



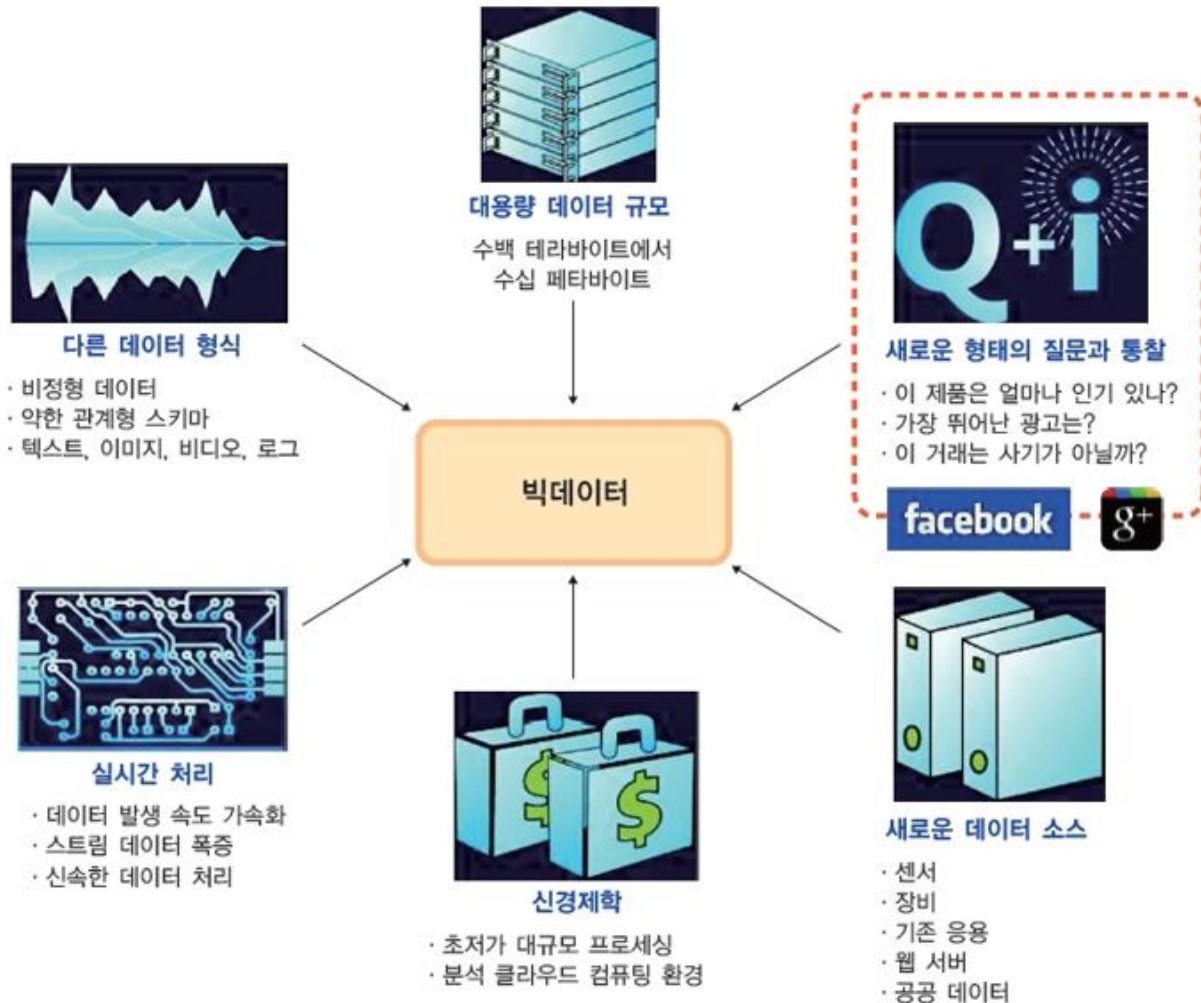
# 클라우드와 AI

무단배포는 금지(본자료는 학습용으로제작)

# 클라우드 컴퓨팅 기술



## 빅데이터(Big Data) 속성과 처리 특징



# 클라우드 컴퓨팅 기술



## 빅데이터(Big Data) 처리 과정과 기술



# 클라우드 컴퓨팅 기술



## • 빅데이터(Big Data) 처리 과정별 기술 영역

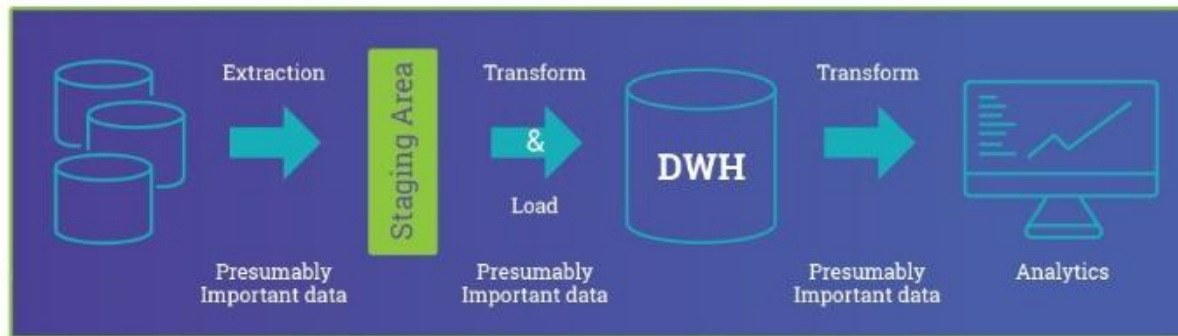
| 과정 | 영역                                      | 개요  |
|----|---|---|
| 생성 | 내부 데이터                                  | 데이터베이스 Database, 파일 관리 시스템 File Management System |
|    | 외부 데이터                                  | 인터넷으로 연결된 파일, 멀티미디어, 스트림                          |
| 수집 | 크롤링 Crawling                            | 검색 엔진의 로봇을 사용한 데이터 수집                             |
|    | ETL Extraction, Transformation, Loading | 소스 데이터의 추출 · 전송 · 변환 · 적재                         |
| 저장 | NoSQL 데이터베이스                            | 비정형 데이터 관리  |
|    | 스토리지 Storage                            | 빅데이터 저장   |
|    | 서버 Server                               | 초경량 서버  |
| 처리 | 맵리듀스 MapReduce                          | 데이터 추출  |
|    | 프로세싱 Processing                         | 다중 업무 처리  |
| 분석 | NLP Neuro Linguistic Programming        | 자연어 처리  |
|    | 기계 학습 Machine Learning                  | 기계 학습으로 데이터의 패턴 발견                                |
|    | 직렬화 Serialization                       | 데이터 간의 순서화  |
| 표현 | 가시화 Visualization                       | 데이터를 도표나 그래픽적으로 표현                                |
|    | 획득 Acquisition                          | 데이터의 획득 및 재해석                                     |

# 클라우드 컴퓨팅 기술



- ETL vs ELT

## ETL



출처 : <https://blog.panoply.io/etl-vs-elt-the-difference-is-in-the-how>

## ELT



출처 : <https://blog.panoply.io/etl-vs-elt-the-difference-is-in-the-how>

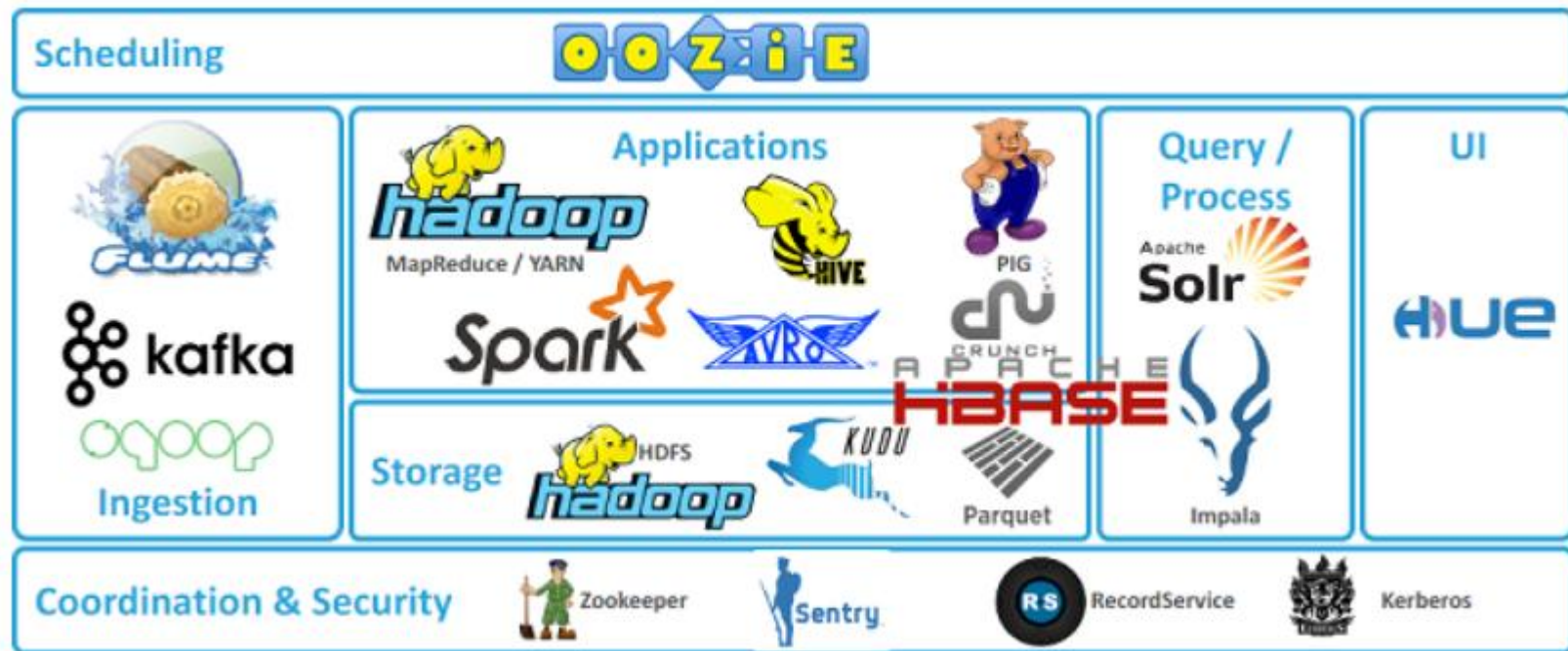


# 클라우드 컴퓨팅 기술



## • Hadoop(하둡)

- 여러 대의 컴퓨터 클러스터에서 대규모 데이터 세트를 분산처리해주는 프레임워크
  - 단일 서버에서 수천 대의 머신으로 확장할 수 있도록 설계
  - 하둡 분산처리시스템(HDFS)과 맵리듀스(MapReduce) 프레임워크로 시작
  - 여러 데이터 저장, 실행엔진, 프로그래밍 및 데이터 처리 같은 하둡 생태계 전반으로 포함하는 의미로 확장



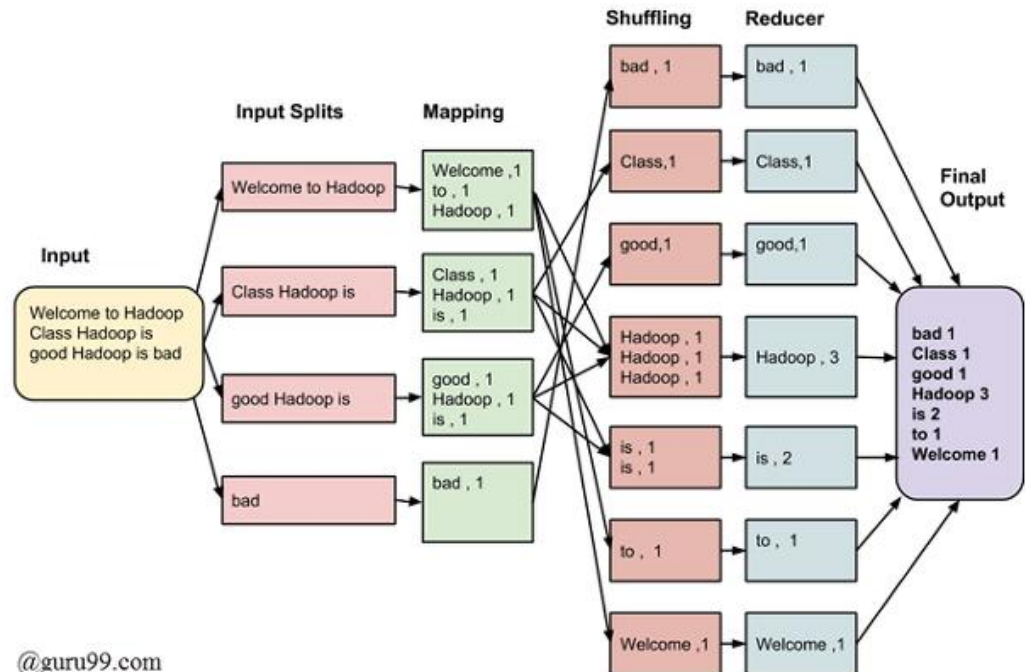
# 클라우드 컴퓨팅 기술



## • 맵리듀스(MapReduce)

### ■ 구글에서 대용량 데이터 처리를 분산 병렬 컴퓨팅에서 처리하기 위한 목적으로 제작한 소프트웨어 프레임워크 (2004년 발표)

- 맵리듀스는 Input, Splitting, Mapping, Shuffling, Reducing, Final Result 과정으로 진행
- 장점 : 단순하고 사용이 편리, 유연, 저장 구조와 독립적, 확장성이 높음
- 단점 : 복잡한 연산이 어렵고 기존 DBMS가 제공하는 스키마, 질의 언어, 인덱스 등의 기능을 미지원함  
상대적으로 성능이 낮음



# 클라우드 컴퓨팅 동향



## • 국내 사업자 서비스 동향

자료: SW · 컴퓨팅 클라우드컴퓨팅, TTA, 2018

| 사업자         | 주요 현황  |
|-------------|--|
| KT          | <ul style="list-style-type: none"><li>- 산업별 특화 솔루션사와의 제휴, 서비스 결합을 통해 산업별로 클라우드 서비스를 제공하기 위한 체계 구축</li><li>- 이를 위해 티맥스소프트, 한컴, 비트컴퓨터, 이베스트, بانک웨어글로벌, 핸디소프트, SAS코리아 등과 협력 중</li><li>- KT의 인프라에 각 사의 SW를 클라우드 서비스 형태로 제공 중</li><li>- KT그룹 KT DS는 오픈소스 컴플라이언스 관리 솔루션 '오빅스 컴파스(OBICs COMPASS)'를 출시</li><li>- 하이브리드 및 멀티 클라우드 구현을 위한 'Vmware on KT Cloud' 출시</li></ul> |
| SK C&C      | <ul style="list-style-type: none"><li>- SW 기업들과의 협력을 강화하고 국내 소프트웨어 기업인 톨론, 새움소프트, 달리웍스 등과 SK C&amp;C의 기업형 클라우드 서비스 '클라우드 제트'에 국산 소프트웨어를 SaaS 형태로 제공</li><li>- '클라우드 제트'를 IBM의 소프트레이어, 알리바바의 클라우드 등 전 세계 40개 이상의 클라우드 데이터센터의 서비스를 자유롭게 제공하여 하이브리드 클라우드 구축이 가능하도록 지원</li><li>- 멀티 클라우드 인프라 배포, 관리 지원</li></ul>   |
| LG CNS      | <ul style="list-style-type: none"><li>- 자사의 스마트팩토리 솔루션에 Microsoft Azure를 도입하여 기존 시스템통합 형태로 제공하던 스마트 팩토리 솔루션을 서비스 상품으로 제공</li><li>- 클라우드 마켓을 제공하는 방식의 협력 모델을 제공하여, 백화점 방식의 클라우드 마켓플레이스인 '매시업플러스'를 선보였으며, SW 업체들이 매시업플러스에 입점해 SaaS를 제공 중</li><li>- 멀티 클라우드 기반 AI 빅데이터 플랫폼 'DAP(Data Analytics &amp; AI Platform '답')'를 출시하고 AI 빅데이터 사업을 확대 중</li></ul>               |
| 코스콤         | <ul style="list-style-type: none"><li>- 정부기관이나 기업 등에서 R&amp;D 사업을 쉽게 추진할 수 있도록 서비스형 인프라 (IaaS)와 서비스형 플랫폼(PaaS)을 합친 통합 플랫폼 개발환경인 '클라우드 R&amp;D 존(Zone)'의 상용서비스를 개시</li><li>- '케이(K) 파스-타'를 통해 서버, 스토리지 같은 인프라부터 운영체제(OS), 미들웨어, 프레임워크 같은 플랫폼까지 전체적인 개발, 테스트, 운용 환경을 제공 중</li></ul>  |
| 네이버비즈니스 플랫폼 | <ul style="list-style-type: none"><li>- 자사의 클라우드 서비스인 '네이버 클라우드 플랫폼'에 네이버 API 중심의 신규</li><li>- 상품 14개(네이버 클로바, 파파고 등 인공지능 API 포함)를 추가하여 제공 중</li></ul>   |



# 클라우드 컴퓨팅 동향



## • 국외 사업자 서비스 동향

| 사업자            | 주요 현황  |
|----------------|--|
| IBM            | <ul style="list-style-type: none"><li>- 2017년 4월, 미국 내 데이터센터 4개(텍사스 댈러스, 워싱턴 DC 등)를 추가로 건설하여 상반기 가동</li><li>- IBM 클라우드의 테슬라 P100 기반 서비스 출시</li></ul>   |
| Amazon         | <ul style="list-style-type: none"><li>- 2017년 10월, AWS의 가격 책정 방식을 변화하여 EC2 서비스를 초(秒) 단위로 과금</li><li>- 2017년 3월, 머신러닝을 통한 계산대 없는 식료품점인 'Amazon고(Amazon Go)'의 시범 개점 시작</li><li>- Amazon AI 서비스(Amazon 렉스, Amazon 폴리, Amazon 리코그니션) 제공</li><li>- 클라우드 커넥티드 기기 연결 확장을 지원하는 AWS Greengrass 솔루션 서비스 제공</li></ul> |
| Microsoft      | <ul style="list-style-type: none"><li>- 사물인터넷과 머신러닝, 빅데이터 등을 한번에 제공하는 SaaS형 인텔리전트 IoT 솔루션 'IoT 센트럴' 제공</li><li>- 애저 쿠버네티스 서비스(Azure Kubernetes Services)를 통해 Docker를 구현한 컨테이너(Container) 엔진과 컨테이너 관리를 위한 쿠버네티스 기반 컨테이너 오케스트레이션 등의 도구들을 제공</li></ul>  |
| Salesforce.com | <ul style="list-style-type: none"><li>- 2017년 4월, 일본에 두 번째 데이터 센터(고베 지역)을 열고 아시아 태평양 지역에 클라우드 서비스 제공</li><li>- 2017년 3월, IBM과 글로벌 전략 파트너십을 맺고 인공지능을 통한 조인트 솔루션(joint solution)'을 허용</li><li>- 고객관계관리(CRM) 도구뿐만 아니라 자사의 전체 서비스에 AI를 확장하기 위해 자사의 서비스 클라우드 고객 플랫폼에 '아인스타인 인공지능'을 추가</li></ul>                 |
| Google         | <ul style="list-style-type: none"><li>- 클라우드 부문의 비즈니스 강화를 위해 의료, 금융, 소매 등의 사업에서 전문화된 인력을 고용하고 SAP SE 등 소프트웨어 업체들과의 제휴를 추진하는 등의 사업 확대를 진행 중</li><li>- 2017년, 최대 데이터 과학자 연결 플랫폼인 '캐글(Kaggle)'을 인수하여 인공지능(AI) 기술 발전관련 계획 발표</li></ul>   |

# 클라우드 컴퓨팅 동향



## • 제조업에서의 클라우드

### 기존의 클라우드

- 제조업에서는 독립 네트워크 내에서 데이터를 수집 · 저장 · 분석 등 활용하기 때문에 클라우드가 필요하지 않음
- 제조업에서는 클라우드가 외부에 데이터가 공유된다는 측면에서 부정적임

### 기존의 클라우드

- 제조업에서도 수집된 데이터를 유기적으로 통합할 수 있는 기술이 요구됨
- IT 자원의 공동 사용으로 새로운 비즈니스 창출, 비용 절감 등을 위한 기술이 필요함



제조업에서 기존의 부정적인 인식이 축소되며, 새로운 가치를 창출하기 위하여 클라우드 컴퓨팅 기술이 활성화 되고 있음

데이터 분석, 인공지능 기술 적용에 유리

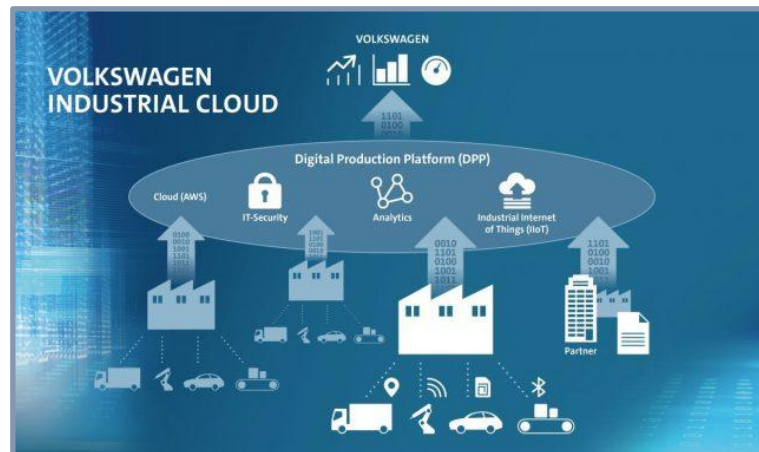
설비 유지보수 등의 비용 절감 효과 기대

# 클라우드 컴퓨팅 동향

## 클라우드 구축 사례

### ■ 폭스바겐 산업용 클라우드(Volkswagen Industrial Cloud)

- 아마존 웹 서비스(AWS)와의 클라우드 공동 개발
- 폭스바겐 그룹의 122개 시설 내 모든 기계, 공장 및 시스템의 데이터 결합
- 해당 아키텍처는 향후 디지털 생산 플랫폼(Digital Production Platform, DPP)으로 확대 가능
  - ✓ 시스템과 공장 간의 데이터 교환을 표준화하고 단순화 하는 역할
  - ✓ 공정 최적화(자재 흐름의 효율적인 제어, 기계 장비 최적화 작동 등)를 위한 새로운 전망 제시
- 가치 사슬에 있는 회사 통합 및 산업 파트너사 간의 네트워크 구축이 목표
- 공장 내외부의 물품 이동 추적 시스템 등을 시작으로 클라우드 서비스 개발 및 투입 목표





THANK  
YOU