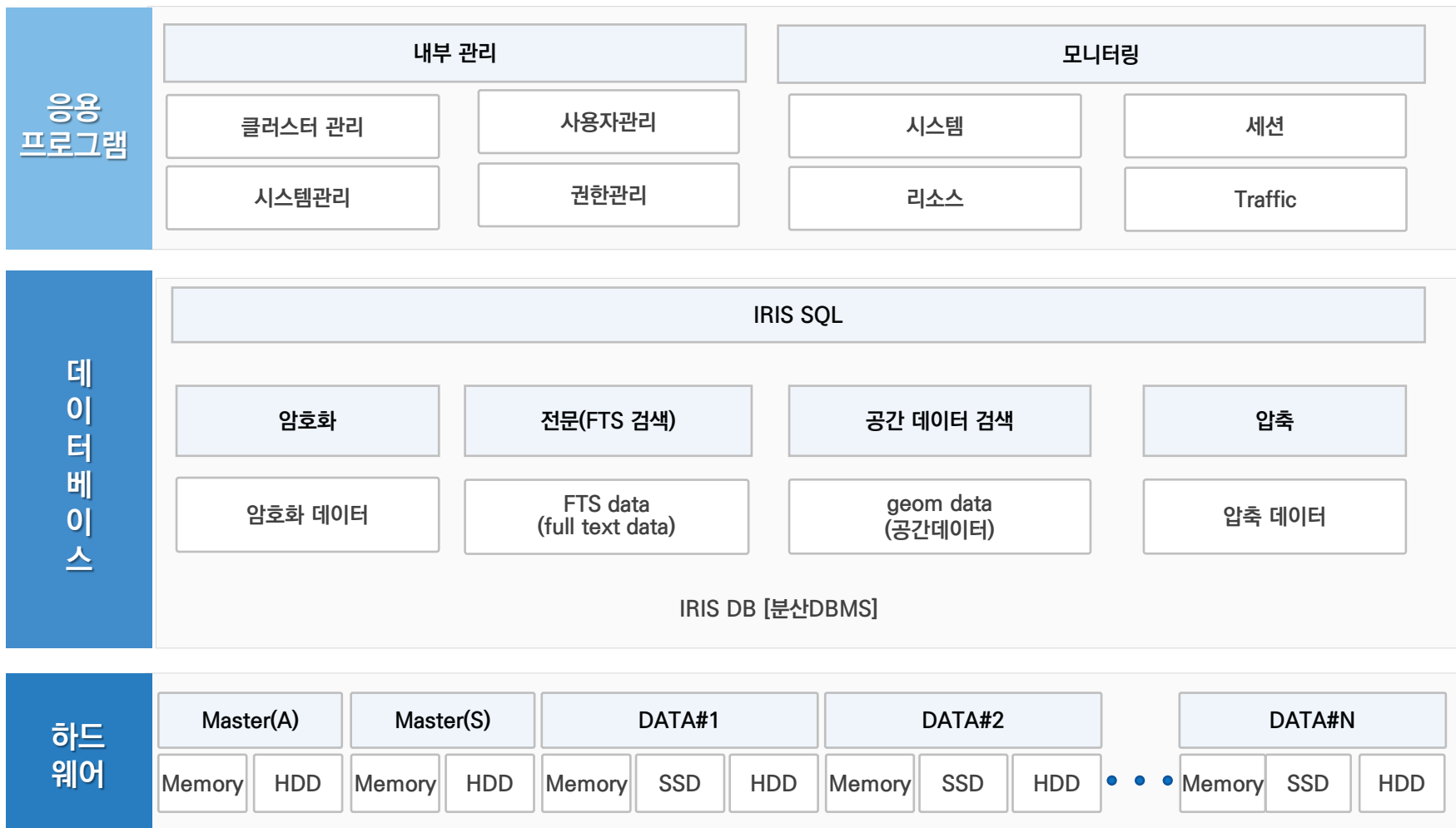


# IRIS Enterprise DB 개요

## 대용량 데이터 처리에 최적화 된 Node 구성



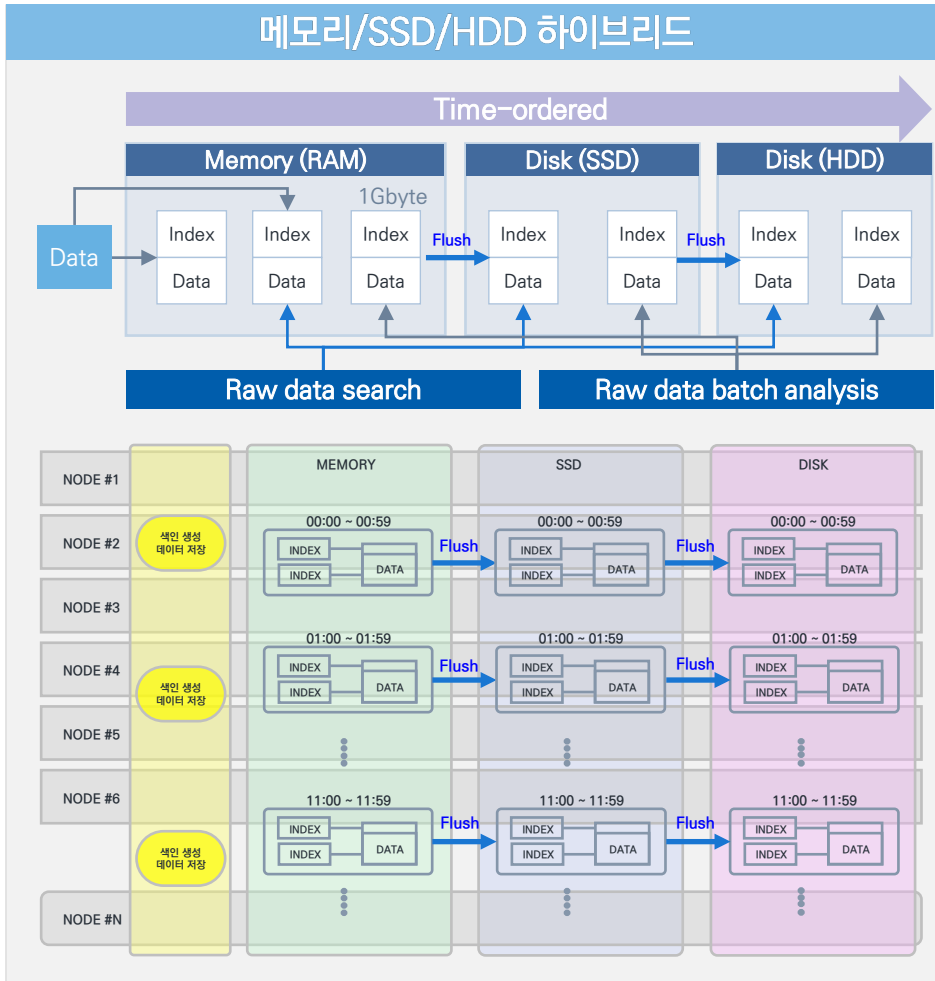
# IRIS Enterprise DB 아키텍처





## 주요특징 - 메모리/SSD/HDD 하이브리드

### 메모리/SSD/HDD 하이브리드 기능을 통한 대용량 실시간 데이터의 고속처리



### 대용량 실시간 데이터 처리

#### 실시간 데이터 처리 절차

- 실시간으로 메모리에서 1GB 단위로 인덱스 와 데이터를 묶어서 1개의 블록 파일로 저장
- 최근 시간 데이터는 메모리에서 Indexing 하고 완료된 데이터는 블록 파일 단위로 SSD, DISK 로 순차적으로 이동 되어 저장

#### 코드단위 실시간 색인 처리

- 고속 검색을 위하여 메모리 상에서 레코드 단위로 인덱스 처리 하며 단일 노드에 저장

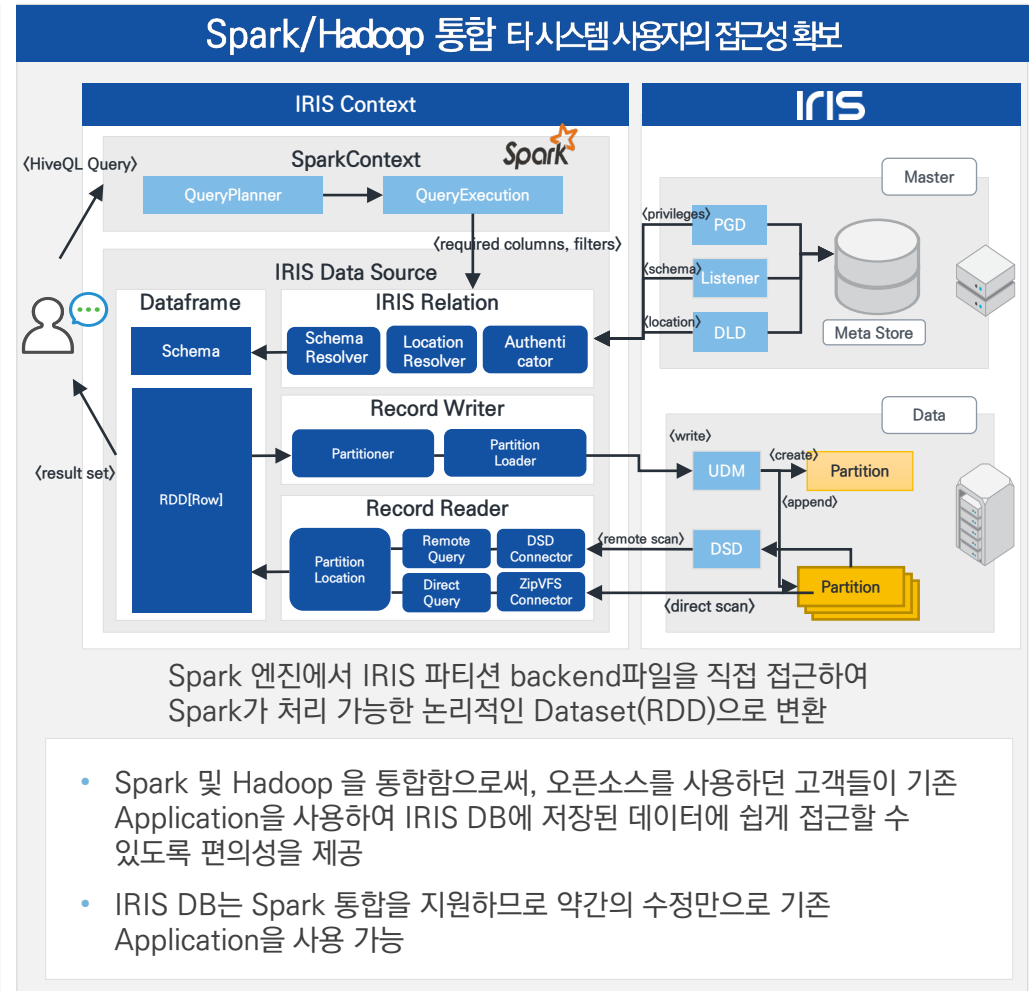
#### in-memory 기반 실시간 데이터 처리

- 각 노드의 메모리 허용 범위 내에서 메모리를 디스크와 동일하게 사용 가능
- I/O 대기시간 최소화, 거의 실시간 검색 성능을 제공
- 1 코어당 초당 최대 1만 레코드 처리, 대형 고객 사이트에서 초당 1백만 ~ 1천만 레코드의 실시간 색인을 성능저하 없이 처리, 충분히 검증되고 안정화된 시스템

- 일일 데이터 처리량 1,000억건, 3개월 8조건의 데이터 처리 클러스터 운영중
- 2PB 이상, 다수의 빅데이터 클러스터 운영중

## 주요특징 - Spark 및 Hadoop 지원 및 대용량 병렬 처리

### 파티션 레벨의 분산 병렬처리로 고속 저장 및 처리, Spark 및 Hadoop 통합으로 사용자 편의성 제공



# 주요특징 - 시스템 관리

IRIS DB는 관리자가 클러스터를 통합 환경을 관리할 수 있는 Web GUI 및 CLI 환경을 제공.

### 시스템 관리 화면 및 CLI

The GUI interface displays various system metrics and node status. Key components include:

- GUI 관리 환경**: Main dashboard with CPU usage graphs and node status indicators.
- 부하가 높은 노드 리스트**: Table listing nodes with high load, including columns for node name, CPU usage, memory, and disk.
- DB 요약정보**: Summary statistics for the database, such as total size (64.1GB) and data size (10.8TB).
- 리소스 사용 현황**: Resource usage charts and tables for CPU, memory, and disk.

---

### iPlus 접속환경[CLI]

```

[iris@master ~]$ iplus test
Password test
Connecting to IRIS(test@127.0.0.1:5050).
Connected to IRIS.
Enter "help" for instructions
iplus> table list
Ret : +OK Success
TABLE_NAME SCOPE RAM_EXP_TIME DSK_EXP_TIME KEY_STRING ...
=====
0 row in set
0.3649 sec
iplus>
    
```

The diagram illustrates the architecture where the **iPlus Client** (consisting of **iPlus**, **iLoader**, and **API**) connects to the **IRIS Cluster**, which includes a **Master Node** and multiple **Data Nodes**.

시스템 관리기능	
GUI 관리기능	세부 내용
통합모니터링	<ul style="list-style-type: none"> <li>각 노드의 자원(CPU, 메모리, 디스크 등) 상태 감시</li> <li>쿼리 에러 발생 노드 감시 세션 상태 감시</li> </ul>
DB 테이블 관리	<ul style="list-style-type: none"> <li>DB 테이블의 정보 조회 및 신규 테이블 생성</li> <li>쿼리 실행</li> </ul>
프로세스 관리	<ul style="list-style-type: none"> <li>각 노드에서 실행되는 프로세스의 동작 관리</li> </ul>
리소스 상태 관리	<ul style="list-style-type: none"> <li>각 노드별 자원의 상태관리(CPU, 메모리, 디스크, 테이블 등)</li> <li>각 노드의 자원 사용률 통계화, 자원 사용률, 히스토리 관리</li> </ul>
세션 관리	<ul style="list-style-type: none"> <li>쿼리 세션 현황 관리, 쿼리 실행 정보 관리</li> </ul>
ETL 상태 관리	<ul style="list-style-type: none"> <li>모비젠 ETL을 운영할 경우, 해당 ETL의 상태 관리</li> </ul>
사용자 관리	<ul style="list-style-type: none"> <li>접속중인 사용자의 접근 권한 등 사용자 정보 관리</li> </ul>
CLI 명령어	세부 내용
iPlus CLI 명령어	<ul style="list-style-type: none"> <li>iPlus CLI를 통해 사용할 수 있는 명령어, 명령/쿼리 실행 기록 등 다양한 명령어 지원</li> </ul>
정보 조회/수정 명령어	<ul style="list-style-type: none"> <li>IRIS DB의 테이블 관리, 세션 관리, 파일 관리, 계정 관리 등을 수행하기 위한 명령어</li> </ul>
시스템 정보 조회 명령어	<ul style="list-style-type: none"> <li>IRIS DB 시스템의 정보(각 노드별 리소스 사용, 테이블 사용 등 관련 통계) 조회를 위한 명령어</li> </ul>
노드 명령어	<ul style="list-style-type: none"> <li>각 노드의 상태 정보와 활성화/비활성화 기능 명령어</li> </ul>
프로세스 명령어	<ul style="list-style-type: none"> <li>각 노드의 다양한 프로세스의 상태를 조회, 각 데몬 프로세스의 시작/종료를 관리</li> </ul>



A New Experience  
In Big Data Analytics

# iS IRIS Analyzer

다양한 분석도구를 제공하는 분석가 편의성 중심 솔루션

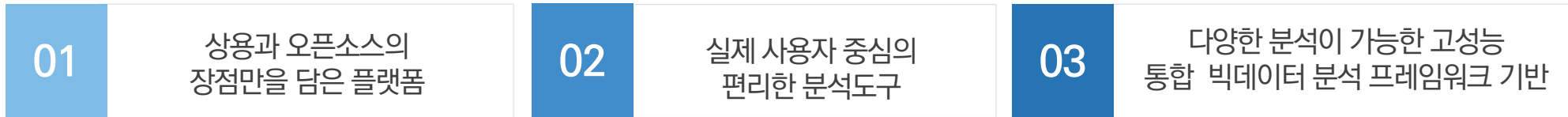


# IRIS Analyzer 개요

## IRIS 개발 배경

대용량 실시간 데이터를 다루는 복잡한 빅데이터 분석 과정과 어려운 분석 도구 사용법이라는 장애에 막혀 고객들은 빅데이터 분석 시스템 도입하고도 활용하기가 어려웠습니다. 그래서 트렌드에 한 발 앞서 지난 수년간 사용자 편의성을 고려한 빅데이터 분석 시스템을 연구하고 개발해 IRIS Analyzer를 출시하였습니다.

## IRIS 개발 방향



### 최신 트렌드 반영

#### 수요자 트렌드

##### 빅데이터 플랫폼 활용도 저하

- 활용 계획없는 빅데이터 플랫폼 도입으로 실제 활용이 어려움
- 오픈소스 기반 빅데이터 플랫폼 도입으로 최적화 및 운영에 어려움, 상용 오픈소스의 높은 OPEX 지출 불가피

##### 데이터 분석가 중심

- 빅데이터 시대 - 데이터 분석가의 역할 중요성 증대
- 데이터 분석가 채용 확대를 통해 빅데이터 분석 업무에 대응
- 다양한 오픈소스 프레임워크를 활용, 데이터 분석가 중심의 분석 도구의 수요가 점차 증가

#### 시장 및 기술 트렌드

##### 분석 도구 트렌드 변화

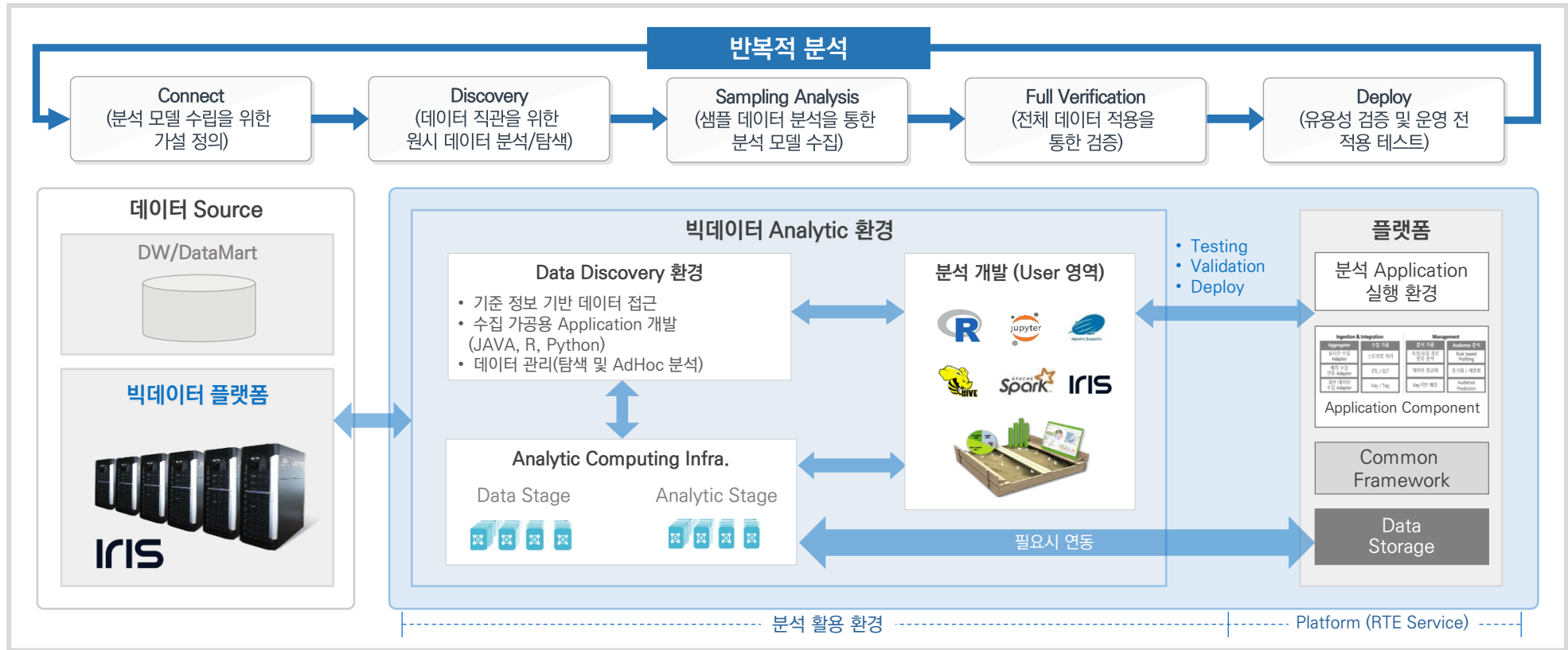
- 분석가나 일반 사용자 대상 셀프서비스 분석 가능한 환경의 분석도구 수요가 증가
- 활용 가능한 라이브러리를 갖춘 오픈소스 분석 도구가 고비용 상용 분석도구를 대체
- 쿼리를 통한 분석 중심 → 인공지능, 머신러닝 고차원 분석으로 변화

##### 빅데이터 기술 트렌드

- 머신, 웨어러블, 센서 등 사물인터넷이 생성하는 대용량 데이터를 처리할 기술과 성능 보장 요구됨
- “빅데이터 분석과 관련된 현상들은 많이 기업이 높은 기대감을 갖고 있는 반면 이를 처리할 역량은 모자란, 낮은 성숙도라 할 수 있다” - Jorgen Heizenberg, Gartner

# IRIS Analyzer - 데이터 분석 Flow 지원

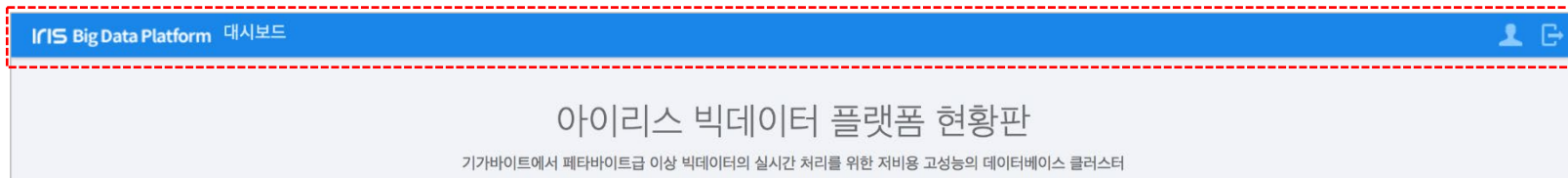
IRIS는 분석가의 실제 업무 환경에 최적화하여 개발된 분석 프로세스를 적용하였습니다.



분석 절차	데이터	분석 데이터 량	도구 사용	업무처리
기존 3~5 단계 분석 절차를 1단계로 간소화 하였습니다.	2주 이상 걸린 분석시간을 2분 이하로 단축하였습니다.	신속한 대용량 데이터 변환이 가능합니다. (PB → GB)	사용자 중심의 오픈소스 및 상용 BI 툴을 통합하였습니다.	개발자 도움 없이 직접 분석 데이터 추출이 가능합니다.

# 주요특징 - 사용자 친화적인 UI / UX로 직관적인 화면 구성

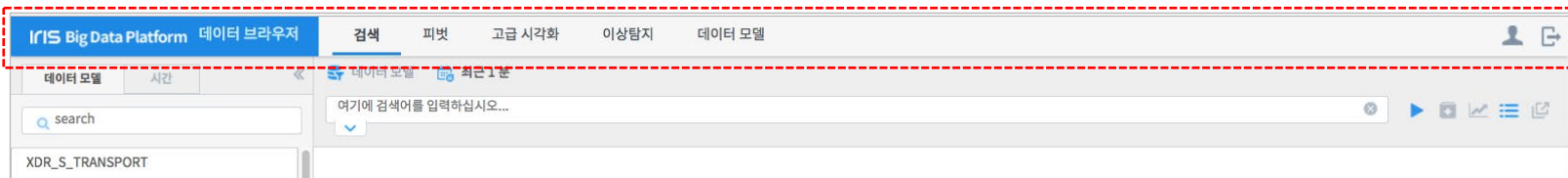
## 1단 메뉴구성



## 어디에서든지 다른 메뉴 이동



## 주요 메뉴 항상 표기



## 주요특징 - 고속 데이터 검색 및 직관적인 시각화 지원

**IRIS Big Data Platform** 데이터 브라우저

데이터 모델    시간

search

- XDR\_S\_TRANSPORT
- XDR\_LTE\_CALL\_KPI
- SYSLOG**
- SYSLOG\_R1MIN
- BATTING
- ATB\_SCHEDULE\_HISTORY

필드    데이터 개수

search

모두 선택    선택 해제

- DATETIME (100+)
- HOST (52)
- FACILITY (10)
- PRIORITY (8)
- LEVEL (8)
- # LEVEL\_INT (8)
- PROGRAM (100+)
- MSG (100+)
- RAW (100+)
- TAG (39)

검색    피벗    고급 시각화    이상탐지    데이터 모델

SYSLOG    오늘

검색 키워드 및 명령어 입력

이벤트(75,790)    통계    시각화

검색결과 데이터 시계열 트렌드 출력

데이터    테이블    목록 개수 20

#	DATETIME	HOST	FACILITY	PRIORITY	LEVEL
> 1	2017/07/04 15:02:52	VPN-5F	authpriv	wa	
> 2	2017/07/04 15:02:51	IVY-01	authpriv	i	
> 3	2017/07/04 15:02:50	IVY-01	authpriv	info	info
> 4	2017/07/04 15:02:50	IVY-01	authpriv	info	info
> 5	2017/07/04 15:02:50	IVY-01	authpriv	info	info
> 6	2017/07/04 15:02:12	VPN-5F	authpriv	warning	warning
> 7	2017/07/04 15:02:12	dn03	authpriv	info	info
> 8	2017/07/04 15:02:10	dn01	authpriv	info	info
> 9	2017/07/04 15:02:09	dn03	authpriv	info	info
> 10	2017/07/04 15:02:09	Sunghoon-Mac-Pro	kern	notice	notice
> 11	2017/07/04 15:02:06	dn01	authpriv	info	info
> 12	2017/07/04 15:02:01	ans02	cron	info	info
> 13	2017/07/04 15:02:01	ans41	cron	info	info
> 14	2017/07/04 15:02:01	ans41	cron	info	info

검색결과 원시데이터 출력



# 주요특징 - 다차원 분석을 위한 웹기반 피벗 기능 및 다양한 피벗 차트 지원

데이터 모델 및 컬럼선택

스프레드시트, 다양한 차트 형태의 피벗팅 결과

피벗 컨트롤 선택

다양한 피벗 차트 지원

IRIS Big Data Platform 데이터 브라우저

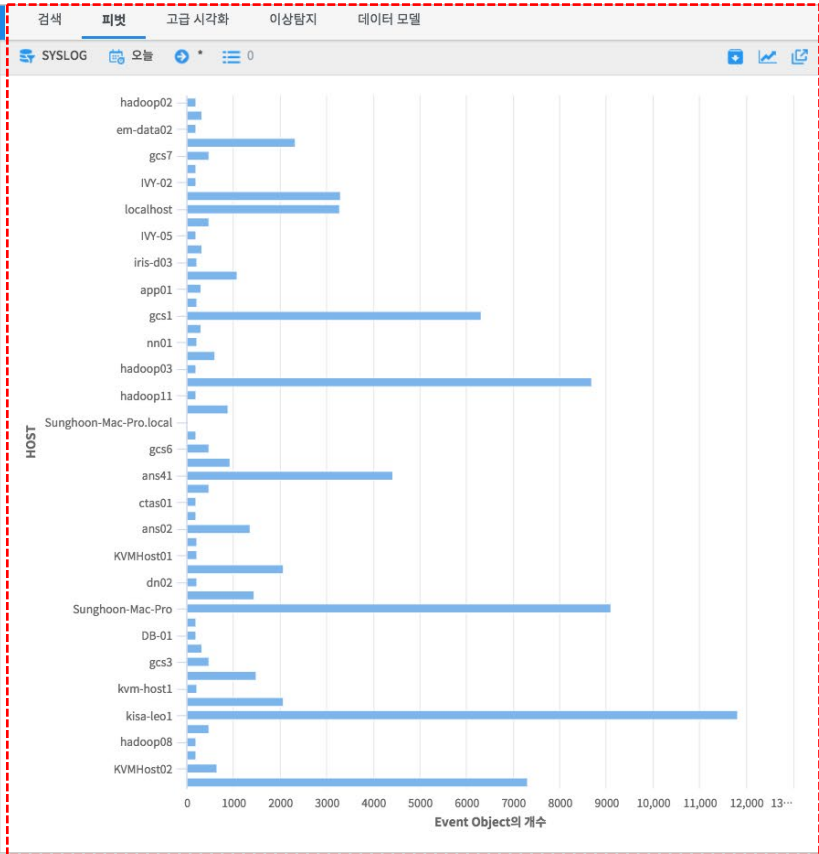
데이터 모델 시간

search

- XDR\_S\_TRANSPORT
- XDR\_LTE\_CALL\_KPI
- SYSLOG**
- SYSLOG\_RIMIN
- BATTING
- ANIMP\_20160223
- QUEUE
- SMTP
- ATB\_SCHEDULE\_HISTORY

필드

- DATETIME
- T RAW
- T HOST
- T FACILITY
- T PRIORITY
- T LEVEL
- # LEVEL\_INT
- T TAG
- T PROGRAM
- T MSG



필터

+ 추가

Y 축(막대)

필드 HOST

레이블 선택사항  표시

정렬 기본값

최대 막대 수 500

X 축(막대 너비)

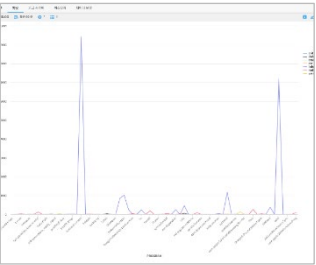
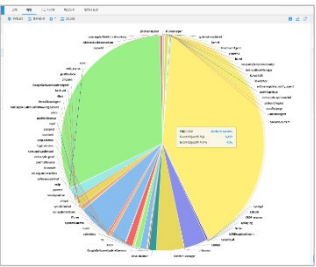
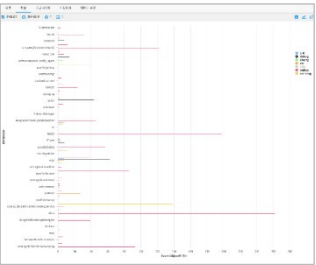
필드 Event Object의 개수

레이블 선택사항  표시

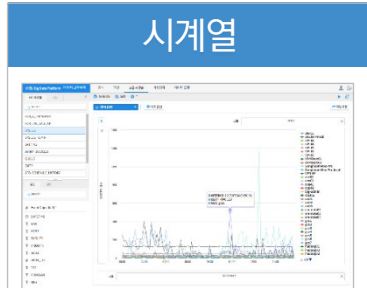
색상(범례)

선택

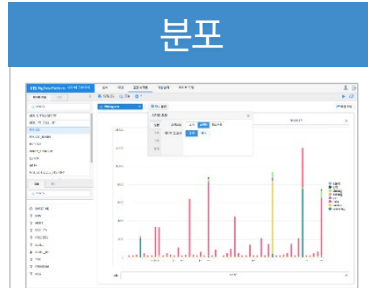
▶ 실행



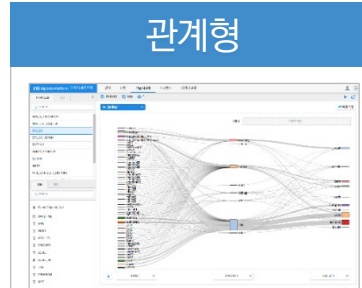
## 주요특징 - 고급 시각화를 위한 다양한 종류의 차트 지원



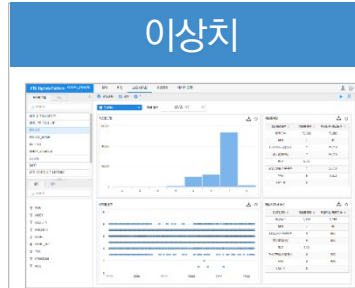
Line Plot



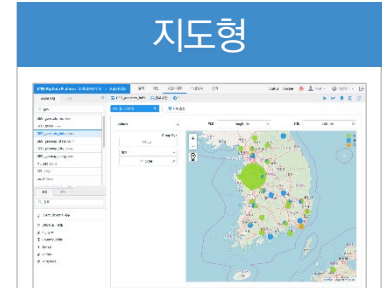
Histogram



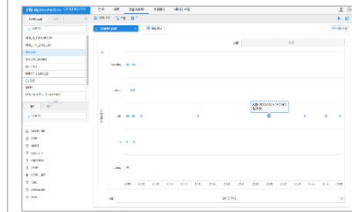
Sankey



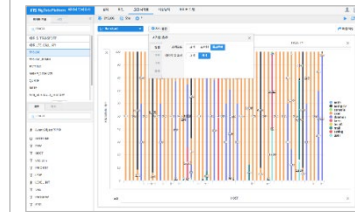
Outlier



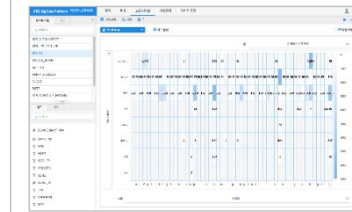
클러스터맵



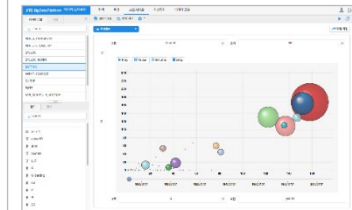
Scatter Plot



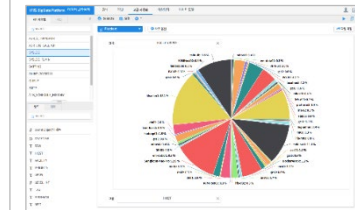
Bar Chart



Heatmap



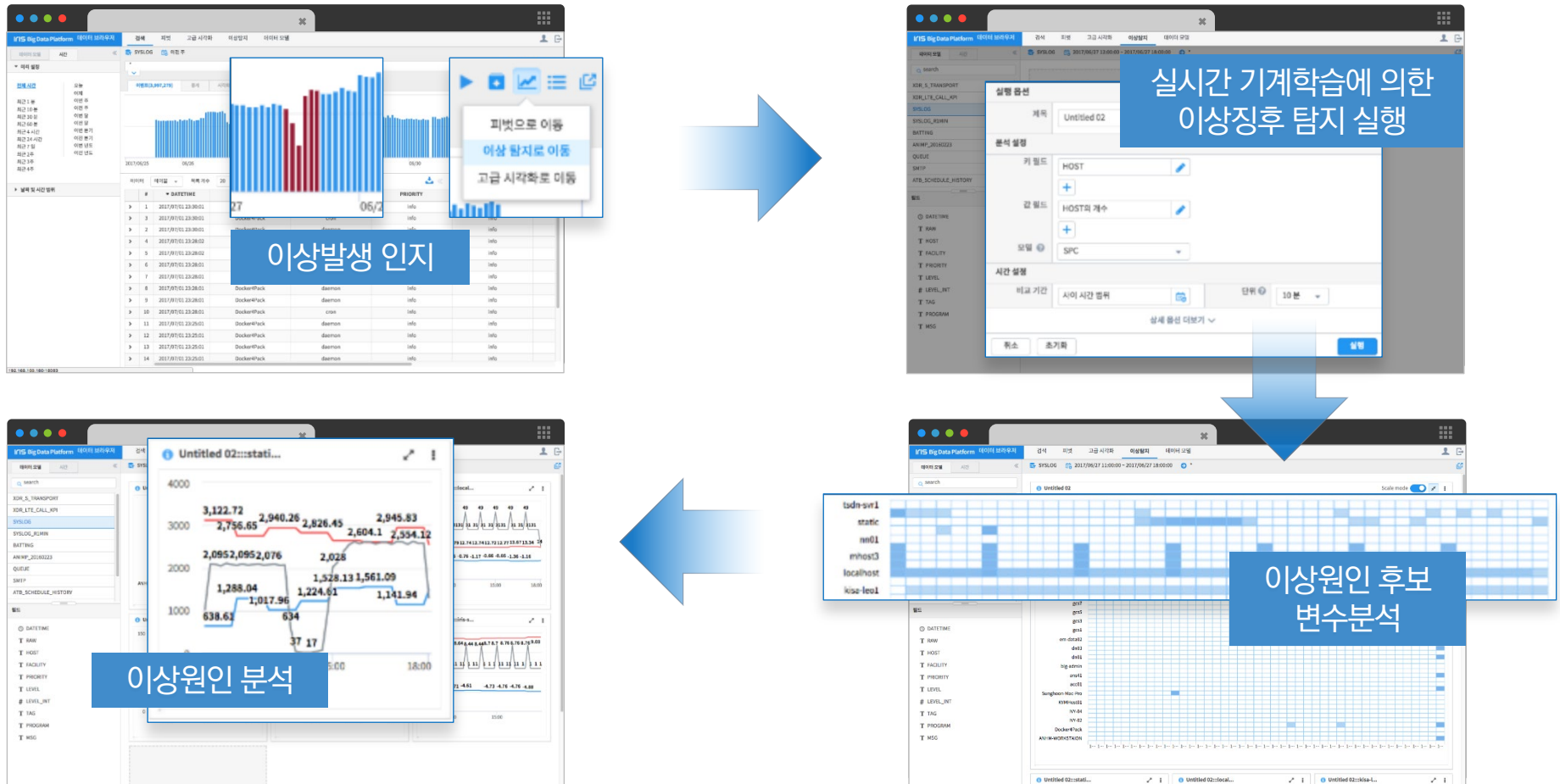
Motion Chart



Pie Chart

## 주요특징 - 이상징후 탐지 기반 고급분석 Flow

내장된 실시간 기계학습을 활용한 이상 징후 탐지 및 분석은 Flow에 따라 사용자가 쉽고 빠르게 이상원인에 대한 분석이 가능하도록 User-friendly UI로 구성되어 있음





A New Experience  
In Big Data Analytics

# **IRIS** Studio

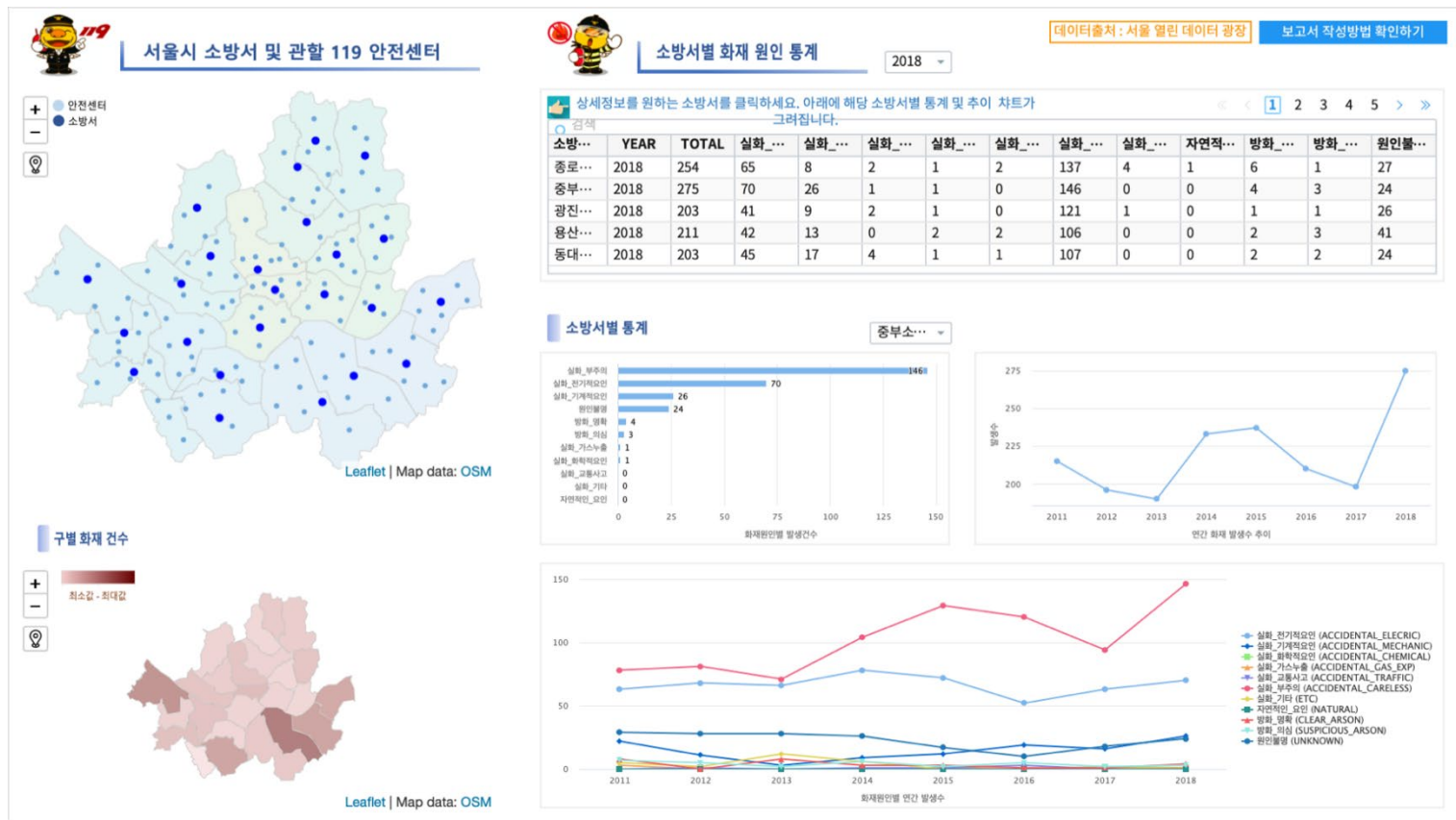
데이터 분석 화면 및 보고서 제작이 가능한  
DIY(Do it yourself) 분석 화면/보고서 제작 도구



# IRIS Studio 개요

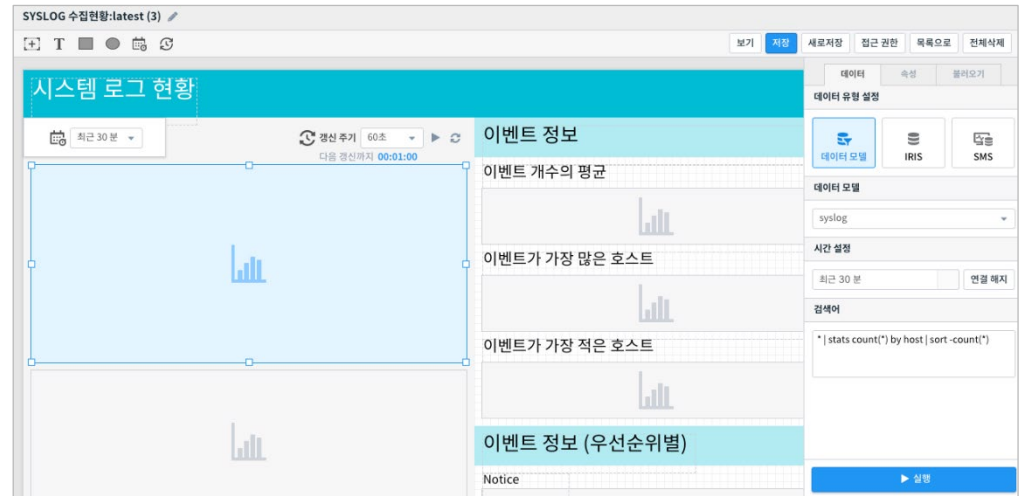
## IRIS Studio 제품 개요

IRIS Studio는 다양한 데이터 시각화 결과를 웹 문서로 작성할 수 있는 도구를 제공합니다. 분석 결과를 보고서로 작성할 수 있으며 해당 분석기능을 활용하여 데이터소스 및 데이터 탐색 범위, 보고서 갱신 주기 등을 동적으로 변경할 수 있는 Dash Board 디스플레이로서 활용할 수 있습니다.



## 주요특징 - GUI기반 대시보드 편집기를 통한 자유로운 대시보드 생성

- 영역, 텍스트, 도형 등 캔버스 내 자유로운 컴포넌트 배치, 크기 조절 기능
- 데이터를 차트/그리드로 출력하고 옵션 기능으로 커스터마이징
- 검색/피벗/고급시각화 에서 저장한 시각화 객체 불러오기
- 시간 설정 컴포넌트를 통해 일괄적으로 시간 설정 및 변경
- 주기 설정 컴포넌트를 통해 지정한 시간 주기로 데이터 요청/수신



시각화 유형

이름	유형
TCP_ACCEPT_특정IP_조회결과	피벗
udp rcv bar chart	피벗
udp send bar chart	피벗
file io bar chart	피벗
tcp accept bar chart	피벗
tcp connect bar chart	피벗
TCP_CONNECT_DAY_TREND	고급 시각화

저장된 시각화

이름	유형
TCP_ACCEPT_특정IP_조회결과	피벗
udp rcv bar chart	피벗
udp send bar chart	피벗
file io bar chart	피벗
tcp accept bar chart	피벗
tcp connect bar chart	피벗
TCP_CONNECT_DAY_TREND	고급 시각화

시간 설정

사용자 정의  사용자 지정

선택된 시간

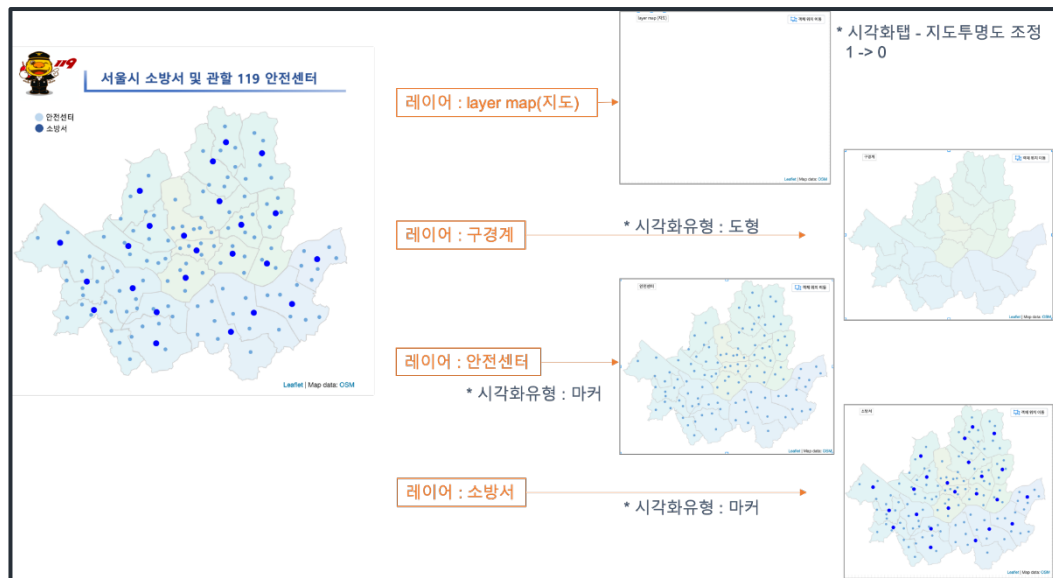
- 최근 10 분
- 최근 30 분
- 최근 60 분
- 최근 4 시간
- 최근 24 시간

적용 대상

- report-component-box53
- report-component-box56
- report-component-box55
- report-component-box58
- report-component-box66
- report-component-box83
- report-component-box62

# 주요특징 - 지도 기반의 데이터 시각화 표현

Map View 기능을 통해 지도상에서 데이터를 다층의 시각화로 표현 가능



IRIS Studio를 활용한 Map View 시각화 기능

메뉴바에서 지도 클릭 한후 캔버스에 드로우 합니다.

**1** 레이어 선택

**2** 3개의 레이어 추가

**3** 구경계 레이어 선택

**4** 구경계 레이어 데이터는 데이터모델 SEOUL\_GU\_WGS84에서 가져옵니다.

**5** 자동 실행

**6** 실행

실행버튼을 눌러 데이터를 가져옵니다.

순서	이름	표시	삭제
1	소방서	<input type="checkbox"/>	<input type="checkbox"/>
2	안전센터	<input type="checkbox"/>	<input type="checkbox"/>
3	구경계	<input type="checkbox"/>	<input type="checkbox"/>
4	layer map (지도)	<input checked="" type="checkbox"/>	<input type="checkbox"/>



# IRIS Discovery Service

공통 명령어로 데이터 검색부터 분석까지 해주는  
Analysis Middleware 서비스

A New Experience  
In Big Data Analytics

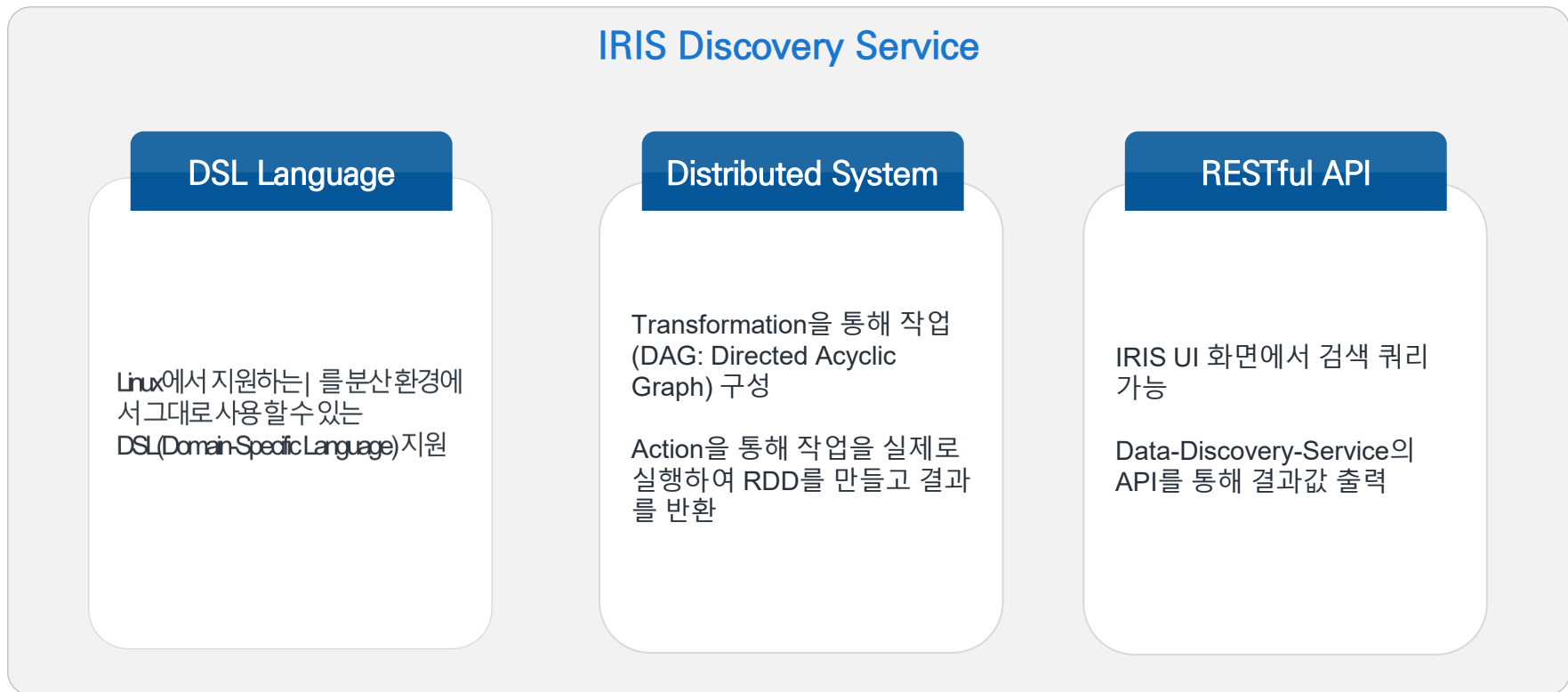




# IRIS Discovery Service 개요

## IRIS Discovery Service 제품 개요

IRIS Discovery Service는 분석 미들웨어(Analysis Middleware)를 제공하는 서비스로서 여러 데이터소스에 있는 데이터를 사용자가 동일한 명령어로 검색 및 분석을 할 수 있는 인프라 환경을 제공합니다.

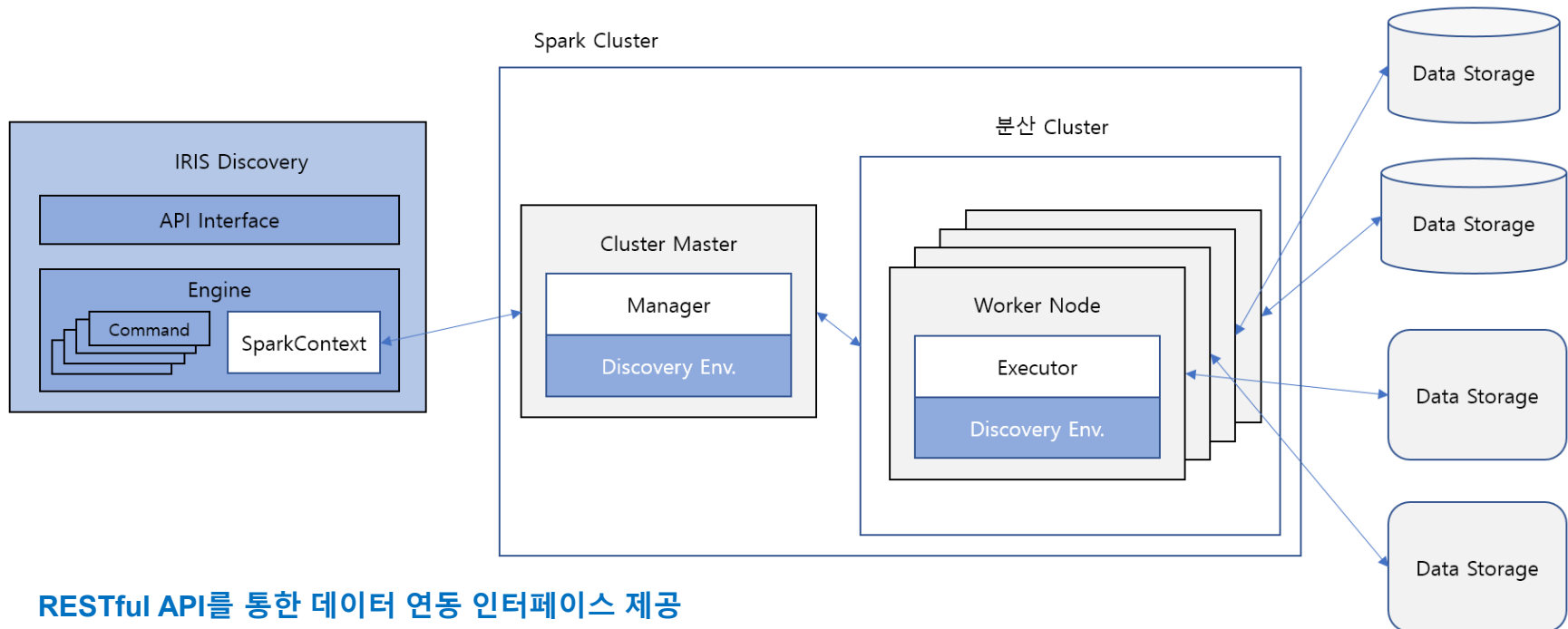


# IRIS Discovery Service 아키텍처

## IRIS Discovery Service Architecture

### 분산환경에서 데이터 처리 지원

분산구조에서 대량의 데이터에 대한 실시간 검색 및 분석이 가능한 환경 제공



### RESTful API를 통한 데이터 연동 인터페이스 제공

- RESTful API를 통해 다양한 사용자 애플리케이션 개발이 용이
- RESTful API기반의 분석도구를 통해 분석가가 직접 데이터에 접근하여 검색 및 분석 가능



# IRIS 적용 기대효과 및 분야

A New Experience  
In Big Data Analytics

## IRIS 적용 분야

### 적용 분야



### 대용량 로그 데이터의 실시간 분석

- 통신 서비스 품질의 실시간 모니터링, 대규모 IT 시스템의 통합 관리
- 개별 고객별 패턴 분석 등 다양한 형태의 대용량 로그 데이터 기반의 실시간 분석
- 실시간 로그 분석을 통한 네트워크 품질 예측 및 사전 대응



### 금융/보안 데이터 분석

- 개인정보 유출에 대한 실시간 모니터링
- 불법 거래 (fraud) 탐지 등 실시간 금융 보안
- 사이버 위협 정보 및 이상 징후 실시간 데이터 처리



### IoT 센서데이터 분석

- 센서 데이터 기반 시계열 데이터의 실시간 분석  
(예시: 스마트 팩토리, 제조 품질 관리, 부품 고장 예측 등)

## IRIS 기대효과

### 도입 효과



### 분석 업무 생산성 향상

- 분석가의 실질적인 업무 중심의 직관적인 UI 구성 및 분석
- 도구간의 자유로운 데이터 공유로 분석 업무 시간 단축으로 생산성 향상에 기여



### 획기적인 TCO 절감

- 개발 요소 최소화, 분석 업무 시간의 획기적인 단축
- 유지 관리 비용의 절감을 통한 전반적인 TCO 절감 효과 기대



### Self-service 분석 환경 제공

- 분석가가 개발자의 도움 없이 빅데이터 저장소 (Data Lake) 로부터 자유로운 데이터 검색 및 분석 업무 수행 가능 (예시: 데이터 브라우징)



A New Experience  
In Big Data Analytics

**iS IRIS** 주요 레퍼런스



# KOTRA 무역·투자 빅데이터 플랫폼 구축 (1/2)

## KOTRA\_무역.투자 빅데이터 플랫폼 구축



## KOTRA 무역·투자 빅데이터 플랫폼 구축 (2/2)

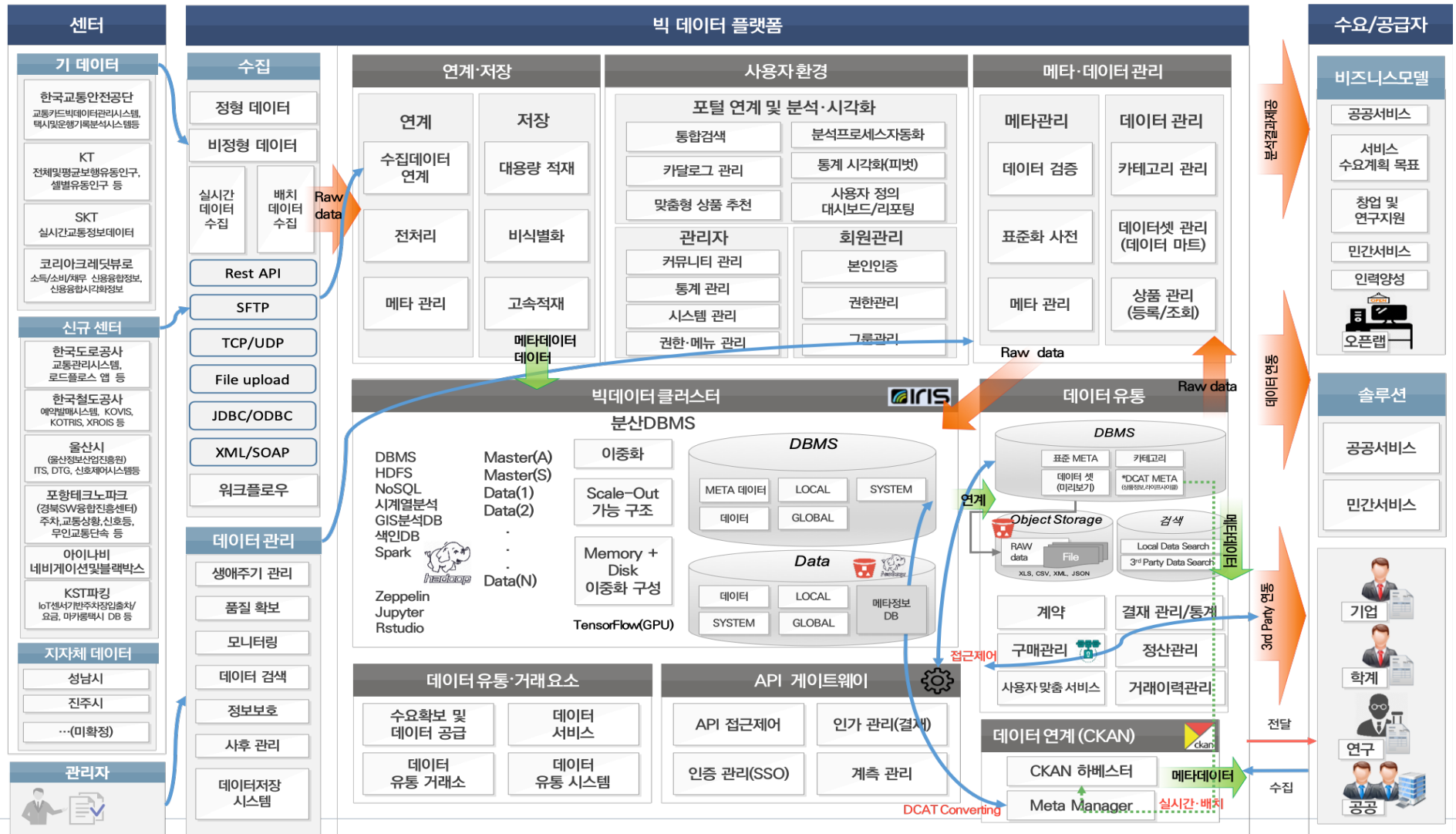
### KOTRA\_무역.투자 빅데이터 플랫폼 구축\_사업범위





# 한국교통연구원 국가 교통 빅데이터 플랫폼 개발(1/2)

## 한국교통연구원\_국가 교통 빅데이터 플랫폼 개발



# 한국교통연구원 국가 교통 빅데이터 플랫폼 개발(2/2)

## 한국교통연구원\_ 국가 교통 빅데이터 플랫폼 개발\_사업범위



### 빅데이터 플랫폼 인프라 구축

- 데이터 연계 및 수집 저장 체계 수립 및 구현
- 빅데이터 플랫폼 운영 관리 기능 개발
- 데이터 처리 및 가공 기능 시스템화
- 데이터 분석을 위한 오픈소스 기반 분석 환경 구축

### 내·외부 데이터 연계·수집·저장

- 데이터 종류 및 특성에 따른 데이터 수집
- 수집 스케줄러에 의한 자동화 수집 및 제어
- 수집 데이터의 분리 및 전처리
- 적재 데이터의 상태, 유효성 관리
- 데이터 보호를 위한 비식별화 조치
- 데이터 사용자 접근 통제

### 표준화 품질관리 체계 구축

- 데이터 표준화 체계 수립
- 개방형 데이터 활용을 위한 DCAT기반 데이터 활용
- 빅데이터 플랫폼 기반 데이터 개방 및 활용 정보 연계
- 데이터 품질관리 요소 도출 및 관리



### 데이터 유통 시스템 구축

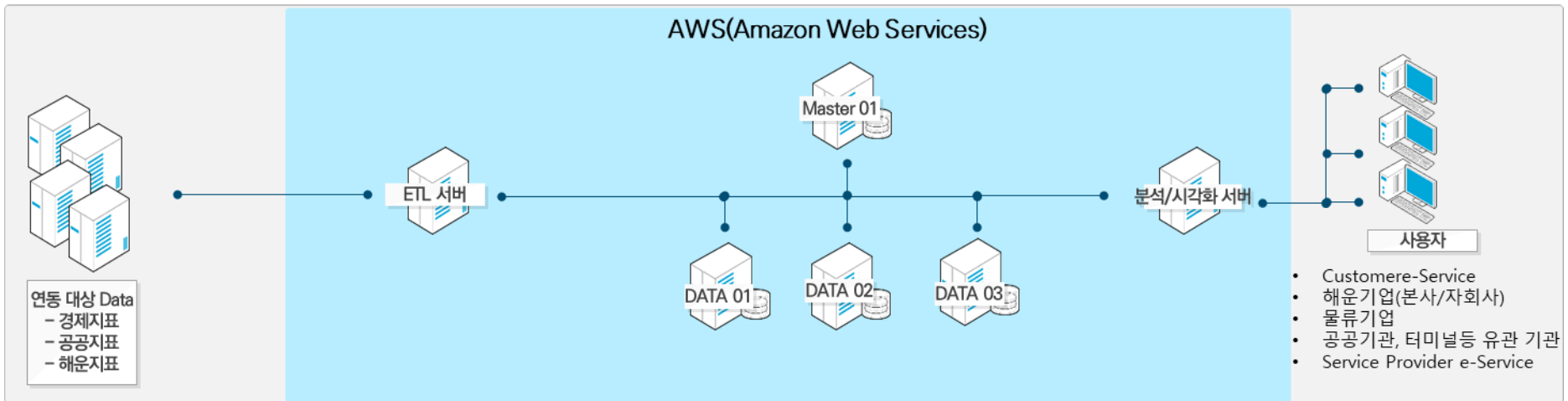
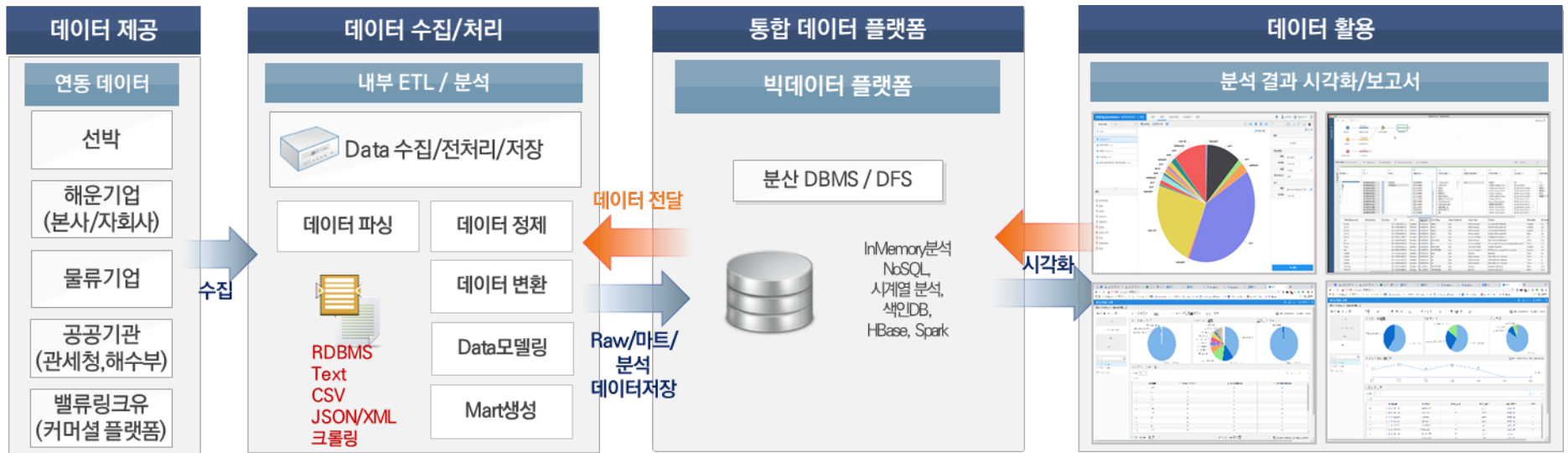
- CKAN 및 DCAT 기반의 플랫폼
- 상품 데이터의 사용자 검색 및 하베스팅 연동
- 데이터 거래를 위한 상품 데이터 관리 체계 구축
- 데이터 카탈로그, 검색 환경, 데이터 관리 환경 제공
- 공급자, 관리자, 소비자용 포털 기능 개발

### 민간 클라우드 활용

- 공공 클라우드 기준을 만족하는 유통 시스템 개발
- 클라우드 서비스 활용 네트워크 적용
- DMZ 및 Private 망 분리를 통한 시스템 아키텍처링
- 온프레미스 시스템과 데이터 운영 연계 최적화

# 밸류링크유 해운·물류 빅데이터 분석 플랫폼 개발(1/2)

## 밸류링크유\_해운·물류 빅데이터 분석 플랫폼 구축



# 밸류링크유 해운·물류 빅데이터 분석 플랫폼 개발(2/2)

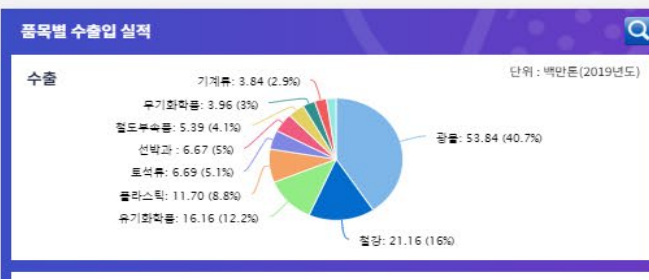
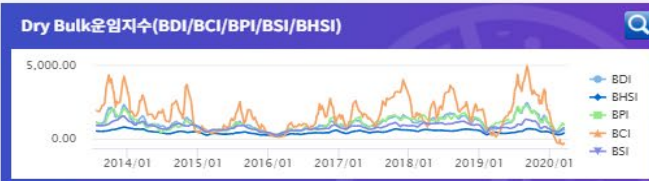
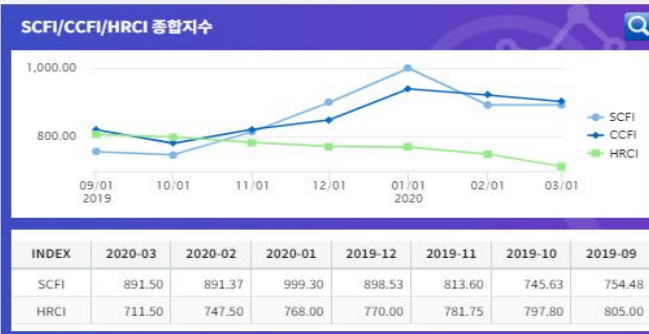
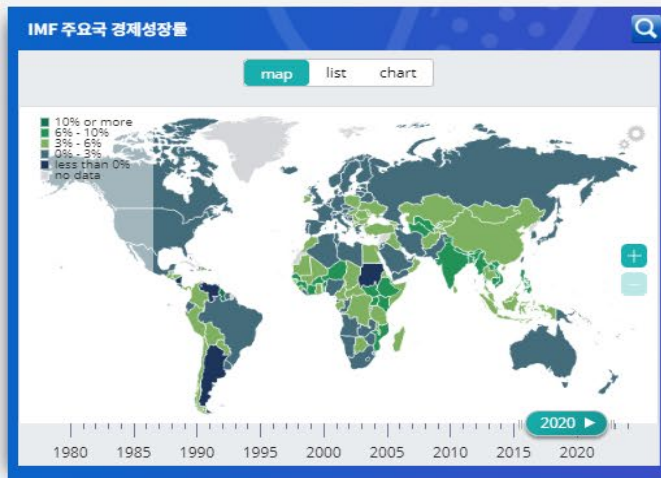
## 밸류링크유\_해운·물류 빅데이터 분석 플랫폼 구축\_ 주요 경제지표 및 해운지표

### 주요국 환율

국가/원화	2020-03-17	2020-03-16	2020-03-13	2020-03-12	2020-03-11
미국 USD	1,263.73	1,249.49	1,239.31	1,228.12	1,212.35
유로 EUR	1,398.20	1,398.72	1,386.40	1,383.30	1,376.40
일본 JPY (100)	1,180.67	1,179.82	1,166.57	1,182.65	1,154.73
중국 CNY	185.19	183.70	182.63	180.67	179.90

### Libor 금리

기간	2020-03	2020-02	2020-01	2019-12	2019-11	2019-10	2019-09
3개월	0.98	1.69	1.83	1.91	1.90	1.98	2.13
6개월	0.94	1.69	1.85	1.90	1.91	1.96	2.05
현은 기준...		1.25	1.25	1.25	1.25	1.25	1.50



### 국내 주요 경제 지표

INDEX	2020-03	2020-02	2020-01	2019-12
경기선행지수				99.56
BSI(실적)		65.00	76.00	74.00
BSI(전망)	69.00	77.00	73.00	71.00
소비자 심리지수		96.90	104.20	100.40

### 국제 유가

유종	2020-03-17	2020-03-16	2020-03-13	2020-03-12	2020-03-11
Brent	28.73	30.05	33.85	33.22	35.79
Dubai	30.83	31.49	33.93	32.69	34.58
WTI	26.95	28.70	31.73	31.50	32.98

### 자유항 연료유가

Oil/Date	2020-03-17	2020-03-16	2020-03-13	2020-03-12	2020-03-11
VLSFO	303.50	304.00	334.00	329.00	357.00
IFO380	210.50	208.00	215.00	212.50	228.00
Diff	93.00	96.00	119.00	116.50	129.00



# 수원시청 스마트 시민소통 정책결정시스템 공공 빅데이터 플랫폼 구축(1/2)

## 수원시청\_스마트 시민소통 정책결정시스템 공공 빅데이터 플랫폼 구축



## 수원시청 스마트 시민소통 정책결정시스템 공공 빅데이터 플랫폼 구축(2/2)

수원시청\_스마트 시민소통 정책결정시스템 공공 빅데이터 플랫폼 구축\_사업범위



데이터를 통한 스마트 행정과 시민 협치 시정 구현을 위한

수원시 공공 데이터 수집·관리부터 개방·활용까지 포함하는 사업입니다.

“ 분산된 시정정보 연계·수집 및 활용 빅데이터 플랫폼 구축 ”

01



데이터 관리 체계  
구축  
[메타데이터/데이터맵]

02



공공 빅데이터  
플랫폼구축  
[수집·연계·저장·분석]

03



시정데이터  
활용 체계 구축  
[데이터 개방 포털 구축]

04

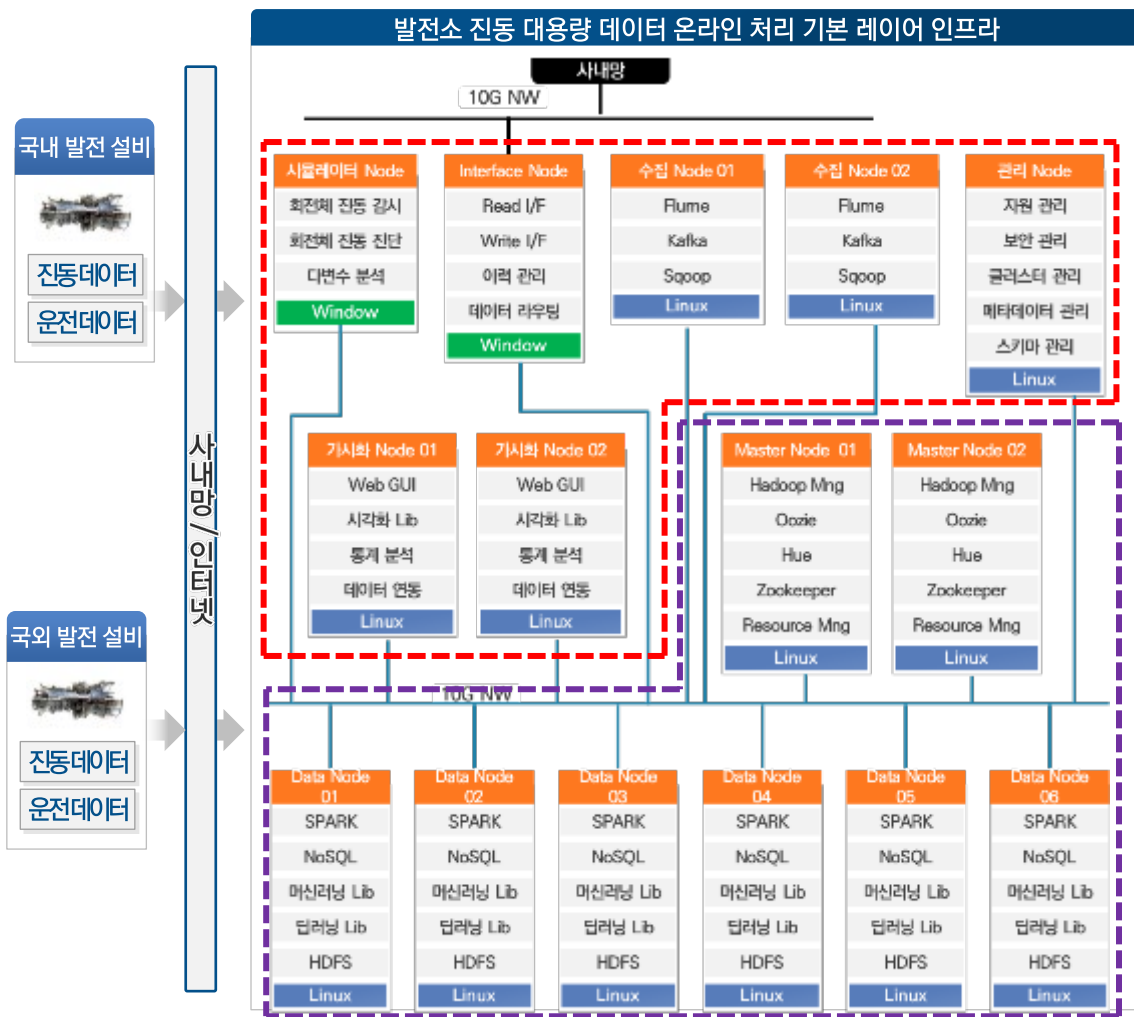


스마트 시민소통  
정책결정시스템  
[데이터 분석 시각화]

데이터 기반 과학적 행정구현 및 데이터 개방·활용으로  
스마트 행정 및 시민 중심 정책결정 지원

# 발전 설비 진동 데이터 분석

발전설비 운용 비용 절감 및 효율성 향상을 위해, 발전설비 데이터 수집 및 분석플랫폼을 구축하여, 발전설비 진동 빅데이터 수집 및 공유 분석 결과의 이해도를 높여주는 다양한 시각화 라이브러리 및 발전 설비 운영 기술 R&D에 필요한 다양한 분석 알고리즘과 대용량 진동 데이터의 시뮬레이션 기능 등을 제공 하는 연구 인프라 구축 및 운영



**01**

빅데이터 H/W 서버 도입 및 시스템 구축

**02**

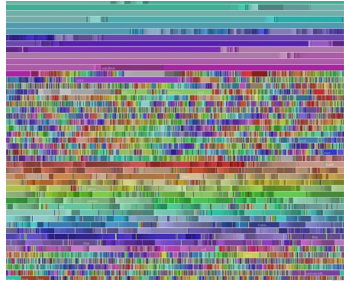
빅데이터 플랫폼 S/W 도입 및 시스템 적용

**03** 빅데이터 기술 기반 진동데이터 연구인프라

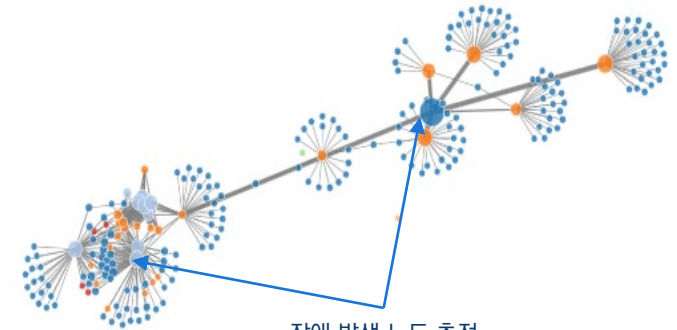
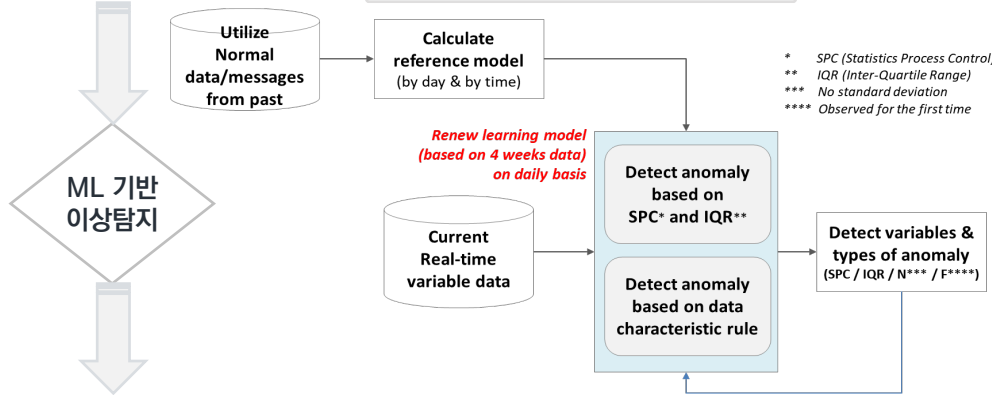
- 진동 빅데이터 인터페이스
- 진동 빅데이터 시뮬레이터
- 진동 빅데이터 분석/시각화
- 진동 빅데이터 수집
- 진동 빅데이터 저장.처리
- 인프라 관리

# 이동통신 로그데이터 기반 망관제 및 서비스 품질 분석

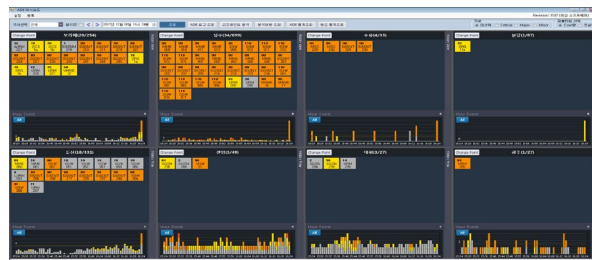
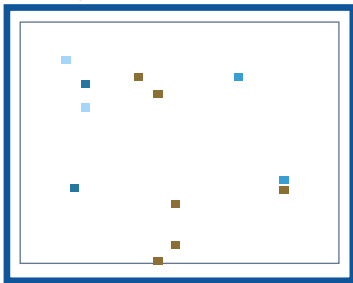
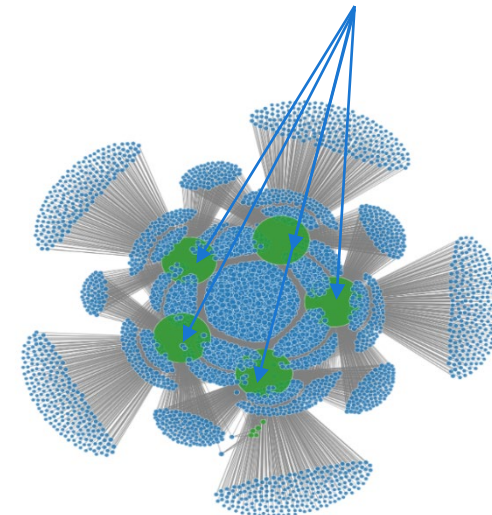
3G-LTE-5G 이동통신망 장비로 부터 입력되는 로그 메시지(Fault, Alarm, Status 등)를 빅데이터 처리하고, ML 기반 이상탐지 알고리즘 적용으로 운영자가 인지하기 힘든 이상(장애) 상황을 감시하고 원인분석(RCA<sup>Root Cause Analysis</sup>)



10초 주기 1,200만건 로그  
※ 60만 이동통신장비 × 200개 Matrix



장애 발생 노드 추적 (RCA 시각화)



빅데이터에서 의미 있는 장비와 Matrix 이상탐지 알람 발생



# WiFi-AP 측위 보정 (1/2)

## ◆ 이상 AP 분석 (k-means)

AP별로 일정기간 동안 데이터가 수집된 위치를 Clustering 하여 AP의 이동 여부를 분석  
수집된 전체 AP를 대상으로 일 1회 분석 하고 이동한 AP의 bssid, 이동한 최초 시점을 분석 결과로 제공  
K-means Clustering 활용 하며 분석 결과 History 및 조회 기능 합니다.

- K-means 라이브러리: Spark ML-lib
- 위경도 스케일 공식

## ◆ Outlier 분석 알고리즘

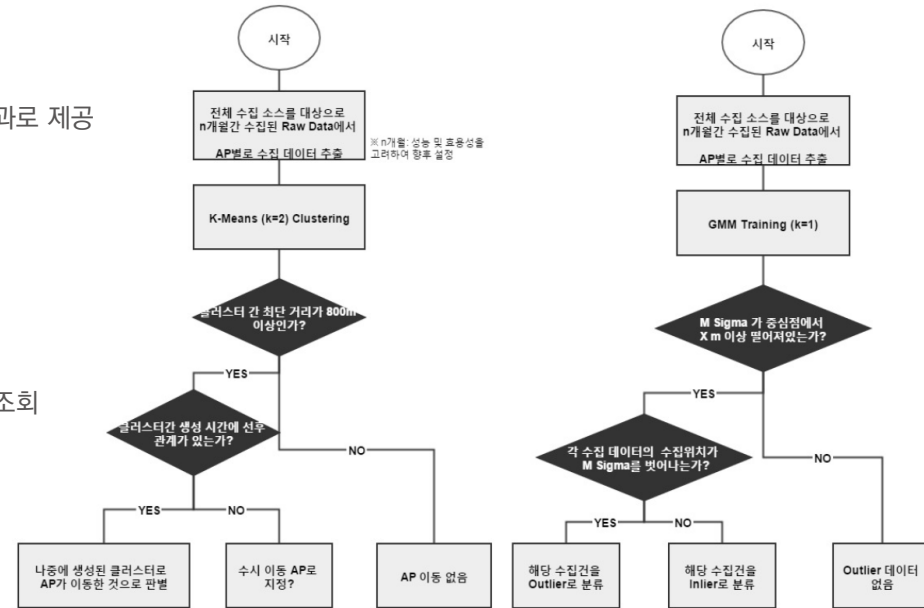
AP별로 일정기간 동안 데이터가 수집된 위치를 기준으로 분석하여 Outlier 데이터를 판별  
수집된 전체 AP를 대상으로 일 1회 분석 하고 Outlier 데이터의 bssid, 수집 시점을 분석 결과 조회

- Gaussian Mixture Model 라이브러리: Spark ML-lib
- 위경도 스케일 공식
- Threshold (3-sigma & NLL)

## ◆ Pairwise 분석 알고리즘

AP별로 동일 단말에서 함께 수집되는 AP 목록을 1회 동시 수집될 때 마다 Scoring 하여 동시 수집 AP Set이 크게 달라지는 시점을 분석하여 AP의 이동 여부를 판별  
수집된 전체 AP를 대상으로 일 1회 분석 하여 이동한 AP의 bssid, 이동한 최초 시점을 분석하고 분석 결과 History 및 조회 기능 제공

- 위경도 스케일 공식
- Threshold (3-sigma & NLL)
- 두 개의 클러스터간 거리 구하는 라이브러리



### Input Scaling 및 reconstruction 방법

위도 / 경도 동일 (= X)

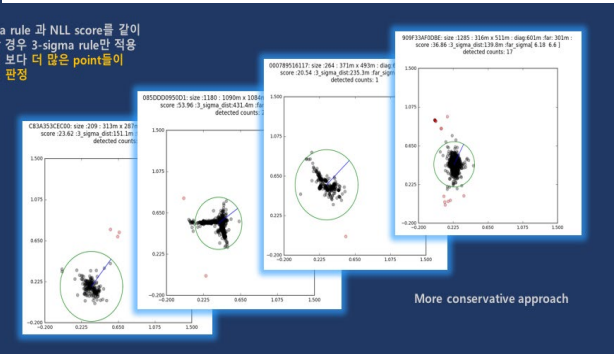
	lat	lng
count	2593379	2593379
mean	36.66811	127.5379
std	1.93962	0.9264924
max	33.8272	126.3273
25%	35.84009	126.9761
50%	37.35043	127.0868
75%	37.54553	128.5487
max	38.38266	130.8997

AP 별

$$X_{scaled} = 100,000(X - \min(X))/diag$$

$$X = X_{scaled} * diag / 100,000 + \min(X)$$

	Scaled lat	Scaled lng
count	2593379	2593379
mean	0.368624	0.9358223
std	0.2447242	0.2955577
min	0	0
25%	0.1533307	0.09803364
50%	0.31063	0.2882716
75%	0.5825542	0.5486149
max	1.3	1.2

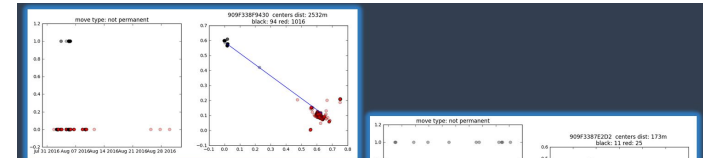
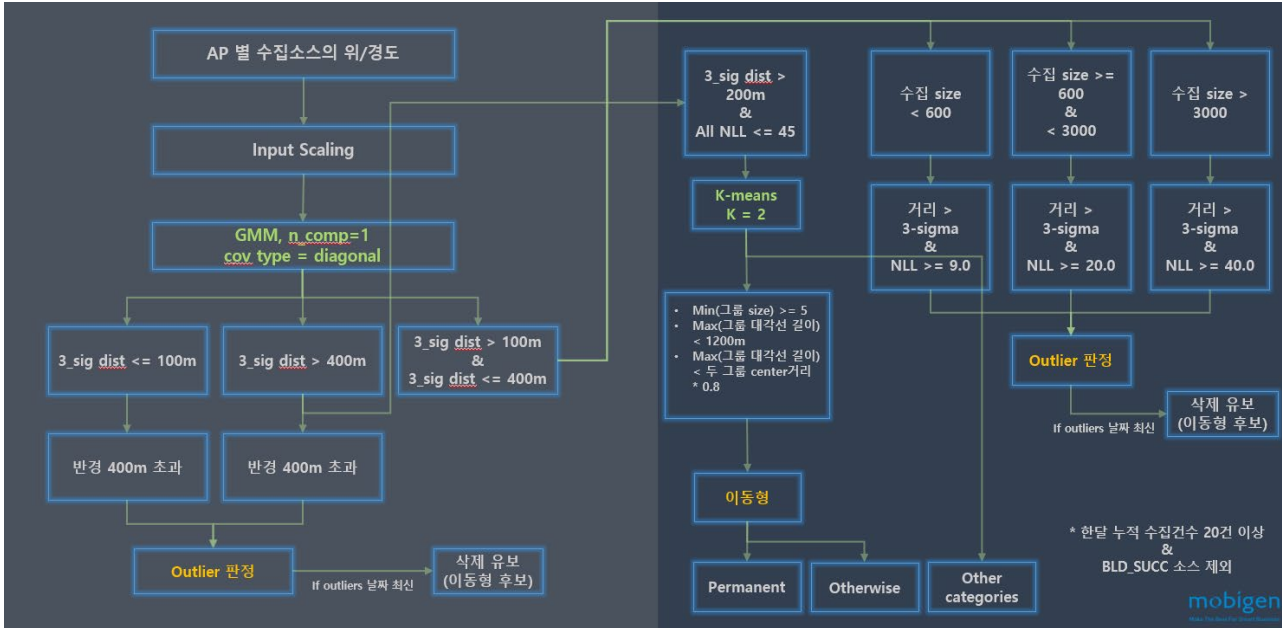


# WiFi-AP 측위 보정 (2/2)

데이터 분석을 통한 이상값 (이동형 AP, Outlier, low-quality data 등)을 정제 하여 적용함으로써 pCell DB 생성의 정확성/신뢰성 향상

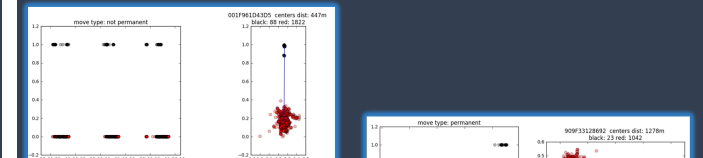
## Final Flow Chart

- Input Data Scaling 및 3-sigma 크기에 따른 Outlier 차별 적용
- 데이터 Scaling을 통한 NLL 기준 차별 적용
- 수집 데이터 Size와 Cluster 그룹의 크기에 따른 판단 기준 적용



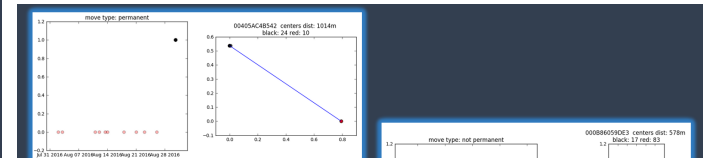
고유 AP 3015개의 sample 데이터에서 총 28건 (약 1%) 검출

알망의 이동형 조건에 걸리지 않는 AP에 대해서는 GMM의 outlier detection 시에 검출되는 outlier의 시간 정보를 조사하여 전부 최선일 시, 삭제를 하지 않고 이동형 후보로 지정



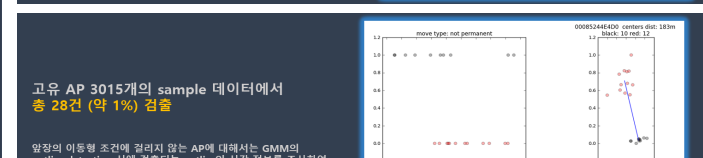
고유 AP 3015개의 sample 데이터에서 총 28건 (약 1%) 검출

알망의 이동형 조건에 걸리지 않는 AP에 대해서는 GMM의 outlier detection 시에 검출되는 outlier의 시간 정보를 조사하여 전부 최선일 시, 삭제를 하지 않고 이동형 후보로 지정



고유 AP 3015개의 sample 데이터에서 총 28건 (약 1%) 검출

알망의 이동형 조건에 걸리지 않는 AP에 대해서는 GMM의 outlier detection 시에 검출되는 outlier의 시간 정보를 조사하여 전부 최선일 시, 삭제를 하지 않고 이동형 후보로 지정



고유 AP 3015개의 sample 데이터에서 총 28건 (약 1%) 검출

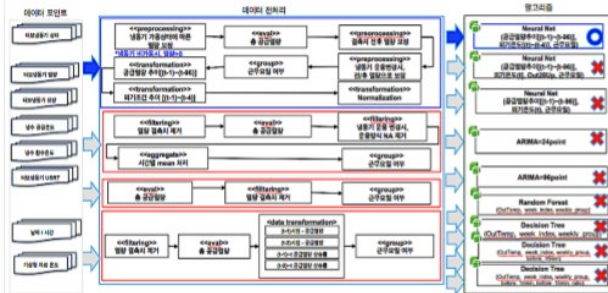
알망의 이동형 조건에 걸리지 않는 AP에 대해서는 GMM의 outlier detection 시에 검출되는 outlier의 시간 정보를 조사하여 전부 최선일 시, 삭제를 하지 않고 이동형 후보로 지정

# 데이터 센터, 빌딩 등의 공조기 운용 최적화

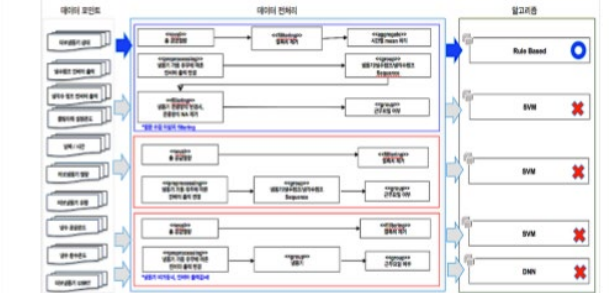
Data Center 전력 및 환경 데이터 분석을 통해 선정된 Item을 기반으로 **분석 모델에 대한 상세 설계 및 구현**을 진행한다

이를 기반으로 분석 결과 검토 및 Review 협의를 통해 분석 모델의 고도화 및 결과에 대한 유효성 검증을 수행하여 에너지 절감 효과를 도출한다.

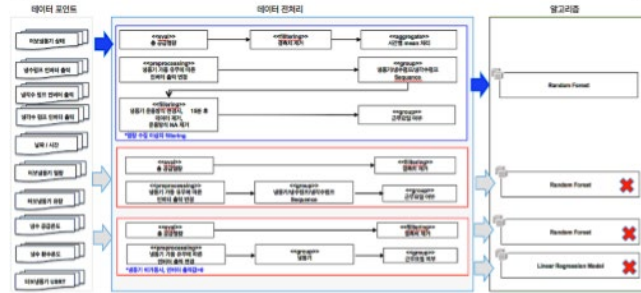
## 공급열량 예측 모델



## 운영패턴 예측 모델



## 전력 사용량 예측 모델



• **공급열량 예측 모델**  
Deep Neural Net (DNN) / Arima / Random Forest / Decision Tree

• **전력사용량 예측 모델**  
Linear Regression / Random Forest

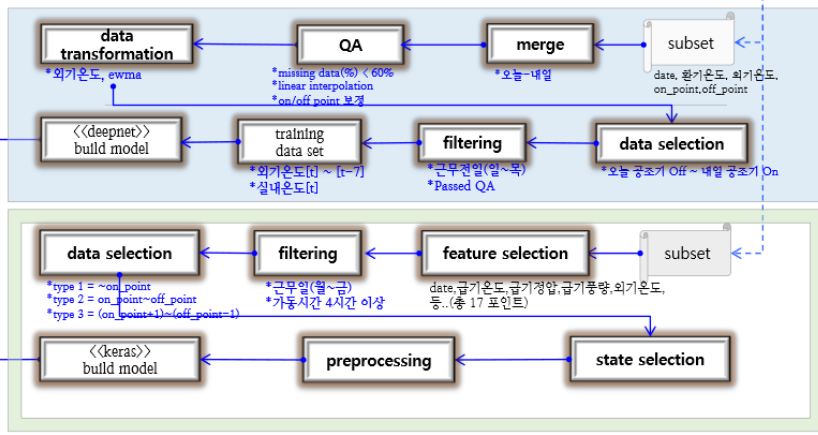
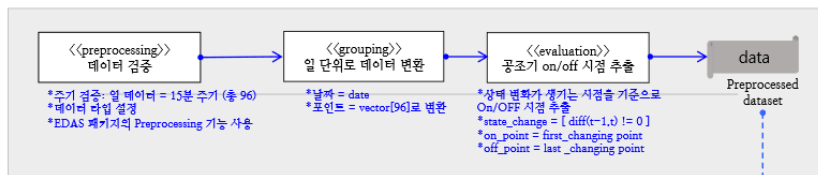
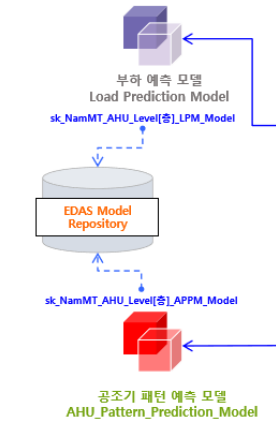
• **운영패턴 예측 모델**  
Rule Base / SVM / DNN

• **Simulator**  
Linear search

• **외기냉방**  
Rule Base

• **검증**: CVRMSE, -최대예측오차, 주간/야간

ML 기반 모델은 기본적으로 미래의 부하를 예측, HVAC 구성 장비(예 냉동기)의 과거 패턴(부하처리 방식)을 기반으로 장비의 패턴을 예측, 그리고 주어진 환경 (예: 운용지침)을 기반으로 최적점을 찾아내는 시뮬레이터로 구성





A New Experience  
In Big Data Analytics

감사합니다.

