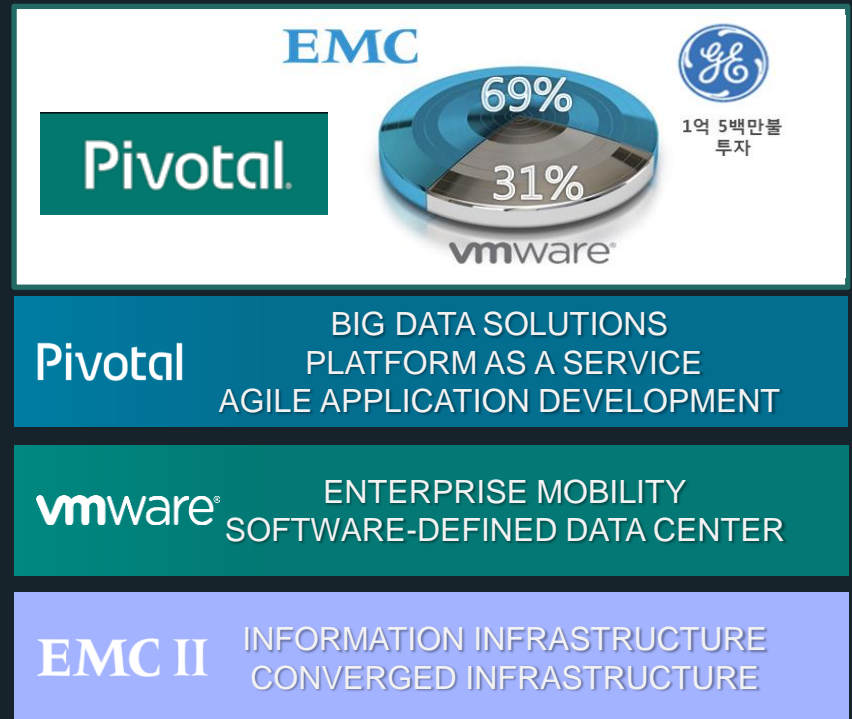


Agenda

- Pivotal 회사 소개
- Pivotal 레퍼런스
- 제조 사례 : General Electronic
- 유통 사례 : 월마트
- IT 사례 : CoreLogic
- 금융 사례 : China CITIC Bank
- 빅데이터 구축 시 고려사항

Pivotal 회사소개

- Pivotal : 2013년 4월 EMC, VMware, GE 의
합작으로 설립된 합작회사
- **Vision** : 새로운 시대에 새로운 플랫폼의
기치로 빅데이터, **Paas, Data Science,**
Agile Development 을 제공
- 주요 솔루션 : Pivotal Greenplum, Pivotal,
HD, Gemfire, Spring, Pivotal Cloud
Foundry, RabbitMQ, Redis 제품을 포함한
Pivotal Big Data suite가 있음.
- 포춘지 2000 기업 중 1000개 이상의
고객을 보유하고 있음.



Pivotal 해외 레퍼런스

Internet of Things



PHILIPS

SIEMENS



MONSANTO



Insurance

Allstate



Liberty Mutual
INSURANCE



Telecommunications



Financials

DISCOVER



Automotive / Industrial / Manufacture



SAIC MOTOR



TOSHIBA
Leading Innovation >>>

Media



Education



Banking



Retail/Delivery



Government



STATE OF ISRAEL
MINISTRY OF DEFENSE



GOVERNMENT OF DUBAI

Pivotal 국내 레퍼런스

제조

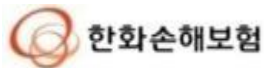


삼성SDI



금융

삼성생명



인터넷

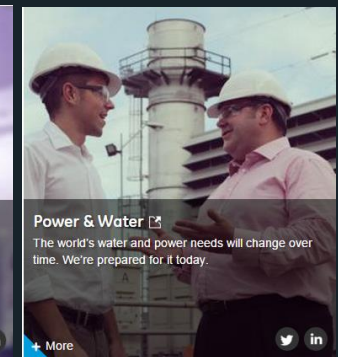
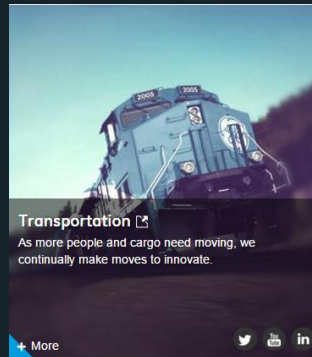
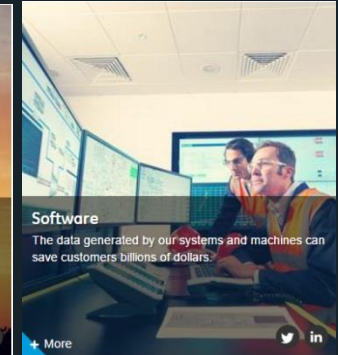


Entertainment



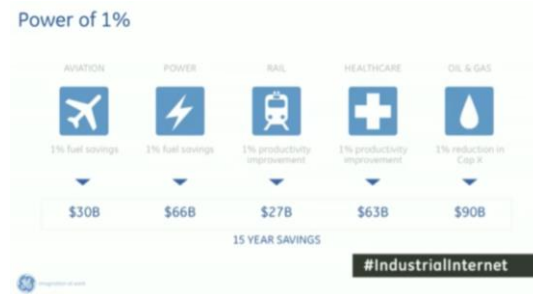
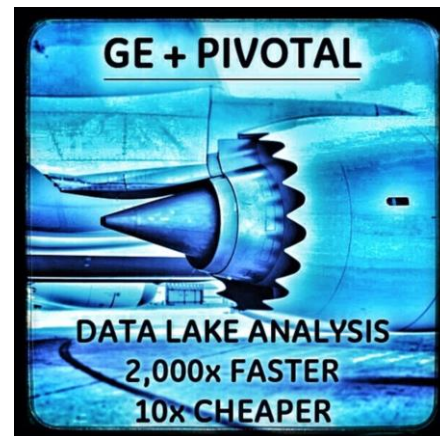
제조 사례 : General Electronic

- GE는 미국의 다국적 대기업 그룹으로 포춘 500대 기업에 선정
- 주요사업 : 항공, 의료, 운송, Capital, 석유 및 가스, 에너지 관리, 발전소, 플랫폼 등
- 오픈 소스의 기술 도입 후 2013년 10월에 산업인터넷 최강자 선언을 하며 산업 인터넷 플랫폼 Predix 공개
- 제트엔진부터 MRI 까지 산업 전 분야에 사용될 산업인터넷 애플리케이션 개발/관리
- 2014년부터 프레딕스플랫폼을 활용할 수 있도록 협력사, 고객사에 기술지원



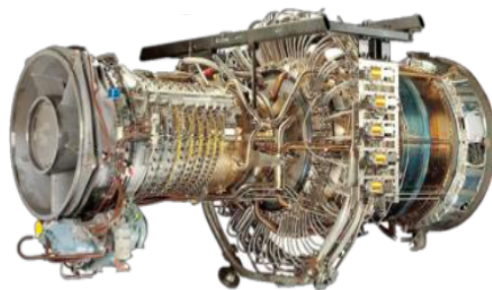
기술의 진화

- Tracked 3 million+ flights
- 340TB as of Aug 2014
- 2000 times faster, 10x cost saving
- Analysis Job : 30 days → 20 min.



<http://www.gereports.com/post/94170227900/angling-in-the-data-lake-ge-and-pivotal-pioneer>

GE 터빈



PHYSICAL

Multiple materials in complex systems... hundreds of patents in each product

Operating with high efficiency & durability in extreme conditions

+

ANALYTICAL

1 gas turbine compressor blade monitoring potential: 500 gigabytes per day



Key messages

- 1 Combination of physical & analytical drives **next evolution of services**
- 2 Customers will **value data** because it improves their **efficiency**
- 3 **GE is well positioned to lead** this wave of change
- 4 **Investors will benefit** through services growth & margins + product share + IT-driven efficiency



GE Transportation



SERVICE ROBOTICS



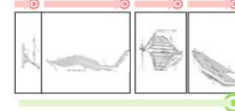
- Productivity, safety**
- ✓ Remote monitoring
- ✓ Tough environment

SENSORS/ANALYTICS



- On-site oil analysis**
- ✓ Predict remaining life
- ✓ Prevent road failures

AUTOMATED DECISIONING



- Optimize yard ops**
- ✓ Reduce dwell 10%
- ✓ Depart on time up 50%



**Impact: for GE ... \$/IB, productivity;
for customers ... velocity, productivity**

GE Aviation

“Analytics”

“Physics”



1,000+ software engineers
GE Software COE

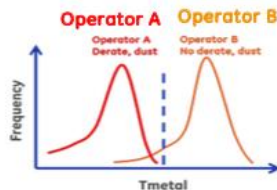


6,300 applied engineers
GE Aviation



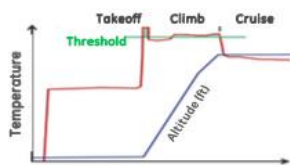
Data

Secure,
continuous,
accessible



Statistics

Unlock trends,
segment



Physics

Apply engine
expertise



**Customer and
GE benefits**

- ↑ Time-on-wing
- ↑ Reliability and availability
- ↓ Fuel cost
- ↓ Repair and overhaul

Pivotal™

GE 풍력 발전

초당 10만개 이상 센서 이벤트 처리

고속 데이터 수집 및 동시에 실시간 및 배치 분석



초당 10만개 이상
센서 데이터 전송



Fast Data

비동기 방식 데이터 기록



In-Memory
DBMS



Big Data

SQL on
Hadoop

리포팅

심층 분석

빌링

운영데이터웨어하우징

- 밀리세컨드 또는 그 이하의 응답성 보장
- 실시간 분산 이벤트 처리
- 고성능이 보장된 **In-memory** 가용성

GE Oil & gas

GE O&G의
년 6~8% 지속적인
성장으로
설비 증설 시 Downtime
발생 등으로 난제에
직면 함



GE Software, Data,
Analytics 으로 난관을
극복하여, 계획되지
않은 다운 시간을
없앴을 있었음.

\$800B+

Annual industry CAPEX
(6-8% growth per year)



\$150MM

Average annual cost for
mid-size LNG facility due to
unplanned downtime



5-10%

Range of O&G industry's
estimated downtime



Customer challenges



Increasing complex operations



Tighter regulations



Skills gap



Machines with data potential

Equipment
Availability





Equipment
Reliability

Operational
Efficiency

Production
Optimization

Industrial Internet ... turning data into actionable information
Enable no unplanned downtime

GE Transforming Services

80's	<p>Big iron</p>  <ul style="list-style-type: none"> Partners & field technicians
90's	<p>Repair</p>  <ul style="list-style-type: none"> Science applied to repairs & upgrades
00's	<p>CSA</p>  <ul style="list-style-type: none"> Risk sharing service contract
Today	<p>Analytics</p>  <ul style="list-style-type: none"> Asset & operations optimization

Driving customer outcomes

	<u>Power of 1%</u>	<u>Annual impact^{-a)}</u>
Airlines	1% fuel	\$2-3B
Utility	1% fuel	\$4-5B
Oil & Gas	1% uptime	\$5-7B
Healthcare	1% productivity	\$4-5B
Transportation	1 mile velocity	\$1-2B

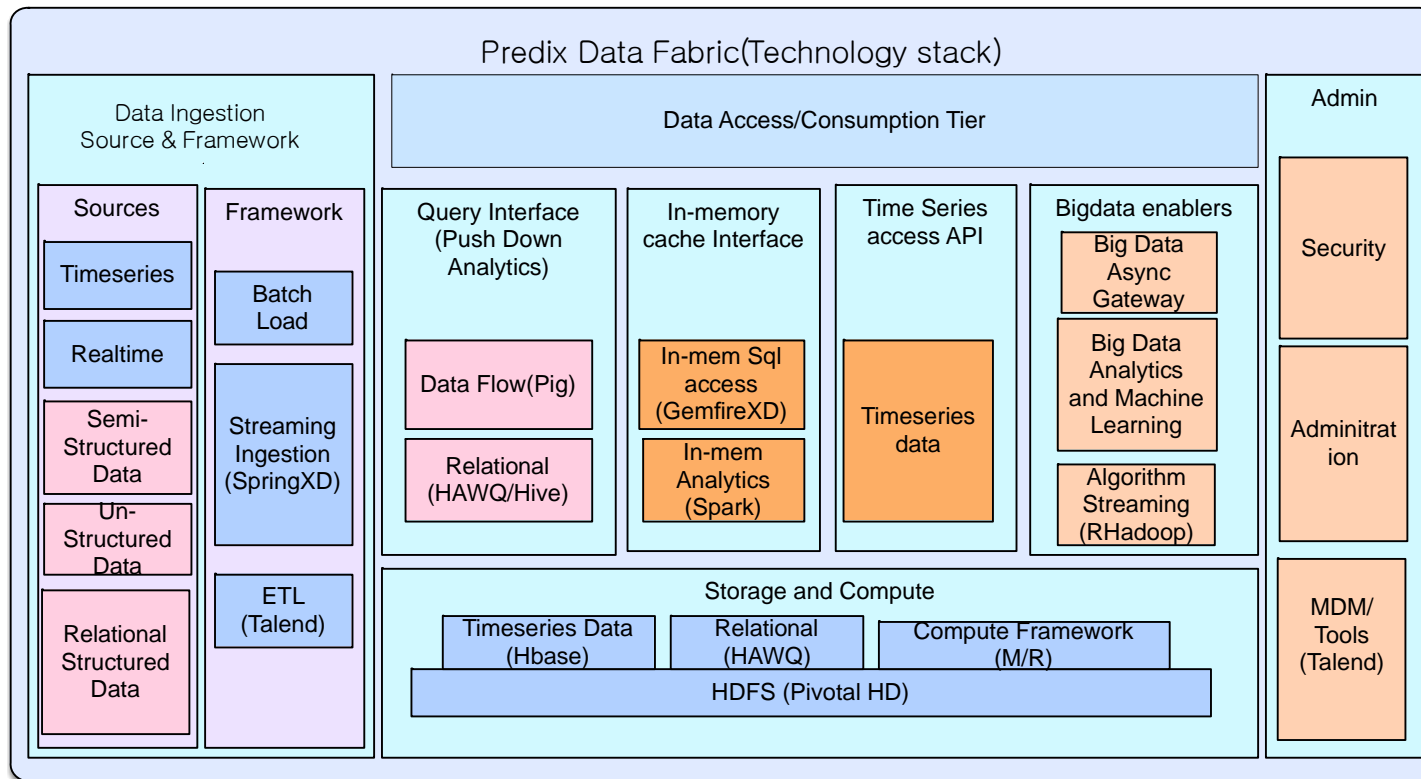
Small changes in ... fuel, maintenance, uptime, productivity, flexibility, safety, labor effectiveness ... have a big impact on profitability

GE Predix Data Fabric

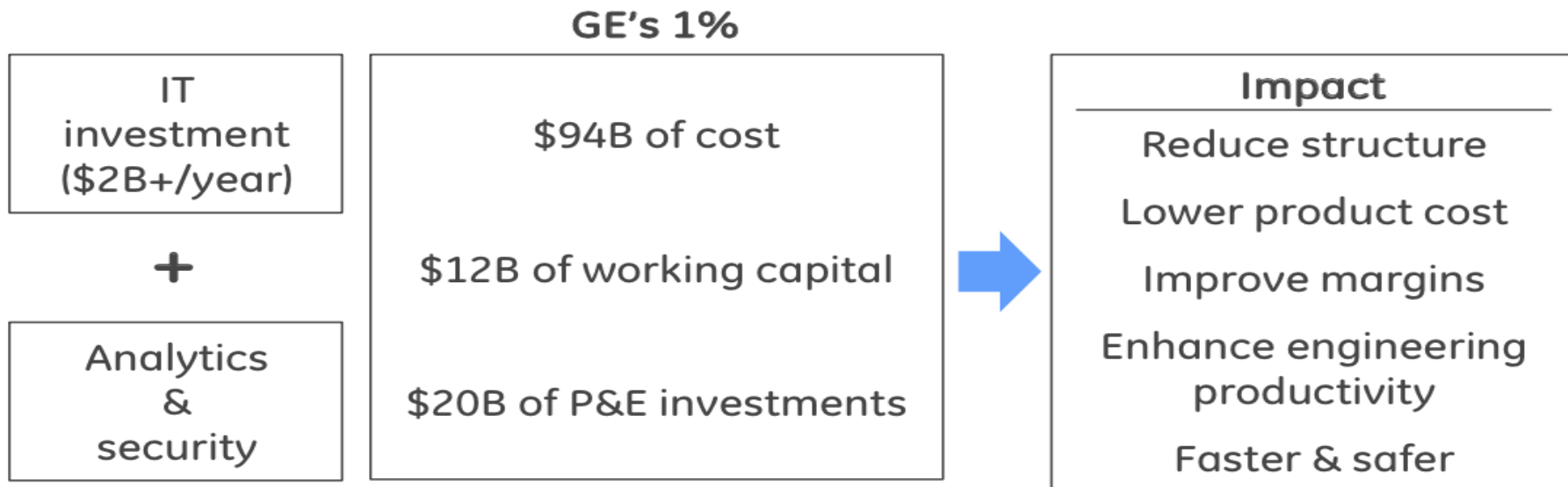
Available

Roadmap

Future



GE의 1% 효과



Every Industrial company → a Software & Analytical company

유통 사례: 월마트

- 미국에서 가장 큰 소매업체
- 주요사업 : 월마트 스토어, 샘스클럽, 월마트 슈퍼센터스토어, 하이퍼마켓, 버드즈웨어하우스 대리점을 포함한 할인 매장을 운영한다.
- 2015년 현재 미국의 50개 주와 전세계 27개국에 걸쳐 있으며, 미국 내 점포수는 월마트 스토어스가 11,000개, 샘즈클럽이 621개, 월마트 인터내셔널이 6242개에 달함.



[Walmart U.S.](#)

Walmart has stores in 50 states and Puerto Rico offering low prices on the broadest assortment of products through a variety of formats.

[Walmart International](#)

Walmart operates in 26 countries outside the U.S., offering low prices through retail units under several banners around the globe.

[Sam's Club](#)

Sam's Club is a warehouse membership format with locations across the U.S. and around the world.



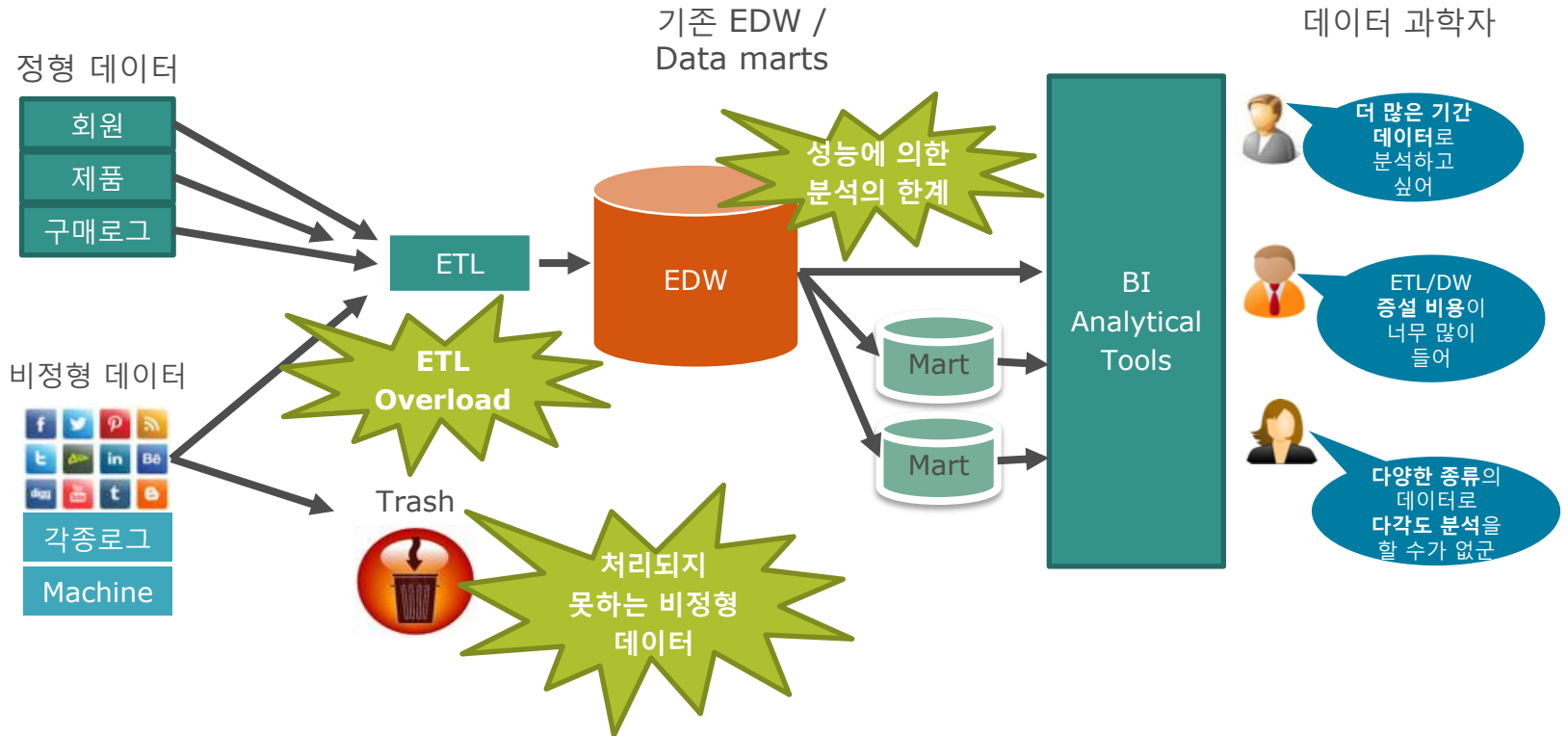
[Walmart Global eCommerce](#)

Based in California's Silicon Valley, Walmart Global eCommerce drives our online, mobile and social innovation, including Walmart.com and @WalmartLabs.

[Interactive Map](#)

Where in the world is Walmart? See where we operate and get key data, such as the number of associates and community giving totals by state.

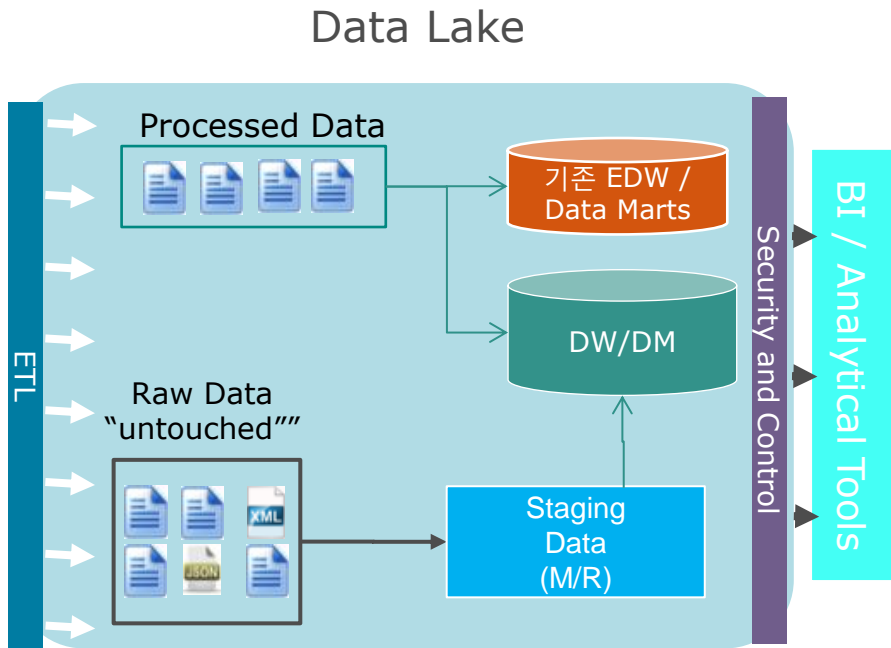
월마트



전통적 데이터
Sources



새로운 데이터
Sources/Form



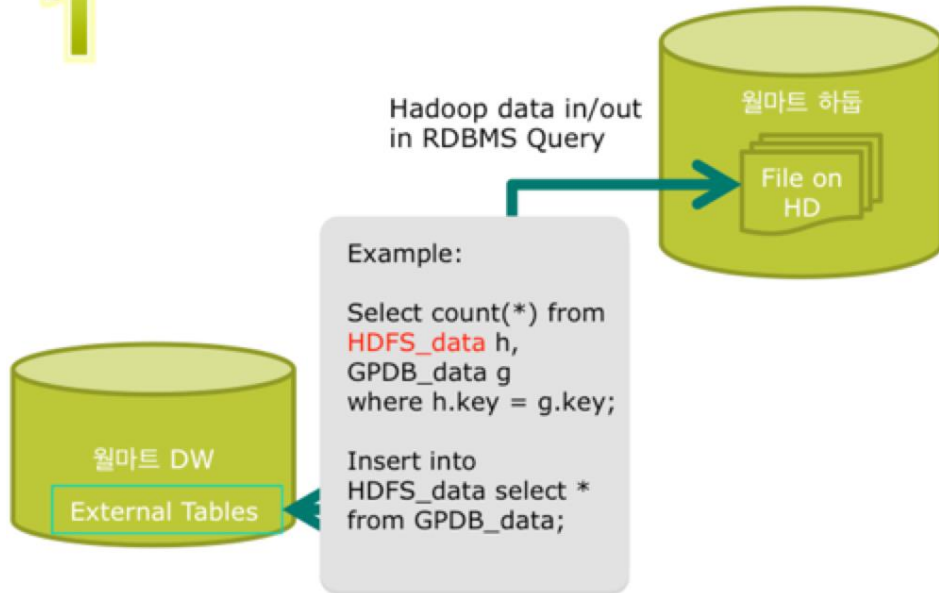
현업 사용자



모든 정형/비정형 데이터를 하둡 HDFS에 저장 및 처리
높은 확장성, 높은 신뢰성, 비용 경쟁력 재고

월마트 : Hadoop과 DB 연동

1

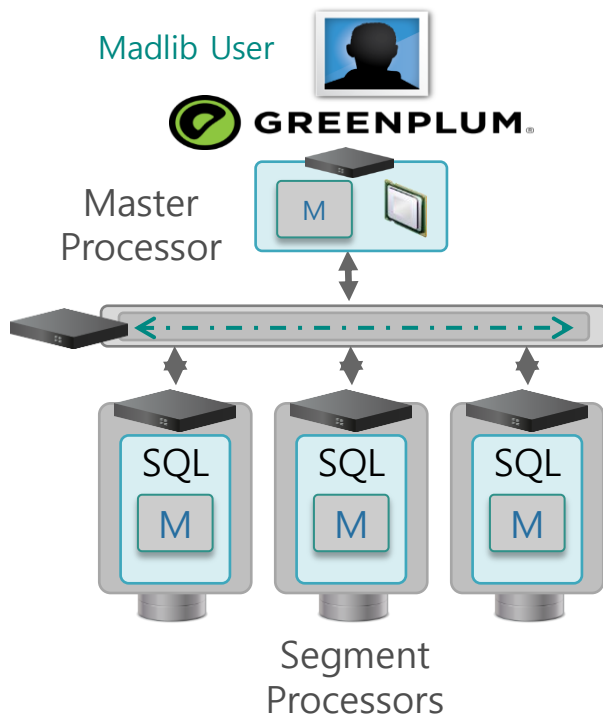


2

Faster, Easier
With SQL Interface

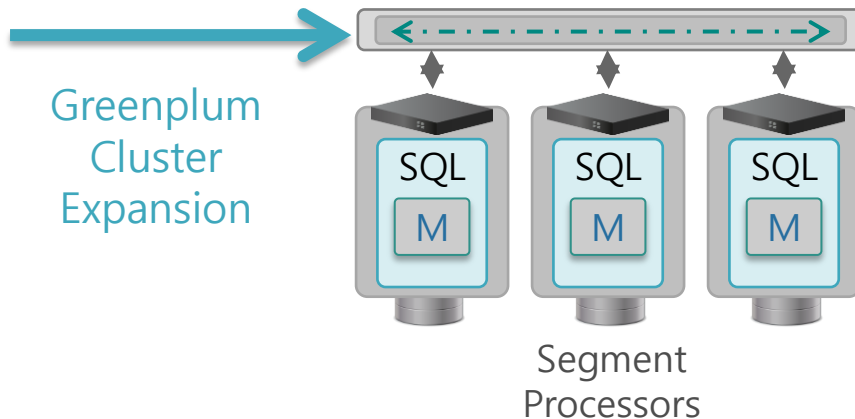


월마트 : 고급분석을 위한 Madlib 활용



MADlib Advantages

- Algorithms designed to leverage partitioned data set
- Algorithms designed to leverage Greenplum Database's scale-out architecture



월마트 : Data Lake 및 분석 고도화 효과

대용량 병렬처리 시스템으로 Madlib를 활용한 멤버십 분석

- 126,203번의 선형 회귀분석 반복
- 6,380번의 로지스틱 회귀분석 반복
- 최종 모델링 단계 : 탄력적 순-회귀분석



680만 회원으로부터 도출된 385,000 가지의 분류

- 3TB 데이터셋으로 부터 조 단위의 행 처리를 통한 프로세싱과 분류 정리 작업



IT 사례: CoreLogic (정보서비스 제공업체)



부동산, 금융산업, 소비자에서 발생하는 거래정보를 저장, 분석하여 기업과 정부 기관에 제공하는 미국 소재의 법인

- 40년간의 축적된 데이터 및 거래 정보를 기반으로 하여 부동산, 소비자 신용, 모기지, 주택, 사기, 자연 재해, 재해 분석 및 예측
- 공공 및 민간 부분에서 4만 이상의 고객

Independent Audits & Reviews

A new way to manage pre-funding and post closing quality control >



Qualified Mortgage & Ability to Repay

Solutions to meet your QM & ATR challenges >



Mortgage Origination Solutions

Solutions to solve your loan origination challenges >



Home Equity Lending Solutions

Are you prepared for the resurgence? >



Flash Flood Risk Score

Account for flash flood risk beyond riverine flooding >



Capital Markets Solutions

Expanded Loan Level Data, predictive analytics, expert advisory services >



CoreLogic IT Landscape

540+ Products and 5000+ applications, components, services, and tools identified across CoreLogic

Business Criticality

80% of Applications are Mission Critical or Important

Users & Use Cases

2M+ professional users
Sub-second to multi-day transactions

Life Cycle Stage

48% of Applications in Maintain or in Maintain w/enhancements stage

Technology

63% of Applications utilize JAVA (42%) or .NET (21%)

330 - Applications



3681 - Components



1009 - Data Stores



832 - Other

- 532 - Tools
- 300 - Externals

Technologies

Mainframe

IIS7



CoreLogic 빅데이터 아키텍처

다양한 방법의 최종사용자 Access



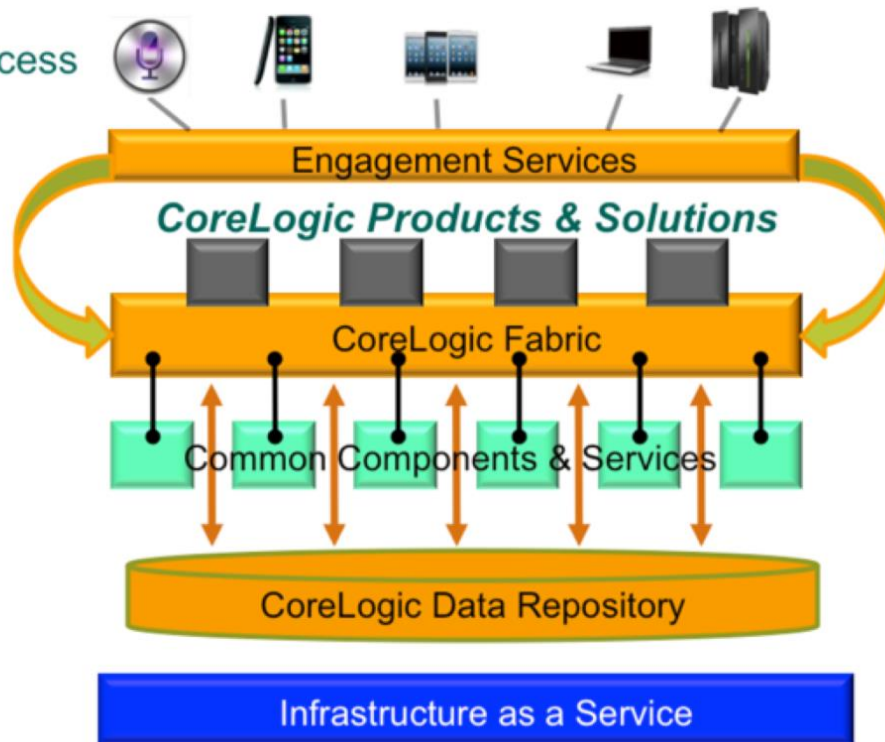
Single Sign-On

솔루션 모듈화
(통합 또는 분리)

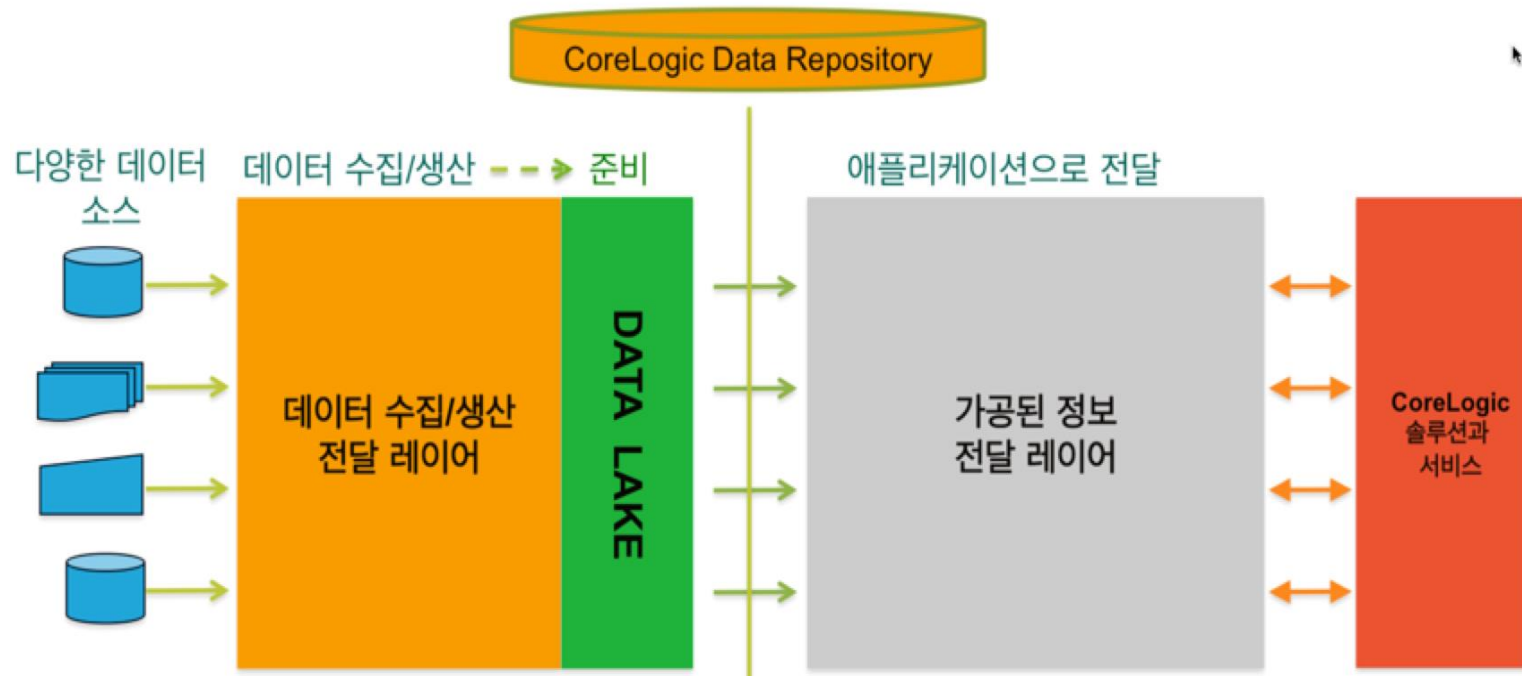
재활용 가능한 서비스

데이터 및 분석 제공

확장성, 유연성, 효율성 제공



CoreLogic 빅데이터 흐름



CoreLogic

파트너십을 통한 최신 기술 수용

PaaS + Data Lake + 개발협업 + 데이터 과학

PaaS



- 오픈소스기반 서비스 모델
- 다양한 IaaS 지원
- 시장 지배적인 PaaS 채택
- Private PaaS 도입



개발협업

LABS



- Extreme Agile 활용
- Pair Programming 방법
- Test Driven Development
- 풍부한 경험 활용

Data Lake

- 최신 기술의 데이터 처리 및 분석 도구 도입
- 강력한 데이터과학자팀 지원

CoreLogic & Pivotal Collaboration



Pivotal™

금융 사례: China CITIC Bank



- 중국 CITIC 은행은 중국의 개혁과정과 자금 조달에 참여한 최초의 상업 은행
- 본사 : 베이징, 중국 78지점과 사업 영역이 전세계 130개국
- 중국 신용 붐과 소비자 대출의 대규모 증가로 신용카드 시장에서 최근 강자로 매김, 신용카드 1억 4,000만 건 발급



이슈 사항

- 1 부정 카드 사용 등의 미션 크리티컬 분석 시 신속성 요구
- 2 시장 점유율 높이기 위한 전략으로 고수준의 타겟 마케팅
- 3 개별 분석 시스템의 통합 요구 사항 발생 (중앙집중, 확장성 필요)
- 4 빠른 모델 검증 및 반영을 위한 MPP 아키텍처의 DBMS 필요

Pivotal Greenplum 도입 효과

1 전화 판매 센터의 생산성 향상

- 신용카드 판매 센터의 아웃 바운드 마케팅
- 적재, 쿼리 병렬 개선으로 부서 사용자의 생산성, 수익 증가

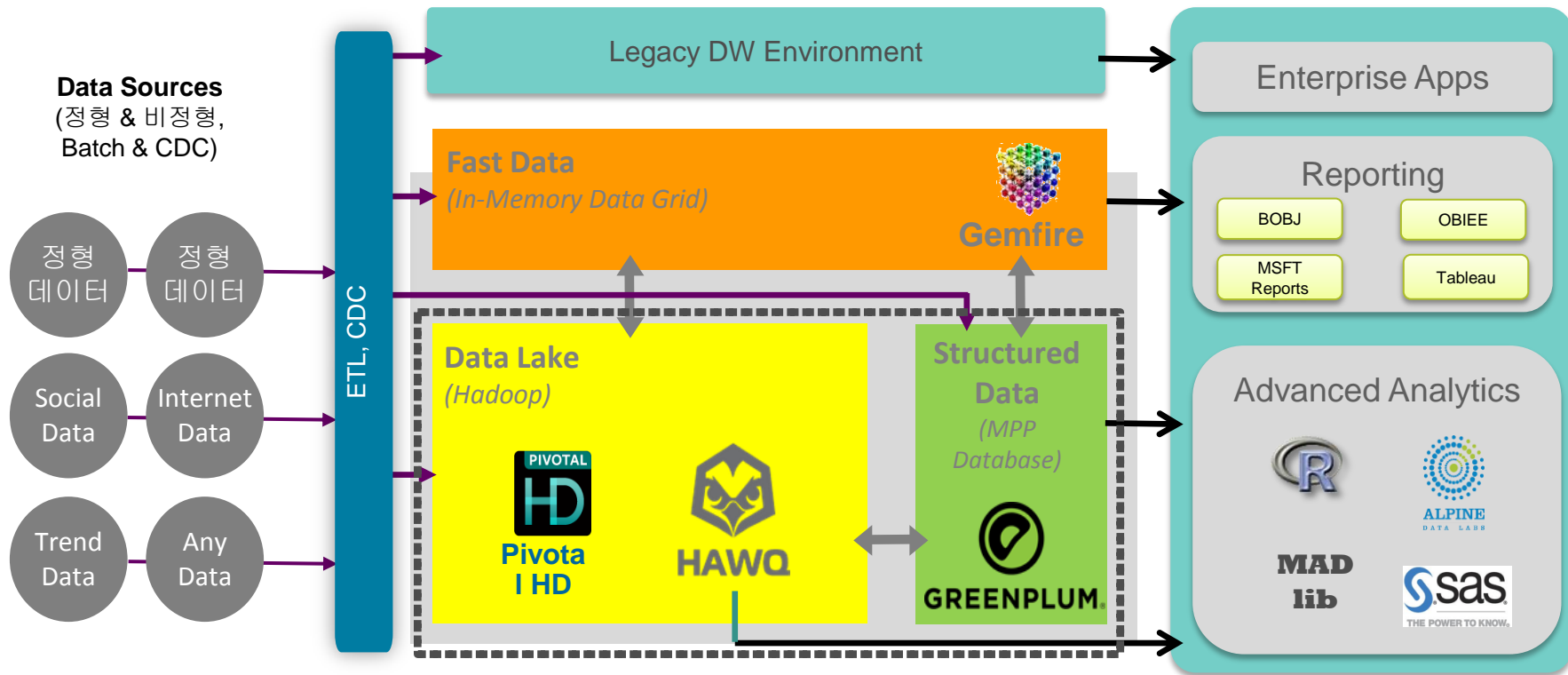
2 최적화된 마케팅 캠페인 및 가속화된 Data 분석

- 1286 마케팅 캠페인을 위한 배치 구성 시간을 86 % 줄임, (마케팅 캠페인 구성 시간이 2주에서 2~3일로 단축)
- 새로운 모델 개발 등의 분석 프로세스를 2~3개월에서 1개월 미만으로 단축

3 위험 감소

- 실시간 고객 데이터 분석으로 구매 행동 분석 후 개별 고객의 신용 한도 즉시 조정

CITIC Bank: Data Lake



Big Data 구축 시 고려사항

기술적 관점

- 최근 트렌드를 반영하는 기술 접목
- In-Memory 기반 실시간 분석
- 병렬처리 기반 대용량 분석
- BI/DW 포함하는 전체 아키텍처 관점
- 빅데이터 확장 가능한 인프라 구성
- 하둡 기반의 다양한 솔루션 지원 (SQL on Hadoop)
- 프로젝트 상황에 맞는 유연한 인프라 구성
- Cloud 환경에 완벽한 빅데이터 솔루션 제공

비즈니스 관점

- 분석 및 예측이 가능한 분석 고도화
- 기업 내부에 데이터 과학자 양성
- 파트너십을 통한 최신 기술 수렴
 - Data Lake, 데이터 과학, 개발협업
- Vendor Lock 방지를 위한 Open Source 기반의 솔루션
- 엔터프라이즈 운영에 필요한 기술 지원

추진
전략

Big Data 분석을 위해 대규모의 프로젝트보다는

1. 소규모 프로젝트로 시작하여 데이터 과학자와 협업
2. 분석 고도화 효과 검증 후에 단계적으로 규모 확대

Pivotal Big Data Suite

애자일한 통합 빅데이터 플랫폼

Advanced Analytics

최고의 빅데이터 분석툴

Pivotal
Greenplum
Database

Pivotal
HAWQ

Apps at Scale

확장성있는 어플리케이션
개발

Redis

Rabbit MQ

Pivotal BDS on PCF
클라우드
+ 빅데이터

Pivotal Cloud Foundry
INCLUDED

Data Processing

Spring XD

Spark

리얼타임 데이터 스트리밍
처리

Pivotal HD & Open Data Platform

Commodity Hardware

어떠한 클라우드, 인프라에서라도
지원

Pivotal Cloud Foundry

Pivotal™

Pivotal

A NEW PLATFORM FOR A NEW ERA